J. Biochem. 2009;145(3)365–375
doi:10.1093/jb/mvn174

Structural Divergence of Cysteine-Rich Secretory Proteins in Snake Venoms†

Yukiko Matsunaga[1], Yasuo Yamazaki[1], Fumiko Hyodo[1], Yusuke Sugiyama[1], Masatoshi Nozaki[2] and Takashi Morita[1,*]

[1]Department of Biochemistry, Meiji Pharmaceutical University, 2-522-1 Noshio, Kiyose, Tokyo 204-8588; and [2]Okinawa Prefectural Institute of Health and Environment, 2003 Ozato, Ozato, Nanjo, Okinawa 901-1202, Japan

Received October 31, 2008; accepted December 9, 2008; published online December 23, 2008

Cysteine-rich secretory proteins (CRISPs) are expressed in spermatocytes and granules of neutrophils in mammals, and are associated with sperm maturation and host defense. Related proteins have recently been recovered in snake venoms, and some of the snake venom-derived CRISPs exhibit ion channel blocking activity. Here we isolated and identified two novel CRISPs (kaouthin-1 and kaouthin-2) from the venom of *Naja kaouthia* (Elapidae), and cloned the encoding cDNAs. Kaouthin-1 and kaouthin-2 were classified into two broad sister groups of Elapidae, the Asian species and the marine/Australian species, respectively. Sequence comparisons reveal that the high-frequency variable regions among snake venom CRISPs define a continuous line on the molecular surface of the N-terminal pathogenesis-related protein-1 (PR-1) domain and the C-terminal cysteine-rich domain (CRD). Snake venom proteins generally display efficient molecular diversity around functionally key regions, suggesting that the PR-1 domain of CRISPs is important for the recognition of target molecules.

Key words: CRISP, *Naja kaouthia*, pathogenesis-related protein, cysteine-rich domain, ion channel.

Abbreviations: CNG, cyclic nucleotide-gated; CRD, cysteine-rich domain; CRISP, cysteine-rich secretory protein; ELISA, enzyme-linked immunosorbent assay; PR-1, pathogenesis-related protein 1; PsTx, pseudechetoxin; RACE, rapid amplification of cDNA ends.

INTRODUCTION

The cysteine-rich secretory protein (CRISP) family is a large group of secreted proteins of molecular mass 20–30 kDa (*1*, *2*). In mammals, CRISP members have been assigned to four classes (CRISP-1 to CRISP-4) (*3*, *4*). CRISP-1 is secreted by proximal epididymis in an androgen-dependent manner, associates with the sperm surface during maturation, and is latterly involved in gamete fusion (*5*). CRISP-2 is expressed only in the testis and secreted from spermatocytes, which mediates the interaction between spermatocytes and Sertoli cells (*6*). The distribution of CRISP-3 is wider than other classes, including saliva, granules of neutrophils, plasma and thymus, indicating a possible role of CRISP-3 in innate host defense (*5*, *7*). CRISP-3 has been reported as a specific ligand of alpha 1B-glycoprotein in human plasma (*8*). CRISP-4 is expressed exclusively in the epithelium in an androgen-dependent manner (*4*). Although endogenous CRISPs have been steadily identified, their physiological significance is poorly understood.

We have identified and characterized snake venom-derived CRISPs from the venoms of elapids, colubrids and vipers, which comprise a new group of snake venom proteins that are widely distributed among snake venoms (*1*, *2*, *9*, *10*). Several CRISPs have also been recovered from the venoms of lizards as well as snakes, which are a monophyletic venom toxin exclusive of related non-venom proteins that evolved from a single early recruitment event before the separation of snakes and lizards (*11–13*). Crystal structures of snake venom CRISPs reveal that they are composed of three structurally distinct regions: the N-terminal pathogenesis-related protein 1 (PR-1) domain, the hinge region and the C-terminal cysteine-rich domain (CRD) (*14*). Some snake venom CRISPs have been shown to block various cation channels such as cyclic nucleotide-gated (CNG) channels, voltage-gated Ca^{2+} channels, voltage-gated K^+ channels and Ca^{2+}-activated K^+ channels (*15–18*). The CRD of snake venom CRISPs contains a structurally flanking six-cysteine repeat motif, similar to that of the K^+-channel-blocking toxins from sea anemones and scorpions, and is reported to be a possible site for interaction with ion channels (*14*, *19*, *20*). In contrast to this, we recently found that the CRD of a CNG-channel-blocking CRISP (Pseudechetoxin, PsTx) alone does not block CNG channels at all (*21*). These data strongly indicate that the interaction between CRISPs and ion channels is not simply mediated by CRDs.

In this report, we have identified two novel CRISPs, kaouthin-1 and kaouthin-2 in the venom of *Naja kaouthia* (Thai cobra, Elapidae). Kaouthin-1 and

*To whom correspondence should be addressed.
E-mail: tmorita@my-pharm.ac.jp
†Nucleotide sequence data reported are available in the GenBank databases under the accession numbers, EU938339 and EU938340.

kaouthin-2, which share 63% amino-acid identity, were divided into the two Elapidae sister clades of the Asian species and the marine/Australian species by phylogenetic analysis. Snake venom proteins are generally known to vary within their functionally important regions, resulting in the rapid evolution of various biological functions while maintaining a conserved structural scaffold. We have demonstrated that the high frequent divergence among snake venom CRISPs occurs along a continuous line on the molecular surface of the protein. Our results strongly suggest that a functional region of CRISPs should also be located on a linear region of the PR-1 domain that recognizes and interacts with target molecules.

MATERIALS AND METHODS

Materials—Lyophilized venom of *N. kaouthia* was purchased from Kentucky Reptile Zoo (Slade, KY, USA). Other lyophilized venoms were from the Japan Snake Institute (Gunma, Japan), Venom Supplies Pty. (Tanunda, South Australia), and S.A. venom Suppliers CC. (Louis Trichardt, South Africa). The Viperidae or Elapidae snake venom-derived CRISP-specific anti-ablomin or anti-pseudecin antisera were prepared as described previously; ablomin and pseudecin are CRISPs from the venoms of *Agkistrodon blomhoffi* (the formal name is *Gloydius blomhoffi*, viper) and *Pseudechis porphyriacus* (elapid), respectively (*15*, *16*). FPLC columns of Superdex 75 pg gel-filtration, SP Sepharose High Performance, and Hi-Trap Heparin were obtained from GE Healthcare (Uppsala, Sweden), and the HPLC COSMOSIL 5C18 AR-300 column was from Nacalai Tesque (Kyoto, Japan). Lysyl endopeptidase was purchased from Seikagaku Co. (Tokyo, Japan). Other chemicals of analytical grade were purchased from Sigma, GE Healthcare, Wako Pure Chemical Industries (Osaka, Japan) and Kanto Chemical Co. (Tokyo, Japan).

Purification of Kaouthin-1, Kaouthin-2 and Other Snake Venom CRISPs—Three hundred milligrams of lyophilized venom of *N. kaouthia* was dissolved in 20 ml of 50 mM Tris–HCl buffer, pH 8.0, containing 0.2 M of NaCl. After centrifugation to remove a small amount of insoluble particulate, the supernatant was applied onto a Superdex 75 pg gel-filtration column with the same buffer (Fig. 1A). The fractions reacted with both anti-ablomin and anti-pseudecin antisera by enzyme-linked immunosorbent assay (ELISA) were pooled and dialysed with 20 mM imidazole–HCl buffer, pH 6.0, and then loaded onto a SP Sepharose High Performance column with same buffer (Fig. 1B). Proteins were eluted with a linear gradient of 0–1 M NaCl. Subsequently, the fractions reacted with anti-ablomin antiserum containing kaouthin-1 were dialysed with 50 mM Tris–HCl buffer, pH 8.0, and applied onto a Hi-Trap Heparin column with same buffer (Fig. 1C). Kaouthin-1 and kaouthin-2 were

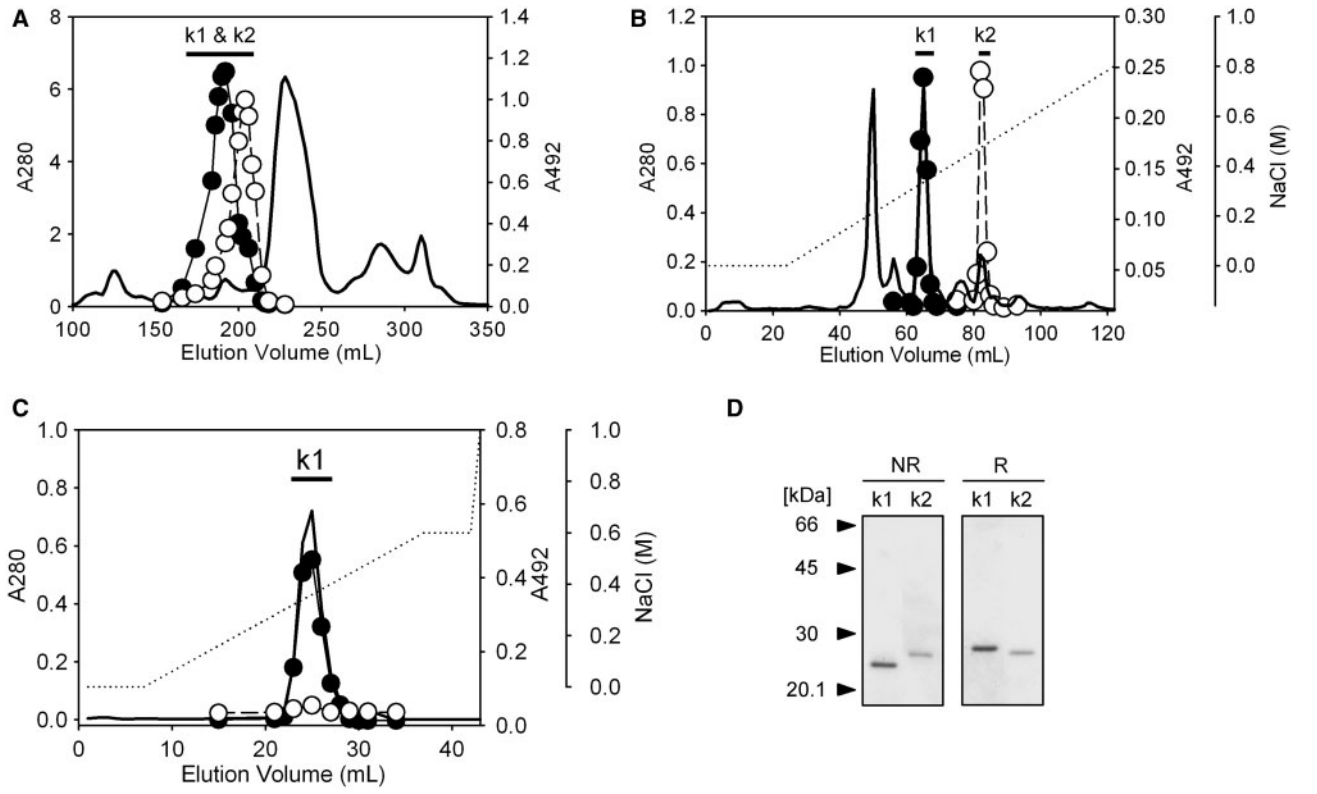

Fig. 1. **Purification of kaouthin-1 and kaouthin-2.** (A–C) Profiles of gel-filtration chromatography (A), cation-exchange chromatography (B) and heparin affinity chromatography (C). Kaouthin-1 and kaouthin-2 (denoted as k1 and k2) were detected by ELISA using anti-ablomin (filled circle) and anti-pseudecin (unfilled circle) antisera, respectively. The pooled fractions are indicated by the bold bar at the top. (D) SDS–PAGE of purified kaouthin-1 and kaouthin-2 was performed without and with reducing agent (NR; non-reducing, R; reducing conditions).

>95% pure as judged by SDS–PAGE (Fig. 1D). Similar purification procedure was used for the isolation of other 11 CRISPs.

Cloning of the cDNAs Encoding Kaouthin-1 and Kaouthin-2—The total RNAs isolated with ISOGEN (Wako Pure Chemical Industries, Osaka, Japan) from the venom gland of *N. kaouthia* were reverse transcribed using Super Script II Reverse Transcriptase (Gibco, MD, USA) to obtain the cDNA. A 5′ and 3′-rapid amplification of cDNA ends (RACE) method was carried out to determine the nucleotide sequence of 5′-end and 3′-end cDNAs using the SMART RACE cDNA amplification kit (Clontech, CA, USA). One degenerate sense primer was designed based on the highly conserved nucleotide sequences encoding the signal peptide among snake venom CRISPs: 5′-CTT GCT GCA GTG CTG C(A/G)A CAG TCT T-3′ for 3′-RACE, and two antisense primers were designed based on the highly conserved nucleotide sequences among Viperidae snake venom-derived CRISPs or Elapidae snake venom-derived CRISPs: 5′-TCC TGG CAG CTA CCT TGT TTC AAC AAA CTG-3′ (corresponding to amino acids 192–201 in kaouthin-1) and 5′-CTT TTG TAC CAA ACT ATC TG(A/G) GT(A/G) (A/T)AA TGG CC-3′ (corresponding to amino acids 95–105 in kaouthin-2) for 5′-RACE. PCR products were subcloned into dephosphorylated pUC19 vector digested with *Hinc*II, and transformed into *E. coli* strain XL-1 Blue. Their complete nucleotide sequences were determined by DNA sequencer LIC-4200L (Aloka, Tokyo, Japan) (Fig. 2A and B).

Amino-Acid Sequence Analysis—Internal amino-acid residues were identified from S-pyridylethylated peptides digested with lysyl endopeptidase. 2.5 nmol of purified kaouthin-1 and kaouthin-2 were reduced for 5 h at room temperature with 20 mM dithiothreitol in 500 µl of buffer solution, containing 0.5 M Tris–HCl, pH 8.5, containing 6 M guanidine hydrochloride and 2 mM EDTA. 3 µl of 4-vinylpyridine was added and alkylated for 30 min at room temperature. After alkylation, the *S*-pyridylethylated proteins were dialysed in 0.1% TFA solution, and lyophilized. Then the proteins were redissolved in 500 µl of 50 mM Tris–HCl, pH 9.5, containing 4 M urea, and digested by lysyl endopeptidase. The digested peptides separated by C18 reverse-phase HPLC with a linear gradient of acetonitrile up to 30%. Eluted peptides were analysed on Applied Biosystems protein sequencers (Models 473A and 477) (Fig. 2A and B, underlined). The N-terminal amino-acid sequencing of kaouthin-1, kaouthin-2 and other snake venom CRISPs (500 pmol) were performed on a Shimadzu protein sequencer PPSQ-21A (Shimadzu, Kyoto, Japan) (underlined in Figs 2A, B and 3).

Phylogenetic Analysis—For neighbour-joining method phylogenetic analysis, we used multiple nucleotide sequence alignment in open reading frame of 24 CRISPs (Fig. 4B and Supplementary Fig. S2). Multiple sequence alignments were constructed using the ClustalW algorithm. The distance matrix for the alignment sequences was calculated by using Kimura's two-parameter method as implemented in a computer program by using Genetyx version 9. Homology searching analysis at the nucleotide and amino-acid levels among snake venom CRISPs was performed by using Genetyx version 9 (Supplementary Table S1). GenBank accession numbers for the nucleotide sequences used are: *N. kaouthia* kaouthin-1 (EU938339); *N. atra* natrin (AY324325, first described as NA-CRVP2 which is not in Genbank database) (*17*, *22*); *Ophiophagus hannah* ophanin (AY181984) (*23*); *N. kaouthia* kaouthin-2 (EU938340); *N. atra* NA-CRVP1 (AY261468) (*22*); *Laticauda semifasciata* latisemin (AF384220) (*16*); *P. australis* PsTx (AY072695) (*24*); *P. porphyriacus* pseudecin (AY072696) (*15*); *Oxyuranus microlepidotus* CRISP-OXY1 (DQ139896) (*13*); *Leioheterodon madagascariensis* CRISP-LEI1 (DQ139894) (*13*); *Philodryas olfersii* CRISP-PHI2 (DQ139898) (*13*); *Liophis poecilogyrus* CRISP-LIO1 (DQ139895) (*13*); *Enhydris polylepis* CRISP-ENH1 (DQ139892) (*13*); *Rhabdrophis t. tigrinus* tigrin (AY093955) (*16*); *Dispholidus typus* CRISP-DIS1 (DQ139889) (*13*); *Telescopus dhara* CRISP-TEL1 (DQ139900) (*13*); *Trimorphodon biscutatus* CRISP-TRI1 (DQ139901) (*13*); *Trimeresurus flavoviridis* (also called *Protobothrops flavoviridis*) triflin (AF384219) (*16*); *P. mucrosquamatus* TM-CRVP (PMU59447) (*25*); *T. jerdonii* (also called *P. jerdonii*) TJ-CRVP (AY261467) (*22*); *Viridovipera stejnegeri* stecrisp (AY423708) (*20*); *A. blomhoffi* (the formal name is *G. blomhoffi*) ablomin (AF384218) (*16*); *Crotalus atrox* catrin (AY181983) (*23*); *A. p. piscivorus* piscivorin (AY181982) (*23*); and human CRISP-1 (NM_001131).

Homology Modelling—The tertiary structures of kaouthin-1, ophanin, tigrin, CRISP-ENH1, ablomin and physiological human CRISPs were constructed by homology modelling based on the crystal structure of triflin (*14*). The tertiary structures of kaouthin-2 and latisemin were constructed by homology modelling based on the crystal structure of pseudecin (*21*). The crystal structure of natrin has been analysed (*17*). All of the structures were constructed with colour electrostatic potential levels (red, −5.0; white, 0; blue, 5.0) by Swiss PDB Viewer version 3.7, and drawn by POV Ray version 3.6.

WEBLOGO Analysis—The sequence logo was generated within a multiple amino-acid sequence alignment among 24 snake venom CRISPs shown in Fig. 4B by using the WEBLOGO program (http://weblogo.berkeley.edu/). The overall height of each stack indicates the sequence conservation at that position (measured in *bits*), whereas the height of symbols within the stack reflects the relative frequency of the corresponding amino acid at that position.

RESULTS

Isolation of Kaouthin-1 and Kaouthin-2 from N. Kaouthia Venom—Snake venom proteins have evolved to acquire unique structures and functions unseen in mammals. To explore the distribution and evolution of snake venom CRISPs, we screened several species of snake venoms using two anti-snake venom-derived CRISP antisera: anti-ablomin and anti-pseudecin antisera (Supplementary Fig. S1). Ablomin and pseudecin are CRISPs from the venoms of *A. blomhoffi* (the formal name is *G. blomhoffi*, viper) and *P. porphyriacus* (elapid), respectively (*15*, *16*). As a result, though most of the

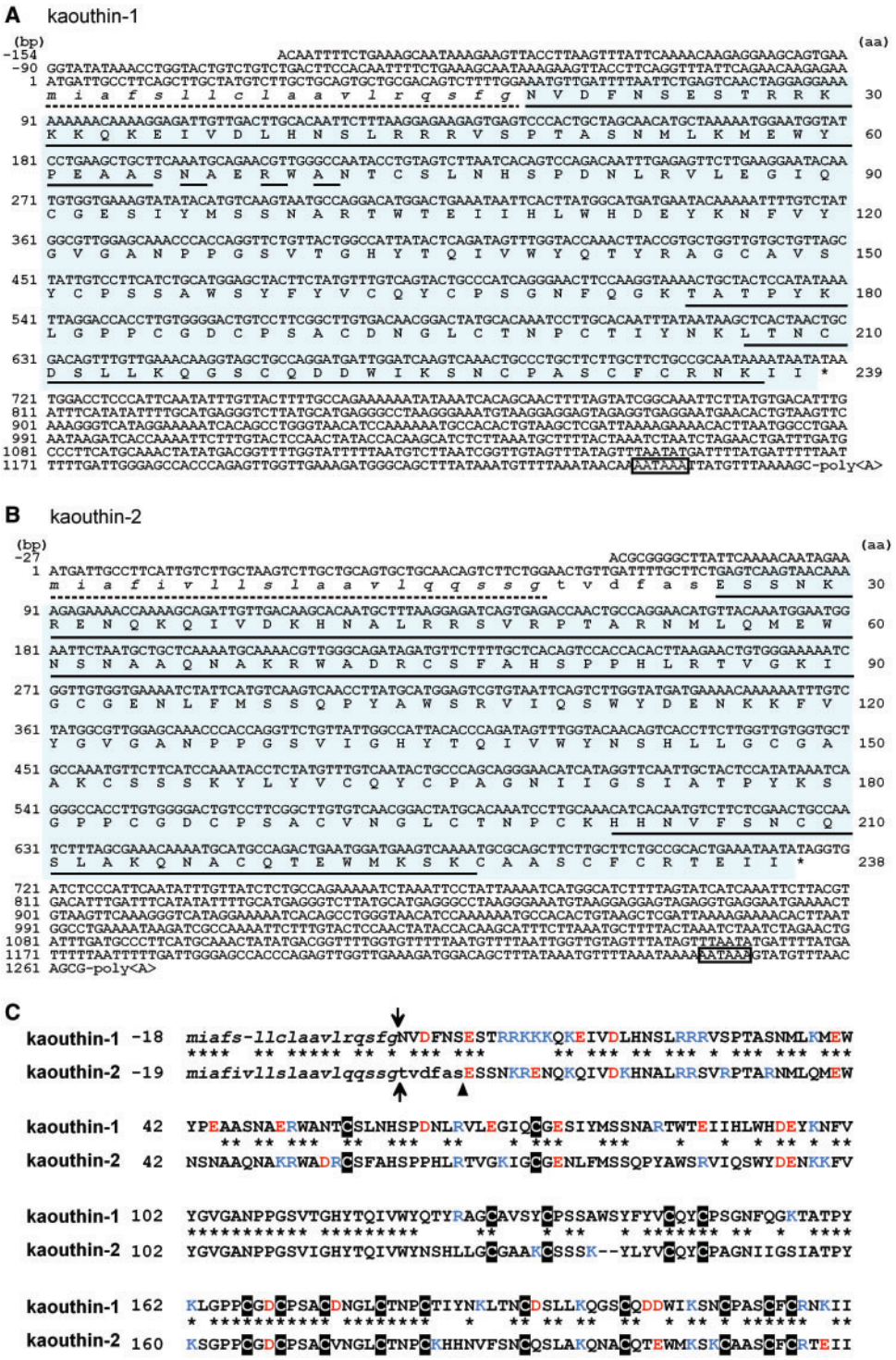

Fig. 2. **Nucleotide and deduced amino-acid sequences of kaouthin-1 and kaouthin-2.** (A, B) The amino-acid sequences are shown in single-letter code beneath the nucleotide sequences encoding kaouthin-1 (A) and kaouthin-2 (B). Nucleotide (left) and amino-acid (right) numbers are shown in the columns at either side. Putative signal peptides are shown in italics-style lowercase and dashed-underlined. The residues which are not determined by N-terminal sequencing analysis of kaouthin-2 are shown in lowercase (B). The mature protein coding region is coloured in blue. The underlined amino acids were identified by sequencing *S*-pyridylethylated peptides from enzymatic digests and N-terminal sequencing analyses of purified proteins. The putative polyadenylation signal is boxed. (C) Comparison of the amino-acid sequences between kaouthin-1 (top) and kaouthin-2 (bottom). Conserved cysteine residues are highlighted. Identical residues of the two kaouthins are marked with asterisks. Basic amino acids (K and R) and acidic amino acids (D and E) are shown in blue and red, respectively. Putative signal peptide cleavage sites and pro-peptide cleavage site are indicated by arrow and arrowhead. Similar proteins to kaouthin-1 and kaouthin-2 were identified by Osipov *et al*. (*40*), although they were partially sequenced. Note that the above reported kaouthin-2-like protein has five additional residues on its N-terminus as compared with kaouthin-2 (*40*).

CRISP	N-terminal sequence of purified protein	*species*	Family	Distribution
Kaouthin-1-type (Asian species clade)				
	1 10 20 30			
kaouthin-1	NVDFNSESTRRKKKQKEIVDLHNSLRRVS	*Naja kaouthia*	Elapidae	Southeast Asia
najanajin	SVDFNSESTRxK	*Naja naja*	Elapidae	India
annuliferin-a	NVDFNSESTRRKKKQKEIVDLHNxLRRxVD	*Naja annulifera*	Elapidae	South Africa
annuliferin-b	NVDFNSESTRRKNKQKEIVDLHNSLRRxVD	*Naja annulifera*	Elapidae	South Africa
mossambin	NVDFNSESTRRKKKQNEIVDLHNSLRRTVN	*Naja mossambica*	Elapidae	South Africa
hematin	SVDFNSESTRREEKQKEIVDLHNxLRDxVx	*Hemachatus haemachatus*	Elapidae	South Africa
okinavin	SVDFDSESPRKPxIQNEIVDLHNPLRRxVN	*Ovophis okinavensis*	Viperidae	Japan/Okinawa
Kaouthin-2-type (marine/Australian species clade)				
	1 10 20			
kaouthin-2	------ESSNKRENQKQIVDKHNALRRSVR	*Naja kaouthia*	Elapidae	Southeast Asia
notescatin-a	--------SNKKDYQKEIVDKHNALRRSVK	*Notechis scatatus*	Elapidae	South-eastern Australia
notescatin-b	--------SNKKDYQKEIVDKHNALRRSVK	*Notechis scatatus*	Elapidae	South-eastern Australia
pseuguttin	--------SNKKNYQKEIVDKHNALRRSVK	*Pseudechis guttutus*	Elapidae	South-eastern Australia
collettin-a	--------SNKKNYQKEIVDKHNALRRSVK	*Pseudechis colletti*	Elapidae	East Australia
collettin-b	----------KKNYQKEIVDKHNALRRSVK	*Pseudechis colletti*	Elapidae	East Australia

Fig. 3. **N-terminal amino-acid sequences of isolated CRISPs.** Residues identical with kaouthin-1 are shaded. Undetermined residues are indicated by 'x', and undetected residues from mature proteins are indicated by '-'. Snake venom-derived CRISPs were classified into two types, kaouthin-1-type (upper side) and kaouthin-2-type (lower side) by using their N-terminal sequences.

venoms reacted with either anti-ablomin or anti-pseudecin antisera, we detected two proteins in the venom of *N. kaouthia* (Elapidae, *Nk* in Supplementary Fig. S1) with different molecular weights (−25 kDa), suggesting that two immunologically distinct CRISP-like proteins are contained in the venom of *N. kaouthia* (Supplementary Fig. S1). Using three successive chromatographic runs, we isolated two proteins from the venom of *N. kaouthia*, which were individually detected by ELISA with anti-ablomin and anti-pseudecin antisera (filled and open circles, respectively in Fig. 1A–C). The two proteins were eluted and separated with 0.3 M and 0.7 M NaCl by a second step of cation-exchange chromatography (Fig. 1B), and named kaouthin-1 and kaouthin-2 (fractions k1 and k2 in Fig. 1), respectively. Kaouthin-2, which reacted with anti-pseudecin antiserum, was completely purified by the second step, while anti-ablomin reacting kaouthin-1 was further purified by heparin affinity chromatography (Fig. 1C). The venom content of kaouthin-1 (0.7% of total venom peptide and protein) is higher than that of kaouthin-2 (0.2%). The purified kaouthin-1 migrated with a M_r of 22 kDa under non-reducing conditions and 25 kDa under reducing conditions by β-mercaptoethanol on SDS–PAGE, though kaouthin-2 was a 24 kDa protein under both non-reducing and reducing conditions (Fig. 1D).

Cloning cDNAs Encoding Kaouthin-1 and Kaouthin-2—To determine the primary structures of kaouthin-1 and kaouthin-2, two distinct degenerate primer sets specific to Viperidae snake venom and Elapidae snake venom-derived CRISPs were designed based on the highly conserved nucleotide sequences among Viperidae snake venom-derived CRISPs or Elapidae snake venom-derived CRISPs (see MATERIALS AND METHODS section). Two cDNAs encoding CRISP-like proteins were cloned from the venom gland cDNAs of *N. kaouthia*; they are 1,410 bp and 1,291 bp in length and encode proteins composing 239 and 238 amino acids, respectively (Fig. 2A and B). The two nucleotide sequences share 85% nucleotide identity, and the encoded proteins share 63% amino-acid identity (Supplementary Table S1), suggesting that the snake venom CRISPs have evolved at the greater substitution rates common for snake venom proteins (*26*). The 1,410 bp cDNA encodes a protein highly homologous to ablomin, while the 1,291 bp cDNA encodes a protein highly protein to pseudecin (Supplementary Table S1), suggesting that they correspond to kaouthin-1 and kaouthin-2 cDNAs, respectively. The pI of kaouthin-1 was 7.8, while that of kaouthin-2 was 8.8 using the pKa values for the predicted amino-acid sequences of kaouthin-1 and kaouthin-2. Number of basic amino-acid residues of kaouthin-1 and kaouthin-2 are comparable, while acidic amino-acid residues of kaouthin-1 is seven residues greater than that of kaouthin-2 (Fig. 2C).

Identification of Kaouthin-1 and Kaouthin-2—*N*-terminal amino-acid sequence analyses of pro-proteins using the signal peptide prediction programs PSORT II revealed that pro-kaouthin-1 and pro-kaouthin-2 should be cleaved at Gly18/Asn19 and Gly19/Thr20 from the initial methionine at sequences qsfg/NVDF and qssg/tvdf, respectively (signal peptides are shown in italic-style lowercase, and signal peptide cleavage sites are indicated by arrows in Fig. 2C). The N-terminal amino-acid sequence of purified kaouthin-1, Asn-Val-Asp-Phe-, was identical to the predicted sequence, whereas that of purified kaouthin-2, Glu-Ser-Ser-Asn-, was different from the predicted sequence by the signal peptide prediction programs, starting at amino acid 26 from the initial methionine, suggesting that pro-kaouthin-2 may be further cleaved at Ser25/Glu26 by a sequence-specific protease after secretion, otherwise the signal peptide cleavage site of kaouthin-2 may be different from kaouthin-1 (the cleaved peptide is shown in lowercase, and the putative cleavage site is indicated by an arrowhead in Fig. 2C). The internal amino-acid residues of purified kaouthin-1 and kaouthin-2 were identified

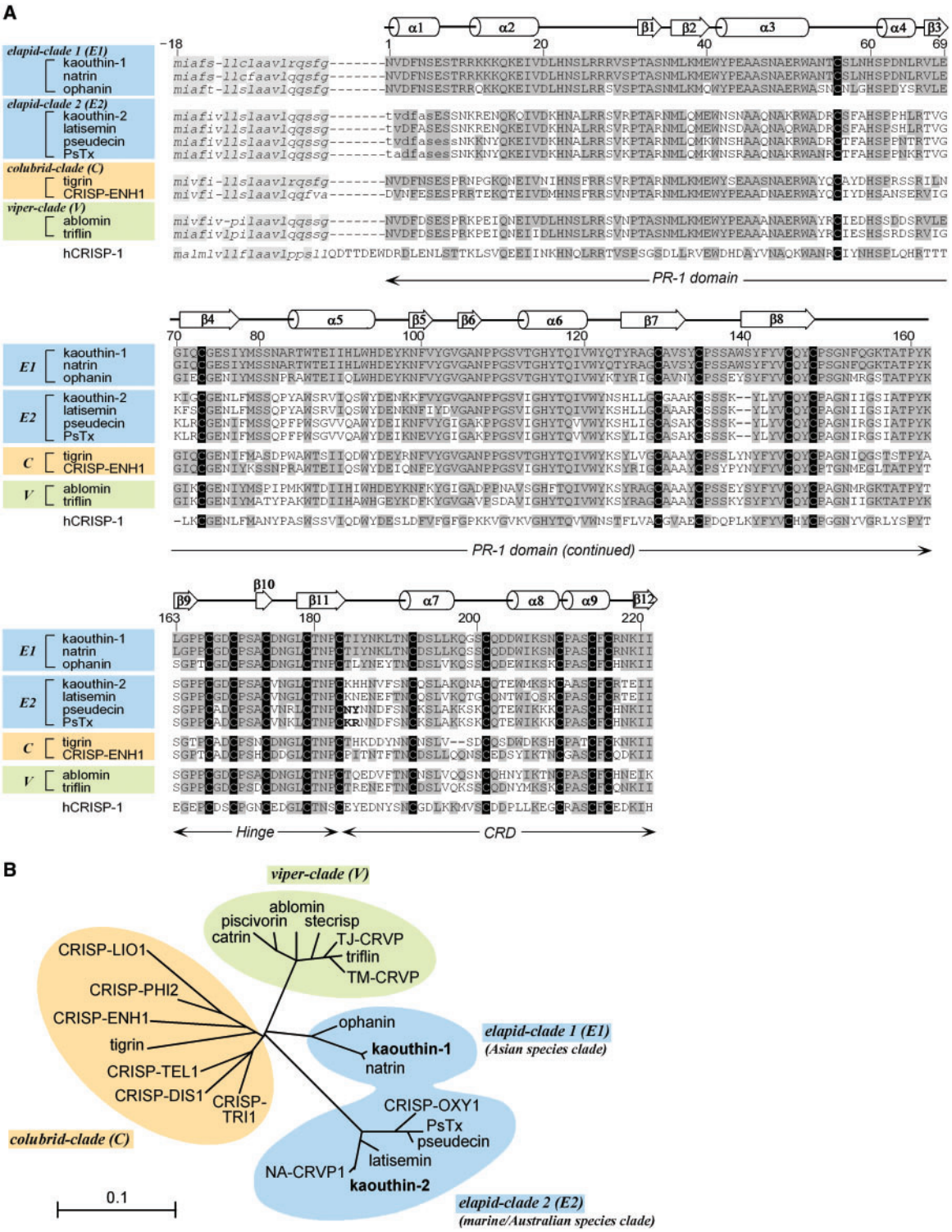

Fig. 4. **Primary structures and phylogenetic tree of snake venom CRISPs.** (A) Amino-acid sequence alignment of 10 representative snake venom CRISPs and human CRISP-1. Identical residues with kaouthin-1 are shaded, and conserved cysteine residues are highlighted. Putative signal peptides are indicated in lowercase italics. Residues which were not detected by N-terminal amino-acid sequencing are indicated in lowercase. Gaps (–) were inserted to maximize similarity. The numbering for kaouthin-1 is indicated above its sequence. The secondary structural elements identified for triflin (*14*) are shown as arrows for β-strands and cylinders for α-helices. The two adjacent residues numbered as 184 and 185 of pseudecin and PsTx are indicated by bold letters. Natrin was first described as NA-CRVP2 (*22*). Kaouthin-1 was quite similar to natrin/NA-CRVP2 at both the nucleotide and amino-acid levels (99% and 99%, respectively, Supplementary Table S1), suggesting that kaouthin-1 should possess similar biological activities with natrin (*17*, *18*). (B) The neighbour-joining method phylogenetic tree was built by using 24 nucleotide sequences in open reading frame. Snake venom CRISPs are divided into four clades (*E1*, *E2*, *C* and *V*). GenBank accession numbers are listed under MATERIALS AND METHODS section.

using *S*-pyridylethylated peptides digested with lysyl endopeptidase (Fig. 2A and B, underlined). All of the analysed peptides completely corresponded with the predicted amino-acid sequences deduced from the nucleotide sequences encoding kaouthin-1 and kaouthin-2. These results indicate that mature kaouthin-1 and kouthin-2 are composed of 221 and 213 amino-acid residues with predicted molecular weights of 24.9 and 23.7 kDa, respectively.

Isolation of 11 CRISPs from the Venoms of Eight Different Species—To further explore the distribution of snake venom CRISPs, we also screened several species of snake venoms that are not reported to contain CRISPs by ELISA using anti-triflin antiserum. As a result, a total of 11 CRISP-like proteins were isolated from eight different snake venoms (Fig. 3). Five CRISPs, named najanajin, annuliferin-a/annuliferin-b, mossambin and hematin, were from the venoms of Asian or African Elapidae snakes classified as cobra: *N. naja* (Indian cobra), *N. annulifera* (banded Egyptain cobra), *N. mossambica* (Mozambique cobra) and *Hemachatus haemachatus* (ringhals). A CRISP named okinavin was from the Japanese Viperidae venom of *Ovophis okinavensis* (Himehabu in Japanese). The other five CRISPs, named notescatin-a/notescatin-b, pseuguttin and collettin-a/collettin-b, were isolated from the venoms of Australian Elapidae snakes: *Notechis scutatus* (common tiger snake), *P. guttatus* (spotted black snake) and *P. colletti* (Collett's snake). According to their N-terminal sequences as determined by amino-acid sequencing, the 11 snake venom-derived CRISPs were classified into two types, kaouthin-1-type and kaouthin-2-type (upper and lower side in Fig. 3). Kaouthin-1-type CRISPs were from the venoms of Asian and African Elapidae and a Japanese Viperidae, and kaouthin-2-type CRISPs were from the venoms of Australian elapids. All kaouthin-2-type CRISPs, including kaouthin-2, lack their N-terminal 6–10 residues predicted to compose the α1 helix. The three elapid snakes, *N. haje annulifera*, *N. scutatus* and *P. colletti*, also contained two CRISPs in their venoms: annuliferin-a/annuliferin-b, notescatin-a/notescatin-b and collettin-a/collettin-b, which were eluted and separated with different NaCl concentrations by cation-exchange chromatography in a similar manner to kaouthin-1/kaouthin-2 (Fig. 1B). However, the N-terminal amino-acid sequences of the three pairs, except for kaouthin-1/kaouthin-2, were highly homologous or identical to each other (Fig. 3).

DISCUSSION

Distribution of Snake Venom CRISPs—Snake venom proteins, which possess unique structures and functions unseen in mammals, are plentiful tools for understanding the underlying the physiological mechanisms of their targets (*2*, *27*). Elapid snake toxins are generally known to act on the nervous system; they inhibit synaptic transmission by specifically and potently blocking a variety of ion channels, including acetylcholine receptors (*28*, *29*). In contrast, viper and colubrid snake toxins not only act on the nervous system, but also the blood coagulation system and the vascular system (*30*, *31*).

In this study, we identified two structurally distinctive CRISPs, kaouthin-1 and kaouthin-2, in the venom of *N. kaouthia* (Thai cobra, Elapidae). Snake venom CRISPs are distributed in the venoms of widespread species of elapids, colubrids and vipers (Fig. 4). The primary structures of 10 representative snake venom CRISPs containing kaouthin-1 and kaouthin-2 are aligned in Fig. 4A. Phylogenetic trees in Fig. 4B and Supplementary Fig. S2 using 24-nucleotide sequences revealed that snake venom CRISPs are divided into four clades: two elapid-clades (blue groups E1 and E2 in Fig. 4), a colubrid-clade (yellow group C in Fig. 4), and a viper-clade (green group V in Fig. 4). Kaouthin-1 and kaouthin-2 fall into two elapid-clades: the Asian species clade (group E1 in Fig. 4) and marine/Australian species clade (group E2 in Fig. 4), respectively. Molecular phylogenetic studies based on cytochrome *b* and 16S rRNA mtDNA genes of elapid snakes agree that the marine and the Australian-Melanesian species are collectively monophyletic, similar to the elapid-clade 2 including kaouthin-2 (*32*). In similar phylogenetic studies, elapid snakes are broadly divisible into two sister clades: the Asian/African/American species clade, and the marine/Australian–Melanesian species clade, in a similar manner to snake venom CRISPs (*33*). It is suggested that snake venom CRISP would be useful for elucidating the distribution of venomous snake species. Similar to kaouthin-1 and kaouthin-2, two chromatography-distinctive CRISPs were found in various snake venoms; however, their primary structures are highly similar or identical to each other (Fig. 3) (*13*, *23*). Two structurally distinctive snake venom CRISPs, NA-CRVP1 and natrin/NA-CRVP2, have been reported from the venom of *N. atra* (Chinese cobra) (*17*, *22*). These findings suggest that genus *Naja* in Asia would be the unique species which having two sister elapid-clades CRISPs in their venoms.

Functional Mapping in a Surface Plot of CRISPs—Snake venom proteins are generally known to acquire unique functions via rapid evolution of non-structural residues (*2*, *9*, *10*, *13*, *27*, *34*, *35*). Various snake venom CRISPs have been characterized as affecting specific cation channels; pseudecin and PsTx block olfactory and retinal CNG channels (*15*, *24*, *36*), ablomin and triflin exhibit voltage-gated Ca^{2+} channel blocker-like properties (*16*), and natrin blocks the high conductance Ca^{2+}-activated K^+ (BK_{Ca}) channel currents and voltage-gated K^+ (Kv1.3) channel currents (*17*, *18*). Tertiary structure models of kaouthin-1, kaouthin-2, tigrin and human CRISPs were constructed by homology modelling based on the crystal structures of triflin and pseudecin (Fig. 5) (*14*, *21*). The surface charge distribution of snake venom CRISPs was diversified and classified into four groups similar to their phylogenetic division, whereas those of human CRISPs were highly similar overall despite their relatively lower amino-acid identity (54–72%) compared to snake venom CRISPs (Fig. 5, Supplementary Table S1), suggesting that there are differentiated functions of snake venom CRISPs. The C-terminal CRD of snake venom CRISPs is considered to be the interactive site for ion channels, because the CRD possesses a similar conserved motif with six cysteine

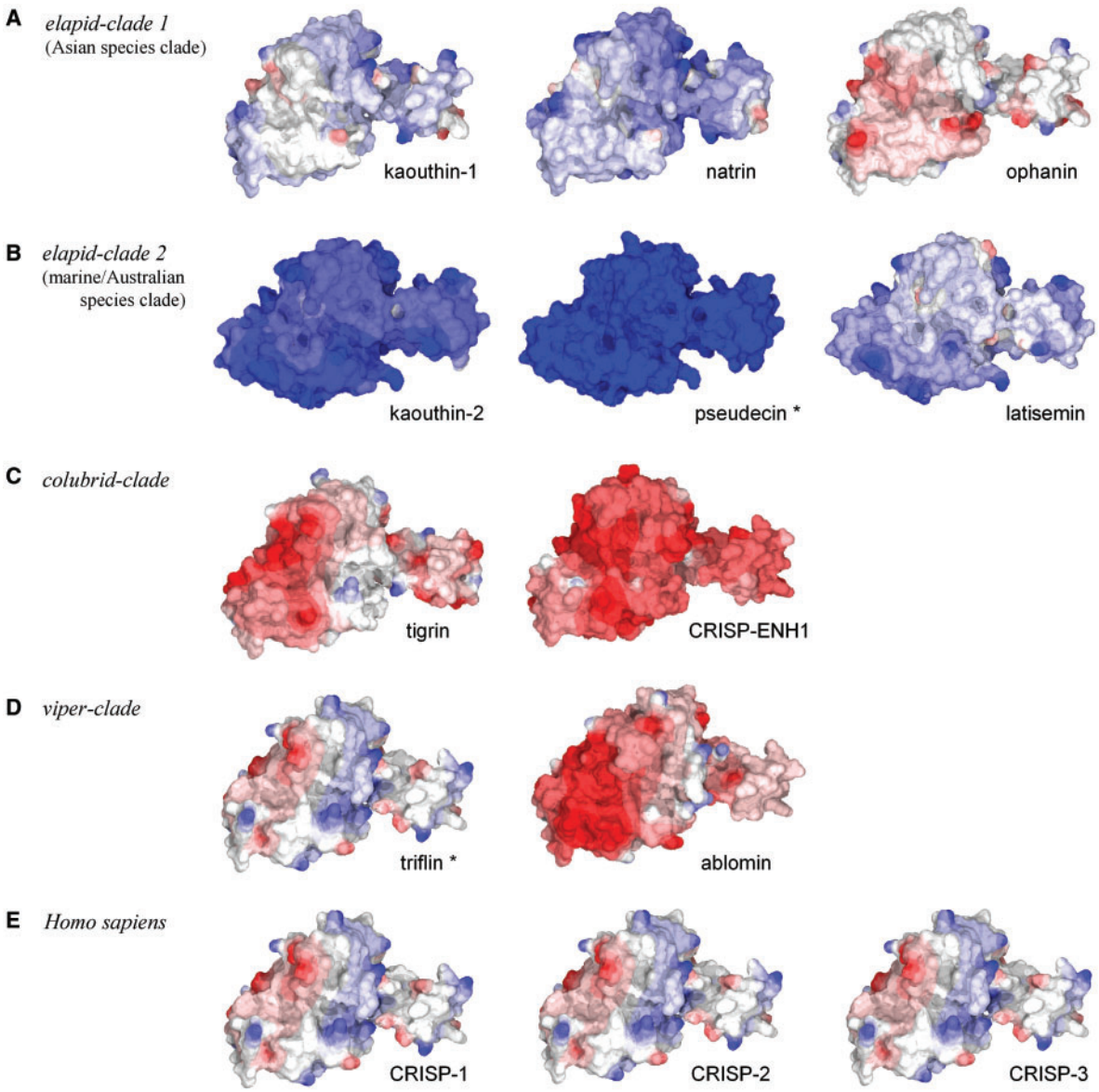

Fig. 5. **The surface electrostatic representation of snake venom CRISPs and human CRISPs.** The tertiary structures of kaouthin-1 (A, left), natrin (A, centre) (*17*), ophanin (A, right), kaouthin-2 (B, left), pseudecin (B, centre) (*21*), latisemin (B, right), tigrin (C, left), CRISP-ENH1 (C, right), triflin (D, left) (*14*), ablomin (D, right) and three human CRISPs (E). Negatively charged regions are represented in red, and positively charged regions in blue. The homology structures were constructed with colour levels (red, −5.0; white, 0; blue, 5.0) using Swiss PDB Viewer. Asterisk denotes the crystal structures of PsTx (*21*) and stecrisp (*20*) were highly similar to the crystal structures of pseudecin and triflin, respectively. The surface potentials of snake venom CRISPs are diversified in a similar manner to their phylogenetic division, whereas those of human CRISPs are highly similar overall.

residues similar to BgK, ShK, kaliotoxin and margatoxin (from the venoms of sea anemones and scorpions), which are peptide blockers for voltage-gated K^+ channels (*14*, *19–21*, *37*, *38*). However, we recently found that a synthesized C-terminal sub-domain corresponding to the CRD of PsTx (Pro182 to Ile221 in Fig. 4A, Pro180 to Ile219 in PsTx) is not sufficient for the inhibition of CNG channels (*21*), indicating that other interactions between CRISPs and ion channels are required for blockage. In this way, although several snake venom CRISPs have been shown to target distinct ion channels, the functionally key regions that recognize and interact with these target proteins are poorly understood. A sequence logo depicted in Fig. 6A indicates the sequence conservation and the relative frequency of each amino-acid position. The sequence logo reveals that the primary structures among snake venom CRISPs are diversified markedly around the α4-β3-β4-α5 and CRD regions (Fig. 6A). All 16 high-frequency varied residues (orange residues, bits <1.6 in Fig. 6A) are exposed on the molecular surface, and compose a continuous line on the PR-1 domain and CRD (orange area in Fig. 6B).

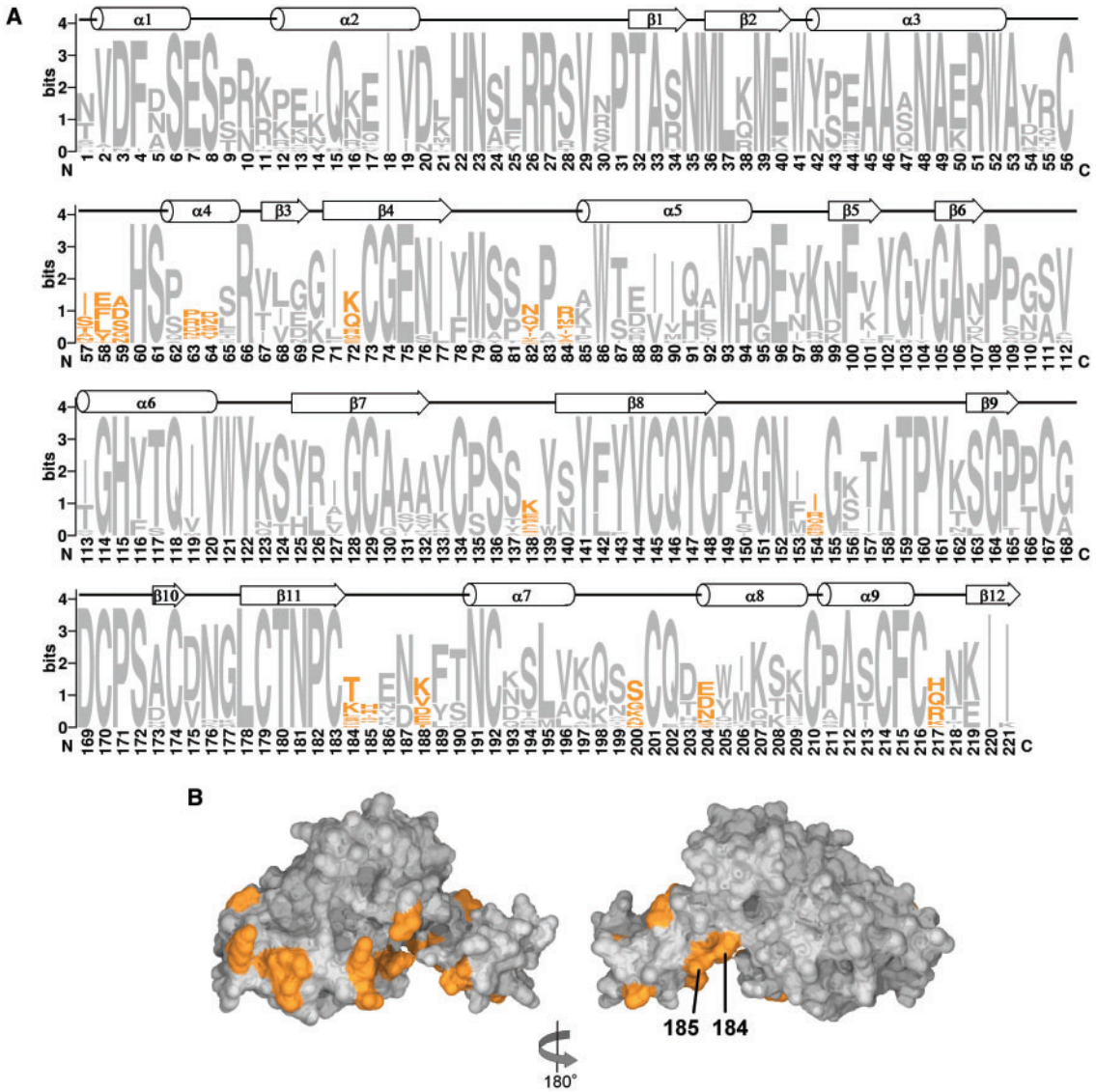

Fig. 6. **Variable residues of snake venom CRISPs.** (A) Amino-acid sequence diversity among 24 snake venom CRISPs shown in Fig. 2B. Variable residues (bits <1.6) are indicated in orange. The secondary structural elements identified for triflin (*14*) are shown as arrows for β-strands and cylinders for α-helices. (B) Surface potential maps among snake venom CRISPs. The tertiary structure of triflin, which is the first discovered functional snake venom CRISP (*16*), is shown in the standard orientation (left) and a 180° rotation (right) about a vertical axis. Two adjacent residues numbered as 184 and 185 are indicated. All of the highly variable residues (bits <1.6 in Fig. 4A) define a continuous line on the molecular surface (orange area), suggesting that this is an important region for recognizing and interacting with various target proteins.

In our previous report, we found that though pseudecin and PsTx are highly homologous proteins (>97% identity), they display 30-fold different inhibitory effects on CNG channels (*15*). The structure comparison reveals that the significant amino-acid differences between pseudecin and PsTx (N184/Y185 and K184/R185 written in bold letters in Fig. 4A, N182/Y183 and K182/R183 in pseudecin and PsTx) are located around the concave surface formed between the PR-1 domain and CRD (*14*, *39*). The two amino acids were also localized on the diversified line structure of snake venom CRISPs (numbered as 184 and 185 in Fig. 6B), suggesting that the PR-1 domain should account for the role of ligand recognition and interaction in the selectivity of snake venom CRISPs in addition to CRD.

In conclusion, we demonstrated that kaouthin-1 and kaouthin-2, which we isolated from the venom of *N. kaouthia*, are classified into two different elapid-clades CRISPs from the Asian species and the marine/Australian species, respectively. Our results also revealed that CRISP is the common toxin component in several snake venoms. The diversified residues among snake venom CRISPs define a continuous line on the molecular surface and most likely confer the toxins with selectivity and affinity for many target proteins. Our findings will be useful to explore the distribution of

venomous snakes and to elucidate the interaction between CRISPs and receptors.

SUPPLEMENTARY DATA

Supplementary data are available at *JB* online.

ACKNOWLEDGEMENTS

We thank Drs Jung Sup Lee (Chosun University, Korea) and Chun Qi Ma (our laboratory) for purification of proteins. We thank Drs Michihisa Toriba and Atsushi Sakai (Japan Snake Institute, Gunma, Japan) for extracting the venom glands. We also thank Satsuki Hori, Kahori Ono, Koji Takani and Nao Kamei for technical assistance in early stage of this study.

CONFLICT OF INTEREST

None declared.

REFERENCES

1. Yamazaki, Y. and Morita, T. (2004) Structure and function of snake venom cysteine-rich secretory proteins. *Toxicon* **44**, 227–231
2. Yamazaki, Y. and Morita, T. (2007) Snake venom components affecting blood coagulation and the vascular system: structural similarities and marked diversity. *Curr. Pharm. Des.* **13**, 2872–2886
3. Kratzschmar, J., Haendler, B., Eberspaecher, U., Roosterman, D., Donner, P., and Schleuning, W. D. (1996) The human cysteine-rich secretory protein (CRISP) family. Primary structure and tissue distribution of CRISP-1. CRISP-2 and CRISP-3. *Eur. J. Biochem.* **236**, 827–836
4. Jalkanen, J., Huhtaniemi, I., and Poutanen, M. (2005) Mouse cysteine-rich secretory protein 4 (CRISP4): a member of the Crisp family exclusively expressed in the epididymis in an androgen-dependent manner. *Biol. Reprod.* **72**, 1268–1274
5. Haendler, B., Kratzschmar, J., Theuring, F., and Schleuning, W. D. (1993) Transcripts for cysteine-rich secretory protein-1 (CRISP-1; DE/AEG) and the novel related CRISP-3 are expressed under androgen control in the mouse salivary gland. *Endocrinology.* **133**, 192–198
6. Maeda, T., Sakashita, M., Ohba, Y., and Nakanishi, Y. (1998) Molecular cloning of the rat Tpx-1 responsible for the interaction between spermatogenic and Sertoli cells. *Biochem. Biophys. Res. Commun.* **248**, 140–146
7. Udby, L., Cowland, J. B., Johnsen, A. H., Sorensen, O. E., Borregaard, N., and Kjeldsen, L. (2002) An ELISA for SGP28/CRISP-3, a cysteine-rich secretory protein in human neutrophils, plasma, and exocrine secretions. *J. Immunol. Methods* **263**, 43–55
8. Udby, L., Sorensen, O. E., Pass, J., Johnsen, A. H., Behrendt, N., Borregaard, N., and Kjeldsen, L. (2004) Cysteine-rich secretory protein 3 is a ligand of alpha1B-glycoprotein in human plasma. *Biochemistry.* **43**, 12877–12886
9. Mackessy, S. P. (2002) Biochemistry and pharmacology of colubrid snake venoms. *J. Toxicol. – Toxin Rev.* **21**, 43–83
10. 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
11. Mochca-Morales, J., Martin, B. M., and Possani, L. D. (1990) Isolation and characterization of helothermine, a novel toxin from *Heloderma horridum horridum* (Mexican beaded lizard) venom. *Toxicon* **28**, 299–309
12. Fry, B. G. (2005) From genome to "venome": molecular origin and evolution of the snake venom proteome inferred from phylogenetic analysis of toxin sequences and related body proteins. *Genome Res.* **15**, 403–420
13. Fry, B. G., Vidal, N., Norman, J. A., Vonk, F. J., Scheib, H., Ramjan, S. F., Kuruppu, S., Fung, K., Hedges, S. B., Richardson, M. K., Hodgson, W. C., Ignjatovic, V., Summerhayes, R., and Kochva, E. (2006) Early evolution of the venom system in lizards and snakes. *Nature* **439**, 584–588
14. Shikamoto, Y., Suto, K., Yamazaki, Y., Morita, T., and Mizuno, H. (2005) Crystal structure of a CRISP family Ca^{2+}-channel blocker derived from snake venom. *J. Mol. Biol.* **350**, 735–743
15. Yamazaki, Y., Brown, R. L., and Morita, T. (2002) Purification and cloning of toxins from elapid venoms that target cyclic nucleotide-gated ion channels. *Biochemistry* **41**, 11331–11337
16. Yamazaki, Y., Koike, H., Sugiyama, Y., Motoyoshi, K., Wada, T., Hishinuma, S., Mita, M., and Morita, T. (2002) Cloning and characterization of novel snake venom proteins that block smooth muscle contraction. *Eur. J. Biochem.* **269**, 2708–2715
17. Wang, J., Shen, B., Guo, M., Lou, X., Duan, Y., Cheng, X. P., Teng, M., Niu, L., Liu, Q., Huang, Q., and Hao, Q. (2005) Blocking effect and crystal structure of natrin toxin, a cysteine-rich secretory protein from *Naja atra* venom that targets the BKCa channel. *Biochemistry* **44**, 10145–10152
18. Wang, F., Li, H., Liu, M. N., Song, H., Han, H. M., Wang, Q. L., Yin, C. C., Zhou, Y. C., Qi, Z., Shu, Y. Y., Lin, Z. J., and Jiang, T. (2006) Structural and functional analysis of natrin, a venom protein that targets various ion channels. *Biochem. Biophys. Res. Commun.* **351**, 443–448
19. Alessandri-Haber, N., Lecoq, A., Gasparini, S., Grangier-Macmath, G., Jacquet, G., Harvey, A. L., de Medeiros, C., Rowan, E. G., Gola, M., Menez, A., and Crest, M. (1999) Mapping the functional anatomy of BgK on Kv1.1, Kv1.2, and Kv1.3. Clues to design analogs with enhanced selectivity. *J. Biol. Chem.* **274**, 35653–35661
20. Guo, M., Teng, M., Niu, L., Liu, Q., Huang, Q., and Hao, Q. (2005) Crystal structure of the cysteine-rich secretory protein stecrisp reveals that the cysteine-rich domain has a K^+ channel inhibitor-like fold. *J. Biol. Chem.* **280**, 12405–12412
21. Suzuki, N., Yamazaki, Y., Brown, R. L., Fujimoto, Z., Morita, T., and Mizuno, H. (2008) Crystal structure of pseudechetoxin and pseudecin, two snake venom cysteine-rich secretory proteins that target cyclic nucleotide-gated ion channels: implications for movement of the C-terminal cysteine-rich domain. *Acta. Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **D64**, 1034–1042
22. Jin, Y., Lu, Q., Zhou, X., Zhu, S., Li, R., Wang, W., and Xiong, Y. (2003) Purification and cloning of cysteine-rich proteins from *Trimeresurus jerdonii* and *Naja atra* venoms. *Toxicon* **42**, 539–547
23. Yamazaki, Y., Hyodo, F., and Morita, T. (2003) Wide distribution of cysteine-rich secretory proteins in snake venoms: isolation and cloning of novel snake venom cysteine-rich secretory proteins. *Arch. Biochem. Biophys.* **412**, 133–141
24. Brown, R. L., Haley, T. L., West, K. A., and Crabb, J. W. (1999) Pseudechetoxin: a peptide blocker of cyclic nucleotide-gated ion channels. *Proc. Natl Acad. Sci. USA* **96**, 754–759
25. Chang, T. Y., Mao, S. H., and Guo, Y. W. (1997) Cloning and expression of a cysteine-rich venom protein from *Trimeresurus mucrosquamatus* (Taiwan habu). *Toxicon* **35**, 879–888

26. Ogawa, T., Oda, N., Nakashima, K., Sasaki, H., Hattori, M., Sakaki, Y., Kihara, H., and Ohno, M. (1992) Unusually high conservation of untranslated sequences in cDNAs for *Trimeresurus flavoviridis* phospholipase A2 isozymes. *Proc. Natl Acad. Sci. USA* **89**, 8557–8561
27. Yamazaki, Y. and Morita, T. (2008) Snake venoms and other toxic components affecting thrombosis and Hemostasis in *Recent Advances in Thrombosis and Hemostasis 2008* (Tanaka, K. and Davie, E. W., eds.) pp. 462–482, Springer, Japan
28. Harvey, A. L., Bradley, K. N., Cochran, S. A., Rowan, E. G., Pratt, J. A., Quillfeldt, J. A., and Jerusalinsky, D. A. (1998) What can toxins tell us for drug discovery? *Toxicon* **36**, 1635–1640
29. Tsetlin, V. (1999) Snake venom alpha-neurotoxins and other 'three-finger' proteins. *Eur. J. Biochem.* **264**, 281–286
30. Calvete, J. J., Marcinkiewicz, C., Monleon, D., Esteve, V., Celda, B., Juarez, P., and Sanz, L. (2005) Snake venom disintegrins: evolution of structure and function. *Toxicon* **45**, 1063–1074
31. Morita, T. (2005) Structures and functions of snake venom CLPs (C-type lectin-like proteins) with anticoagulant-, procoagulant-, and platelet-modulating activities. *Toxicon* **45**, 1099–1114
32. Keogh, J. S., Shine, R., and Donnellan, S. (1998) Phylogenetic relationships of terrestrial Australo-Papuan elapid snakes (subfamily Hydrophiinae) based on cytochrome b and 16S rRNA sequences. *Mol. Phylogenet. Evol.* **10**, 67–81
33. Slowinski, J. B. and Keogh, J. S. (2000) Phylogenetic relationships of elapid snakes based on cytochrome b mtDNA sequences. *Mol. Phylogenet. Evol.* **15**, 157–164
34. Kini, R. M. (2004) Platelet aggregation and exogenous factors from animal sources. *Curr. Drug Targets Cardiovasc. Haematol. Disord.* **4**, 301–325
35. Doley, R., Pahari, S., Mackessy, S. P., and Kini, R. M. (2008) Accelerated exchange of exon segments in Viperid three-finger toxin genes (*Sistrurus catenatus edwardsii*; Desert Massasauga). *BMC Evol. Biol.* **8**, 196
36. Brown, R. L., Lynch, L. L., Haley, T. L., and Arsanjani, R. (2003) Pseudechetoxin binds to the pore turret of cyclic nucleotide-gated ion channels. *J. Gen. Physiol.* **122**, 749–760
37. Gilquin, B., Braud, S., Eriksson, M. A., Roux, B., Bailey, T. D., Priest, B. T., Garcia, M. L., Menez, A., and Gasparini, S. (2005) A variable residue in the pore of Kv1 channels is critical for the high affinity of blockers from sea anemones and scorpions. *J. Biol. Chem.* **280**, 27093–27102
38. Lange, A., Giller, K., Hornig, S., Martin-Eauclaire, M. F., Pongs, O., Becker, S., and Baldus, M. (2006) Toxin-induced conformational changes in a potassium channel revealed by solid-state NMR. *Nature* **440**, 959–962
39. Suzuki, N., Yamazaki, Y., Fujimoto, Z., Morita, T., and Mizuno, H. (2005) Crystallization and preliminary X-ray diffraction analyses of pseudechetoxin and pseudecin, two snake-venom cysteine-rich secretory proteins that target cyclic nucleotide-gated ion channels. *Acta. Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.* **61**, 750–752
40. Osipov, A. V., Levashov, M. Y., Tsetlin, V. I., and Utkin, Y. N. (2005) Cobra venom contains a pool of cysteine-rich secretory proteins. *Biochem. Biophys. Res. Commun.* **328**, 177–182