Maejo Int. J. Sci. Technol. **2018**, *12*(03), 187-198

Maejo International Journal of Science and Technology

ISSN 1905-7873
Available online at www.mijst.mju.ac.th

Full Paper

Sequence analysis and 3-dimensional molecular modelling of Heteroscorpine from the venom of *Heterometrus* spp.

Paroonkorn Incamnoi [1,2], Rina Patramanon [2], Sompong Klaynongsruang [2], Arunrat Chaveerach [3], Wandee Bunyatratchata [4], Nunthawun Uawonggul [2, 5], Sophida Sukprasert [2, 6], Teerawat Namonsai [2], Kowit Noikotr [7], Prapenpuksiri Rungsa [2], Jureerut Daduang [8] and Sakda Daduang [2, 9, *]

[1] Department of Chemistry, Faculty of Engineering, Rajamangala University of Technology Isan Khon Kaen Campus, Khon Kaen, 40000, Thailand
[2] Protein and Proteomics Research Centre for Commercial and Industrial Purposes (ProCCI), Department of Biochemistry, Faculty of Science, Khon Kaen University, Khon Kaen 40002, Thailand
[3] Department of Biology, Faculty of Science, Khon Kaen University, Khon Kaen 40002, Thailand
[4] Department of Microbiology, Faculty of Science, Khon Kaen University, Khon Kaen 40002, Thailand
[5] Department of Chemistry, Faculty of Science, Nakhon Phanom University, Nakhon Phanom, 48000, Thailand
[6] Chulabhorn International College of Medicine, Thammasat University, Rangsit Centre, Pathumthani 12121, Thailand
[7] Department of Biology, Faculty of Science, Ramkhamhaeng University, Bangkok, 10240, Thailand
[8] Department of Clinical Chemistry, Faculty of Associated Medical Sciences, Khon Kaen University, Khon Kaen, 40002, Thailand
[9] Division of Pharmacognosy and Toxicology, Faculty of Pharmaceutical Sciences, Khon Kaen University, Khon Kaen 40002, Thailand

* Corresponding author, tel./fax: +66-43-342911; e-mail: sakdad@kku.ac.th

Received: 24 August 2017 / Accepted: 9 March 2018 / Published: 7 September 2018

Abstract: Based on the recently identified sequence of Heteroscorpine-1 (HS-1), the genes encoding the scorpine family of toxins from three *Heterometrus* spp. in Thailand were amplified by polymerase chain reactions (PCR). All three genomic DNAs encoded a precursor protein consisting of 95 amino acid residues, and these proteins share a highly

Maejo Int. J. Sci. Technol. **2018**, *12*(03), 187-198

conserved structure with other known scorpion long-chain K^+-channel toxins. Based on the data obtained from our specimens, all scorpine family genes possess a 5′ UTR and a region encoding the putative signal peptide. The six cysteine positions are highly conserved in invertebrate defensins and scorpion K^+ channel toxins, suggesting the presence of a common three-dimensional structure with a cysteine-stabilised αβ (CSαβ) motif. Our scorpine gene family contains an intron that is completely identical to the known scorpion K^+-channel blockers in size, consensus junctions, putative branch point, and A + T abundance. We used α-toxin (PDB i.d. 4hhf.1.A), toxin II (PDB i.d.1aho.1.A) and defensin (PDB i.d. 2b68.1) as templates in our analysis of 3-dimensional structures. Residues Asn48 – Lys55 of the mature Heteroscorpine sequence form the α-helix, residues Lys60 – His64 form β-sheet 1, and residues Lys67 – Gly71 form β-sheet 2. Heteroscorpine is predicted to have this structure due to the presence of 3 disulphide bridges. The structure completely follows the typical Cys pattern of invertebrate defensins ($CX_{4\text{-}16}CX_2HCX_{6\text{-}9}GX_1CX_{4\text{-}9}CX_1C$). Thus, the scorpine family of toxins is likely to use the same mechanism to penetrate the bacterial cell wall as defensin peptides from other insect families.

Keywords: defensin, *Heterometrus* spp., Heteroscorpine-1 (HS-1), scorpine, scorpion Venom

INTRODUCTION

Scorpion venom comprises many bioactive peptides including antimicrobial peptides, ion channel-targeted neurotoxins, lipolysis activating peptides, proteases and protease inhibitors [1-5]. Neurotoxins are the main components of scorpion venom [6]. *Heterometrus* scorpions are commonly found in South and Southeast Asia. In Thailand they are called giant, black or Asian forest scorpions. These scorpions are approximately 10 cm long and have a deep-black body. Different species of *Heterometrus* are found in different parts of Thailand. In the present investigation *H. laoticus* from northeastern and northern Thailand and *H. cimrmani* from southern Thailand were studied. Our group has identified and characterised a scorpine toxin from the venom of *H. laoticus* and named it Heteroscorpine-1 (HS-1). The toxin exhibits potential bactericidal activity [7].

HS-1 is a long-chain neurotoxin and a member of the scorpine family with potassium channel blocking and defensin activity. It comprises 95 amino acid residues. Six cysteine residues were predicted to form 3 disulphide bridges. Alignment of the HS-1 amino acid sequence with Panscorpine and Opiscorpine showed 82% and 78% similarity respectively, while alignment with many potassium channel blockers showed approximately 40% similarity [8]. The present study successfully identifies 2 genes in the scorpine family that are similar to *HS-1*, from two other sources of scorpions: *HS-2* from the venom of *H. laoticus* from northern Thailand and *HS-3* from the venom of *H. cimrmani* from southern Thailand. Three-dimensional (3D) structures of these Heteroscorpines were computationally modelled to clarify plausible structures that are necessary for toxin function.

Maejo Int. J. Sci. Technol. **2018**, *12*(03), 187-198

MATERIALS AND METHODS

Preparation of Genomic DNA

Northeastern *H. laoticus* and northern *H. laoticus* and southern *H. cimrmani* were cold-shocked at 4°C for at least 3 hr. Tail sections were then dissected and the tail muscles were removed with scissors and forceps. Genomic DNA was extracted from 50 mg of the tail muscles of each species by phenol–chloroform extraction and ethanol precipitation [9]. The sample was digested with 2 mL of 400 μg/mL proteinase K (Sigma, USA) at 55°C overnight. An equal volume of phenol saturated buffer was added and gently mixed. After centrifuging at 3000 rpm for 20 min., an equal volume of lysis buffer (10 mM Tris, pH 8.0; 150 mM NaCl; 10 mM EDTA, pH 8.0; and 0.5% sodium dodecyl sulphate) saturated with phenol/chloroform (1:1) was added to the sample. DNA in the upper phase was precipitated with 2 volumes of cold ethanol and washed with 95% ethanol, followed by air drying. Approximately 1 mg of the genomic DNA was resuspended in TE buffer (10 mM Tris, pH 8.0 and 10 mM EDTA, pH 8.0) and stored at -30°C until use.

Primer Design and Polymerase Chain Reactions (PCR)

Degenerate and non-degenerate primers were designed based on the amino acid sequence of HS-1 [7]. The PCR reaction mixture contained 2 ng of genomic DNA template, 0.4 mM dNTPs, 0.2 mM forward primer, 0.2 mM reverse primer, 4 mM $MgCl_2$ and 1.25 units of *Taq* polymerase (Fermentas, USA). PCR reactions were performed as follows: an initial denaturation step at 94°C for 5 min., 40 cycles of denaturation at 94°C for 30 sec., annealing at 45-55°C for 30 sec., elongation at 72°C for 2 min., and a final elongation step at 72°C for 7 min.

DNA Sequencing

After the DNA was resolved on an agarose gel, PCR products were purified using a QIAquick gel extraction kit (Qiagen, USA) according to the manufacturer's instructions. The DNA was eluted with a specific volume of TE buffer before being subsequently cloned into the pGEM easy vector (Promega, USA). The nucleotide sequence of the construct was verified using the automated dideoxy method and the MegaBACE DNA Analysis System (Amersham Biosciences, UK), at the Biomolecular Analysis Service Unit, Department of Biochemistry, Faculty of Medicine, Khon Kaen University.

Bioinformatics

All bioinformatics tools used in this study are freely accessible online. The DNA sequence was translated into the corresponding amino acids using Translate Program [10]. The nucleotide (*HS-1*, *HS-2*, *HS-3*, *Opiscorpine*, *KTX2* and BmTXKS2) and amino acid sequence (HS-1, HS-2 and HS-3) were aligned using the online software BLAST [11] and ClustalW [12]. The introns were determined by sequence alignment with the complete sequence of Opiscorpine [8]. The sequence similarity was determined using the Protein BLAST program [11]. The construction and analysis of the phylogenetic tree and sequence comparisons were performed using the CLUSTAL-X software from the PHYLIP Package and the TREE-VIEW program from Glasgow University [13, 14]. Cysteine patterns were determined using the DiANNA 1.1 program [15].

Maejo Int. J. Sci. Technol. **2018**, *12*(03), 187-198

Homology Modelling

The 3D homology models of HS-1, HS-2 and HS-3 were established using known 3D structural coordinates as templates [16]. An automated mode of comparative homology modelling in the Swiss-Model server [17, 18] was used.

RESULTS AND DISCUSSION

Genomic DNA Analysis

The total genomic DNA was extracted from the tail sections of northeastern *H. laoticus*, northern *H. laoticus* and southern *H. cimrmani*. PCR products were sequenced; the *HS-1*, *HS-2* and *HS-3* genes were 1443, 1433 and 1445 bp respectively. All three genes contained two exons and one intron, similar to other genes in the scorpine family [7, 8]. After an analysis using BLAST and ClustalW, *HS-1* showed 98% similarity to *HS-2* and 96% similarity to *HS-3*, while *HS-2* showed 96% similarity to *HS-3*. The base pair variations were mainly located in the intron region; a few variants were located in the exon region (Figure 1).

The three toxin genes encoded 95-amino-acid peptides. The toxin gene structures contained four parts: 5′ UTR, ORF, intron and 3′ UTR (Figure 2). The nucleotide sequence at the 5′ splice site contained AAAA and TTTT elements (Figure 1), indicating a splicing recognition site [19]. The flanking regions of the initiation codon **ATG** of *HS-1*, *HS-2* and *HS-3* were AAG**ATG**A, AAG**ATG**A and AAT**ATG**A respectively. *HS-1* and *HS-2* showed similarity in flanking regions, but had different flanking regions from *HS-3*. The AA(A/C)**ATG**A sequence is conserved among scorpion toxins acting on ion channels (Na^+, K^+ and Cl^-), including genes from *Drosophila* spp. [20]. However, this study clearly defines these sequences as AA(G/T)**ATG**A, with a G/T substituting for A/C. These sequences are relatively uncommon in the mechanism of translation initiation.

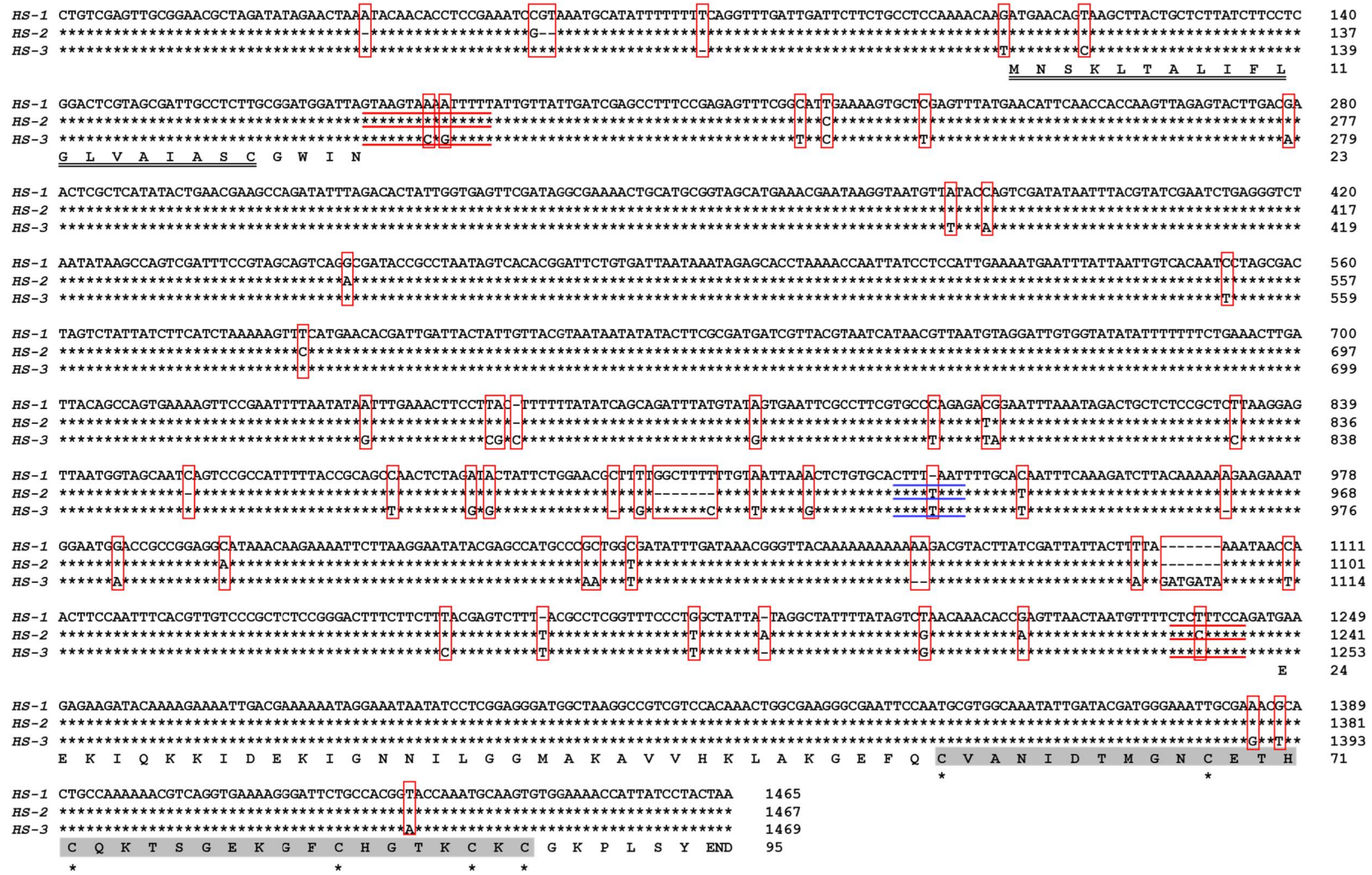
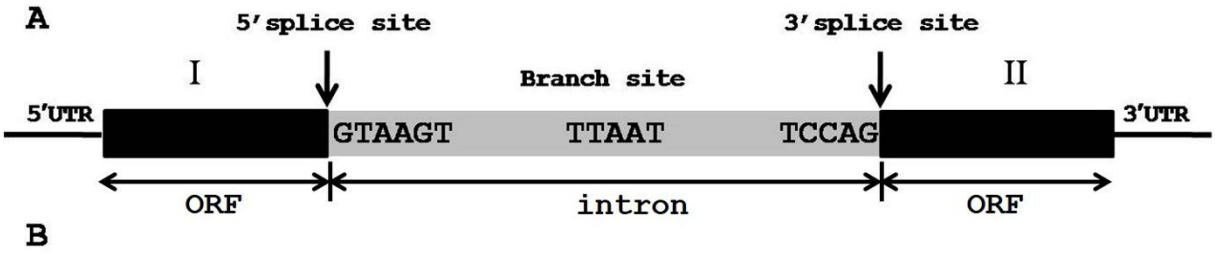


Figure 1. The nucleotide sequence of genomic DNA of *Heterometrus* spp. encoding Heteroscorpine-1, -2 and -3 (HS-1, -2 and -3) precursors. Substituted nucleotides are boxed. Splicing-factor binding regions are underlined in red; the branch site is underlined in blue. Flanking regions of the initiation codon **ATG** in *HS-1*, *-2* and *-3* are AAG**ATG**A, AAG**ATG**A and AAT**ATG**A, respectively. The deduced amino acid sequence is presented below the nucleotide sequence of exon I and II. The signal peptide is underlined. The typical Cys pattern of invertebrate defensins ($CX_{4-16}CX_2HCX_{6-9}GX_1CX_{4-9}CX_1C$) is highlighted in grey.

Maejo Int. J. Sci. Technol. **2018**, *12*(03), 187-198



Gene	5′splice site	Branch site	3′splice site	Ref.
HS-1	gtaagtaaaa	ctttaat	ctctttccag	This study
HS-2	gtaagtaaaa	ttttaat	ctccttccag	This study
HS-3	gtaagtacag	ttttaat	ctctttccag	This study
Opiscorpine	ggaagtagtc	gtatt	ttctttccag	[8]
KTX2	gtaattacga	cattaat	aatattttag	[14]
BmTXKS2	gtaagttaat	ctttaat	ttaaatttag	[15]
Consensus	gtaagtnnnn	ynttaat	nnnattttag	

Figure 2. Organisation of the genomic sequences of *HS-1*, *HS-2* and *HS-3*: (A) model of the three genes. Exons I and II are highlighted with black boxes; (B) their introns are compared to a member of the scorpine family (*Opiscorpine*) and K^+ channel blockers (*KTX2* and *BmTXKS2*). For "Consensus", y is T or C and n is A, T, G or C.

Intron Analysis

Adenine (A) and thymine (T) are enriched in throughout intron sequences and could help the splicing factors locate 3' splice sites. These bases may play a non-specific role in limiting the secondary structures or reducing the likelihood that AG dinucleotides will occur in this region [21]. The intron has a consensus GT-AG splice junction and a branch site region that shares sequence similarity (Figure 2A). *HS-1*, *HS-2* and *HS-3* introns were compared to other scorpion K^+ channel toxin genes (*Opiscorpine*, *KTX2* and *BmTXKS2*). Introns of the *HS* toxins have the same 5' splicing donor (5'A|gtaagt3'), similar to *BmTXKS2* but different from *Opiscorpine* (5'A|ggaagt3') and *KTX2* (5'A|gtaatt3') (Figure 2B, right panel). Surprisingly, the toxin genes differ in the 3' splice sites (*HS-1*, 5′ctctttccag|A3′; *HS-2*, 5′ctccttccag|A3′; *HS-3*, 5′ctctttccag|A3′) (Figure 2B, upper panel), indicating that position 7 upstream of the 3′ splice site is less conserved. However, a putative branch site (5′ttttaat3′) (Figure 2B, middle panel) that is typically responsible for the identification of the 3′ splice site is similar to the consensus sequences of the animal branch sites (PyNPyPyPuAPy) [22] (Figure 2B).

Maejo Int. J. Sci. Technol. **2018**, *12*(03), 187-198

Amino Acid Sequence Alignments

Nucleotide sequences in the exon region of the three toxins were translated. The sequences encode a 95-amino-acid peptide with seven conserved cysteine residues that potentially form three disulphide bridges and completely match the typical Cys pattern of invertebrate defensins ($CX_{4-16}CX_2HCX_{6-9}GX_1CX_{4-9}CX_1C$) [7]. Although nucleotides vary at some positions between *HS-1*, *HS-2* and *HS-3*, the deduced amino acids of the three toxins show 100% similarity to each other after an analysis with ClustalW (Figure 3). Since the nucleotide substitution occurs at the third position of the codon, the protein sequence is not altered [3, 7] (Figure 1). The deduced amino acids of HS were protein-blasted for similarity. The result showed 43–82% similarity to the scorpion long-chain K^+-channel toxins such as the scorpine precursor Panscorpine, the Opiscorpine-3 precursor, Tityus scorpine, TsTXKβ, BmTXKβ2 and AaTXKβ [7].

```
HS-1  MNSKLTALIFLGLVAIASCGWINEEKIQKKIDEKIGNNILGGMAKAVVHKLAKGEFQCVA 60
HS-3  MNSKLTALIFLGLVAIASCGWINEEKIQKKIDEKIGNNILGGMAKAVVHKLAKGEFQCVA 60
HS-2  MNSKLTALIFLGLVAIASCGWINEEKIQKKIDEKIGNNILGGMAKAVVHKLAKGEFQCVA 60
      ************************************************************
HS-1  NIDTMGNCETHCQKTSGEKGFCHGTKCKCGKPLSY 95  100%
HS-3  NIDTMGNCETHCQKTSGEKGFCHGTKCKCGKPLSY 95  100%
HS-2  NIDTMGNCETHCQKTSGEKGFCHGTKCKCGKPLSY 95  100%
      *********************************
```

Figure 3. Alignment of deduced amino acid sequences of HS-1, HS-2 and HS-3 using the program ClustalW. * indicates identical residues in all sequences of the alignment.

Comparative Analysis

The HS-1 amino acid sequence was used to construct a phylogenetic tree containing K^+ channel blockers and defensins, the sequences of which are available online. The peptide toxins specific for the scorpine precursor branched into distantly related clusters, whereas peptides specific for the K^+ channel toxins are more dispersed. Heteroscorpine (HS) is clustered with K^+ channel toxins including the Opiscorpine-2 precursor, Opiscorpine-3 precursor, Opiscorpine-4 precursor and scorpine-like peptide 1 [8, 23]. However, HS is evolutionarily closely related to the scorpine precursor. All of the peptides share seven conserved cysteines with scorpion K^+-channel toxins. Additionally, the residues involved in the formation of β-strands are mostly conserved.

The unrooted phylogenetic tree of HS, K^+-channel blockers and defensins was analysed based on primary structures. A neighbour-joining tree was constructed based on the sequence alignment of K^+-channel blockers. As shown in the phylogenetic tree, HS clearly shares higher homology with the scorpine precursor, clusters in the same branch with PANIM scorpine, and shows a very close relationship with Opiscorpine-2, SCRP2OPICA, SCRP4OPICA, SCRP3OPICA and SCRP1OPICA (Figure 4A). The full-length sequence contains seven cysteines: Cys19, 58, 68, 72, 82, 87 and 89. The toxin forms three disulphide bridges, viz. Cys58–Cys68, Cys72–Cys87 and Cys82–Cys89 (Figure 4B).

Maejo Int. J. Sci. Technol. **2018**, *12*(03), 187-198

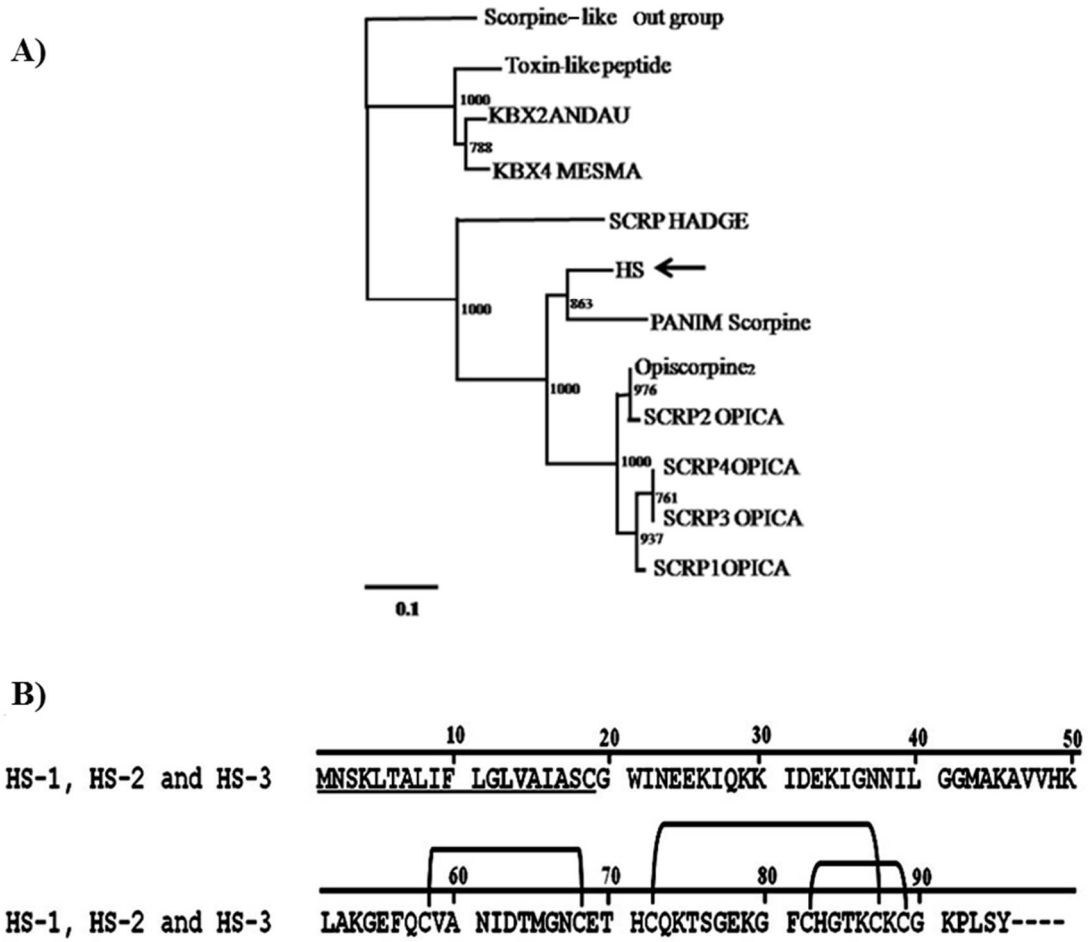


Figure 4. Phylogenetic relationship and disulphide bridge analysis: (A) For the rooted phylogenetic tree of HS-1, HS-2 and HS-3, a neighbour-joining tree is constructed by aligning the sequences of the following K^+-channel blockers: PANIM_Scorpine (P56972.1) [31], Opiscorpine2 (AAQ94353.1) [32], SCRP1OPICA (Q5WR03.1) [8], SCRP2OPICA (Q5WR01.1) [8], SCRP3OPICA (Q5WQZ7.1) [8], SCRP4OPICA (Q5WQZ9.1) [8], SCRPHADGE (Q0GY40) [33], KBX2ANDAU (P69939) [34], KBX4MESMA (Q9N661) [35], toxin-like peptide (ABP35519.1) [36] and scorpine-like peptide (AAW72464.1) [37]. HS is indicated by an arrow. (B) Cysteine patterns are shown. The signal peptide is underlined. The 3 disulphide bridges are Cys58–Cys68, Cys72–Cys87 and Cys82–Cys89.

The 3D structure of HS was generated with molecular modelling using the structures of the proteins with the highest amino acid sequence identities as a template. The model used established 3D structures in the Protein Data Bank. The sequence of HS is 44.00% identical to α-toxin from the scorpion *Odonthobuthus doriae* [16], 40.00% identical to toxin II from the scorpion *Androctonus australis* Hector [24] and 35.48% identical to *Cg*-Def, a defensin from the oyster *Crassostrea gigas* [25]. Residues Asn48–Lys55 of the mature sequence of HS show an α-helical structure, residues Lys60–His64 show a β-sheet-1 structure, and residues Lys67–Gly71 show a β-sheet-2 structure (Figure 5). The helical structure of residues Asn48–Lys55 (Figure 6) is presumed to disrupt the cell membrane and inhibit the activity of microorganisms. HS-1 is a small cationic antimicrobial peptide

Maejo Int. J. Sci. Technol. **2018**, *12*(03), 187-198

[8]. As shown in the molecular model, the HS-1 fold is similar to that of the K^+-channel toxins and defensins. HS has been predicted to inhibit microorganisms by the 'carpet model' [26].

In this model many antimicrobial peptides such as ovispirin [27] accumulate on a bilayer surface and are oriented parallel ('in-plane') to the membrane surface [28]. The peptides are electrostatically attracted to the anionic phospholipid head groups at numerous sites, covering the surface of the membrane in a carpet-like manner. At high concentrations, surface-oriented peptides are thought to disrupt the bilayer, similar to a detergent, eventually leading to micelle formation [29, 30]. At a critical point, the peptides form transient toroidal holes in the membrane, which disintegrates and forms micelles after disruption of the bilayer curvature [31].

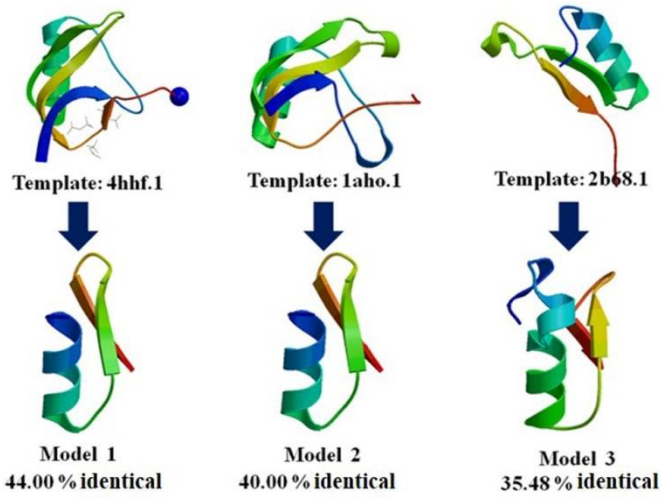


Figure 5. 3D structures of HS-1 based on three templates: α-toxin from the *Odonthobuthus doriae* (PDB i.d. 4hhf.1.A, left panel), toxin II from *Androctonus australis* Hector (PDB i.d. 1aho.1.A, middle panel) and a defensin from oyster *Crassostrea gigas* (PDB i.d.2b68.1, right panel) (resolution 0.96–1.80 Å). Templates are selected using template alignment mode and constructed in automated mode using the Swiss-Model server.

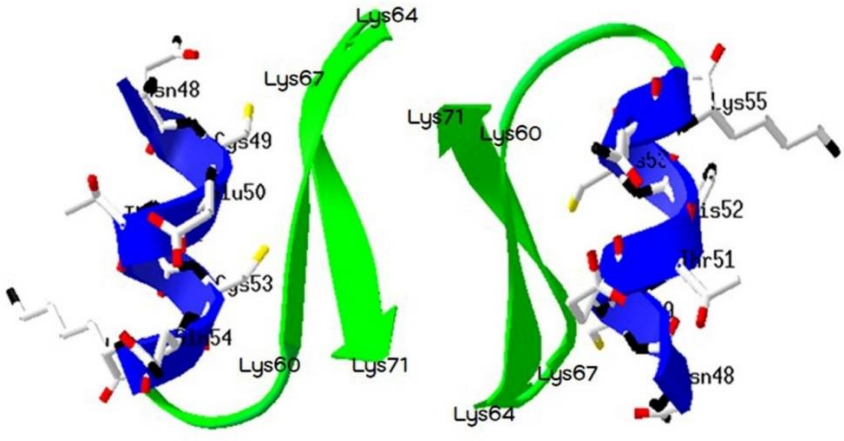


Figure 6. Analysis of HS structure using α-toxin from *Odonthobuthus doriae* (PDB i.d. 4hhf.1.A) as template. HS-1 (mature sequence) comprises one α-helix, a region involved in anti-microorganism activity (residues Asn48–Lys55), and two β-sheets: β-sheet 1 (Lys60–His64) and 2 (Lys67–Gly71). Blue ribbon = α-helix; green ribbon = β-sheet. The 3D structure is constructed with the Swiss-PdbViewer program.

Maejo Int. J. Sci. Technol. **2018**, *12*(03), 187-198

CONCLUSIONS

We have reported the genomic DNA and amino acid sequences of a group of toxins, HS, from the scorpion *Heterometrus*. An analysis of the sequence alignment of HS-1, HS-2 and HS-3 toxins shows 100% similarity between the proteins. The 3D structures share a common dense scaffold formed by one α-helix and two β-strands stabilised by 3 disulphide bridges. These structures exhibit a common minimal motif, named "cysteine-stabilised-αβ (CSαβ)." According to the structural analysis, the C-terminus of the peptide is important; it possesses two structural and functional domains. The first domain is the α-helix that exhibits cytolytic or antimicrobial activity, similar to insect defensins. The second domain is the CSαβ motif, which blocks voltage-gated potassium channels. A comparative analysis of amino acid sequences reveals that the conserved residues of HS align with many members of the K^+ long-chain neurotoxin and defensin families.

ACKNOWLEDGEMENTS

This work was supported by a TRF-CHE Research Grant for Mid-Career University Faculty that was jointly funded by Thailand Research Fund (TRF) and the Commission on Higher Education (CHE), Ministry of Education, with additional support from Khon Kaen University Research Fund. The study was also partially supported by CHE through a grant from Strategic Scholarships for Frontier Research Network for Thai Doctoral Students (CHE-PhD-THA-SUPV) to P.I. and a postdoctoral fellowship to N.U. The authors thank Professor Yukifumi Nawa, the Publication Clinic, Khon Kaen University for critically reviewing the manuscript.

REFERENCES

1. C. Goudet, C. W. Chi and J. Tytgat, "An overview of toxins and genes from the venom of the Asian scorpion *Buthus martensi Karsch*", *Toxicon*, **2002**, *40*, 1239-1258.
2. L. D. Possani, B. Becerril, M. Delepierre and J. Tytgat, "Scorpion toxins specific for Na^+-channels", *Eur. J. Biochem*., **1999**, *264*, 287-300.
3. R. C. R. de la Vega and L. D. Possani, "Current views on scorpion toxins specific for K^+-channels", *Toxicon*, **2004**, *43*, 865-875.
4. R. C. R. de la Vega and L. D. Possani, "Overview of scorpion toxins specific for Na^+ channels and related peptides: Biodiversity, structure-function relationships and evolution", *Toxicon*, **2005**, *46*, 831-844.
5. N. Soudani, J. Gharbi-Chihi, N. Srairi-Abid, C. M.-E. Yazidi, R. Planells, A. Margotat, J. Torresani and M. E. Ayeb, "Isolation and molecular characterization of LVP1 lipolysis activating peptide from scorpion *Buthus occitanus tunetanus*", *Biochim. Biophys. Acta*, **2005**, *1747*, 47-56.
6. S. Zhu and B. Gao**,** "Molecular characterization of a new scorpion venom lipolysis activating peptide: Evidence for disulfide bridge-mediated functional switch of peptides", *FEBS Lett*., **2006**, *580*, 6825-6836**.**
7. N. Uawonggul, S. Thammasirirak, A. Chaveerach, T. Arkaravichien, W. Bunyatratchata, W. Ruangjirachuporn, P. Jearranaiprepame, T. Nakamura, M. Matsuda, M. Kobayashi, S. Hattori and S. Daduang, "Purification and characterization of Heteroscorpine-1 (HS-1) toxin from *Heterometrus laoticus* scorpion venom", *Toxicon*, **2007**, *49*, 19-29.

8. S. Zhu and J. Tygat, "The scorpion family of defensins: Gene structure, alternative polyadenylation and fold recognition", *Cell Mol. Life Sci*., **2004**, *61*, 1751-1763.
9. J. Sambrook, E. F. Fritsch and T. Maniatis, "Molecular Cloning: A Laboratory Manual", 2nd Edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY), **1989**, pp.9.16-9.19.
10. SIB Swiss Institute of Bioinformatics, "ExPASy SIB bioinformatics resource portal", **2011**, http://br.expasy.org (Accessed: December 2015).
11. National Center for Biotechnology Information, "Basic local alignment search tool", **2006**, https://blast.ncbi.nlm.nih.gov/Blast.cgi (Accessed: December 2015).
12. European Bioinformatics Institute (EMBL-EBI), "Multiple sequence alignment", **2005**, http://www2.ebi.ac.uk/clustalw/ (Accessed: December 2015).
13. University of Washington, "PHYLogeny inference package (PHYLIP)", **1980**, http://evolution.genetics.washington.edu/phylip.html (Accessed: December 2015).
14. R. D. Page, "TreeView: An application to display phylogenetic trees on personal computers", *Comput. Appl. Biosci.*, **1996**, *12*, 357-358.
15. R. Russell, "Protein chemistry", http://molbiol-tools.ca/Protein_Chemistry.htm (Accessed: December 2015).
16. T. Durek, I. Vetter, C.-I. Wang, L. Motin, O. Knapp, D. J. Adams, R. J. Lewis and P. F. Alewood, "Chemical engineering and structural and pharmacological characterization of the alpha-scorpion toxin OD1", *ACS Chem. Biol*., **2013**, *8*, 1215-1222.
17. Protein Structure Bioinformatics Group, "Swiss Model", http://www.swissmodel. expasy.org (Accessed: December 2015).
18. T. Schwede, J. Kopp, N. Guex and M. C. Peitsch, "SWISS-MODEL: An automated protein homology-modeling server", *Nucleic Acids Res*., **2003**, *31*, 3381-3385.
19. M. L. Delabre, P. Pasero, M. Marilly and P. E. Bougis, "Promoter structure and intron-exon organization of a scorpion α-toxin gene", *Biochem.*, **1995**, *34*, 6729-6736.
20. D. R. Cavener, "Comparison of the consensus sequence flanking translational start sites in *Drosophila* and vertebrates", *Nucleic Acids Res*., **1987**, *15*, 1353-1361.
21. C. Csank, F. M. Taylor and D. W. Martindale, "Nuclear pre-mRNA introns: Analysis and comparison of intron sequences from *Tetrahymena thermophila* and other eukaryotes", *Nucleic Acids Res*., **1990**, *18*, 5133-5141.
22. B. Lewin, "Genes VI", Oxford University Press, Oxford, **1997**, pp.891-893.
23. E. F. Schwartz, E. Diego-Garcia, R. C. R. de la Vega and L. D. Possani, "Transcriptome analysis of the venom gland of the Mexican scorpion *Hadrurus gertschi* (Arachnida: Scorpiones)", *BMC Genomics*, **2007**, *8*, 119.
24. G. D. Smith, R. H. Blessing, S. E. Ealick, J. C. Fontecilla-Camps, H. A. Hauptman, D. Housset, D. A. Langs and R. Miller, "*Ab initio* structure determination and refinement of a scorpion protein toxin", *Acta Crystallogr. D Biol. Crystallogr*., **1997**, *53*, 551-557.
25. Y. Gueguen, A. Herpin, A. Aumelas, J. Garnier, J. Fievet, J. M. Escoubas, P. Bulet, M. Gonzalez, C. Lelong, P. Favrel and E. Bachère "Characterization of a defensin from the oyster *Crassostrea gigas*. Recombinant production, folding, solution structure, antimicrobial activities, and gene expression", *J. Biol. Chem*., **2006**, *281*, 313-323.
26. Y. Pouny, D. Rapaport, A. Mor, P. Nicolas and Y. Shai, "Interaction of antimicrobial dermaseptin and its fluorescently labeled analogues with phospholipid membranes", *Biochem.*, **1992**, *31*, 12416-12423.

27. S. Yamaguchi, D. Huster, A. Waring, R. I. Lehrer, W. Kearney, B. F. Tack and M. Hong, "Orientation and dynamics of an antimicrobial peptide in the lipid bilayer by solid-state NMR spectroscopy", *Biophys. J.*, **2001**, *81*, 2203-2214.
28. B. Bechinger, "The structure, dynamics and orientation of antimicrobial peptides in membranes by multidimensional solid-state NMR spectroscopy", *Biochim. Biophys. Acta*, **1999,** *1462*, 157-183.
29. Y. Shai, "Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by alpha-helical antimicrobial and cell non-selective membrane-lytic peptides", *Biochim. Biophys. Acta*, **1999**, *1462*, 55-70.
30. A. S. Ladokhin and S. H. White, "'Detergent-like' permeabilization of anionic lipid vesicles by melittin", *Biochim. Biophys. Acta*, **2001**, *1514*, 253-260.
31. Z. Oren and Y. Shai, "Mode of action of linear amphipathic α-helical antimicrobial peptides", *Biopolymers*, **1998**, *47*, 451-463.
32. S. Zhu and J. Tytgat, "A family of naturally-occurring hybrids of Cecropin-insect defensin from *Opistophthalmus Carinatus* (Scorpionidae): Alternative polyadenylation, gene structure, fold recognition and evolutionary origin", Unpublished data.
33. S. Y. Zhu, W. X. Li and X. C. Zeng, "Precursor nucleotide sequence and genomic organization of BmTXKS1, a new scorpion toxin-like peptide from *Buthus martensii* Karsch", *Toxicon*, **2001**, *39*, 1291-1296.
34. C. Legros, B. Ceard, P. E. Bougis and M. F. Martin-Eauclaire, "Evidence of a new class of scorpion toxins active against K^+ channels", *FEBS Lett.*, **1998**, *431*, 375-380.
35. S. Zhu, W. Li, X. Zeng, D. Jiang, X. Mao and H. Liu, "Molecular cloning and sequencing of two 'short chain' and two 'long chain' K(+) channel-blocking peptides from the Chinese scorpion *Buthus martensii* Karsch", *FEBS Lett.*, **1999**, *457*, 509-514.
36. S. Zhu, B. Gao, A. Aumelas, M. D. C. Rodriguez, H. Lanz-Mendoza, S. Peigneur, E. Diego-Garcia, M. F. Martin-Eauclaire, J. Tytgat and L. D. Possani, "MeuTXKbeta1, a scorpion venom-derived two-domain potassium channel toxin-like peptide with cytolytic activity", *Biochim. Biophys. Acta*, **2010**, *1804*, 872-883.
37. E. Diego-Garcia, C. V. Batista, B. I. Garcia-Gomez, S. Lucas, D. M. Candido, F. Gomez-Lagunas and L. D. Possani, "The Brazilian scorpion *Tityus costatus* Karsch: Genes, peptides and function", *Toxicon*, **2005**, *45*, 273-283.