




## RESEARCH ARTICLE

# Structural analysis of an *Asterias rubens* peptide indicates the presence of a disulfide-directed $\beta$ -hairpin fold

Rozita Takjoo<sup>1</sup>, David T. Wilson<sup>1</sup> , Justine Le Quilliec<sup>1,2</sup> , Casey A. Schmidt<sup>1</sup>, Guangzu Zhao<sup>1</sup>, Michael J. Liddell<sup>3</sup> , Naeem Y. Shaikh<sup>4</sup>, Kartik Sunagar<sup>4</sup>, Alex Loukas<sup>1</sup>, Michael J. Smout<sup>1</sup> and Norelle L. Daly<sup>1</sup>

- 1 Australian Institute of Tropical Health and Medicine, James Cook University, Cairns, Australia
- 2 Univ Brest, École Supérieure d'Ingénieurs en Agroalimentaire de Bretagne atlantique, Plouzané, France
- 3 College of Science and Engineering, James Cook University, Cairns, Australia
- 4 Evolutionary Venomics Lab, Centre for Ecological Sciences, Indian Institute of Science, Bangalore, India

## Keywords

cell proliferation assay; disulfide-directed  $\beta$ -hairpin; molecular evolution; NMR spectroscopy; peptide synthesis

## Correspondence

N. L. Daly, Australian Institute of Tropical Health and Medicine, James Cook University, Cairns, Qld 4878, Australia  
 Tel: +61 (07) 4232 1815  
 E-mail: [norelle.daly@jcu.edu.au](mailto:norelle.daly@jcu.edu.au)

Rozita Takjoo and David T. Wilson contributed equally to this article.

(Received 24 May 2024, revised 19 October 2024, accepted 6 November 2024)

doi:10.1002/2211-5463.13931

Edited by Alberto Alape-Girón

Sea stars are an abundant group of marine invertebrates that display remarkably robust regenerative capabilities throughout all life stages. Numerous proteins and peptides have been identified in a proteome study on the coelomic fluid (biofluid) of the common sea star *Asterias rubens*, which appear to be involved with the wound-healing response in the organism. However, the three-dimensional structure and function of several of these injury-responsive peptides, including the peptide KASH2, are yet to be investigated. Here, we show that the KASH2 peptide adopts a disulfide-directed  $\beta$ -hairpin fold (DDH). The DDH motif appears to be evolutionarily related to the inhibitor cystine knot motif, which is one of the most widespread disulfide-rich peptide folds. The DDH motif was originally thought to be restricted to arachnids, but our study suggests that as a result of convergent evolution it could also have originated in sea stars. Although the widely conserved DDH fold has potential cross-phyla wound-healing capacity, we have shown that KASH2 does not enhance the proliferation of human fibroblasts, a simple method for wound-healing re-epithelialisation screening. Therefore, additional research is necessary to determine the role of KASH2 in the sea stars.

Sea stars are invertebrates with exceptional powers of regeneration and the capability to regrow lost limbs [1]. Several molecules have been shown to be involved in regeneration, including the serine protease calpain [2], neurotransmitters (monoamines), neuropeptides (substance P, SALMFamides 1 and 2) and growth-

factor-like molecules [3], and a recent study suggested that there could be many more [4].

Proteomic analysis of cell-free coelomic fluid (biofluid) of the common sea star *Asterias rubens* highlighted several proteins that were upregulated in response to injury and, therefore, are potentially

## Abbreviations

Acm, acetamidomethyl; ACN, acetonitrile; BPP, Bayesian posterior probability; CHCA,  $\alpha$ -cyano-4-hydroxycinnamic acid; DDH, disulfide-directed  $\beta$ -hairpin; DMEM/F12, Dulbecco's modified Eagle medium/nutrient mixture F-12; DMF, dimethylformamide; FA, formic acid; FBS, fetal bovine serum; Fmoc, fluorenylmethoxycarbonyl; HCTU, *O*-(1*H*-6-chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate; ICK, inhibitory cystine knot; MALDI-TOF, matrix-assisted laser desorption ionisation time-of-flight; ML, maximum likelihood; NMR, nuclear magnetic resonance; RP-HPLC, reversed-phase high-performance liquid chromatography; RP-HPLC/MS, reversed-phase high-performance liquid chromatography mass spectrometry; SPPS, solid-phase peptide synthesis; TFA, trifluoroacetic acid; TIPS, triisopropyl silane; Trt, trityl.

involved in defence, cell migration and wound healing [4]. In particular, these studies followed changes that occurred in response to puncture wounds and blood loss. The majority of the upregulated proteins were classified as pattern-recognition receptors or peptidase inhibitors. However, there were also several uncharacterised proteins upregulated, including the disulfide-rich proteins KARTESH (KASH1 and KASH2) and CHUPA (CHU1, CHU2 and CHU3) (Fig. 1). These are novel proteins that are expressed with a signal peptide, an acidic region and the mature peptide region containing four-cysteine residues. Convertase cleavage sites between the acidic region and the cysteine-rich regions result in the predicted mature peptides for the KARTESH peptides containing 28–30 residues and the CHUPA peptides containing 35–48 residues.

Several homologues of the KARTESH and CHUPA proteins are present in other echinoderms, but not in any other phyla. The cysteine residues are conserved throughout the predicted mature peptides, but there is variation in the residue type and the number of residues in the inter-cysteine loops. This phenomenon of families of peptides with inter-cysteine loop sequence variations combined with conservation of the cysteine residues is prevalent throughout nature, including in the venom peptides of cone snails, spiders and sea anemones, with the former containing numerous peptides with four-cysteine residues [5–13].

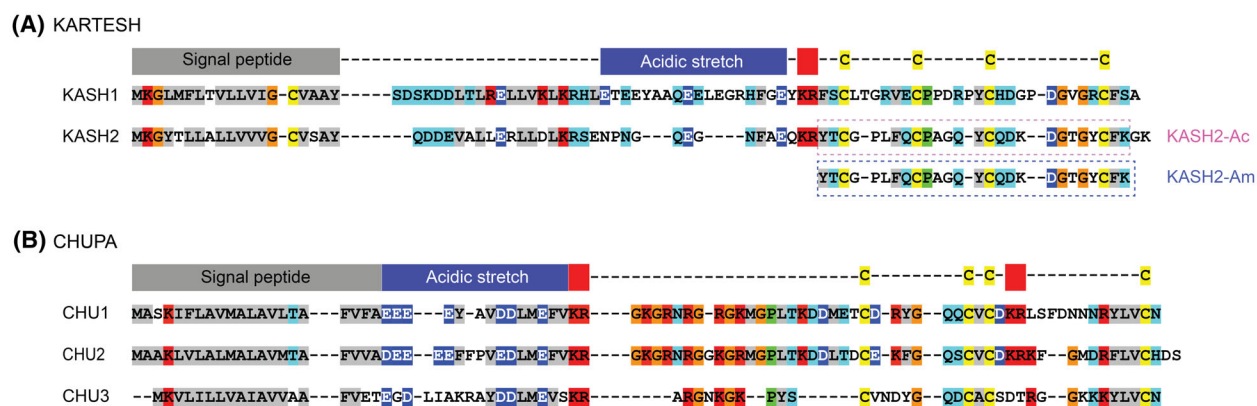
The upregulation of the uncharacterised sea star proteins in response to injury warrants further investigation. In particular, the determination of the three-dimensional structures might provide insight into

their function and evolution. Here, we have determined the three-dimensional structure of KASH2 using NMR spectroscopy. Two constructs of the peptides were studied: one with the predicted sequence with a C-terminal acid (KASH2-Ac) and the other with a C-terminal amide (KASH2-Am), given the conserved amidation signal (Xaa-Gly-Lys/Arg) at the C terminus. The sequences of the KASH2 peptides are given in Fig. 1. Intriguingly, we found that the three-dimensional solution structure contains an ancestral protein fold termed the disulfide-directed  $\beta$ -hairpin (DDH), initially defined by Wang *et al.* [12], and previously thought to be restricted to arachnids. As we have previously seen cross-phyla wound-healing activity from cysteine-rich peptides [14–16], we explored whether the potential for therapeutic human wound-healing properties is associated with this DDH-fold peptide, similar to what we have observed for granulins derived peptides from the human parasitic liver fluke *Opisthorchis viverrini*, with a human fibroblast proliferation assay [14,15,17,18]. However, neither the acid nor amide versions of KASH2 showed any effect on cell proliferation, suggesting that it might be phylum-specific or it might not be directly related to wound healing in the sea star.

## Materials and methods

### Peptide synthesis and purification

The KASH2 acid form, KASH2-Ac and CHUPA1 peptides were synthesised by manual solid-phase peptide synthesis



**Fig. 1.** Sequence alignments of peptides from the *Asterias rubens* transcriptome predicted to display wound-healing properties. (A) KARTESH sequences. (B) CHUPA sequences. Positions of predicted signal peptide, acidic stretch and dibasic convertase cleavage sites are highlighted. The sequences of the synthetic acid and amide versions of KASH2 are also boxed. The mass spectrometry proteomics data and search database associated with these sequences were previously deposited at the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier [PXD010228](https://doi.org/10.1002/2211-5463.13931). The codes for the sequences are as follows: Kartesh-2 UniParc:UPI000FECAB46; Kartesh-1 UniParc:UPI000FEC03F; Chupa-3 UniParc:UPI000FECB650; Chupa-2 UniParc:UPI000EB785BA; Chupa-1 UniParc:UPI000EB634EE.

(SPPS) using Fmoc-SPPS chemistry and were assembled on 2-chlorotriyl chloride resin (Auspep, Tullamarine, VIC, Australia). Amino acids were activated using *O*-(1*H*-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HCTU; Peptides International, Louisville, KY, USA) in peptide grade dimethylformamide (DMF; Auspep). Peptide was cleaved from the resin using a mixture of 95% trifluoroacetic acid (TFA; Auspep)/2.5% triisopropyl silane (TIPS; Sigma-Aldrich, St. Louis, MO, USA)/2.5% H<sub>2</sub>O at room temperature for 2 h [14,16], and the TFA was removed by evaporation. Cleaved peptide was precipitated with ice-cold diethyl ether. Diethyl ether was removed by filtration, and the precipitated peptide was dissolved in 50% acetonitrile/50% H<sub>2</sub>O/0.1% TFA (v/v/v) and subsequently lyophilised. The resulting crude peptide was purified with reversed-phase high-performance liquid chromatography (RP-HPLC) (Agilent 1260 Infinity; Agilent Technologies, Mulgrave, VIC, Australia) on a C<sub>18</sub> preparative column (Phenomenex Jupiter 10 µm, C<sub>18</sub>, 300 Å, 250 mm × 21.2 mm) using an appropriate gradient of solvents A (0.05% TFA in water) and B (0.045% TFA, 90% acetonitrile in water). UV absorbance was monitored at 214 and 280 nm. The purity of the peptides was assessed using analytical RP-HPLC on a C<sub>18</sub> analytical column (Agilent Eclipse Plus C<sub>18</sub>, 3.5 µm, 4.6 × 100 mm or Phenomenex Aeris 3.5 µm PEPTIDE XB-C18, 100 Å, 150 × 2.1 mm) using a gradient of 0–50% solvent B in 50 min, 50–90% solvent B in 5 min, 90–0% solvent B in 5 min at 1 mL·min<sup>-1</sup> or a gradient of 0–60% solvent B in 60 min at 0.25 mL·min<sup>-1</sup>, respectively. The KASH2 peptide with a C-terminal amide, KASH2-Am, was purchased from GenScript (GenScript, Piscataway, NJ, USA), with the cysteine residues in the reduced form.

### Disulfide formation

Both the C-terminal acid and amide versions of KASH2 were oxidised by dissolving lyophilised, purified peptide in 100 mM ammonium bicarbonate (pH 8.2) for 24 h at room temperature. Oxidised peptides were acidified, filtered and purified using RP-HPLC on a C<sub>18</sub> preparative column (Phenomenex Jupiter 10 µm C<sub>18</sub> 300 Å, 250 × 21.2 mm). The purity of the peptides was assessed using analytical RP-HPLC on the same C<sub>18</sub> analytical column and gradient as above. The peptide masses were analysed by matrix-assisted laser desorption ionisation time-of-flight (MALDI-TOF) mass spectrometry using a 5800 MALDI TOF/TOF spectrometer (SCIEX, Framingham, MA, USA). Samples (0.5 µL) were spotted onto a 384-well stainless-steel target plate with 0.5 µL of  $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA; Sigma-Aldrich, St. Louis, MO, USA) matrix. Analysis used reflector positive mode with a scan range of *m/z* 800–4500, averaged over 2000 laser shots. Samples were also analysed using RP-HPLC/mass spectrometry (RP-HPLC/MS) on a Shimadzu LCMS-2020 mass spectrometer combined with a Shimadzu Prominence HPLC system

(Shimadzu, Kyoto, Japan). Samples were injected onto a RP-HPLC column (Phenomenex Aeris 3.6 µm PEPTIDE XB-C18 100 Å, 2.1 × 150 mm; Phenomenex, Torrance, CA, USA) at 30 °C and eluted using a 1% gradient of RP-HPLC/MS solvent B (90% acetonitrile [ACN; OPTIMA LCMS grade, Thermo Fisher Scientific, Scoresby, Vic., Australia]/0.09% formic acid [FA; Sigma-Aldrich]/water) in RP-HPLC/MS solvent A (0.1% FA) from 0% to 60% LC/MS solvent B over 60 min at a flow rate of 0.250 mL·min<sup>-1</sup>. UV absorbance was monitored at 214 and 280 nm, and mass spectra were collected in positive ion mode with a scan range of *m/z* 250–2000, detector voltage of 1.15 kV, nebulising gas flow of 1.5 L·min<sup>-1</sup> and drying gas flow of 3.0 L·min<sup>-1</sup>. Shimadzu LABSOLUTIONS v5.96 software (Shimadzu) was used to collect and analyse data.

Although the native connectivity of KASH2 is not known, to determine the disulfide connectivity present in the major isomer in the one-step oxidative folding of the C-terminal acid version, selective protection of the cysteine residues was carried out. Side-chain protection with acetamidomethyl (Acm) groups was used for Cys II and IV, and trityl (Trt) protecting groups were used for Cys I and Cys III in the synthesis. Following TFA cleavage and deprotection in the presence of scavenger (TIPS), the disulfide bond between Cys I and Cys III was formed by air oxidation in ammonium bicarbonate (pH 8.2) for 24 h, and the peptide was purified using RP-HPLC (as above). The second disulfide bond was formed following removal of the Acm protecting group from Cys II and Cys IV. Cys-Acm peptide was dissolved in 50% aqueous acetic acid (~0.5 mg·mL<sup>-1</sup>), 1 M HCl (0.1 mL·mg<sup>-1</sup> of peptide) was then added, followed immediately by I<sub>2</sub> in 100% aqueous acetic acid until the solution was yellow. The flask was flushed with nitrogen and sealed. The reaction was left at room temperature for 2 h and then quenched by adding 1 M ascorbic acid solution dropwise until the solution was colourless. Finally, the quenched reaction was diluted with RP-HPLC solvent A until the percentage of acetic acid was ~10%. Oxidised peptide was purified using a C<sub>12</sub> semi-preparative column (Phenomenex Jupiter Proteo C<sub>12</sub>, 4 µm, 90 Å, 250 × 10.0 mm) or a C<sub>18</sub> semi-preparative column (Phenomenex Aeris 5 µm PEPTIDE XB-C18, 100 Å, 250 × 10.0 mm). The purity of the peptides was assessed using analytical RP-HPLC on a C<sub>18</sub> analytical column (Agilent Eclipse Plus C<sub>18</sub>, 3.5 µm, 4.6 × 100 mm or Phenomenex Aeris 3.5 µm PEPTIDE XB-C18, 100 Å, 150 × 2.1 mm) using a gradient of 0–50% solvent B in 50 min, 50–90% solvent B in 5 min, 90–0% solvent B in 5 min at 1 mL·min<sup>-1</sup> or a gradient of 0–60% solvent B in 60 min at 0.25 mL·min<sup>-1</sup>, respectively.

### NMR spectroscopy and structure determination

NMR spectra were recorded at 290 K on Bruker Avance III 600 MHz spectrometer (Bruker, Karlsruhe, Germany) equipped with a TCI cryoprobe. Samples were prepared

from lyophilised peptides at concentrations of approximately 0.2 mM in 90% H<sub>2</sub>O/10% D<sub>2</sub>O (D<sub>2</sub>O; Cambridge Isotope Laboratories, Woburn, MA, USA). Two-dimensional spectra including <sup>1</sup>H-<sup>1</sup>H TOCSY, <sup>1</sup>H-<sup>1</sup>H NOESY, <sup>1</sup>H-<sup>1</sup>H DQF-COSY, <sup>1</sup>H-<sup>15</sup>N HSQC and <sup>1</sup>H-<sup>13</sup>C HSQC were used for assignment. TOCSY and NOESY spectra were collected using mixing times of 80 and 200 ms, respectively. A series of one-dimensional proton and TOCSY spectra were recorded following dissolution of the peptides in 100% D<sub>2</sub>O to monitor the rate of exchange of the amide protons. All spectra were analysed using TOPSPIN (Bruker, Billerica, MA, USA) and assigned using CCPNMR based on the approach described by Wüthrich [19,20]. The  $\alpha$ H secondary chemical shifts were determined by subtracting random coil <sup>1</sup>H NMR chemical shifts from the experimental  $\alpha$ H chemical shifts [21]. The 2D NOESY spectra were assigned and an ensemble of structures calculated using the program CYANA [22]. A total of 100 initial structures were calculated using the CYANA program. Torsion-angle restraints predicted using TALOS-N were used in the structure calculations [23]. The disulfide-bond connectivities were included in the final calculations of KASH2-Ac and KASH2-Am. A final ensemble of the 20 lowest energy structures with no distance violations > 0.5 Å or angle violations > 4° was prepared. Structures were visualised using MOLMOL [24]. The structures and chemical shifts of KASH2-Ac (PDB ID: 8V2M, BMRB ID: 31128) and KASH2-Am (PDB ID: 8V2U, BMRB ID: 31129) have been deposited into the Protein Data Bank and the Biological Magnetic Resonance Data Bank.

### Sequence retrieval and alignment

A DDH nucleotide dataset was assembled using sequences from the NCBI-NR database (<https://www.ncbi.nlm.nih.gov/>). KASH2, a peptide from the common sea star (*A. rubens*) and associated with wound-healing and adopting the DDH fold, was used as query sequences for BLAST searches (<https://blast.ncbi.nlm.nih.gov/BlastAlign.cgi>). Sequences obtained were manually curated before alignment. Translated sequences were aligned in MEGA7 using the MUSCLE algorithm [25,26]. Universally conserved cysteines conforming to the DDH motif were used as guides to refine the alignment. Regions with gaps in the alignment of > 50% were trimmed prior to phylogenetic analyses.

### Phylogenetic analysis

The evolutionary relationship between DDH sequences was determined by subjecting the nucleotide dataset to Bayesian and maximum likelihood (ML) analyses. The MPI version of MRBAYES 3.2.6 [27,28] was implemented to perform Bayesian inferences. Analyses were run for a minimum of

two hundred million generations using 12 Markov chains across four runs, sampling every 100th tree. Twenty-five per cent of the sampled trees were discarded as burn-in. The log-likelihood score for each of the saved trees was plotted against the number of generations to assess if the log-likelihood scores of the analyses had reached asymptote. Bayesian posterior probability (BPP) was used to evaluate the branch node support. ML analyses were performed using IQ-TREE v1.6.12 [29,30] with an Edge-proportional partition model coupled with MODELFINDER for tree reconstruction using the best-fit partitioning scheme. Bootstrap was used to evaluate branch node support and analyses were run for 100 bootstrap replicates. All phylogenetic trees were midpoint rooted and were visualised using FIGTREE v1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>).

### Human skin normal fibroblast cells

The human skin normal fibroblast cell line 1BR.3.GN was obtained from a European Collection of Authenticated Cell Cultures (ECACC, Porton Down, UK). The 1BR.3.GN cells were grown and maintained in complete media: Dulbecco's modified Eagle medium/Nutrient Mixture F-12 (DMEM/F12) with GlutaMAX (Life Technologies, Melbourne, Vic., Australia) containing 1 × antibiotic/antimycotic and 1 × GlutaMAX, supplemented with 10% fetal bovine serum (FBS) (Gibco, Glasgow, UK) at 37 °C and 5% CO<sub>2</sub>. Cell proliferation assays were performed with low-nutrient media: DMEM/F12 low-nutrient media supplemented with 0.5% FBS and 1 × antibiotic/antimycotic.

### The real-time xCELLigence cell proliferation assay

Cells were seeded at 5000 cells per well in 150 µL of complete media in E-plates (ACEA Biosciences, San Diego, CA, USA) and grown overnight while monitoring with an xCELLigence SP system (ACEA Biosciences), which monitors cellular events in real time by measuring electrical impedance across gold microelectrodes integrated into the base of tissue culture plates. Cells were washed three times with low-nutrient media prior to the addition of 150 µL of low-nutrient media and incubated for a minimum of 6 h before further treatment. Treatments were prepared at 8.5 × concentration and added to each of six replicate wells in 20 µL, for a final 1 × concentration in a total volume of 170 µL. The xCELLigence system recorded cell indices at intervals of 1 h for 5–6 days following treatment. Readings for the cell index were normalised before treatment, and cell proliferation ratios characterised the relative numbers of cells compared to control cells. GRN<sub>P4A</sub> was used as a positive control in the cell proliferation assay, and a scrambled version of this peptide (ALPYSPSGHGGRSSGTDSVSYQTR), referred to as

Gsp, was used as a negative control. GRN<sub>P4A</sub> is a liver fluke-derived granulin peptide that significantly accelerates proliferation of fibroblasts [15]. Day 4 cell proliferation rates were compared between treatment and control wells, and a one-way ANOVA test was used for multiple comparisons with Holm–Sidak’s correction, using GRAPHPAD PRISM 9.0 (GraphPad Software, San Diego, CA, USA).

## Results

Although several peptides and proteins have been shown to be upregulated in *A. rubens* in response to injury [4], the current study focussed on KASH2 because marine-derived cysteine-rich peptides have been shown to have potential in drug design and structural studies could provide insight into molecular evolution. Furthermore, detailed studies were carried out on KASH2 because our preliminary results on the CHUPA3 peptide indicated that the peptide did not fold into a well-defined structure *in vitro* and therefore would be difficult to discern the most likely native fold.

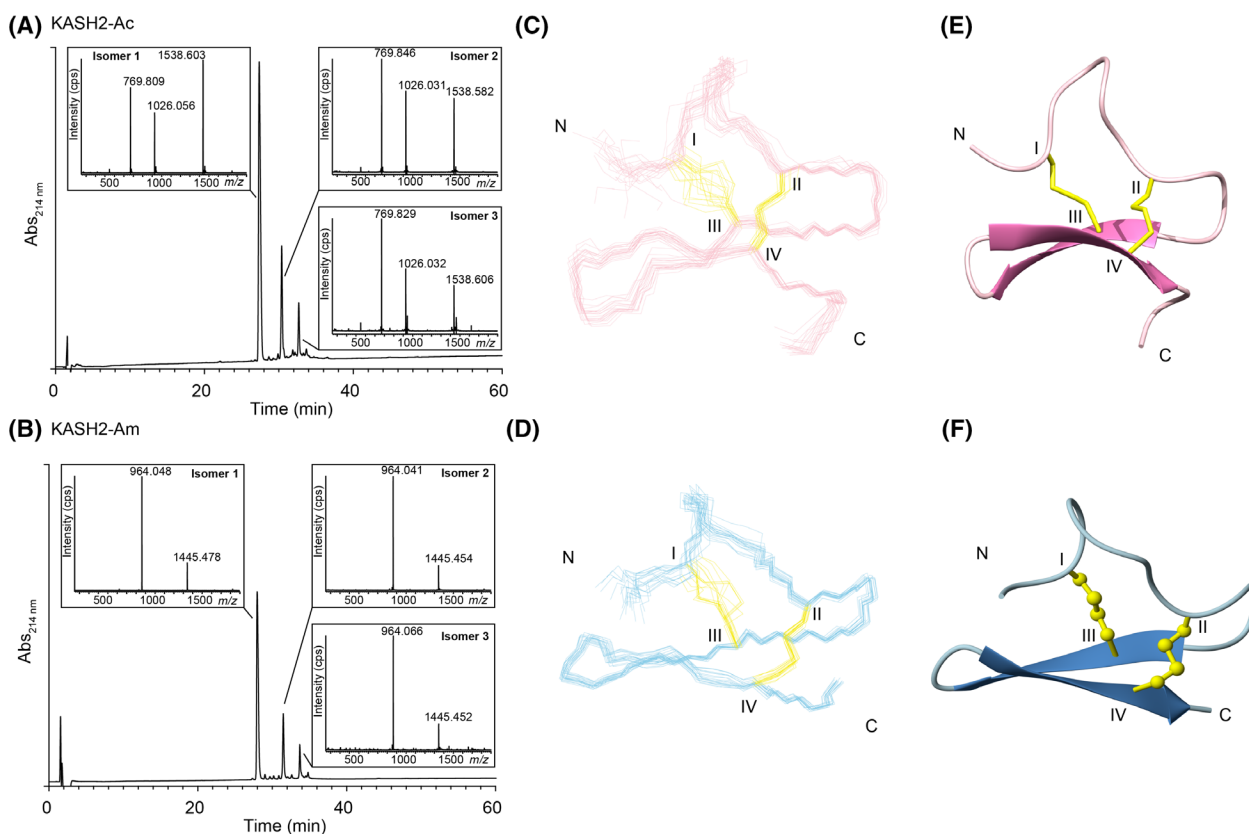
### Peptide synthesis

The sequence corresponding to the expected KASH2 and CHUPA3 mature peptides (Fig. 1) were synthesised using Fmoc-SPPS. It is possible that KASH2 has an amidated C terminus based on the conserved sequence corresponding to an amidation signal, and consequently, the peptide was also synthesised with a C-terminal amide (this peptide has 26 residues rather than 28 for the acid peptide). This is in contrast to KASH1, which appears to have a C-terminal acid based on the absence of the conserved motif associated with amidation. The crude peptides were purified using RP-HPLC, and mass analysis was carried out using MALDI mass spectrometry. Oxidative folding was carried out in a single-step oxidation reaction using ammonium bicarbonate. The oxidation of CHUPA-3 did not produce well-resolved isomers and further study was not carried out on this peptide. By contrast, similar results were observed for both KASH2-Ac and KASH2-Am, and the analytical RP-HPLC trace of the oxidation reactions contained three main peaks at different retention times but the same mass (confirmed by RP-HPLC/MS analysis), indicating the presence of three isomers with different disulfide-bond connectivities, as shown in Fig. 2A,B. The earliest eluting peak, subsequently referred to as isomer 1, was the major isomer, most likely representing the lowest energy/most stable isomer. For both KASH2-Ac and KASH2-Am, all three isomers were subsequently purified using RP-HPLC for further analysis.

### NMR spectroscopy and structure determination

The structures of the KASH2 isomers were analysed using NMR spectroscopy. The pH of the NMR sample solution was ~ 4.5. Spectra also recorded at higher pH (~ 5.5) had greater amide exchange (Fig. S1), which prevented the assignment of several resonances and were not involved in further analysis. For both peptides, the first peak (isomer 1) in the HPLC trace had significant dispersion in the amide region indicative of a well-structured peptide (Fig. S2), in contrast to the other isomers, which had significant overlap in this region. The three-dimensional structure of isomer 1 was determined for both KASH2-Ac and KASH2-Am using torsion-angle dynamics in the CYANA program (see Figs S3–S12 for 2D NMR data). The major element of the secondary structure was a  $\beta$ -hairpin braced by two disulfide bonds. The proline residues appear to be in trans conformation based on the  $\alpha$ - $\delta$  NOEs, and the <sup>13</sup>C chemical shifts for the beta and gamma carbon atoms (Fig. S13), which are consistent with trans isomers based on the study by Schubert *et al.* [31]. These <sup>13</sup>C chemical shifts for isomers 1 of KASH2-Ac and KASH2-Am are given in Table S1. The preliminary structures and slow exchange data were used to infer the hydrogen bonds (Table S2), which were then included in the structure refinements.

Structures were initially calculated with no disulfide-bond restraints, and based on the inter-cysteine distances measured in MOLMOL, the most likely disulfide bond was between Cys II to Cys IV. Consequently, the most likely connectivity involved Cys I–Cys III, Cys II–Cys IV. Structures calculated with this connectivity are shown in Fig. 2C–F. The refinement statistics for the final structures are provided in Table S3. Analysis of the three-dimensional structure of KASH2 indicates that it contains the DDH motif based on the disulfide connectivity with respect to the  $\beta$ -hairpin [12]. Although both forms of the peptide have similar structures, the peptide containing the C-terminal amide has more NOE restraints and consequently provided a more refined structure. Analysis of the NOE restraints between the two structures indicated that the increased number of NOEs occurs throughout the molecule, but in particular involves residues 16 and 18. Therefore, it appears that the C-terminal amide is having an influence on the overall structure, but it should be noted that this peptide is shorter than the acid version, and this might also be having an impact on the structural definition. A more well-defined structure is possibly linked to greater stability, suggesting that the C-terminal amide is present in the native peptide, but this is speculation at this stage.



**Fig. 2.** Oxidative folding of KASH2. RP-HPLC/MS analysis of the oxidation reactions of (A) KASH2-Ac and (B) KASH2-Am. Three main peaks were observed for both peptides, with the first peak being the predominant form. The masses for each of the peaks are given as insets to the figure and show the same mass for each of the isomers within a reaction, indicating they represent disulfide-bond isomers. (C) The three-dimensional structures of the 20 lowest energy structures of KASH2-Ac (PDB: 8V2M), and (D) KASH2-Am (PDB: 8V2U), determined based on NMR data. (E) Ribbon format of KASH2-Ac (PDB: 8V2M), and (F) KASH2-Am (PDB: 8V2U). Cysteines are numbered with Roman numerals.

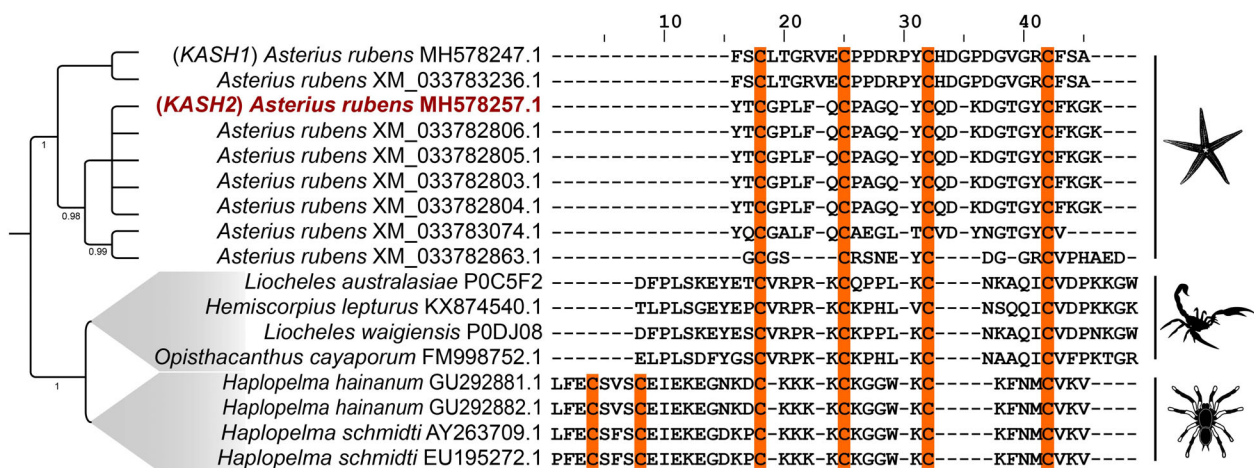
### KASH2 synthesis using selective protection of the cysteine residues

To confirm the disulfide connectivity present in isomer 1 of KASH2, selective protection of the cysteine residues was used to direct the folding to form the predicted disulfide connectivity (Cys I–Cys III and Cys II–Cys IV) (Fig. S14). The experiments were done on the C-terminal acid form. In the synthesis of the peptide, side-chain protection was used for Cys II and IV (Acm) and protecting groups were used for Cys I and Cys III (Trt). Following deprotection, the disulfide bond between Cys I and Cys III was formed by air oxidation and the second disulfide bond was formed following removal of the protecting groups from Cys II and Cys IV. The fully oxidised peptide was subsequently purified using RP-HPLC and shown to have a similar retention time and similar NMR spectra to KASH2-Ac isomer 1 (Fig. S2). A comparison of the chemical shifts is given in Fig. S15.

### Phylogenetic analyses of a DDH peptide identified from the common sea star *A. rubens*

The DDH is a part of the disulfide-rich peptide class of proteins, which includes other members such as the inhibitory cystine knot (ICK) [7,32,33] and the disulfide-stabilised antiparallel  $\beta$ -hairpin stack [7]. The various disulfide linkages in these peptides play a crucial role in determining their native conformation and biological function. Examples of these different classes of disulfide-rich peptides have been found in arachnid toxins [33,34] and have also been detected in other organisms including plants and fungi [35]. However, KASH2 is the first peptide to contain the DDH fold reported from outside the arachnid lineage and warrants further investigation to understand the evolutionary origin and diversification of disulfide-rich peptides.

In our in-depth searches against various sea star genomes and other sequences from public repositories, we could not retrieve homologues for KASH2. Very



**Fig. 3.** The Bayesian phylogeny and sequence alignment of DDH peptides from *Asterias rubens* and Arachnida. This figure depicts the representative Bayesian relationship between sea star (*A. rubens*) and arachnid DDH sequences along with, sequence alignment of the mature DDH peptides from *A. rubens* and Arachnida. Here, the KASH2 peptide from *A. rubens* is shown in red, while the cysteine residues involved in the formation of the DDH fold are highlighted in orange.

limited positive hits in certain sea star genomes turned out to be characterised by very poor coverage, and therefore, we could not use these in our analyses. The only exceptions were the sequences that we retrieved from the *A. rubens* genome.

To understand their evolutionary history, we performed phylogenetic reconstructions of all known peptides from arachnids and sequences retrieved from *A. rubens* containing the DDH motif. We observed that the DDH-containing sequences identified from *A. rubens* form a distinct clade from the known arachnid DDH sequences (Figs S16 and S17, with highlights shown in Fig. 3). When compared and aligned with the known DDH sequences from arachnids, we found that the sequences from *A. rubens* are relatively short in length, and share very little sequence similarity with their arachnid DDH counterparts, except for the four corresponding cysteine residues. The three-dimensional structure of KASH2 is consistent with the DDH motif, as it adopts the typical disulfide bridging [C1–C3] [C2–C4] pattern, similar to U<sub>1</sub>-LITX-Lw1a identified from the scorpion *Liocheles waigiensis* (Fig. 3).

Based on our phylogenetic analyses and the sequence comparisons of DDH-containing peptides, we conclude that the presence of DDH motifs found in the European or common sea star is likely a result of convergent evolution (Figs S16 and S17).

### Cell proliferation monitoring in real time using xCELLigence

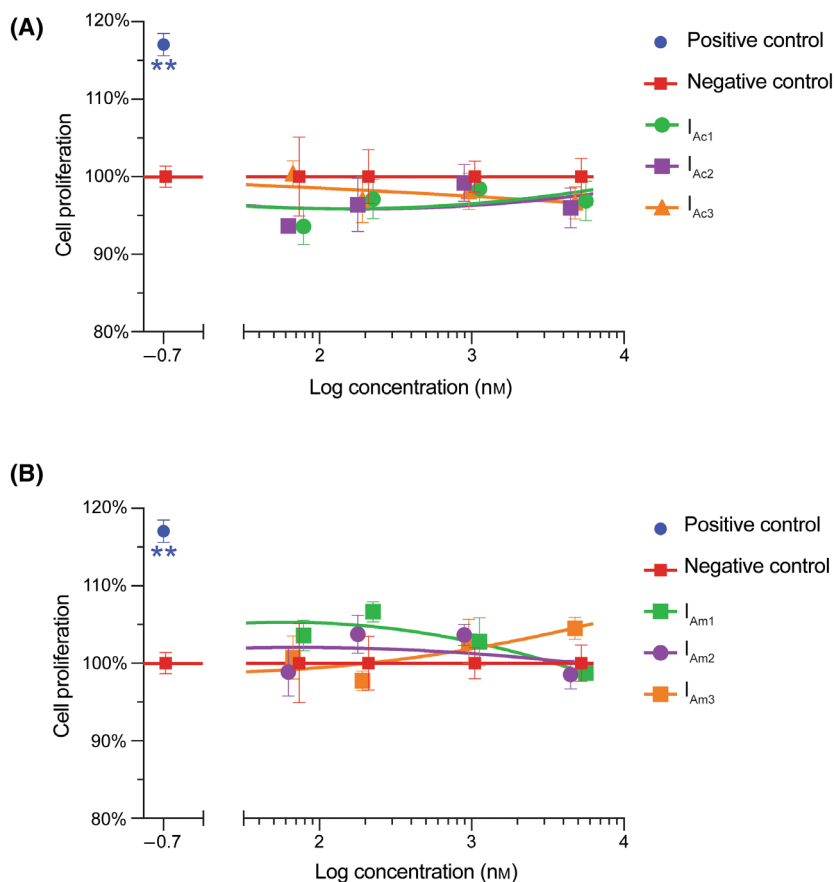
The effect of both KASH2-Ac and KASH2-Am peptides on the stimulation of an *in vitro* cell

proliferation assay of 1BR.3.GN fibroblasts was carried out using an xCELLigence system at the range of concentrations (0.32–1000 nM) at peak cell index (~120 h/5 days). The KASH2 peptides did not show significant cell proliferation on human fibroblast cell growth over the concentration range tested. The positive control, GRN<sub>p4A</sub>, showed approximately 115% cell proliferation relative to the negative control as shown in Fig. 4 and was consistent with previous results [15].

### Discussion

Several proteins from *A. rubens* have recently been shown to be upregulated in response to injury, including proteins with no known function or structural information [4]. Determination of the three-dimensional structure of one of the mature peptides predicted from the identified precursor proteins has provided insight into the molecular evolution of a disulfide-rich structural motif.

The synthetic version of KASH2 with a C-terminal acid produced three isomers during the oxidative folding reaction, which most likely correspond to the three possible disulfide connectivities for a four-cysteine containing peptide. For small disulfide-rich peptides, it is quite common for the earliest eluting peak on RP-HPLC to correspond to the natively folded peptide. This appears to be the case for KASH2, as the earliest eluting peak was the major form and displays a well-defined structure. Preliminary structural analyses of this isomer suggested that the disulfide connectivity was Cys I–Cys III and Cys II–Cys IV, and this was subsequently



**Fig. 4.** The influence of KASH2 peptides on the cell proliferation of human fibroblasts. Human fibroblast cell line was incubated with the (A) three KASH2-Ac and (B) KASH2-Am isomers and analysed using an xCELLigence system over 120 h. Cell proliferation is plotted relative to Gsp (scrambled granulin) controls with five replicates and pooled from four separate experiments ( $n = 4$ ). Data as analysed with a one-way ANOVA with Holm–Sidak multiple comparison against Gsp control: \*\* $P < 0.01$ , unlabelled = not significant.

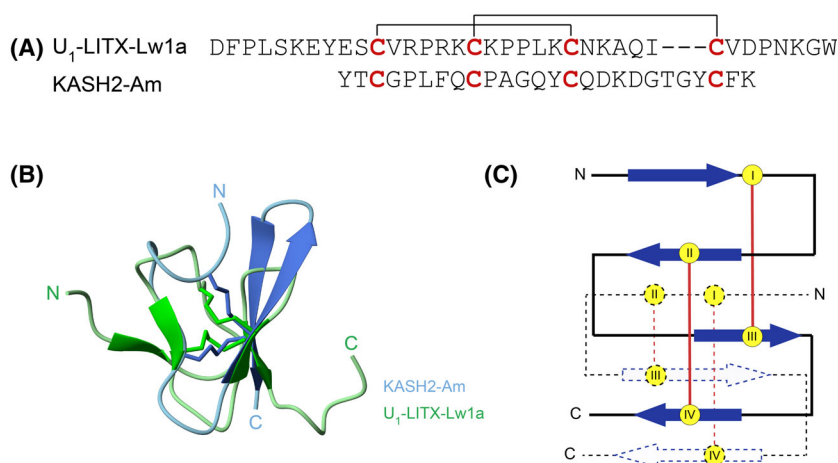
confirmed by selective protection/deprotection oxidation of the cysteine residues.

Similar results were observed in the oxidative folding of the KASH2 peptide with the C-terminal amide with two less residues than the C-terminal acid form. The structures of the major isomers of KASH2-Ac and KASH2-Am peptides were very similar in overall fold, but the amidated version was more well-defined than the acid version. Peptides with amidated C-termini are widespread throughout nature, particularly in disulfide-rich peptides from cone snail venom. C-terminal amidation of an  $\alpha$ -conotoxin has previously been shown to stabilise the secondary structure [36], similar to the results observed in the current study. For the  $\alpha$ -conotoxin, it was suggested that this amide stabilisation played a role in receptor binding rather than interacting directly with the binding site.

Analysis of the three-dimensional structure of KASH2 with the VAST [37] structural similarity program indicated that KASH2 does not have significant structural similarity to any characterised peptides. However, many toxin peptides are known to contain disulfide-rich motifs, and we noticed that KASH2 has

the same disulfide connectivity [I–III, II–IV] and similar fold to the scorpion toxin U<sub>1</sub>-LITX-Lw1a [10] as shown in Fig. 5. U<sub>1</sub>-LITX-Lw1a has more residues than KASH2 and contains an additional  $\beta$ -hairpin, but the RMSD between residues 17–23, 28–30 in U<sub>1</sub>-LITX-Lw1a, and 9–15, 23–25 in KASH2 is 0.532 ( $\text{\AA}$ ), highlighting the similarity between the peptides. The single-domain DDH fold was first reported in the native toxin peptide, J-ACTX-Hv1c, from the Australian funnel-web spider [8,12] and has subsequently been shown in U<sub>1</sub>-LITX-Lw1a [10].

In terms of structural evolution of peptides such as KASH2, the Arachnida and Asteroidea have been phylogenetically separated for nearly 560 million years. It is likely that the two lineages adopted the DDH scaffold subsequent to their divergence from the most recent common ancestor. Further analyses, such as gene synteny and the identification of similar disulfide-rich peptides along the phylogenetic tree leading to Arachnida and Asteroidea, would help in delineating the molecular phylogeny more precisely. However, with the existing evidence, the independent origin of DDH in sea



**Fig. 5.** Conserved structural fold between sea star *Asterias rubens* and scorpion *Liocheles waigiensis* peptides. (A) Alignment of the sequences with cysteine connectivities shown with lines. (B) Overlay of the three-dimensional structures of U<sub>1</sub>-LITX-Lw1a PDB ID: 2KYJ and KASH2-Am, superimposed over the cysteine residues. (C) Graphical representation of the DDH motif, with the  $\beta$ -sheets shown as arrows and cysteine residues labelled with Roman numerals. The KASH2 DDH motif is shown using a dashed line.

star and arachnids seems to be the most parsimonious explanation.

There has been speculation regarding the evolution of the ICK motif, with one hypothesis being that the DDH motif is an ancestral fold of the ICK motif [12]. However, a subsequent study suggested that although the structural data of U<sub>1</sub>-LITX-Lw1a support an ancestral link between the ICK and DDH motifs, molecular evolution data and the apparent restriction of single-domain DDH peptides to arachnids suggested that DDH is a derived ICK and that this derivation has occurred on multiple occasions [33]. Consequently, our finding of a single-domain DDH peptide in sea stars might have implications for future studies on the evolution of the ICK fold.

Sea stars have been shown to have remarkable abilities to regenerate tissue [38,39]. Our team explored evolutionary conserved structural elements across phyla that have human wound-healing potential. As part of our exploration of cross-phyla structure–function relationships, it was of interest to investigate the biological activity of KASH2 in terms of human cell proliferation. The lack of activity of KASH2 on the proliferation of human fibroblast rules out the cross-phyla structural wound-healing aspect. It is possible that KASH2 is either not directly related to wound healing in the sea star or is part of a proteome that activates receptors not involved in human wound healing. In order to further explore the potential bioactivities associated with this peptide, collection of coelomic fluid (containing coelomocytes) from live animals and measuring cell growth *in vitro* immediately after collection and after

addition of KASH2 might be considered as a preliminary screen. However, detailed investigations would be required to understand the actual functions of KASH2 in wound healing in *A. rubens*.

## Conclusions

Overall, we have identified the first DDH structural motif peptide characterised from the sea star *A. rubens*. We have significantly expanded the sequence diversity known to adopt this fold by placing DDH in both the Echinodermata and Arthropoda phyla. Although KASH2 was upregulated in response to injury, we show here that it did not have any effect on the cell proliferation of human fibroblasts. There are numerous other roles it might play in injury in *A. rubens*, which remain to be determined. Disulfide-rich scaffolds, such as the DDH fold, have previously been shown to be useful as scaffolds in drug design and agricultural applications [9,40–42], and the KASH2 peptide might add to the library of peptides available for these applications.

## Acknowledgements

This research was partially supported by the James Cook University NMR facility funded by the Australian Research Council (NLD, AL) (LE160100218), an Australian Research Council Future Fellowship (FF110100226) awarded to NLD, funding from the Merchant Foundation (AL, MJS, NLD), and National Cancer Institute (NCI), US National Institutes of

Health (NIH) award R01CA164719 (MJS). RT would like to thank James Cook University for a PhD scholarship. KS was supported by DBT/Wellcome Trust India Alliance Fellowship grant IA/I/19/2/504647.

## Conflict of interest

The authors declare no conflict of interest.

## Data accessibility

All data are available in the main text or the [Supporting Information](#). The coordinates of the structures of the KASH2-Ac and KASH2-Am peptides are available from the Protein Data Bank and the Biological Magnetic Resonance Data Bank with accession numbers (PDB ID: [8V2M](#), BMRB ID: 31128) and (PDB ID: [8V2U](#), BMRB ID: 31129), respectively.

## Author contributions

RT and NLD contributed to conceptualisation. RT, NLD, DTW and MJS contributed to methodology. RT, NLD, DTW, JLQ, MJS, CAS, NYS and KS contributed to investigation. RT and NLD contributed to visualisation. NLD, DTW, MJS and MJL contributed to supervision. RT and NLD contributed to writing—original draft. RT, NLD, DTW, MJS, AL, CAS, NYS, KS and MJL contributed to writing—review and editing.

## References

- Gupta H, Srivastava A, Menon GR, Agrawal CS, Chumber S and Kumar S (2008) Comparison of interrupted versus continuous closure in abdominal wound repair: a meta-analysis of 23 trials. *Asian J Surg* **31**, 104–114.
- Franco CF, Soares R, Pires E, Santos R and Coelho AV (2012) Radial nerve cord protein phosphorylation dynamics during starfish arm tip wound healing events. *Electrophoresis* **33**, 3764–3778.
- Thorndyke MC and Carnevali C (2001) Regeneration neurohormones and growth factors in echinoderms. *Can J Zool* **79**, 1171–1208.
- Shabelnikov SV, Bobkov DE, Sharlaimova NS and Petukhova OA (2019) Injury affects coelomic fluid proteome of the common starfish, *Asterias rubens*. *J Exp Biol* **222**, jeb198556.
- Chen J, Xu Y, San M, Cao Z, Li W, Wu Y and Chen Z (2015) Cloning and genomic characterization of a natural insecticidal peptide LaIT1 with unique DDH structural fold. *J Biochem Mol Toxicol* **29**, 207–212.
- Cremonese CM, Maiti M, Peigneur S, Cassoli JS, Dutra AA, Waelkens E, Lescrinier E, Herdewijn P, de Lima ME, Pimenta AM *et al.* (2016) Structural and functional elucidation of peptide Ts11 shows evidence of a novel subfamily of scorpion venom toxins. *Toxins (Basel)* **8**, 288.
- Pineda SS, Chin YK, Undheim EAB, Senff S, Mobli M, Dauly C, Escoubas P, Nicholson GM, Kaas Q, Guo S *et al.* (2020) Structural venomomics reveals evolution of a complex venom by duplication and diversification of an ancient peptide-encoding gene. *Proc Natl Acad Sci USA* **117**, 11399–11408.
- Shu Q, Lu SY, Gu XC and Liang SP (2002) The structure of spider toxin huwentoxin-II with unique disulfide linkage: evidence for structural evolution. *Protein Sci* **11**, 245–252.
- Smith JJ, Herzig V, King GF and Alewood PF (2013) The insecticidal potential of venom peptides. *Cell Mol Life Sci* **70**, 3665–3693.
- Smith JJ, Hill JM, Little MJ, Nicholson GM, King GF and Alewood PF (2011) Unique scorpion toxin with a putative ancestral fold provides insight into evolution of the inhibitor cystine knot motif. *Proc Natl Acad Sci USA* **108**, 10478–10483.
- Sunagar K, Undheim EA, Chan AH, Koludarov I, Munoz-Gomez SA, Antunes A and Fry BG (2013) Evolution stings: the origin and diversification of scorpion toxin peptide scaffolds. *Toxins (Basel)* **5**, 2456–2487.
- Wang X, Connor M, Smith R, Maciejewski MW, Howden ME, Nicholson GM, Christie MJ and King GF (2000) Discovery and characterization of a family of insecticidal neurotoxins with a rare vicinal disulfide bridge. *Nat Struct Biol* **7**, 505–513.
- Shafee T, Mitchell ML and Norton RS (2019) Mapping the chemical and sequence space of the ShKT superfamily. *Toxicon* **165**, 95–102.
- Bansal PS, Smout MJ, Wilson D, Cobos Caceres C, Dastpeyman M, Sotillo J, Seifert J, Brindley PJ, Loukas A and Daly NL (2017) Development of a potent wound healing agent based on the liver fluke granulin structural fold. *J Med Chem* **60**, 4258–4266.
- Dastpeyman M, Bansal PS, Wilson D, Sotillo J, Brindley PJ, Loukas A, Smout MJ and Daly NL (2018) Structural variants of a liver fluke derived granulin peptide potently stimulate wound healing. *J Med Chem* **61**, 8746–8753.
- Takjoo R, Wilson D, Bansal PS, Loukas A, Smout MJ and Daly NL (2020) Folding of truncated granulin peptides. *Biomolecules* **10**, 1152.
- Smout MJ, Laha T, Mulvenna J, Sripra B, Suttiprapa S, Jones A, Brindley PJ and Loukas A (2009) A granulin-like growth factor secreted by the carcinogenic liver fluke, *Opisthorchis viverrini*, promotes proliferation of host cells. *PLoS Pathog* **5**, e1000611.

- 18 Smout MJ, Sotillo J, Laha T, Papatpremsiri A, Rinaldi G, Pimenta RN, Chan LY, Johnson MS, Turnbull L, Whitchurch CB *et al.* (2015) Carcinogenic parasite secretes growth factor that accelerates wound healing and potentially promotes neoplasia. *PLoS Pathog* **11**, e1005209.
- 19 Wuthrich K (2003) NMR studies of structure and function of biological macromolecules (Nobel Lecture). *J Biomol NMR* **27**, 13–39.
- 20 Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J and Laue ED (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696.
- 21 Wishart DS, Bigam CG, Yao J, Abildgaard F, Dyson HJ, Oldfield E, Markley JL and Sykes BD (1995)  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$  chemical shift referencing in biomolecular NMR. *J Biomol NMR* **6**, 135–140.
- 22 Guntert P (2004) Automated NMR structure calculation with CYANA. *Methods Mol Biol* **278**, 353–378.
- 23 Shen Y and Bax A (2015) Protein structural information derived from NMR chemical shift with the neural network program TALOS-N. *Methods Mol Biol* **1260**, 17–32.
- 24 Koradi R, Billeter M and Wuthrich K (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* **14**, 51–55, 29–32.
- 25 Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792–1797.
- 26 Kumar S, Stecher G and Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* **33**, 1870–1874.
- 27 Altekar G, Dwarkadas S, Huelsenbeck JP and Ronquist F (2004) Parallel Metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. *Bioinformatics* **20**, 407–415.
- 28 Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA and Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* **61**, 539–542.
- 29 Nguyen LT, Schmidt HA, von Haeseler A and Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* **32**, 268–274.
- 30 Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A and Jermini LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* **14**, 587–589.
- 31 Schubert M, Labudde D, Oschkinat H and Schmieder P (2002) A software tool for the prediction of Xaa-Pro peptide bond conformations in proteins based on  $^{13}\text{C}$  chemical shift statistics. *J Biomol NMR* **24**, 149–154.
- 32 Nielsen LD, Foged MM, Albert A, Bertelsen AB, Soltoft CL, Robinson SD, Petersen SV, Purcell AW, Olivera BM, Norton RS *et al.* (2019) The three-dimensional structure of an H-superfamily conotoxin reveals a granulin fold arising from a common ICK cysteine framework. *J Biol Chem* **294**, 8745–8759.
- 33 Undheim EA, Mobli M and King GF (2016) Toxin structures as evolutionary tools: using conserved 3D folds to study the evolution of rapidly evolving peptides. *Bioessays* **38**, 539–548.
- 34 Zhu LM, Gao B and Zhu SY (2015) Origin of neurotoxins from defensins. *Sheng Li Xue Bao* **67**, 239–247.
- 35 Zhu S, Darbon H, Dyason K, Verdonck F and Tytgat J (2003) Evolutionary origin of inhibitor cystine knot peptides. *FASEB J* **17**, 1765–1767.
- 36 Ho TNT, Lee HS, Swaminathan S, Goodwin L, Rai N, Ushay B, Lewis RJ, Rosengren KJ and Conibear AC (2021) Posttranslational modifications of alpha-conotoxins: sulfotyrosine and C-terminal amidation stabilise structures and increase acetylcholine receptor binding. *RSC Med Chem* **12**, 1574–1584.
- 37 Gibrat JF, Madej T and Bryant SH (1996) Surprising similarities in structure comparison. *Curr Opin Struct Biol* **6**, 377–385.
- 38 Dai Y, Prithviraj N, Gan J, Zhang XA and Yan J (2016) Tissue extract fractions from starfish undergoing regeneration promote wound healing and lower jaw blastema regeneration of zebrafish. *Sci Rep* **6**, 38693.
- 39 Ben Khadra Y, Ferrario C, Di Benedetto C, Said K, Bonasoro F, Carnevali MD and Sugni M (2015) Wound repair during arm regeneration in the red starfish *Echinaster sepositus*. *Wound Repair Regen* **23**, 611–622.
- 40 King GF (2011) Venoms as a platform for human drugs: translating toxins into therapeutics. *Expert Opin Biol Ther* **11**, 1469–1484.
- 41 King GF and Hardy MC (2013) Spider-venom peptides: structure, pharmacology, and potential for control of insect pests. *Annu Rev Entomol* **58**, 475–496.
- 42 Windley MJ, Herzig V, Dziemborowicz SA, Hardy MC, King GF and Nicholson GM (2012) Spider-venom peptides as bioinsecticides. *Toxins (Basel)* **4**, 191–227.

## Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Fig. S1.** KASH2-Ac at pH 5.5  $^1\text{H}$  NMR (600 MHz, 90%  $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$  v/v) spectra.

**Fig. S2.** KASH2-Ac, -Am  $^1\text{H}$  NMR (600 MHz, 90%  $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$  v/v) spectrum.

**Fig. S3.** KASH2-Ac isomer 1  $^1\text{H}$ - $^1\text{H}$  TOCSY NMR (600 MHz, 90%  $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$  v/v) spectrum.

**Fig. S4.** KASH2-Ac isomer 1  $^1\text{H}$ - $^1\text{H}$  NOESY NMR (600 MHz, 90%  $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$  v/v) spectrum.

**Fig. S5.** KASH2-Ac isomer 1  $^1\text{H}$ - $^1\text{H}$  COSY NMR (600 MHz, 90%  $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$  v/v) spectrum.

**Fig. S6.** KASH2-Ac isomer 1  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR (600 MHz, 90%  $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$  v/v) spectrum.

**Fig. S7.** KASH2-Ac isomer 1  $^1\text{H}$ - $^{13}\text{C}$  HSQC NMR (600 MHz, 90%  $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$  v/v) spectrum.

**Fig. S8.** KASH2-Am isomer 1  $^1\text{H}$ - $^1\text{H}$  TOCSY NMR (600 MHz, 90%  $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$  v/v) spectrum.

**Fig. S9.** KASH2-Am isomer 1  $^1\text{H}$ - $^1\text{H}$  NOESY NMR (600 MHz, 90%  $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$  v/v) spectrum.

**Fig. S10.** KASH2-Am isomer 1  $^1\text{H}$ - $^1\text{H}$  COSY NMR (600 MHz, 90%  $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$  v/v) spectrum.

**Fig. S11.** KASH2-Am isomer 1  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR (600 MHz, 90%  $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$  v/v) spectrum.

**Fig. S12.** KASH2-Am isomer 1  $^1\text{H}$ - $^{13}\text{C}$  HSQC NMR (600 MHz, 90%  $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$  v/v) spectrum.

**Fig. S13.**  $^1\text{H}$ - $^1\text{H}$  NOESY NMR spectra (600 MHz, 90%  $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$  v/v) of KASH2 peptides showing the NOEs consistent with the trans conformation.

**Fig. S14.** Characterisation by RP-HPLC and RP-HPLC/MS of KASH2-Ac produced with selective protection of the cysteine residues.

**Fig. S15.**  $\alpha\text{H}$  Secondary chemical shifts for KASH2 peptides.

**Fig. S16.** A Bayesian phylogeny of DDH motifs.

**Fig. S17.** Maximum likelihood phylogeny of the DDH motif.

**Table S1.** Analysis of the backbone conformation of proline residues in KASH2 isomers.

**Table S2.** Hydrogen bond restraints for KASH2.

**Table S3.** Structural statistics for KASH2.