Supporting Information

Neurotoxic and cytotoxic peptides underlie the painful stings of the tree nettle
_Urtica ferox_

Jing Xie,1# Samuel D. Robinson,1# Edward K. Gilding,1# Sina Jami,1 Jennifer R. Deuis,1 Fabian B. H. Rehm,1 Kuok Yap,1 Lotten Ragnarsson,1 Lai Yue Chan,1 Brett R. Hamilton,2 Peta J. Harvey,1 David J. Craik,1,3 Irina Vetter,1,4* Thomas Durek1,3*

1Institute for Molecular Bioscience, The University of Queensland, Brisbane, QLD 4072, Australia.

2Centre for Microscopy and Microanalysis, The University of Queensland, Brisbane, QLD 4072, Australia.

3Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, The University of Queensland, Brisbane, QLD 4072, Australia.

4School of Pharmacy, The University of Queensland, Brisbane, QLD 4102, Australia.

Material included in this pdf file:

- Supporting Figures S1-S8
- Supporting Tables S1-S4
Figure S1: Peptide mass fingerprinting of β/δ-Uf2a fragments after reduction and alkylation with iodoacetamide and digestion with trypsin. S-carbamidomethylated cysteines are labelled as C$_{cam}$. The precursor ion masses of digested fragments are shown in Da.
Figure S2: RP-HPLC analysis of synthetic Δ-Uf1a. A) Folding of the synthetic polypeptide chain monitored by HPLC. Fully reduced polypeptide (top), crude folding mixture after 4h (middle) and the purified, fully oxidized Δ-Uf1a (bottom). B) Co-elution studies of synthetic and plant-derived Δ-Uf1a. Native peptide isolated from trichomes (top), synthetic peptide (middle), and co-injection of equal amounts of synthetic and plant-derived peptides (bottom). Separations were carried out with a Phenomenex Luna Omega C18 column (50 x 2.1 mm, 1.6 μm, 100 Å) at 40˚C, a flow rate of 0.8 mL/min and using a linear gradient of 5-75% buffer B (90% CH3CN, 0.05% TFA in water) in buffer A (0.05% TFA in water) over 16 min.
Figure S3: NMR structure of Δ-Uf1a. Overlay of the 20 lowest energy structures. Disulfide bonds are shown as yellow sticks.
Table S1. Statistical analysis of Δ-Uf1a NMR structures

| Experimental restraints | Value |
|-------------------------|-------|
| total no. distance restraints | 512   |
| intraresidue | 121   |
| sequential | 144   |
| medium range, $i-j<5$ | 125   |
| long range, $i-j\geq5$ | 122   |
| hydrogen bond restraints | 38    |
| dihedral angle restraints |       |
| Phi | 24    |
| Psi | 23    |
| Chi | 1     |

| Deviations from idealized geometry | Value |
|-----------------------------------|-------|
| bond lengths (Å) | 0.010 ± 0.001 |
| bond angles (deg) | 1.119 ± 0.047 |
| impropers (deg) | 1.226 ± 0.094 |
| NOE (Å) | 0.014 ± 0.002 |
| cDih (deg) | 0.246 ± 0.073 |

| Mean energies (kcal/mol) | Value |
|--------------------------|-------|
| overall | -1255 ± 39 |
| bonds | 15.1 ± 1.4 |
| angles | 43.0 ± 3.3 |
| improper | 15.7 ± 2.0 |
| van Der Waals | -152.7 ± 6.4 |
| NOE | 0.10 ± 0.03 |
| cDih | 0.38 ± 0.21 |
| electrostatic | -1361 ± 34 |

| Violations | Value |
|------------|-------|
| NOE violations exceeding 0.2 Å | 0     |
| Dihedral violations exceeding 2.0 Å | 0     |

| Rms deviation from mean structure, Å | Value |
|-------------------------------------|-------|
| all backbone atoms | 1.04 ± 0.32 |
| all heavy atoms | 1.63 ± 0.33 |

| Stereochemical quality\(^b\) | Value |
|------------------------------|-------|
| Residues in most favoured Ramachandran region, % | 91.5 ± 2.1 |
| Ramachandran outliers, % | 0.05 ± 0.22 |
| Unfavourable sidechain rotamers, % | 0.15 ± 0.37 |
| Clashscore, all atoms | 7.10 ± 2.28 |
| Overall MolProbity score | 1.66 ± 0.18 |

\(^a\)All statistics are given as mean ± SD.

\(^b\)Molprobity
Figure S4: RP-HPLC and MS analysis of purified β/δ-Uf2a from U. ferox trichomes. A) RP-HPLC of purified β/δ-Uf2a detected via UV absorption at 214 nm. Analysis was carried out with a Phenomenex Jupiter C18 column (150 x 2 mm, 1.8 µm, 300 Å) at a flow rate of 0.15 mL/min and using a linear gradient of 5-60% buffer B (90% CH3CN, 0.05% TFA in water) in buffer A (0.05% TFA in water) over 42 min. B) TIC chromatogram of the same sample run on a Agilent Zorbax C18 HPLC column (100 x 2.1 mm, 1.8 µm, 300 Å) at a flow rate of 0.2 mL/min, a temperature of 60°C and using a gradient of 1-70% buffer B (0.1% formic acid in CH3CN) in buffer A (0.1% formic acid in water) over 71 min. C) HR MS spectrum across the entire peak at tR 33.54 min (Mobs: 6719.22 Da; M calc: 6719.88 Da, most abundant isotope composition). D) Deconvoluted HR-MS spectrum.
Figure S5: The effect of buffer on the $\tau$ of fast inactivation at $\text{Na}_V1.7$ where the DII S1-S2, DII S3-S4, DIV S1-S2 and DIV S3-S4 extracellular loops were replaced by $\text{Na}_V1.8$. The loop substitutions had no effect on the $\tau$ of fast inactivation in the absence of toxin: wildtype $\text{Na}_V1.7$ ($\tau = 1.0 \pm 0.1$ ms), DII S1-S2 ($\tau = 0.8 \pm 0.2$ ms), DII S3-S4 ($\tau = 0.8 \pm 0.2$ ms), DIV S1-S2 ($\tau = 1.2 \pm 0.3$ ms), DIV S3-S4 ($\tau = 0.6 \pm 0.1$ ms), $P = 0.2153$, one-way ANOVA ($n = 5-10$).
Table S2. Primers used for cloning from cDNA or genomic DNA

| Primer name         | Sequence (5' → 3')                           | Target amplicon |
|---------------------|----------------------------------------------|-----------------|
| U_fer_F6_fwd        | ATGGGTGCAATAGTGTTGG                          | Uf2a            |
| U_fer_F6_rev        | CTATTTTCACGGTTCCATTGAAATAG                   |                 |
| U_inc_F6_fwd        | ATGGGCGCAATAGTGTT                          | Ui2a            |
| U_inc_F6_rev        | TTAATTGCACGGTTCCATTGAA                      |                 |
| D_exc_F6_fwd        | ATGAAGACTAGTACAGCTCTG                        | De2a            |
| D_exc_F6_rev        | TTATCTGCATGTCCGGTAGATA                      |                 |
| D_mor_F6_fwd        | ATGAAGAGTACAGGTCACGG                        | Dm2a            |
| D_mor_F6_rev        | TTATACGCAGTCTCCGACTGAA                      |                 |
| U_fer_thionin-1_fwd| ATGGGAAGAAACTGTTATTGTGAG                     | Uf1a            |
| U_fer_thionin-1_rev| TTAGGCAGTTTTCAATAGGTTTTTG                   |                 |

Note: all primers listed in the above table had attB Gateway™ sites incorporated at their 5' termini (5' → 3' attB1: GGG GAC AAG TTT GTA CAA AAA AGC AGG CT and attB2: GGG GAC CAC TTT GTA CAA GAA AGC TGG GT).

Table S3. Sequences of the DII and DIV extracellular loops of Na\textsubscript{V}1.7, Na\textsubscript{V}1.8, and the Na\textsubscript{V}1.7 mutant channels used in this study

| DII S1-S2          | DIV S1-S2            |
|--------------------|----------------------|
| Nav1.7             | AMEHHFMTEEFFKNVL     |
| Nav1.7             | TMMVEKEGQSCHMT       |
| Nav1.8             | AMEHHGMSPTFEAML      |
| Nav1.8             | TMMVETDQQSEEKT       |
| Chimera            | AMEHHGMSPTFEAML      |
| Chimera            | TMMVETDQQSEEKT       |

| DII S3-S4          | DIV S3-S4            |
|--------------------|----------------------|
| Nav1.7             | SLVELFLADVGSGSLVR    |
| Nav1.7             | IET---YFVSP         |
| Nav1.8             | SLEGAVAKGSSLVR      |
| Nav1.8             | LKSLQSF-SPT         |
| Chimera            | SLVEL3VAKGSSLVR     |
| Chimera            | IETLQSF-SPT         |
Figure S6: A) Alignment of urticatoxin and gympietide precursors. Sequences are based on cDNA/gDNA clones identified in this study. B) Pairwise alignment matrix showing sequence identity in %.
Figure S7: M-coffee alignment of predicted mature thionin peptides from transcriptome data. Manual adjustment was performed to align Cys residues. Raw data used for assembly is from NCBI BioProject PRJNA592832 except for U_dio (Urtica dioica), which is from BioProject PRJEB21674. U_fer: Urtica ferox; U_dio: Urtica dioica; U_inc: Urtica incisa; D_exc: Dendrocnide excelsa.
Figure S8: M-coffee alignment of predicted mature urticatoxin peptides from transcriptome data. Manual adjustment was performed to align Cys residues, except for U_incT2.1 DN161_c0_g5_i1 which has additional residues. Raw data used for assembly is from NCBI BioProject PRJNA592832 except for U_dio (Urtica dioica), which is from BioProject PRJEB21674.
Table S4. Sequences of the DIV extracellular loops of Na\textsubscript{V}1.1-Na\textsubscript{V}1.8. Transmembrane segments are shaded grey and in the extracellular loop regions the acidic/basic residues have been highlighted in red/blue, respectively.

|       | DIV S1-S2                     | DIV S3-S4                     |
|-------|-------------------------------|-------------------------------|
| Nav1.1| TMMVET\textbf{EQSE}YVTTLRSR IEK---YFVSPTLFR |                               |
| Nav1.2| TMMVET\textbf{EQSQUE}MNTILYW IEK---YFVSPTLFR |                               |
| Nav1.3| TMMVET\textbf{EQGKY}MTLVLRSR IEK---YFVSPTLFR |                               |
| Nav1.4| TMMVET\textbf{ENQSKL}KVDILYN IQK---YFVSPTLFR |                               |
| Nav1.5| TMMVET\textbf{ENQSKP}EKINILAK IQK---YFFSPTLFR |                               |
| Nav1.6| TMMVET\textbf{TQSK}QSMENILYW IEK---YFVSPTLFR |                               |
| Nav1.7| TMMVEK\textbf{EQSQ}HMTEVLYW IET---YFVSPTLFR |                               |
| Nav1.8| TMMVET\textbf{EQSE}KTKILGK LKSLSYF---PTLFR |                               |