Suppporting information

A Venomics approach coupled to high-throughput toxin production strategies identifies the first venom-derived melanocortin receptor agonists

Steve Reynaud[†,⊥], Justyna Ciolek[⊥], Michel Degueldre[ǁ,ε], Natalie J. Saez[ɡ,Ψ], Ana Filipa Sequeira[ϕ,ʃ], Yoan Duhoo[ɡ,Ψ], Joana L. A. Brás[ϕ,ʃ], Hervé Meudal[¶], Miguel Cabo Díez[ϰ], Victoria Fernández Pedrosa[ϰ], Marion Verdenaud[⊥], Julia Boeri[⊥], Oscar Pereira Ramos[⊥], Frédéric Ducancel[ɾ], Margot Vanden Driessche[⊥], Rudy Fourmy[δ], Aude Violette[δ], Grégory Upert[⊥], Gilles Mourier[⊥], Annette G. Beck-Sickinger[ϓ], Karin Mörl[ϓ], Céline Landon[¶], Carlos M. G. A. Fontes[ϕ,ʃ], Rebeca Miñambres Herráiz[ϰ], Ricardo C. Rodríguez de la Vega[§], Steve Peigneur[ϑ], Jan Tytgat[ϑ], Loïc Quinton[ǁ], Edwin De Pauw[ǁ], Renaud Vincentelli[ɡ], Denis Servent[⊥], Nicolas Gilles[⊥]*

[†]Université Paris-Sud, 15 Rue Georges Clemenceau, Orsay, FR 91405.

[⊥]Université Paris Saclay, CEA, Département Médicaments et Technologies pour la Santé, SIMoS, 91191 Gif-sur-Yvette, France.

[ǁ]Université de Liège, Mass Spectrometry Laboratory, Allée du six Aout 11, Quartier Agora, Liege, BE 4000.

[ε]UCB, Department of Analytical Science Biologicals, Chemin du Foriest, Braine L'Alleud , BE 1420.

[ɡ]Centre National de la Recherche Scientifique, Architecture et Fonction des Macromolécules Biologiques, Campus de Luminy, Marseille, FR 13288.

[Ψ]The University of Queensland, Institute for Molecular Bioscience, Institute for Molecular Bioscience, The University of Queensland, St Lucia, QLD, AUS 4072.

[ϕ]Universidade de Lisboa, CIISA - Faculdade de Medicina Veterinária, Avenida da Universidade Técnica, Lisboa, PT 1300-477.

[∫]NZYTech Lda, Genes & Enzymes, Estrada do Paço do Lumiar, Campus do Lumiar, Edifício E - R/C, Lisboa, PT 1649-038.

[ϰ]Sistemas Genómicos Ltd, Next-Generation Sequencing Laboratory, Ronda de Guglielmo Marconi, 6, Paterna, ES 46980.

[≠] Université Paris Saclay, CEA, Département IDMIT, 18 route du Panorama, 92265 Fontenay-aux-Roses, France.

[δ] Alphabiotoxine Laboratory sprl, Barberie 15, Montroeul-au-bois, BE 7911.

[Ŧ]Universitat Leipzig, Institute of Biochemistry, Leipzig, DE 04103.

[¶]Centre National de la Recherche Scientifique, Centre de Biophysique Moléculaire, rue Charles Sadron, CS 80054 Orléans, FR 45071.

[§]Universite Paris-Sud, Ecologie, Systematique et Evolution, 15 rue Georges Clémenceau, Orsay, FR 91405.

[ȣ]University of Leuven (KU Leuven), Toxicology and Pharmacology, Herestraat 49, Leuven, BE 3000.

Table of contents

Figure S1. LC-MS analysis of five representative toxins with their cystein patterns, produced by solid phase chemical synthesis. S4

Figure S2. Analysis of 4 representative toxins with their cystein patterns, produced by recombinant expression. S6

Figure S3. Quality assessment at each stage of N-TRTX-Preg1a and N-BUTX-Ptr1a production in Fmoc solid-phase chemical synthesis. S8

Figure S4. Activity of 100 μM N-TRTX-Preg1a on ion channels.....S9

Figure S5. Activity of 100 μM N-BUTX-Ptr1a on ion channels.....S10

Table S1. Cysteine pattern diversity for toxins containing an even number of cysteines identified in the Venomics bank.....S11

Table S2. Cysteine pattern diversity for toxins containing an odd number of cysteine residues identified in the Venomics bank. S17

Table S3. Experimental NMR data used for the calculations, and final structural statistics for the ten models representative of the solution structure of N-TRTX-Preg1a. S19

Table S4. Experimental NMR data used for the calculations, and final structural statistics for the ten models representative of the solution structure of N-BUTX-Ptr1a. S20

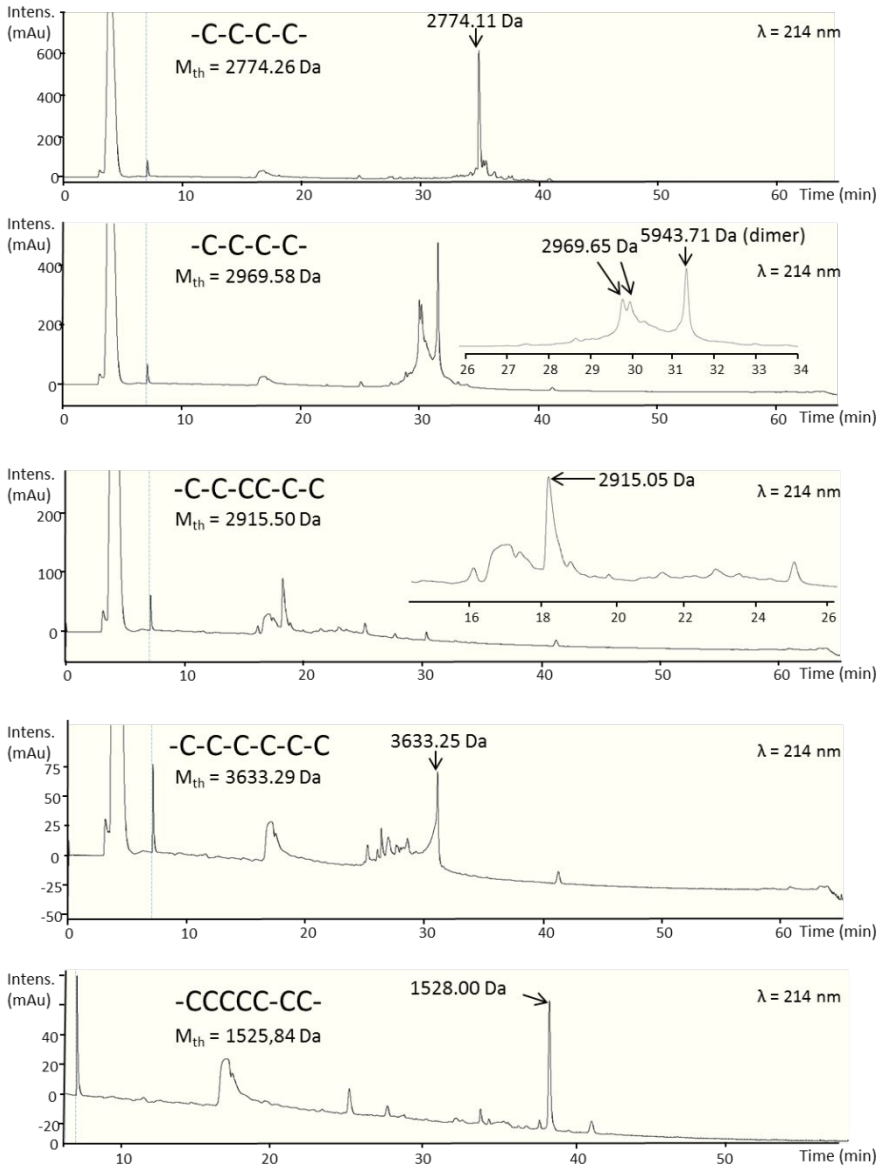


Figure S1. LC-MS analysis of five representative toxins with their cystein patterns, produced by solid phase chemical synthesis. For chemically produced peptides, analysis was performed on Elute Sp Pump HPG 700 (Brucker) coupled with AmaZon SL LC/MS (Bruker) at 214 nm. Column : XBridge® Peptide BEH C18 300Å 3.5µm, 4.6x150 mm (Waters); Gradient : 0-50% CH_3CN/TFA0.1% in 50min ; flow rate : 1mL/min. Oxidation of a two disulfide bridge peptide can lead to one or several oxidized forms included dimers whereas oxidation of a three disulfide bridge peptide often lead to only one oxidized form. For –CCCC-CC- previously never described pattern, produced peptide contains only four of the six possible disulfide bridges. The surface of the peak corresponding to the correct mass was integrated for toxin quantification.

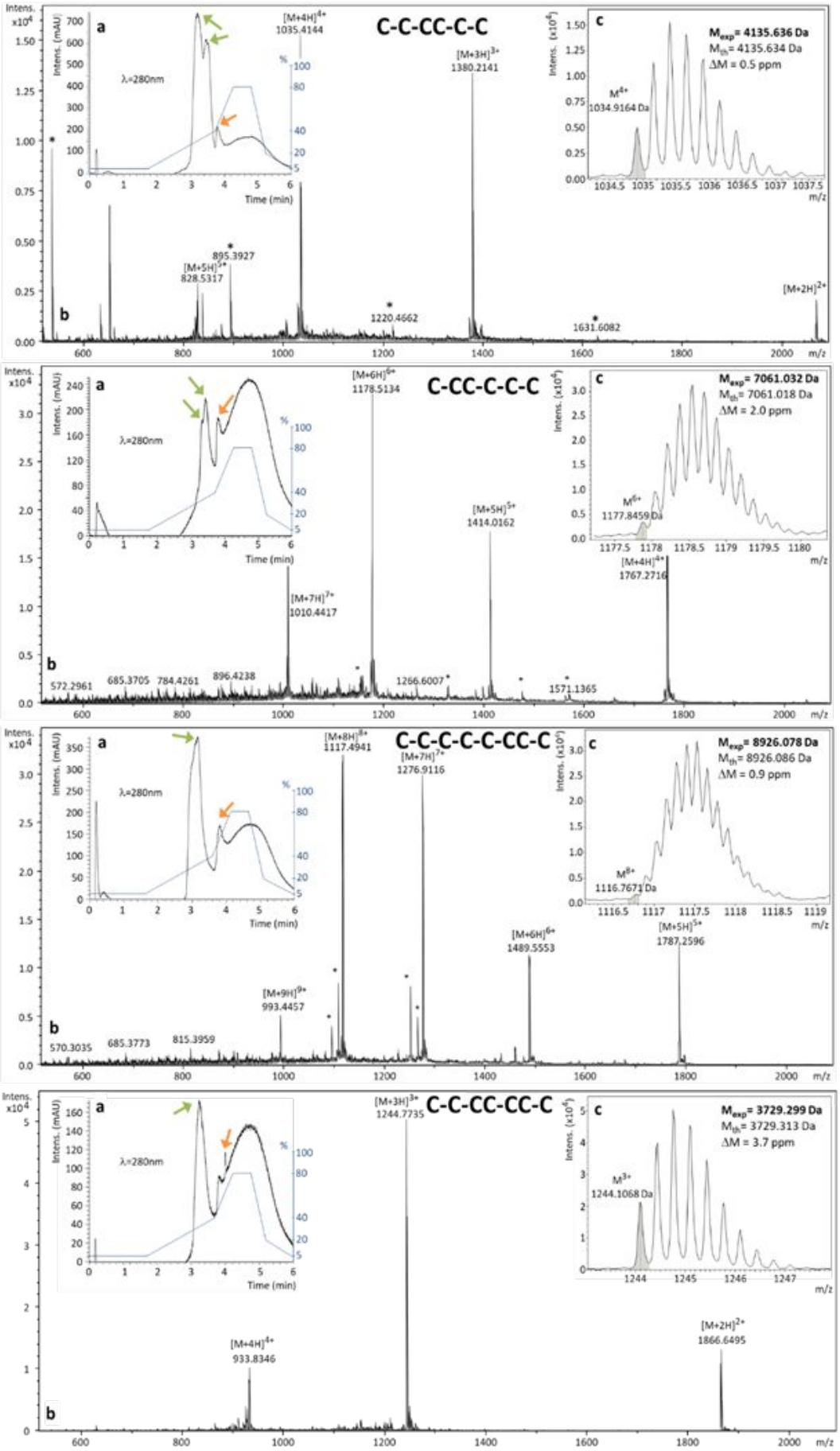

C-C-CC-C-C
a
λ=280nm
Time (min)
[M+4H]4+
1035.4144
[M+3H]3+
1380.2141
[M+5H]5+
828.5317
895.3927
1220.4662
1631.6082
[M+2H]2+
b
c
Mexp= 4135.636 Da
Mth= 4135.634 Da
ΔM = 0.5 ppm
M4+
1034.9164 Da
m/z
C-CC-C-C-C
[M+6H]6+
1178.5134
[M+5H]5+
1414.0162
[M+4H]4+
1767.2716
[M+7H]7+
1010.4417
572.2961
685.3705
784.4261
896.4238
1266.6007
1571.1365
Mexp= 7061.032 Da
Mth= 7061.018 Da
ΔM = 2.0 ppm
M6+
1177.8459 Da
C-C-C-C-C-CC-C
[M+8H]8+
1117.4941
[M+7H]7+
1276.9116
[M+6H]6+
1489.5553
[M+5H]5+
1787.2596
[M+9H]9+
993.4457
570.3035
685.3773
815.3959
Mexp= 8926.078 Da
Mth= 8926.086 Da
ΔM = 0.9 ppm
M8+
1116.7671 Da
C-C-CC-CC-C
[M+3H]3+
1244.7735
[M+4H]4+
933.8346
[M+2H]2+
1866.6495
Mexp= 3729.299 Da
Mth= 3729.313 Da
ΔM = 3.7 ppm
M3+
1244.1068 Da

Figure S2. Analysis of 4 representative toxins with their cystein patterns, produced by recombinant expression. A) Analysis on a UHPLC (Accela High Speed LC) on a Hypersil GOLD column C18 175 Å 1.9 µm, 1x50 mm (ThermoScientific) at 280 nm. The acetronitril gradient is drawn on the graph. The surface of the peak corresponding to the correct mass was integrated for toxin quantification. Green arrows indicate the correct mass. Orange arrows indicates a ghost peak corresponding to the 80% acetonitrile flush. B) mass analysis on a MicroTOF (Bruker, accuracy <5ppm, resolution >10,000). An external calibration was performed using NaI clusters. C) Zoom on the most intense charge state. ppm: parts per million.

A

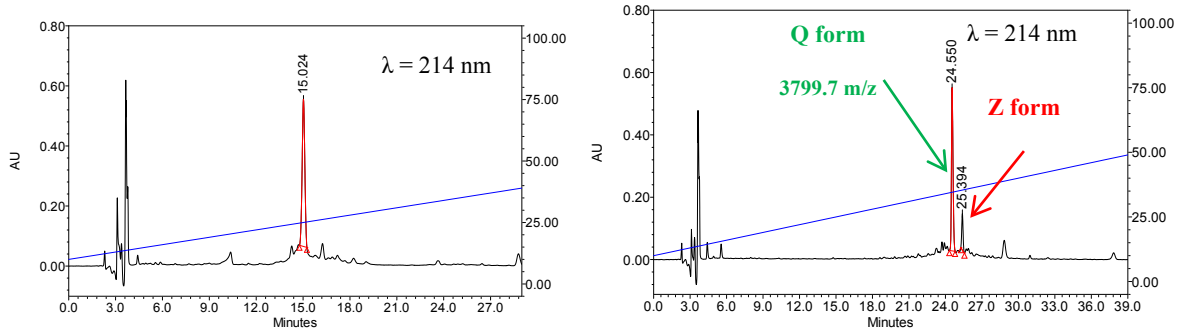
N-TRTX-Preg1a

RCLHAGAACSGPIQKIPCCGTCSRRKCT

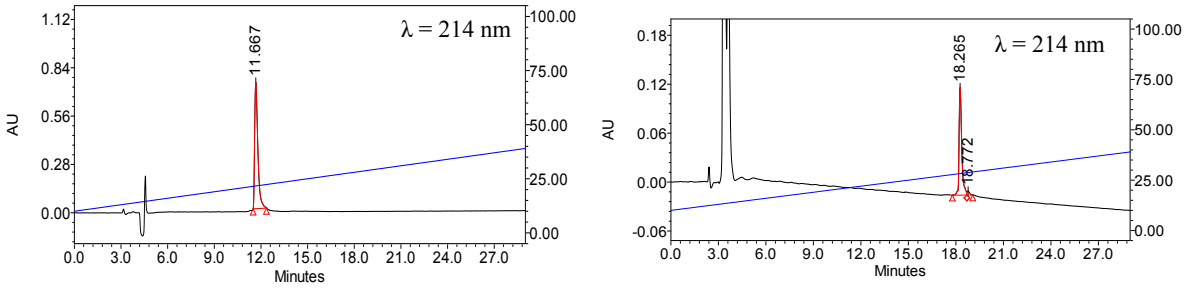
N-BUTX-Ptr1a

QMDMRCSASVECKQKCLKAIGSIFGKCMNKKCKC

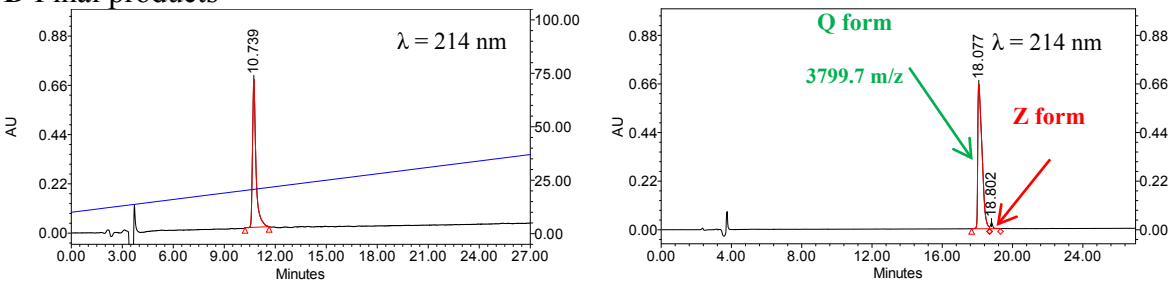
B Crude peptides



C After 22h of oxidation



D Final products



E Mass analysis

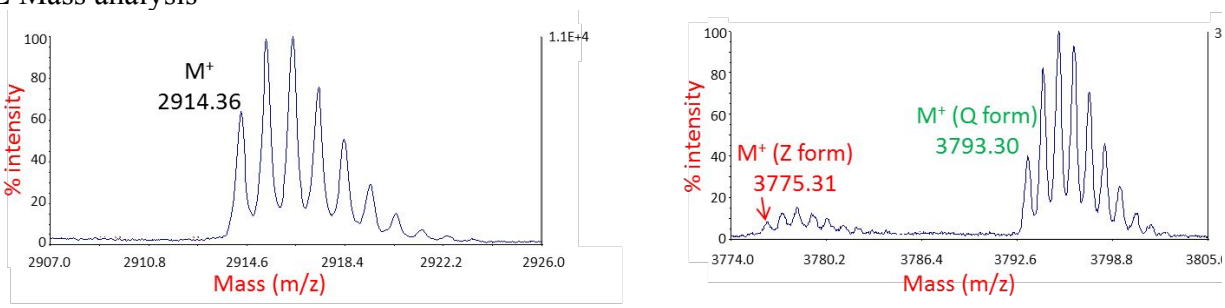


Figure S3. Quality assessment at each stage of N-TRTX-Preg1a (left panel) and N-BUTX-Ptr1a (right panel) production in Fmoc solid-phase chemical synthesis. (A) Peptide sequences with cysteine residues highlighted in red. (B) Analytical HPLC chromatogram of the crude peptide, (C) after 22h of oxidation, and (D) of the final purified product at 214 nm. Column: X-Bridge BEH 300Å 5 µm 4.6 x 250 mm; Gradient: 10-50% CH_3CN/TFA 0.1% in 40 min; flow rate: 1 mL/min. (E) MALDI-TOF spectra of the final pure products. Noted masses correspond to monoisotopic masses. A spontaneous transformation of the N-terminal glutamine (Q form) residue of the peptide N-BUTX-Ptr1a into a pyroglutamic acid (Z form) seems to occur during crude peptide deprotection and can be detected in the final purified product in spite of purification.

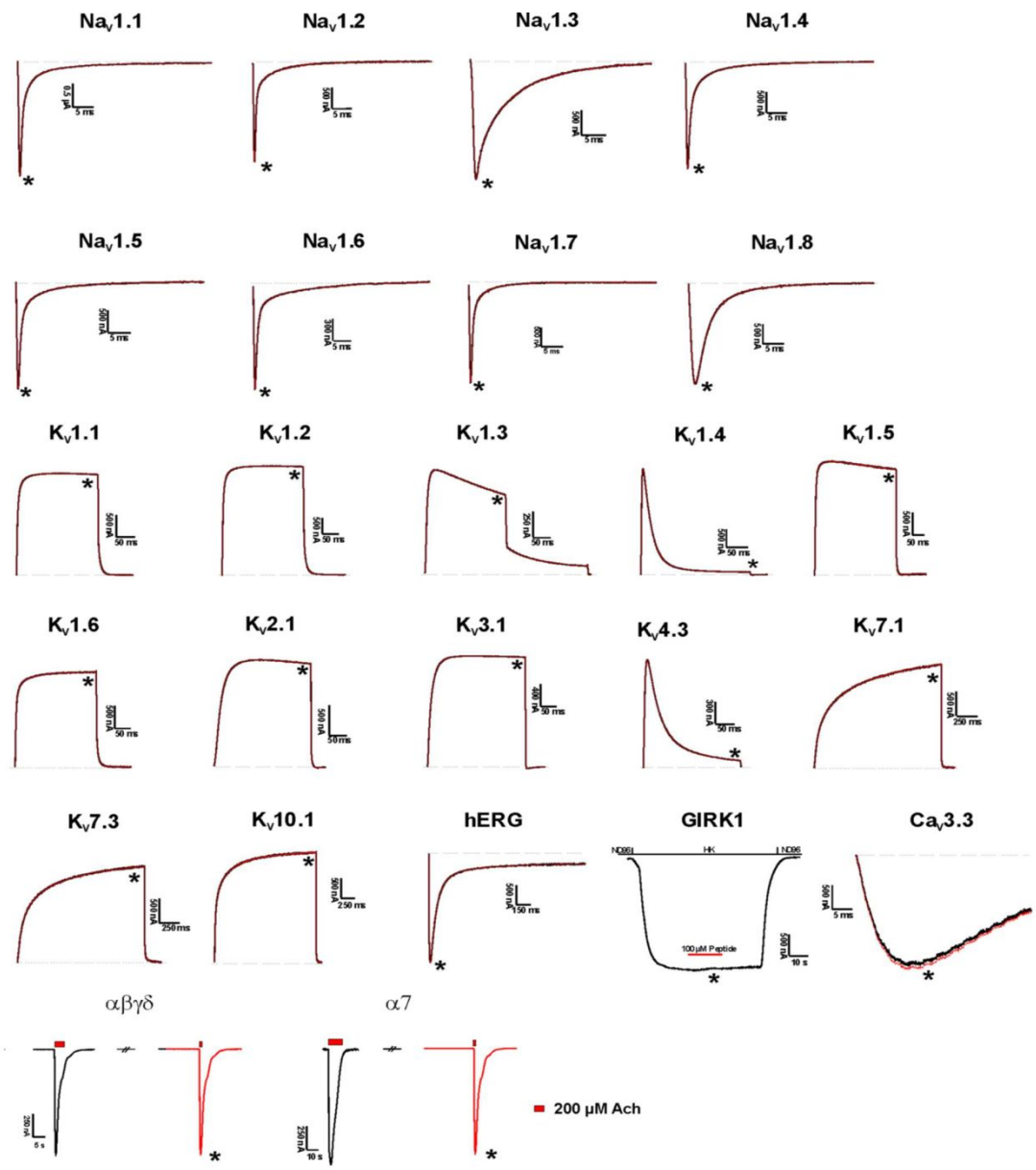


Figure S4. Activity of 100 µM N-TRTX-Preg1a on ion channels. Representative whole cell traces of oocytes expressing cloned Na_v isoforms ($Na_v1.1$-$Na_v1.6$, $Na_v1.8$), K_v isoforms ($K_v1.1$-1.6, $K_v2.1$, $K_v3.1$, $K_v4.3$, $K_v7.1$, $K_v7.3$, $K_v10.1$, hERG), $Ca_v3.3$, GIRK1 or nAChRs (αβγδ and α7). The dotted line indicates the zero-current level. The red * lines indicate the steady-state current peak amplitude in the presence of 100 µM N-TRTX-Preg1a.

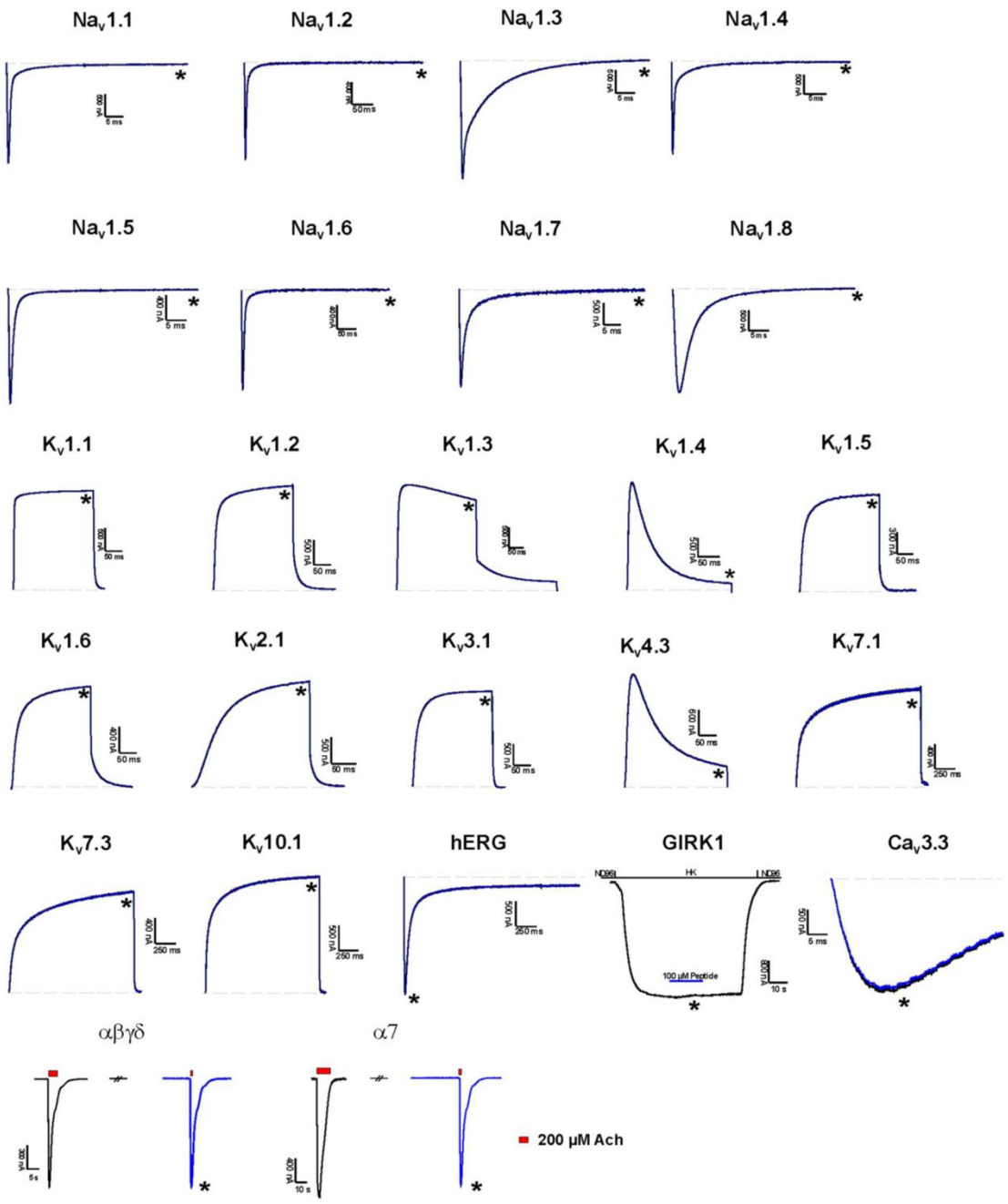


Figure S5. Activity of 100 μM N-BUTX-Ptr1a on ion channels. Representative whole cell traces of oocytes expressing cloned Na_v isoforms ($Na_v1.1$-$Na_v1.6$, $Na_v1.8$), K_v isoforms ($K_v1.1$-1.6, $K_v2.1$, $K_v3.1$, $K_v4.3$, $K_v7.1$, $K_v7.3$, $K_v10.1$, hERG), $Ca_v3.3$, GIRK1 or nAChRs (αβγδ and α7). The dotted line indicates the zero-current level. The blue * lines indicate the steady-state current peak amplitude in the presence of 100 μM N-BUTX-Ptr1a.

Table S1. Cysteine pattern diversity for toxins containing an even number of cysteines identified in the Venomics bank. Newly described patterns are highlighted in grey.

Cysteine framework	N° of Cys	N° of sequences in: Data base	Peptide bank	Taxa in produced bank (N° of species represented)
C-C	2	637	142	Conus (112), Araneae (11), Serpentes (8), Chilopoda (6), Scorpiones (4), Terebridae (1)
CC-CC	4	247	61	Conus (52), Araneae(5), Serpentes (4)
CC-C-C	4	611	106	Conus (92), Araneae (11), Scorpiones (2), Serpentes (1)
C-CCC	4	2	2	Conus (2)
C-CC-C	4	9	1	Conus (1)
C-C-CC	4	63	35	Conus (29), Araneae (5), Serpentes (1)
C-C-C-C	4	764	182	Conus (81), Chilopoda (49), Scorpiones (25), Araneae (21), Terebridae (3), Serpentes (2), Hymenoptera (1)
CCCC-CC	6	10	2	Conus (2)
CCC-C-C-C	6	1	0	
CC-CC-C-C	6	3	0	
CC-C-CC-C	6	2	1	Araneae (1)
CC-C-C-CC	6	450	99	Conus (70), Araneae (26), Serpentes (3)
CC-C-C-C-C	6	10	1	Serpentes (1)
C-CC-CCC	6	1	0	
C-CC-C-C-C	6	20	5	Araneae (4), Conus (1)
C-C-CCC-C	6	1	0	
C-C-CC-CC	6	5	1	Serpentes (1)

C-C-CC-C-C	6	3,739	1,047	Araneae (887), Conus (82), Scorpiones (25), Serpentes (20), Chilopoda (14), Terebridae (12), Hymenoptera (7)
C-C-C-CC-C	6	244	78	Chilopoda (67), Scorpiones (7), Araneae (3), Conus (1)
C-C-C-C-CC	6	83	20	Chilopoda (11), Conus (5), Scorpiones (4)
C-C-C-C-C-C	6	2,199	539	Scorpiones (367), Araneae (76), Conus (33), Serpentes (26), Chilopoda (20), Terebridae (15), Hymenoptera (2)
CC-C-C-C-C-C-C	8	10	2	Conus (1), Scorpiones (1)
C-CCC-C-C-C-C	8	1	0	
C-CC-CC-C-C-C	8	3	0	
C-CC-C-CC-C-C	8	15	7	Araneae (7)
C-CC-C-C-C-C-C	8	11	3	Scorpiones (2), Araneae (1)
C-C-CCCC-C-C	8	1	0	
C-C-CCC-CC-C	8	1	0	
C-C-CC-CC-C-C	8	94	28	Conus (17), Araneae (8), Terebridae (2), Serpentes (1)
C-C-CC-C-CC-C	8	12	3	Araneae (2), Conus (1)
C-C-CC-C-C-CC	8	1	0	
C-C-CC-C-C-C-C	8	438	112	Araneae (68), Terebridae (13), Conus (11), Scorpiones (8), Serpentes (6), Hymenoptera (3), Chilopoda (3)
C-C-C-CCC-C-C	8	29	4	Conus (2), Serpentes (1), Araneae (1)
C-C-C-CC-C-CC	8	2	0	
C-C-C-CC-C-C-C	8	1,544	257	Araneae (190), Scorpiones (27), Chilopoda (17), Conus (10), Hymenoptera (6), Serpentes (5), Terebridae (2)
C-C-C-C-CC-CC	8	4	1	Chilopoda (1)
C-C-C-C-CC-C-C	8	101	43	Scorpiones (23), Serpentes (8), Conus (5), Araneae (4), Chilopoda (2), Terebridae (1)

C-C-C-C-C-CC-C	8	2,128	161	Serpentes (70), Chilopoda (44), Scorpiones (33), Araneae (12),Conus (1), Hymenoptera (1)
C-C-C-C-C-C-CC	8	1	0	
C-C-C-C-C-C-C-C	8	2,560	213	Scorpiones (171), Terebridae (23), Chilopoda (11),Conus (4), Serpentes (2), Araneae (2)
CC-C-C-CC-C-C-C-C	10	338	0	
CC-C-C-C-CC-C-C-C	10	2	0	
CC-C-C-C-C-CC-C-C	10	1	1	Scorpiones (1)
CC-C-C-C-C-C-C-C-C	10	1	1	Scorpiones (1)
C-CC-C-CC-C-C-C-C	10	54	7	Conus (4), Araneae (3)
C-CC-C-C-C-C-C-C-C	10	482	0	
C-C-CC-CC-CC-C-C	10	2	1	Scorpiones (1)
C-C-CC-CC-C-C-CC	10	1	0	
C-C-CC-CC-C-C-C-C	10	13	6	Conus (3), Araneae (1), Serpentes (1), Terebridae (1)
C-C-CC-C-CC-C-C-C	10	57	22	Araneae (20), Serpentes (2)
C-C-CC-C-C-C-CC-C	10	4	0	
C-C-CC-C-C-C-C-C-C	10	224	87	Araneae (46), Scorpiones (26), Chilopoda (8), Serpentes (4), Terebridae (2), Hymenoptera (1)
C-C-C-CCC-C-C-C-C	10	12	4	Araneae (2), Conus (1), Serpentes (1)
C-C-C-CC-C-C-CC-C	10	1	0	
C-C-C-CC-C-C-C-C-C	10	86	17	Araneae (14), Scorpiones (2), Terebridae (1)
C-C-C-C-CCCCC-C	10	1	0	
C-C-C-C-CCC-CC-C	10	1	0	
C-C-C-C-CC-C-C-C-C	10	3	1	Araneae (1)
C-C-C-C-C-CC-C-C-C	10	5	0	
C-C-C-C-C-C-CCC-C	10	1	0	
C-C-C-C-C-C-CC-C-C	10	1	1	Serpentes (1)
C-C-C-C-C-C-C-CC-C	10	605	9	Serpentes (7), Scorpiones (2)

C-C-C-C-C-C-C-C-C-C	10	253	107	Scorpiones (51), Araneae (41), Conus (7), Serpentes (4), Terebridae (4)
CC-C-C-C-CC-C-C-C-C-C	12	5	1	Conus (1)
CC-C-C-C-C-C-CC-C-C-C	12	2	1	Serpentes (1)
CC-C-C-C-C-C-C-C-C-C-C	12	1	1	Serpentes (1)
C-CC-C-C-C-C-CC-C-C-C	12	1	0	
C-CC-C-C-C-C-C-C-C-C-C	12	4	0	
C-C-CC-CC-CC-CC-C-C	12	2	1	Scorpiones (1)
C-C-CC-CC-C-C-C-C-C-C	12	2	0	
C-C-CC-C-C-CC-C-C-C-C	12	5	0	
C-C-CC-C-C-C-C-CC-C-C	12	15	4	Araneae (4)
C-C-CC-C-C-C-C-C-C-C-C	12	3	0	
C-C-C-CC-C-C-C-C-C-C-C	12	144	7	Araneae (7)
C-C-C-C-CCC-C-C-C-C-C	12	3	0	
C-C-C-C-CC-CC-CC-CC	12	1	1	Scorpiones (1)

C-C-C-C-CC-C-C-C-C-CC	12	1	1	Scorpiones (1)
C-C-C-C-CC-C-C-C-C-C-C	12	33	1	Hymenoptera (1)
C-C-C-C-C-CC-C-C-C-C-C	12	245	99	Scorpiones (90), Araneae (7), Chilopoda (2)
C-C-C-C-C-C-C-C-C-CC-C	12	1	0	
C-C-C-C-C-C-C-C-C-C-CC	12	1	0	
C-C-C-C-C-C-C-C-C-C-C-C	12	22	3	Conus (2), Scorpiones (1)
CC-C-C-C-C-C-CC-C-C-C-C	14	3	0	
C-CC-C-C-C-C-C-CC-C-C-C	14	1	0	
C-CC-C-C-C-C-C-C-CC-C-C	14	2	1	Conus (1)
C-C-CC-CC-C-C-C-C-C-C-C	14	11	0	
C-C-C-CC-C-C-C-C-C-CC-C	14	13	3	Conus (2), Araneae (1)
C-C-C-CC-C-C-C-C-C-C-C	14	574	7	Araneae (6), Terebridae (1)
C-C-C-C-CC-C-C-CC-C-C-C	14	1	1	Araneae (1)
C-C-C-C-C-CC-C-CC-C-C-C	14	3	0	

C-C-C-C-C-CC-C-C-C-C-C-C-C	14	2	0	
C-C-C-C-C-C-CC-C-C-C-C-C-C	14	4	0	
C-C-C-C-C-C-C-C-C-C-C-C-C-C	14	14	1	Scorpiones (1)
C-CC-C-C-C-C-C-CC-C-C-C-C-C-C	16	1	0	
C-C-C-C-CC-C-C-C-C-C-C-CC-C-C	16	2	0	
C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C	16	4	2	Scorpiones (1) , Terebridae (1)
C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C	18	1	1	Scorpiones (1)

Table S2. Cysteine pattern diversity for toxins containing an odd number of cysteine residues identified in the Venomics bank. Newly described patterns are highlighted in grey.

Cysteine framework	N° of Cys	N° of sequences in : Data base	N° of sequences in : Peptide bank	Taxa in produced bank (N° of species represented)
CC-C	3	27	3	Araneae (3)
C-CC	3	3	0	
C-C-C	3	37	3	Scorpiones (1), Conus (1), Serpentes (1)
CC-CCC	5	2	0	
CC-C-CC	5	4	1	Conus (1)
CC-C-C-C	5	11	0	
C-CC-C-C	5	17	0	
C-C-CC-C	5	42	1	Araneae (1)
C-C-C-CC	5	3	0	
C-C-C-C-C	5	179	3	Araneae (3)
CC-C-C-CC-C	7	1	0	
C-CC-C-C-C-C	7	3	2	Araneae (2)
C-C-CC-CC-C	7	10	2	Araneae (2)
C-C-CC-C-C-C	7	97	20	Araneae (20)
C-C-C-CC-C-C	7	15	3	Araneae (3)

C-C-C-C-CC-C	7	55	0	
C-C-C-C-C-CC	7	3	0	
C-C-C-C-C-C-C	7	130	2	Chilopoda (1), Scorpiones (1)
CC-C-CC-C-C-C-C	9	2	0	
CC-C-C-CC-C-C-C	9	12	0	
CC-C-C-C-C-CC-C	9	2	0	
C-CC-C-C-C-CC-C	9	9	1	Serpentes (1)
C-CC-C-C-C-C-C-C	9	5	0	
C-C-CC-CCC-C-C	9	1	0	
C-C-CC-CC-C-C-C	9	1	0	
C-C-CC-C-C-C-C-C	9	23	0	
C-C-C-CC-C-C-C-C	9	3	0	
C-C-C-C-CC-C-C-C	9	12	0	
C-C-C-C-C-C-CC-C	9	5	0	
C-C-C-C-C-C-C-CC	9	1	0	
C-C-C-C-C-C-C-C-C	9	14	0	
C-CC-C-C-C-C-CC-C-C	11	1	0	
C-CC-C-C-C-C-C-C-C-C	11	8	0	
C-C-CC-C-C-C-C-C-C-C	11	1	0	
C-C-C-CC-C-C-C-C-C-C	11	1	0	
C-C-C-C-C-C-C-C-CC-C	11	3	0	
C-C-C-C-C-C-C-C-C-CC	11	1	0	
C-C-C-CC-C-C-C-C-C-C-C-C	13	121	0	
C-C-C-C-C-C-C-C-C-C-CC-C	13	6	0	
C-C-C-CC-C-C-C-CC-C-C-C-C-C	15	3	0	

Table S3. Experimental NMR data used for the calculations, and final structural statistics for the ten models representative of the solution structure of N-TRTX-Preg1a.

NMR constraints	
Distance restraints	
Total NOE	546
Unambiguous	454
Ambiguous	92
Hydrogen bonds	10
Dihedral Angle Restraints	38
Disulfide bridges[a]	(C2-C19, C9-C22, C18-C27)
Structural Statistics (6SAA.pdb)	
Average violations per structure	
NOEs ≥ 0.3Å	0
Hydrogen bonds ≥ 0.5Å	0
Dihedrals ≥ 25°	0
Dihedrals ≥ 15°	3
Average pairwise rmsd (Å)	0.18
Ramachandran Analysis	
Most favored region	64.8%
Allowed region	35.2%
Generously allowed	0%
Disallowed	0%
Energies (kcal.mol^{-1})[b]	
Electrostatic	-811 ± 8
Van der Walls	-169 ± 3
Total energy	-608 ± 15
Residual NOE energy	26 ± 5

a : Introduced as covalent bonds; b : values are given as mean ± standard deviation (n=10)

Table S4. Experimental NMR data used for the calculations, and final structural statistics for the ten models representative of the solution structure of N-BUTX-Ptr1a.

NMR constraints	
Distance restraints	
Total NOE	839
Unambiguous	558
Ambiguous	281
Hydrogen bonds	4
Disulfide bridges	ambiguous
Structural Statistics (6SAB.pdb)	
Average violations per structure	
NOEs ≥ 0.3Å	0
Hydrogen bonds ≥ 0.5Å	0
Average pairwise rmsd (Å)	0.40
Ramachandran Analysis	
Most favored region	53.0%
Allowed region	40.3%
Generously allowed	5.0%
Disallowed	1.7%
Energies (kcal.mol^{-1})[a]	
Electrostatic	-1252 ± 46
Van der Walls	-116 ± 79
Total energy	-1056 ± 96
Residual NOE energy	22 ± 6

a : values are given as mean ± standard deviation (n=10)