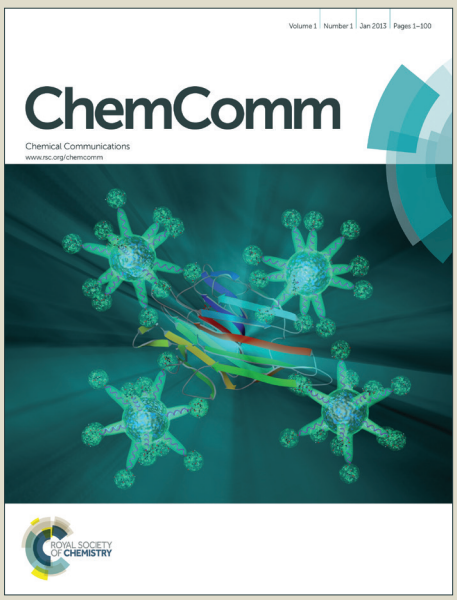
ChemComm

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

Chemical Communications

RSCPublishing

COMMUNICATION

High-Throughput Production of Two Disulphide-Bridges Toxins

Cite this: DOI: 10.1039/x0xx00000x

Grégory Upert,* [a] Gilles Mourier, [a] Alexandra Pastor, [a] Marion Verdenaud, [b] Doria Alili, [a] Denis Servent [a] and Nicolas Gilles* [a]

Received ooth January 2012,
Accepted ooth January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

A quick and efficient production method compatible with high-throughput screening was developed using 36 toxins belonging to four different families of two disulphide-bridges toxins. Final toxins were characterized using HPLC co-elution, CD and pharmacological studies.

Venomous animals have developed an arsenal of small reticulated peptides for their defense and predation. Based on various disulphide-linked scaffolds, these toxins continuously selected and highly refined by the natural evolution process display a low immunogenicity, and remarkable proteolytic resistance[1] associated with a high selectivity and efficacy for a variety of important membrane receptors such as ion channels[2] or G-Protein Coupled Receptors.[3] Due to their low sizes and compact structures, these disulphide rich peptides (DRP) show great interest for pharmaceutical applications, being currently therapeutic candidates or already on the market.[4] Toxinology of the last century was mostly focused to the identification of toxin activities in order to understand venom toxicity. It is estimated that only 0.01% of the animal toxins that represent a likely natural bank of 40 million of peptides are described in dedicated databases.[5] Nowadays, the advent of the complementary transcriptomic[6] and proteomic[7] techniques applied to venom glands and venoms respectively leads to an explosion of the number of toxin sequences. While few thousands of sequences were known at the beginning of the century, hundreds of thousands will be described at the end of this decade.[8] This giant database of DRP sequences represents a source of potential drug candidates that have to be synthesized in order to characterize their biological activities. Consequently, the high-throughput production of a large number of reticulated peptides in sufficient amount is currently the major limiting step between sequencing and biological screening and represents nowadays the next challenge.

Herein, we describe a fast and efficient production of two disulphide bridges toxins compatible with high-throughput screening. 36 toxins from 4 different families with different cysteine patterns were used to optimize and validate our method. These toxins display a large panel of biological activities and are mainly present in the venoms of cone snails and arthropods. The α-conotoxins that have the cysteine pattern C_1C_2-C_3-C_4 with the cysteine connectivity 1-3, 2-4 (family I, **1-15**, Table 1) are potential inhibitors of the nicotinic acetylcholine receptors (nACHR). The χ- conotoxins with the same pattern but the connectivity 1-4, 2-3 (family II, **16-24**, Table 1) inhibit the norepinephrine transporter (NET). Both targets are involved in the pain transmission mechanisms and/or in various neurodegenerative diseases (Alzheimer, Parkinson,…).[3h, 9] The toxins with the pattern C_1-C_2-C_3-C_4 and the 1-3, 2-4 connectivity identical to the apamin from the bee venom (family III, **25-30**, Table 1) have a very recent great interest on Parkinson disease model.[10] The sarafotoxins and their derivatives with the connectivity 1-4; 2-3 (family IV, **30-36**, Table 1) are agonists of endothelin receptors that are involved in cancer and cardiovascular diseases.[11]

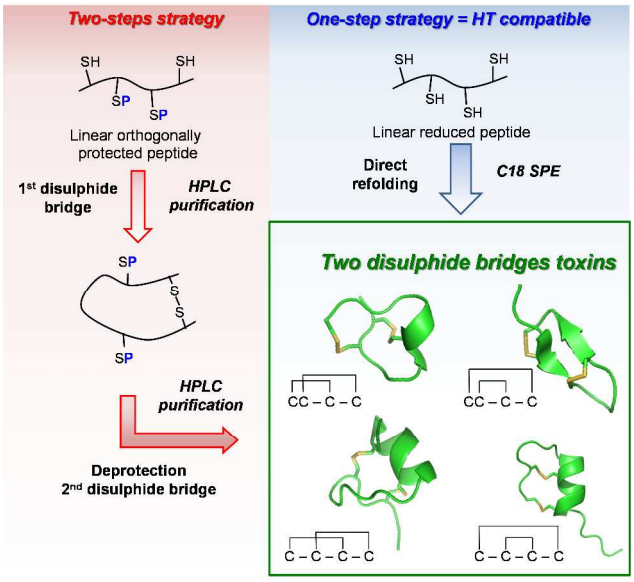


Fig.1 Comparison of the traditional strategy and the high-throughput strategy developed for the synthesis of two disulphide- bridges toxins.

Due to the presence, in these small toxins, of post-translational modifications such as disulphide bonds, amidation or hydroxyproline as

Table 1. Characteristics of the toxins used in this study. The names of the different toxins are available in the supplementary information.

Number	Sequence	Family/Pattern
1	ECCNPACGRHYSC*	
2	GRCCHPACGKNYSC*	
3	ICCNPACGPKYSC*	
4	GCCSTPPCAVLYC*	
5	GCCSLPPCALSNPDYC*	
6	FNWRCCLIPACRRNHKKFC*	
7	RDOCCYHPTCNMSNPQIC*	I
8	CCHPACGKYYSC*	
9	ECCNPACGRHYSCGK*	CC – C – C
10	DGRCCHPACAKHFNC*	
11	NGRCCHPACARKYNC*	
12	ECCHPACGKHFSC*	
13	ZSOGCCWNPACVKNRC*	
14	GCCSHPACNVNNPHIC*	
15	GRCCHPACGKYYSC*	
16	NGVCCGYKLCHOC	
17	VGVCCGYKLCHOC	
18	ZGVCCGYKLCHOC*	II
19	GICCGVSFCYOC	
20	ZTCCGYRMCVOC*	CC – C – C
21	GVCCGVSFCYOC	
22	RCCGYKMCHOC	
23	VCCGYKLCHOC	
24	CCHSSWCKHLC	
25	CNCKAPETALCARRCQQH*	
26	IKCNCKRHVIKPHICRKICGKN*	III
27	ALCNCNRIIIPHMCWKKCGKK*	
28	DCPPHPVPGMHKCVCLKTC	C – C – C – C
29	DGCPPHPVPGMHPCMCTNTC	
30	FPRPRICNLACRAGIGHKYPFCHCR*	
31	CTCKDMTDKECLYFCHQDIIW	
32	CTCNDMTDEECLNFCHQDVIW	IV
33	CSCKDMTDKECLNFCHQDVIW	
34	CSCKDMSDKECLNFCHQDVIW	C – C – C – C
35	CSCADMTDKECLYFCHQDVIW	
36	GHACYRNCWREGNDEETCKERC*	

* C-terminal amidation and O=hydroxyproline

examples, recombinant technologies are not the most adapted strategy of production while the chemical synthesis is, thanks to its high versatility and flexibility. Two disulphide bridges toxins can be obtained following several strategies. First, linear peptides are obtained after solid-phase synthesis and HPLC purification. Then, the disulphide bridges are usually formed stepwise in a regioselective manner using appropriate orthogonal thiol protections e.g. Trt, Acm, Mmt, STmp or StBu groups (Fig. 1).[12] This strategy is time consuming with numerous purification steps and therefore is not compatible with a high-throughput production. Another method is to use selenocysteine's derivatives to form selectively a diselenide bridge instead of the natural one.[13] This method respects the disulfide's connectivity and the biological activity but cannot be applied with toxin sequences resulting from transcriptomic and proteomic data where disulphide's connectivity is not known. The chosen and compatible with HT screening approach consists in the random oxidation of the crude unprotected peptide obtained by solid phase synthesis in a specific oxidant solution. This strategy needs excellent synthetic yields and an appropriate folding solution with the lower formation of non-natural bridges (Fig.1).[12a, 14] To date, no general methods were emphasized using this strategy, the studies focusing on a small number of toxins of a single family, refolded using a very limited number of conditions.[15] Therefore, we developed a general screening method including 60 conditions varying buffers, additives and redox couples on chosen toxins covering a wide range of physico-chemical parameters (length, pI, hydrophobicity (GRAVY index)).

In order to obtain two bridges toxins with high purity without the need of HPLC purification, several synthesis protocols were applied varying amino acids equivalents, nature of the base (Hünig's base or NMM), reaction time or resins for having the fastest and more efficient method keeping reasonable costs. Finally, we used a fast Fmoc strategy consisting in double short coupling steps (3 min) with 5 equiv. of Fmoc protected amino acid using HCTU/NMM as activating system on a Prelude synthesizer (Protein Technologies®).[16] Every steps (Fmoc deprotection, wash, capping) were optimized leading to a time of 19 min for a cycle deprotection/coupling/capping. This method allowed us to synthesize the 36 linear toxins in one week with excellent purities (≥75%).

During the one-step folding, three theoretical combinations of bridges can arise. The formation of secondary misfolded entities can be minimized by tuning the components of the oxidation solution. Nature and pH of the buffer or the red-ox couple used are crucial

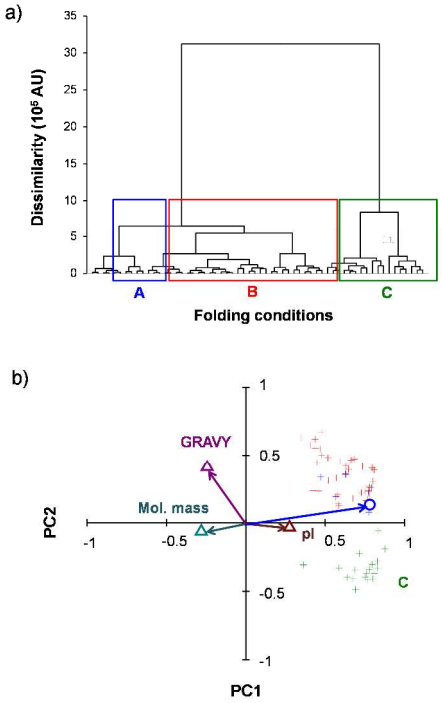


Fig. 2 a) Dendogram obtained using HCA analysis on the different folding conditions including three groups: A (in blue) with the best conditions, B (in red) with conditions giving intermediate folding yields and C (in green) for low yields conditions; b) Representation of the different folding conditions (crosses in blue, red and green for respectively the groups A,B and C with the optimal condition (blue round marker)) and the physico-chemical parameters of the different toxins (triangles) in the principal component space PC1/PC2.

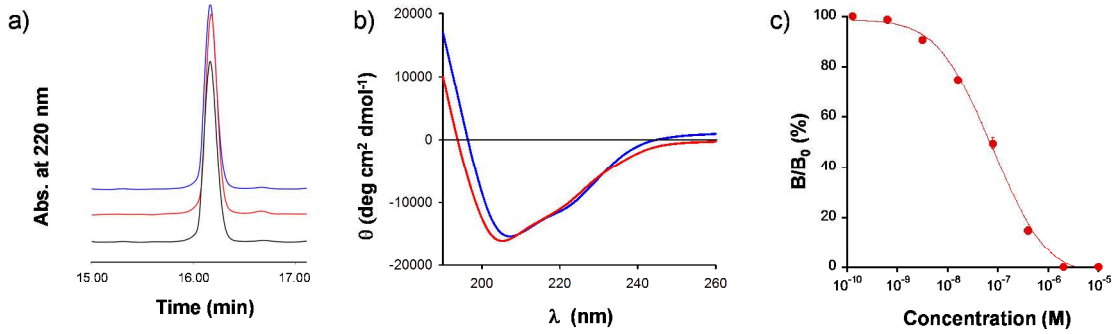


Fig. 3 a) HPLC chromatograms of the synthetic toxin α-conotoxin GI 1 (black trace), the natural form (red trace) and their equimolar mix (blue trace); b) circular dichroism spectrum α-conotoxin GI **1** (in red) and of the α-conotoxin GIA **9** (in blue); c) affinity determination of the toxin α-conotoxin GIA **9** on nAchR (Torpedo).

parameters. 100 mM Tris-HCl and HEPES both containing 1 mM EDTA were used as buffers at pH 8.5 and 7.5 respectively. The chosen additives were 20% acetonitrile (ACN) and glycerol (GOL) for solubilisation purposes, 0.5 M arginine (Arg) and guanidinium chloride (Gdn-Cl) both used as dispersing agents.[17] Glutathione reduced/oxidized (GSH/GSSG) and L-cysteine/L-cystine (Cys/C-C) couples showing slightly different redox potentials (respectively -0.24 V and -0.22 V in water)[18] were used at different ratios (1/0.1, 1/1, 0.1/1 and 1/0.1, 0.1/1 respectively). The temperature used was 4°C in order to slow down the kinetics of the oxidation resulting in the more stable toxin. Each folding conditions were analysed using HPLC for determining the yield of folding of the correct toxin amongst the other isomers.

A hierarchical ascending classification[19] (HCA) that aims at assembling variables between them, divided the conditions in three groups showing the same folding performances for all the toxins (Fig.2a). The first group (A, Fig 2a and S3) contains a few numbers of conditions giving the best folding yields (mean value of 66%). Acetonitrile is the mostly present additive within the conditions of this group with also no additive. The redox couple Cys/C-C in any proportion and the couple GSH/GSSG 1/1 lead to optimum results. The second group (B, Fig.2a and S3) represents a large mix of conditions giving low to reasonable yields (mean value of 53%) into no specific buffers or additives or redox couples stood out. Finally, the last group (C, Fig. 2a and S3) contains conditions showing poor yields (average of 30%) for which the redox couple GSH/GSSG 0.1 mM/1 mM is predominant, the buffer and the additives not playing a role. In conclusion, the combination between HEPES buffer at pH 7.5, ACN as additive and Cys/C-C 1/0.1 as redox couple was chosen since this particular condition gives folding yields of the correct toxin higher than 70% for more than 50% of the disulphide bridges toxins studied. This condition is advantageous in term of easy-to-use and cost. In addition, we applied on the different conditions a principal component analysis[20] (PCA) that aims to compress the number of variables and create a new space for a better comparison of the variables between them (Fig. 2b). pI, GRAVY index and molecular mass of each toxin were included in this new space and lead to the following results: molecular mass negatively correlates with the yields of the optimal condition (vectors in the reverse direction) where the pI varies in the same way (vectors in the same direction) and GRAVY index does not affect the yield (perpendicular vectors) (Fig. 2c). This observation allows a possible prediction of the folding yield using the optimal condition described above. In the universal condition, small peptides (M<2000 g.mol-1) having a pI>7 will be easily folded (yields > 80%) where the larger peptides with lower pI will give lower but still reasonable yields (30-80%). These yields are sufficient enough for the production or the qualitative HT screening of these toxins.

Following the folding, no HPLC separation but only C18 SPE was carried out for the removal of redox couple and salts. For determining the good pairing of the resulting peptides, three different methods were used: i) HPLC co-elution with a reference when it is available or ii) determination of the general structure using circular dichroism analysis and/or iii) characterization of the affinity of the final compound with their respective targets. 16 synthetic and their respective commercially available toxins (compounds **1-7**, **16-18**, **25-27** and **31-33**) highlighted full overlapping in HPLC analysis (Fig. 3a).[24] These toxins were then used as references for the CD analyses. Within each family, the CD spectra of the synthesized toxins were compared with the reference toxins (Fig. 3b).[21] In all cases, they showed the same signature regarding their sequence similarity in particular regarding the size of the loops. In addition, competitive binding assays on nACHR were carried out for the compounds **1-3**, **7**, **13, 15**, **28** and **30** and for NET for the compounds **16** and **18**. Interestingly, all the synthesized toxins showed an activity on their respective targets in agreement with affinity constants reported in the literature.[21] These results demonstrate that our high throughput production protocol is highly efficient and is compatible with HT screening for the determination of biological activities of toxins that have been described only at the sequence level.

Within this study, we developed a fast and efficient method for the synthesis of two disulphide bridges toxins. Our strategy consisting in the fast solid-phase synthesis of the linear peptide bearing or not PTMs followed by a one-step oxidation leads to excellent yields for the synthesis of toxins from four different families having different cysteine patterns and pairings. The screening method for the oxidation step, carried out on a large number of peptides, allowed us to determine a universal condition efficient enough for obtaining a known or unknown toxin having the good structure in a very short time. This method that will be used within the VENOMICS FP7 project[22] is a breakthrough for reproducing synthetically what venomous animals developed over millennia and allows a fastest discovery of new ligands and potential drugs based on the transcriptomic and proteomic studies on venoms.

Notes and references

[a] CEA, DSV, iBiTec-S, Service d'Ingénierie Moléculaire des Protéines (SIMOPRO), CEA Saclay, Gif sur Yvette F-91191, France
E-mail: gregory.upert@cea.fr and nicolas.gilles@cea.fr
[b] CEA, DSV, iBiTec-S, Service de Pharmacologie et d'Immunoanalyse, CEA Saclay, Gif sur Yvette F-91191, France

Support from the European FP7 VENOMICS project is gratefully acknowledged.

† Electronic Supplementary Information (ESI) available: See DOI: 10.1039/b000000x/

1 For a recent review on disulphide rich peptides, see Góngora-Bení, M. Tez, J. Tulla-puche and F. Albericio, *Chem. Rev.*, 2014, **114**, 901.
2 For recent examples on different ionic channels, see (a) C. M. Barber, G. K. Isbister and W. C. Hodgson, *Toxicon* 2013, **66**, 47; (b) Z. Chen, F. Luo, J. Feng, W. Yang, D. Zeng, R. Zhao, Z. Cao, M. Liu, W. Li, L. Jiang and Y. Wu, *PLoS One*, 2013, **8**, e60201; (c) S. Diochot, A. Baron, M. Salinas, D. Douguet, S. Scarzello, A.-S. Dabert-Gay, D. Debayle, V. Friend, A. Alloui, M. Lazdunski and E. Lingueglia, *Nature*, 2012, **490**, 552; (d) S. Peigneur, B. Billen, R. Derua, E. Waelkens, S. Debaveye, L. Béress and J. Tytgat, *Biochem. Pharmacol.*, 2011, **82**, 81; (e) M. Stevens, S. Peigneur and J. Tytgat, *Front. Pharmacol.*, 2011, **2**, 71; (f) E. O. Gracheva, N. T. Ingolia, Y. M. Kelly, J. F. Cordero-Morales, G. Hollopeter, A. T. Chesler, E. E. Sánchez, J. C. Perez, J. S. Weissman and D. Julius, *Nature*, 2010, **464**, 1006.
3 For recent examples on different GPCR receptors, see: (a) C. Petrel, H. G. Hocking, M. Reynaud, G. Upert, P. Favreau, D. Biass, M. Paolini-Bertrand, S. Peigneur, J. Tytgat, N. Gilles, O. Hartley, R. Boelens, R Stöcklin and D. Servent, *Biochem. Pharmacol.*, 2013, **85**, 1663; (b) C. Fruchart-Gaillard, G. Mourier, G. Blanchet, L. Vera, N. Gilles, R. Ménez, E. Marcon, E. Stura and D. Servent, *PLoS One*, 2012, **7**, e39166; (c) C. Marquer, C. Fruchart-Gaillard, G. Letellier, E. Marcon, G. Mourier, S. Zinn-Justin, A. Ménez, D. Servent and B. Gilquin, *J. Biol. Chem.*, 2011, **286**, 31661; (d) K. Näreoja and J. Näsman, *Acta Physiol.*, 2011, 1; (e) C. Rouget, L. Quinton, A. Maïga, C. Gales, G. Masuyer, C. Malosse, J. Chamot-Rooke, R. Thai, G. Mourier, E. De Pauw, N. Gilles and D. Servent, *Br. J. Pharmacol.*, 2010, **161**, 1361; (f) L. Quinton, E. Girard, A. Maiga, M. Rekik, P. Lluel, G. Masuyer, M. Larregola, C. Marquer, J. Ciolek, T. Magnin, R. Wagner, J. Molgó, R. Thai, C. Fruchart-Gaillard, G. Mourier, J. Chamot-Rooke, A. Ménez, S. Palea, D. Servent and N. Gilles, *Br. J. Pharmacol.*, 2010, **159**, 316.
4 G. F. King, *Expert Opin. Biol. Ther.*, 2011, **11**, 1469.
5 (a) F. Jungo, L. Bougueleret, I. Xenarios and S. Poux, *Toxicon*, 2012, **60**, 551; (b) F. Jungo and A. Bairoch, *Toxicon*, 2005, **45**, 293.
6 For recent examples, see: (a) Y. Terrat, K. Sunagar, B. G. Fry, T. N. W. Jackson, H. Scheib, R. Fourmy, M. Verdenaud, G. Blanchet, A. Antunes and F. Ducancel, *Toxins (Basel)*, 2013, **5**, 1948; (b) S. Dutertre, A. Jin, Q. Kaas, A. Jones, P. F. Alewood and R. J. Lewis, *Mol. Cell. Proteomics*, 2013, **12**, 312; (c) M. J. Margres, K. Aronow, J. Loyacano and D. R. Rokyta, *BMC Genomics*, 2013, **14**, 531; (d) D. R. Rokyta, K. P. Wray and M. J. Margres, *BMC Genomics*, 2013, **14**, 394; (e) J. R. Prashanth, R. J. Lewis and S. Dutertre, *Toxicon*, 2012, **60**, 470.
7 For recent examples, see: (a) D. L. Brinkman, A. Aziz, A. Loukas, J. Potriquet, J. Seymour and J. Mulvenna, *PLoS One*, 2012, **7**, e47866; (b) A. Leonardi, D. Biass, D. Kordiš, R. Stöcklin, P. Favreau and I. Križaj, *J. Proteome Res.*, 2012, **11**, 5046; (c) M. E. Peichoto, F. L. Tavares, M. L. Santoro and S. P. Mackessy, *Comp. Biochem. Physiol. Part D. Genomics Proteomics*, 2012, **7**, 361; (d) M. Madrigal, L. Sanz, M. Flores-Díaz, M. Sasa, V. Núñez, A. Alape-Girón and J. J. Calvete, *J. Proteomics*, 2012, **77**, 280–297.
8 G. F. King, M. C. Gentz, P. Escoubas and G. M. Nicholson, *Toxicon*, 2008, **52**, 264.
9 For recent examples see (a) R. J. Lewis, S. Dutertre, I. Vetter and M. J. Christie, *Pharmacol. Rev.*, 2012, **64**, 259; (b) R. J. Lewis, *Toxicon*, 2012, **59**, 524; (c) S. Dutertre and R. J. Lewis, *Biochem. Pharmacol.*, 2006, **72**, 661.
10 (a) D. Alvarez-Fischer, C. Noelker, F. Vulinović, A. Grünewald, C. Chevarin, C. Klein, W. H. Oertel, E. C. Hirsch, P. P. Michel and A. Hartmann, *PLoS One*, 2013, **8**, e61700; (b) A. J. Kallarackal, J. M. Simard and A. M. Bailey, *Behav. Brain Res.*, 2013, **237**, 71.
11 (a) G. Mourier, M. Hajj, F. Cordier, A. Zorba, X. Gao, T. Coskun, A. Herbet, E. Marcon, F. Beau, M. Delepierre, F. Ducancel and D. Servent, *Biochimie*, 2012, **94**, 461; (b) M. A. F. Hayashi, C. Ligny-Lemaire, Z. Wollberg, M. Wery, A. Galat, T. Ogawa, B. H. Muller, H. Lamthanh, Y. Doljansky, A. Bdolah, R. Stöcklin and F. Ducancel, *Peptides*, 2004, **25**, 1243; (c) A. Becker, E. B. Dowdleb, U. Hechler, K. Kausera, P. Donnera and W. Schleuning, *FEBS*, 2000, **315**, 100; (d)A. Bdolah, Z. Wollberg, G. Fleminger and E. Kochva, *FEBS Lett.*, 1989, **256**, 1.
12 (a) T. M. Postma, M. Giraud and F. Albericio, *Org. Lett.*, 2012, **14**, 5468; (b) M. Góngora-Benítez, J. Tulla-Puche, M. Paradís-Bas, O. Werbitzky, M. Giraud and F. Albericio, *Biopolymers*, 2011, **96**, 69.
13 (a) A. Walewska, T. S. Han, M.-M. Zhang, D. Yoshikami, G. Bulaj, and K. Rolka, *Eur. J. Med. Chem.*, 2013, **65**, 144; (b) A. M. Steiner, K. J. Woycechowsky, B. M. Olivera, and G. Bulaj, *Angew. Chem. Int. Ed. Engl.*, 2012, **51**, 5580; (c) T. S. Han, M.-M. Zhang, K. H. Gowd, A. Walewska, D. Yoshikami, B. M. Olivera, and G. Bulaj, *ACS Med. Chem. Lett.*, 2010, **1**, 140 (d) M. Muttenthaler, S. T. Nevin, A. A. Grishin, S. T. Ngo, P. T. Choy, N. L. Daly, S.-H. Hu, C. J. Armishaw, C.-I. A. Wang, R. J. Lewis, J. L. Martin, P. G. Noakes, D. J. Craik, D. J. Adams, and P. F. Alewood, *J. Am. Chem. Soc.*, 2010, **132**, 3514. (e) L. Moroder, H.-J. Musiol, M. Götz and C. Renner, *Biopolymers*, 2005, **80**, 85;
14 (a) M. Reinwarth, D. Nasu, H. Kolmar and O. Avrutina, *Molecules*, 2012, **17**, 12533; (b) L. Moroder and N. Budisa, *Chemphyschem*, 2010, **11**, 1181; (c) C. Boulègue, H.-J. Musiol, V. Prasad and L. Moroder, *Chem. Today*, 2006, **24** (e) L. Moroder, D. Besse, H. J. Musiol, S. Rudolph-Böhner and F. Siedler, *Biopolymers*, 1996, **40**, 207.
15 For examples see: (a) R. Gyanda, J. Banerjee, Y.-P. Chang, A. M. Phillips, L. Toll and C. J. Armishaw, *J. Pept. Sci.* 2013, **19**, 16; (b) A. M. Steiner and G. Bulaj, *J. Pept. Sci.* 2011, **17**, 1; (c) C. J. Armishaw, N. Singh, J. L. Medina-Franco, R. J. Clark, K. C. M. Scott, R. A. Houghten and A. A. Jensen, *J. Biol. Chem.* 2010, **285**, 1809; (d) B. R. Green and G. Bulaj, *Protein Pept. Lett.* 2006, **13**, 67; (e) E. Fuller, B. R. Green, P. Catlin, O. Buczek, J. S. Nielsen, B. M. Olivera, and G. Bulaj, *FEBS*, 2005, **272**, 1727; (f) R. DeLa Cruz, O. Buczek, G. Bulaj, and F. G. Whitby, *J. Pept. Res.*, 2003, **61**, 202.
16 C. A. Hood, G. Fuentes, H. Patel, K. Page, M. Menakuru, J. A. E. H. Park, *J. Pept. Sci.* 2008, **14**, 97.
17 J. T. Chin, S. L. Wheeler, A. M. Klibanov, *Biotechnol. Bioeng.* 1994, **44**, 140; (b) M. Gonnelli, G. B. Strambini, *Biophys. J.* 1993, **65**, 131; (c) K. Tsumoto, M. Umetsu, I. Kumagai, D. Ejima, J. S. Philo, T. Arakawa, *Biotechnol. Prog.* 2004, **20**, 1301; (d) Y. Hagihara, S. Aimoto, A. L. Fink, Y. Goto, *J. Mol. Biol.* 1993, **231**, 180.
18 P. C. Jocelyn, *Eur. J. Biochem.* 1967, **2**, 327.
19 S. C. Johnson, *Psychometrika* 1967, **32**, 241.
20 H. Abdi, L. J. Williams, *Wiley Interdiscip. Rev. Comput. Stat.* 2010, **2**, 433.
21 See Electronic Supplementary Information.
22 For more information on the VENOMICS project, see http://www.venomics.eu

