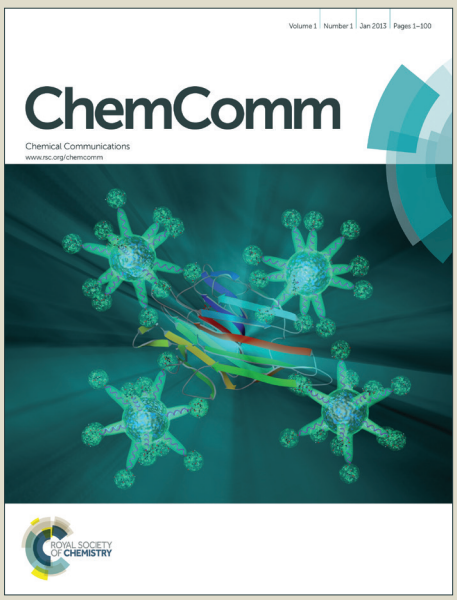
ChemComm

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

Journal Name

RSCPublishing

COMMUNICATION

One-pot hydrazide-based native chemical ligation for efficient chemical synthesis and structure determination of toxin Mambalgin-1

Cite this: DOI: 10.1039/x0xx00000x

Received 00th January 2012,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

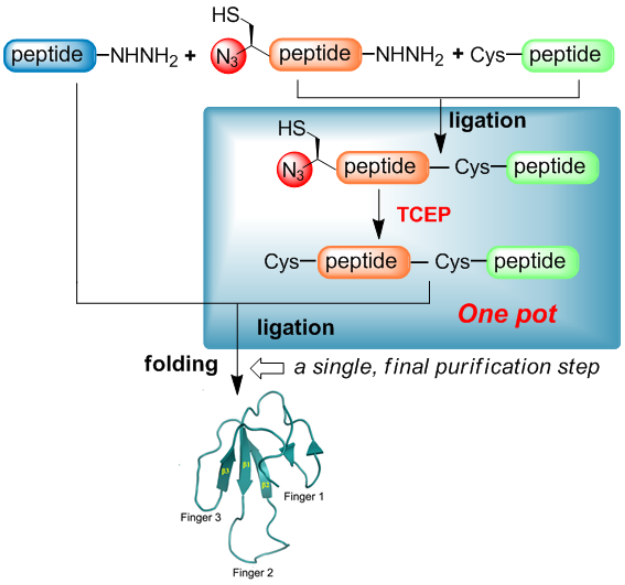
www.rsc.org/

Man Pan,[a,b‡] Yao He,[a,‡] Ming Wen,[a] Fangming Wu,[a] Demeng Sun,[a,b] Sijian Li,[c] Longhua Zhang,[a,*] Yiming Li,[b,c,*] Changlin Tian[a,*]

An efficient one-pot chemical synthesis of snake venom toxin Mambalgin-1 was achieved using an azide-switch strategy combined with hydrazide-based native chemical ligation. Synthetic Mambalgin-1 exhibited a well-defined structure after sequential folding *in vitro*. NMR spectroscopy revealed a three-finger toxin family structure, and the synthetic toxin inhibited human acid-sensing ion channel 1a.

Animal toxins are known to target a wide variety of receptors and ion channels with high affinity and specificity, and they are important pharmacological tools for studying ion channel structure-function relationships, gating mechanisms and tissue localization. Some of these polypeptide toxins have undergone clinical trials and have been developed as venom-based drugs that inhibit neuronal channels involved in pain and other sensory transduction pathways.[1] Mambalgin-1 is a cysteine rich, 57 residue polypeptide isolated from the venom of black mamba snakes, and belongs to the three-finger toxin family. Recently, Mambalgin-1 was demonstrated to abolish pain through inhibition of acid-sensing ion channels (ASICs) either in central or peripheral neurons, and this was achieved without the side effects associated with traditional opioid drugs.[2] Despite this highly potent analgesic effect, the absence of a structure has hindered essential mechanistic studies. An inability to isolate sufficient quantities of pure toxin from black mamba snake venom has been a major bottleneck to structural and functional studies.

In order to facilitate Mambalgin-1 inhibition studies, chemical synthesis may be used to obtain large quantities of highly pure and homogeneous toxin.[3] In this study, we report a one-pot chemical synthesis and ligation approach[4] that enabled the synthesis of full-length Mambalgin-1, and the NMR structure of this cysteine rich protein is also presented.



Scheme 1 One-pot ligation of peptide hydrazides for the synthesis of Mambalgin-1

Mambalgin-1 contains 57 residues, including four intra-molecular disulfide bonds between C3 – C19, C12 – C37, C41 – C49 and C51 – C55.[2] Similar as recently reported, stepwise synthesis of Mambalgin-1 using standard Fmoc-based SPPS initially failed, even when optimized protocols were used (Figure S1 in the Supplementary Information).[5] We next attempted convergent synthesis of the peptide from the C- to the N-terminus using Acm protecting groups and hydrazide-based native chemical ligation (NCL).[6,7] However, a very low yield of final product was obtained using this approach (7.3%) and the process was time-consuming due to the requirement of three HPLC purification steps. To circumvent these problems, we used an azide switch method, which has been used previously to

[a] *School of Life Sciences, University of Science and Technology of China, and High Magnetic Field Laboratory, Chinese Academy of Sciences, Hefei 230026, China. Email: zlhustc@ustc.edu.cn & cltian@ustc.edu.cn*
[b] *Tsinghua-Peking Center for Life Sciences, Tsinghua University, Beijing 100084, China.*
[c] *School of Medical Engineering, Hefei University of Technology, Hefei, Anhui 230009, China, E-mail: lym2007@mail.ustc.edu.cn*
‡These authors contributed equally to this work
†Electronic Supplementary Information (ESI) available: Experimental details. See DOI: 10.1039/b000000x/

control folding of the NPY peptide, which allows a one-pot protecting group-free synthesis by converting the N-terminal Cys residue of the middle segment into an azide.[8] This group is rapidly activated in water using TCEP via intramolecular O-N acyl transfer to generate the native amine group at the N-terminus.[9] To investigate this strategy for peptide ligation, racemization of test and model peptides was initially performed, and this worked well (Figure S2, S3). This strategy was then combined with standard hydrazide-based NCL to synthesize full-length Mambalgin-1. Specifically, the polypeptide chain was divided into three segments containing two ligation sites (Cys19 and Cys41), and Cys19 was switched to an azide group. Ligation can therefore only occur between MA-1[19-40]-$NHNH_2$ and MA-1[41-57]. After reaction, the azide was reduced to an amino group and MA-1[1-18]-$NHNH_2$ was subsequently ligated with MA-1[19-57] (Figure S5). Each segment was synthesized with relative high purity (Figure S4). Without the purification and deprotection of the ACM group, this synthesis of Mambalgins-1 is much more convenient and produced high yields following the final HPLC purification (35% isolated yield, multi-milligram scale and good homogeneity).

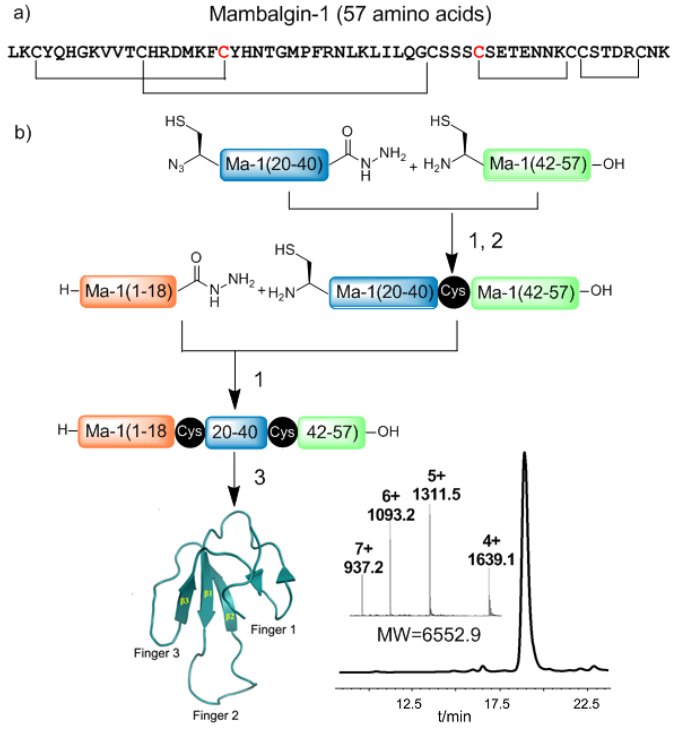


Fig. 1 Synthesis of Mambalgin-1 using the azide switch strategy combined with hydrazide-based native chemical ligation. a) Sequence of Mambalgin-1. b) Procedure for ligation, 1: hydrazide-based ligation, 2: reduction by 5% TCEP, 3: folding. Analytical HPLC (λ = 214 nm) and ESI-MS for the final purified Mambalgin-1 were performed. Observed mass 6552.9 Da, versus calculated mass 6554.6 Da (average isotopes).

We next optimized the *in vitro* folding of the polypeptide to ensure correct formation of the four disulfide bridges. Folding was more efficient under oxidizing conditions containing 10 µM peptide with 1000 µM GSH and 100 µM GSSH, pH = 7.8. Under these conditions, the reaction reached equilibrium in 24 h and the folded Mambalgin-1 was purified by reverse phase HPLC (55% yield). CD spectroscopy confirmed that the synthesized Mambalgin-1 contained well-defined β-sheet structure (Figure S6 in the Supplementary Information).

To confirm that synthetic Mambalgin-1 was folded into the correct biologically active conformation, acid-induced currents with or without application of synthetic Mambalgin-1 were recorded using patch-clamp electrophysiology in CHO cells over-expressing recombinant human ASIC1a (hASIC1a) channels. At a holding potential of -70 mV, the acid current was elicited by a pH drop from 7.4 to 6.0 for 4 s every 3 min through a Y-tube perfusion system that allowed local and rapid changes of solutions. After two stable records of acid currents, synthetic Mambalgin-1 was applied at pH 7.4 for 30 s followed by a 4 s application at pH 6.0. Synthetic Mambalgin-1 inhibited the recombinant hASIC1a channel currents in a concentration-dependent manner, with an IC_{50} of 203 ± 27 nM (n = 3-14). Therefore, further optimizations in peptide refolding and functional studies are necessary to make the physiological conductance data from the synthesized peptide to reach the value of the native toxin isolated from snake venom (hASIC1a in transfected COS-7 cells: 127 nM).[2]

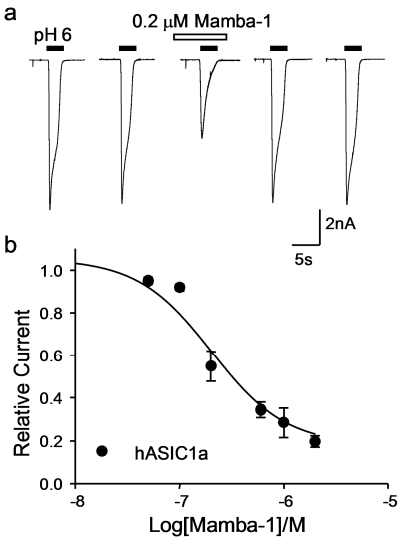


Fig. 2 (a) Synthetic Mambalgin-1 (Mamba-1, in 0.2 µM) reversibly inhibits recombinant human ASIC1a channels in CHO cells. After two stable records of acid currents that were elicited by a pH drop from 7.4 to 6.0 for 4 sec, synthetic Mamba-1 was applied at pH 7.4 for 30 sec followed by another application during a 4 sec pH drop to 6.0. Two additional acid currents were measured after washing to examine the recovery from Mamba-1 treatment; (b) Concentration-response curve showing the inhibition of human ASIC1a channels expressed in CHO cells (n = 3-14) by synthetic Mambalgin-1.

With confirmation that the synthetic Mambalgin-1 was functional and correctly folded, we determined the three-dimensional structure using a set of homo-nuclear two dimensional solution NMR experiments (Table 1 and Figure S8 in the Supplementary Information).[10] Since the 57-residue Mambalgin-1 is large, a relatively high temperature (310 K) was

applied during solution NMR spectra collection. Sufficient resolution was achieved to allow proton resonance assignments (proton chemical shifts were deposited in BMRB database: 19746). A total of 740 distance restraints (including 159 intra-residue, 238 sequential, 102 medium-range and 241 long-range) derived from cross peak intensities in NOESY spectra were applied for structural calculations using Xplor-NIH. The 20 structures with the lowest energy of the 200 calculated structures were applied for convergence analysis. The root-mean-square deviations (RMSD) from the average structure of Mambalgin-1 was 0.306 Å for the backbone atoms and 0.848 Å for all heavy atoms of the residues in secondary structure regions (residues 3-4, 9-10, 18-22, 34-38, and 47-50).

From the Mambalgin-1 structure, four disulfide bridges (Cys3-Cys19, Cys12-Cys37, Cys41-Cys49 and Cys50-Cys55) were observed to stabilize the protein conformation. A typical three-finger toxin structure was observed. Beta-sheets consisting of three anti-parallel β-strands (β1: 18-22; β2: 34-38 and β3: 47-50) are located in the core region of the toxin, while two short fingers (finger 1: 1-17 and finger 3: 39-46) flank the long middle finger (finger 2: 23-33). Very few NOE restraints were assigned to protons of residues in loop2, probably due to high flexibility of the long loop.

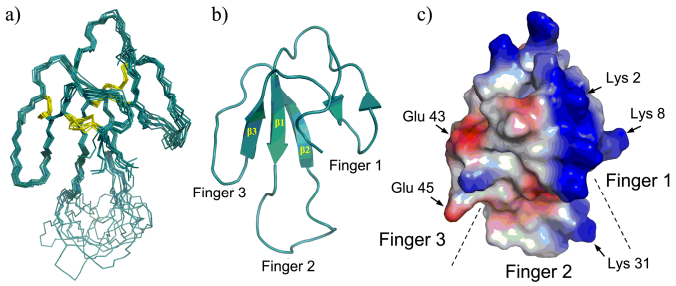


Fig. 3 Structure of synthetic Mamba-1. (a) Backbone alignment of the 20 lowest-energy structures (PDB 2MJY). The four disulfide bonds are colored yellow; (b) Ribbon representation of Mamba-1 showing the typical three-finger toxin structure. The three antiparallel β-strands and three fingers are labeled; (c) Surface diagram of Mamba-1. Positively charged areas are colored blue, negatively charged areas are red, and uncharged areas are white. The figures were generated using program PyMol.

The overall structure of Mambalgin-1 is very similar to the previously determined Mambalgin-2 structure,[5] with an RMSD of 0.51 Å for the backbone atoms in the secondary structure regions (Figure S7). Indeed, there is only a single residue difference between the two toxins (Tyr4 in Mambalgin-1 versus Phe4 in Mambalgin-2). Surface charge distributions are also highly similar, and several clusters of basic or acidic residues map to both structures. There are two negatively charged residues (Glu43 and Glu45) in the third finger, two positively charged residues (Arg28 and Lys31) in the second finger, and a large positively charged cluster in the first finger and the core domain. A stretch of hydrophobic residues (Met25, Phe27, Leu30, Leu32 and Leu34) is observed in close vicinity to the middle finger. The surface charge distribution and three-finger conformation are presumably responsible for the highly potent binding to and inhibition of the ASIC channels by Mambalgin-1.

In conclusion, we devised an efficient one-pot synthesis of the toxin Mambalgin-1 using an azide switch strategy combined with hydrazide-based native chemical ligation. This method provides access to multi-milligram quantities of Mambalgin-1 in a single final purification step. After sequential folding, synthetic Mambalgin-1 was shown to be correctly folded and inhibited human acid-sensing ion channel 1a. NMR spectroscopy revealed a typical three-finger toxin family structure.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (21372058 to Y. M. Li; 31100847 to L. H. Zhang; U1332138 to C. L. Tian). We thank Prof. Lei Liu of Tsinghua University for helpful suggestions, and Prof. Yun Jiang of Kunmin Institute of Zoology for encouraging discussions.

Notes and references

1 (a) A. Baron, S. Diochot, M. Salinas, E. Deval, J. Noël, E. Lingueglia, *Toxicon*, 2013, **75**, 187-204; (b) R.J.P. Dawson, J. Benz, P. Stohler, T. Tetaz, C. Joseph, S. Huber, G. Schmid, D. Hugin, P. Pflimlin, G. Trube, M. G. Rudolph, M. Hennig and A. Ruf *Nat. Commu.* 2012, **3**, 936-944; (c) I. Baconguis and E. Gouaux, *Nature*, 2012, **489**, 400-405.

2 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**, 552-555.

3 (a) D. J. Craik and C. I. Schroeder, *Angew. Chem., Int. Ed.,* 2013, **52**, 3071-3073; (b) A. Ghassemian, C. I. A. Wang, M. K. Yau, R. C. Reid, R. J. Lewis, D. P. Fairlie, P. F. Alewood, D. J. Craik and T. Durek, *Chem. Commun.*, 2013, **49**, 2356-2358; (c) J. E. Jensen, T. Durek, P. F. Alewood, D. J. Adams, G. F. King and L. D. Rash, *Toxicon*, 2009, **54**, 56-61; (d) T. Durek, I. Vetter, C. I. A. Wang, L. Motin, O. Knapp, D. J. Adams, R. J. Lewis and P. F. Alewood, *ACS Chem. Biol.*, 2013, **8**, 1215-1222; (e) R. J. Clark, J. Jensen, S. T. Nevin, B. P. Callaghan, D. J. Adams and D. J. Craik, *Angew. Chem. Int. Ed.*, 2010, **49**, 6545-6548.

4 (a) D. Bang, D. S. B. H. Kent, *Angew. Chem., Int. Ed.*, 2004, **43**, 2534–2538; (b) D. Bang, B. L. Pentelute and S. B. H. Kent, *Angew. Chem. Int. Ed.*, 2006, **45**, 3985; (c) J. S. Zheng, H. K. Cui, G. M. Fang, W. X. Xi and L. Liu, *ChemBioChem*, 2010, **11**, 511-515; (d) R. A. Morales, N. L. Daly, I. Vetter, M. Mobli, I. A. Napier, D. J. Craik, R. J. Lewis, M. J. Christie, G. F. King, P. F. Alewood and T. Durek, *ChemBioChem*, 2010, **11**, 1882-1888.

5 During the preparing of manuscript, related literature of the synthesis and 3D structure of Mambalgin-2 was reported. C. I. Schroeder, L. D. Rash, X. Vila-Farrs, K. J. Rosengren, M. Mobli, G. F. King, P. F. Alewood, D. J. Craik and T. Durek, *Angew. Chem., Int. Ed.,* 2014, **53**, 1017-1020.

6 (a) S. B. H. Kent, *Chem. Soc. Rev.*, 2009, **38**, 338; (b) V. R. Pattabiraman and J. W. Bode, *Nature*, 2011, **480**, 471; (c) L. Raibaut, N. Ollivier and O. Melnyk, *Chem. Soc. Rev.*, 2012, **41**, 7001; (d) P. Siman and A. Brik, *Org. Biomol. Chem.*, 2012, **10**, 5684–5697; (e) P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science*, 1994, **266**, 776–779; (f) L. E. Canne, S. J. Bark and S. B. H. Kent, *J. Am. Chem. Soc.*, 1996, **111**, 5891–5896; (g) V. Y. Torbeev, S. B. H. Kent, *Angew. Chem., Int. Ed.*, 2007, **46**, 1667–1670; (h) K. S. A. Kumar, S. N. Bavikar, L. Spasser, T. Moyal, S. Ohayon and A. Brik, *Angew. Chem. Int. Ed.*, 2011, **50**, 6137-6141; (i) S. Liu, B. L. Pentelute, S. B. H. Kent, *Angew. Chem., Int. Ed.*, 2012, **51**, 993–999; (j) M. Avital-Shmilovici, K. Mandal, Z. P.Gates, N. B. Phillips, M. A. Weiss, S. B. H. Kent, *J. Am. Chem. Soc.*, 2013, **135**, 3173–3185; (k) Y. C. Huang, Y. M. Li, Y. Chen, M. Pan, Y. T. Li, Li, Y. Q. X. Guo and L. Liu, *Angew. Chem. Int. Ed.*, 2013, **52**, 4858-4862.

7 (a) J. S. Zheng, S. Tang, Y. C. Huang and L. Liu, *Acc. Chem. Res.*, 2013, **46**, 2475-2484; (b) G. M. Fang, Y. M. Li, F. Shen, Y. C. Huang, J. B. Li,

Y. Lin, H. K. Cui and L. Liu, *Angew. Chem. Int. Ed.*, 2011, **50**, 7645-7649; (c) J. S. Zheng, H. N. Chang, F. L. Wang and L. Liu, *J. Am. Chem. Soc.*, 2011, **133**, 11080-11083; (d) Y. M. Li, M. Y. Yang, Y. C. Huang, Y. T. Li, P. R. Chen and L. Liu, *ACS Chem. Biol.*, 2012, **7**, 1015-1022; (e) G. M. Fang, J. X. Wang and L. Liu, *Angew. Chem. Int. Ed.*, 2012, **51**, 10347-10350; (f) J. S. Zheng, S. Tang, Y. K. Qi, Z. P. Wang and L. Liu, *Nat. Protoc.*, 2013, **8**, 2483-2495; (g) P. Siman, S. V. Karthikeyan, M. Nikolov, W. Fischle and A. Brik, *Angew. Chem. Int. Ed.*, 2013, **52**, 8059; (h) Y. M. Li, Y. T. Li, M. Pan, X. Q. Kong, Y. C. Huang, Z. Y. Hong and L. Liu, *Angew. Chem. Int. Ed.*, 2014, **53**, 2198-2202.

8 N. Nepomniaschiy, V. Grimminger, A. Cohen, S. DiGiovanni, H. A. Lashuel and A. Brik, *Org. Lett.*, 2008, **10**, 5243-5246.

9 (a) S. D. Santos, A. Chandravarkar, B. Mandal, R. Mimna, K. Murat, L. Saucede, P. Tella, G. Tuchscherer, M. Mutter, *J. Am. Chem. Soc.*, 2005, **127**, 11888–11889; (b) A. Taniguchi, Y. Sohma, M. Kimura, T. Okada, K. Ikeda, Y. Hayashi, T. Kimura, S. Hirota, K. Matsuzaki, Y. Kiso, *J. Am. Chem. Soc.*, 2005, **128**, 696–699.

10 Wuthrich K. 1986. NMR of proteins and nucleic acids. New York: Wiley.

