

RESEARCH

Open Access



Scalable production of recombinant three-finger proteins: from inclusion bodies to high quality molecular probes

Jiang Xu^{1*}, Xiao Lei¹, Ao Li¹, Jun Li², Shuxing Li¹ and Lin Chen^{1*}

Abstract

Background The three-finger proteins are a collection of disulfide bond rich proteins of great biomedical interests. Scalable recombinant expression and purification of bioactive three-finger proteins is quite difficult.

Results We introduce a working pipeline for expression, purification and validation of disulfide-bond rich three-finger proteins using *E. coli* as the expression host. With this pipeline, we have successfully obtained highly purified and bioactive recombinant α -Bungarotoxin, κ -Bungarotoxin, Hannalgesin, Mambalgin-1, α -Cobratoxin, MT α , Slurp1, Pate B etc. Milligrams to hundreds of milligrams of recombinant three finger proteins were obtained within weeks in the lab. The recombinant proteins showed specificity in binding assay and six of them were crystallized and structurally validated using X-ray diffraction protein crystallography.

Conclusions Our pipeline allows refolding and purifying recombinant three finger proteins under optimized conditions and can be scaled up for massive production of three finger proteins. As many three finger proteins have attractive therapeutic or research interests and due to the extremely high quality of the recombinant three finger proteins we obtained, our method provides a competitive alternative to either their native counterparts or chemically synthetic ones and should facilitate related research and applications.

Keywords Three finger protein, *E. coli* recombinant expression, Inclusion body refolding, Disulfide bond formation, Oxidation refolding

Introduction

The snake venom is a large repertoire of digestive enzymes, toxin peptides and compounds. The three-finger neurotoxins (TFNs) are a collection of such toxin peptides used by the snake to kill the prey through binding and blocking ion channels in the

neurological system. The α -neurotoxin binds to and block the muscle type nicotinic acetylcholine receptor ($\alpha_1\beta_1\gamma(\epsilon)\delta$ nAChR) on the neuromuscular junction, leading to paralysis of the prey. The snake venom also contains various toxin peptides that have interesting properties. Some of them have analgesic effect, such as mambalgin-1 [1] and Hannalgesin [2], some of them can bind to receptors in the neurological system, such as MT α and κ -Bungarotoxin, which bind to Muscarinic α_2 B-adrenoceptor [3, 4] and $\alpha_3\beta_2$ nAChR [5, 6], respectively. Although these toxin peptides may have attractive usage in biomedical research, not all toxin peptides are present in large quantities in snake venoms. κ -Bungarotoxin (κ Btx), for example, only takes a very small fraction of the venom, was often

*Correspondence:

Jiang Xu
foxjuly@gmail.com
Lin Chen
linchen@usc.edu

¹ Molecular and Computational Biology, Department of Biological Sciences, University of Southern California, Los Angeles, CA 90089, USA

² Institute of Clinical Medicine, The First Affiliated Hospital of University of South China, Hengyang, Hunan 421001, People's Republic of China



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

contaminated by α -Bungarotoxin (α Btx), leading to inconsistent results in some of the earlier researches [7], and is commercially unavailable now, thus further restricted their study and usage.

Three-finger proteins (TFPs) are a class of proteins (peptides) including TFNs and their mammalian homologue that show similar three-finger structure as TFNs. These mammalian TFP homologues are referred to as the Ly6/uPAR family proteins, some famous players of which include Lynx [8–12], Slurp [13–19], which are proposed to be endogenous modulators of nAChR, the Pate (Prostate and Testis Expression) family [20–23], which play important roles in the capacitation of the sperm and fertilization and CD59, a GPI-anchored membrane protein that protects the cell from complement attack [24–27]. As the detailed biological function of many of these toxin-like proteins (peptides) still largely remains elusive, it is thus desirable to have a reliable production method for the scientific community.

Due to the complex intramolecular disulfide bonds system in TFPs, it is usually impossible to obtain correctly folded TFPs directly from *E. coli* and the expressed recombinant proteins are always in the form of inclusion body (I.B.), and all previous attempts included an additional refolding process [28–30]. Other expression systems were also attempted, such as *Pichia Pastoris* [31–33], or *Eukaryotic* expression systems [34]. While these attempts obtained recombinant three finger proteins (rTFP) and met the end claimed in the research, all these methods suffered from sophisticated post-purification cleavage of fusion tags and low yield. Until now, very few of these recombinant three-finger proteins (rTFPs) have been structurally validated by X-ray diffraction (XRD) protein crystallography studies. There were successful attempts using chemical synthesis, such as muscarinic toxin MT7 and MT1 [35], and the pain-killing toxin Mambalgin [1, 36, 37], but this method suffered from sophisticated synthesizing steps and high costs. As such, a universal, high yield production protocol capable of producing high quality rTFPs is desired.

We previously reported the production of bioactive recombinant α -Bungarotoxin (rec- α Btx) [38], in which we used radioimmunoassay to determine the optimal refolding conditions. For other TFPs, without a valid activity determination method, it is usually hard to tell which refolding condition is optimal for a particular TFP. Here we introduce an efficient, productive pipeline that we used to generate various rTFPs, most of which were structurally and biochemically validated.

Materials and methods

The Buffers and Mediums used are summarized in detail in Additional file 5: Table S1. Restriction enzymes were from Takara or New England Biolabs. All chemicals were from Sigma-Aldrich unless otherwise stated. The overall experimental design is illustrated in Fig. 1.

Vector construction, *E. coli* fermentation and inclusion body extraction

Genes encoding the toxin proteins were codon-optimized to *E. coli*-preferred codons with the Jcat software [39] and synthesized (Genscript Inc, Integrated DNA Technologies Inc) with NdeI (CATATG) or NcoI (CCATGG) site as described before [38] on the 5' end and a termination codon (TAA or TAG) at the 3' end just before the XhoI sites (Additional file 4). The genes were inserted into the NdeI and XhoI sites of pET30b (Novagen) or the NcoI and XhoI sites of pET22b (Novagen), and the reconstructed expression vector was transformed into the *E. coli* cloning host JM109 or Stellar™ (Takara). Positive clone was confirmed by Sanger sequencing and then transformed into expression host BL21(DE3), and was tested for protein expression at 37 °C in 2 ml (or 5 ml) culture in 10 ml (or 20 ml) test tubes, in which IPTG was added to 1 mM at OD₆₀₀ of 0.7, and carried on for 4 h. For analysis on SDS-PAGE gels, 1.5 ml of induced culture were pelleted down, and resuspended in 100 μ l of 2xSDS-PAGE loading dye, boiled for 5 min and run on 15% SDS-PAGE gel. The protein bands of interest from the SDS-PAGE gel were digitized by measuring the pixels' value sum of each band (referred to band amount hereafter) from the inverted grey scale picture of the gel graph with Image J [40]. For rTFPs with observable expression on SDS-PAGE, *E. coli* cells were further fermented either with a home-made 5 L bioreactor or a BioFlo3000 10 L Bioreactor (New Brunswick Scientific). Both bioreactors' fermentation temperature was set to 37 °C. For BioFlo3000 10 L bioreactor, the volume of the base fermentation medium was 6 L, and the stir speed was looped to DO (dissolved oxygen), which was set to 20%, with a minimum speed set to 500 rpm and maximum speed set to 1000 rpm, air flow was set to 10 L/min. The pH value of the medium was automatically kept at around 7 with addition of concentrated ammonium hydroxide solution throughout the fermentation process. For the home-made 5 L bioreactor, the volume of the base fermentation medium was 3 L, and a fixed stirring speed of 1000 rpm was used, air flow was set to 5 L/min. The pH value of the medium was kept at around 7 with addition of concentrated ammonium hydroxide solution after the first DO peak, at 40 min interval by manually sampling the culture and measuring the pH value

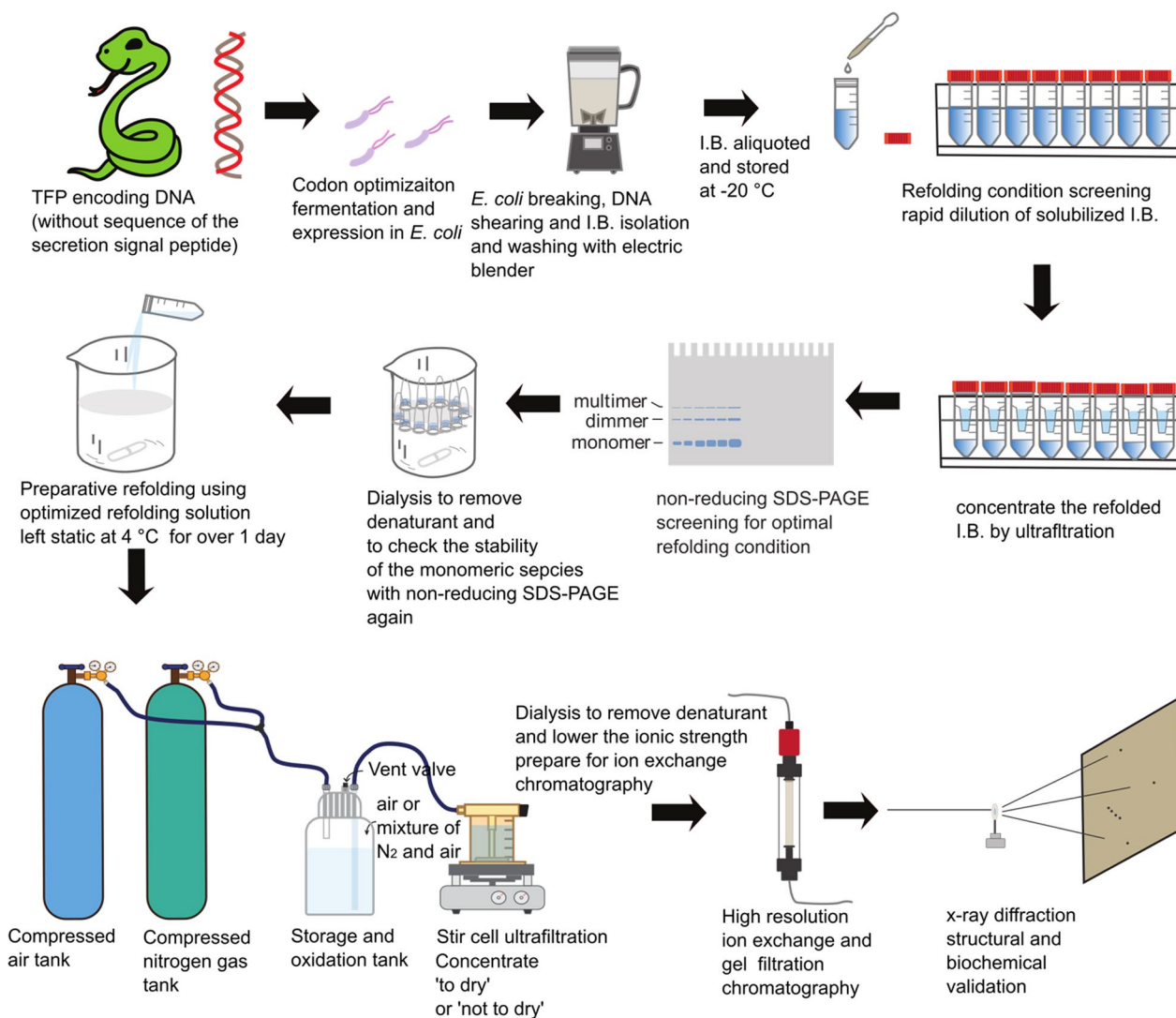


Fig. 1 Experimental flow chart illustration

with pH test strips. Culture density was monitored by measuring the OD_{600} of the diluted culture (with 1xPBS), so the measured OD_{600} value was less than 0.3, at 1 h or 20 min interval, when the culture's OD_{600} was below or above 10, respectively.

Typically, 4~10 freshly transformed *E. coli* colonies were inoculated into 250 ml/500 ml (for 5 L home-made bioreactor or 10 L BioFlo3000 bioreactor, respectively) of 2xYT medium supplemented with 0.5% glucose and 100 µg/ml ampicillin or 50 µg/ml kanamycin according to the pET plasmid used. The *E. coli* culture was incubated in a shaking incubator at 37 °C/250 rpm to an OD_{600} value of 0.7, then inoculated into the base fermentation medium and fermented using the fed-batch strategy with addition of 40% glucose to 1% each time DO

started to peak, until the OD_{600} of the culture reached 19~22. IPTG was then added to 1 mM to induce protein expression for 4 h, after adding IPTG, feed medium was switched to 1/10 culture volume of GYT each time DO started to peak. The growth of *E. coli* dramatically slowed down after addition of IPTG, as could be observed by the increased time interval between each feed and decreased increment rate of OD_{600} of the culture. Normally, 160~200 or 350~400 g of bacteria pellets (wet weight) could be obtained with the 5-L home-made bioreactor or BioFlo3000 10 L bioreactor, respectively. The *E. coli* pellets were collected and stored at -20 °C as 50 g aliquots. To obtain the inclusion bodies, 200 g of bacteria was thawed in 1 L of lysis buffer supplemented with 2 mg of chicken egg lysozyme per gram of bacteria

pellets of was then added and mixed well using a bench-top electric blender (KitchenAid KSB1575ER Blender, or Philips HR2094/00 blender) at their top-speed. The mixture was incubated on ice for 1 h and sheared with the blender at top-speed for 60 s and cooled in the cold room for 15 min, the shearing process was repeated twice until the solution become less sticky, which was then centrifuged at 10,000g/4 °C/15 min. The supernatant was discarded, and the pellets were subjected to a new round of resuspension-shearing-centrifugation process until the pellets became compact and supernatant turned from turbid to translucent. The pellets were finally resuspended evenly in 1~2 L of lysis buffer (the exact volume depends on the total amount of crude I.B., which should be kept within 1 g/tube (wet weight) with the help of the electric blender and aliquoted to 50 ml/tube in 50 ml conical tubes, pellet down by centrifugation at 8000g/10 °C/15 min and stored at -20 °C until use.

I.B. solubilization and refolding screen

To solubilize the I.B., a solubilization buffer containing 50 mM Tris-HCl (pH 8.0), 8 M urea or 6 M guanidine-HCl and 5 mM 2-ME was used. The choice of the solubilization buffer was based on the solubilization effect and contaminating protein level. Taken α Btx for example, this toxin refolded poorly in the presence of contaminating proteins and its I.B. was solubilized well with a solubilization buffer containing 8 M urea. So, after solubilization with 50 mM Tris base, 8 M urea, and 5 mM 2-ME, and centrifuged for 28,000g/10 min/4 °C to get rid of insoluble bacteria debris, the pH of the supernatant was adjusted to 8.5 with concentrated HCl solution and further absorbed with Q Sepharose FF media (2 ml of solution/ml of Q media) equilibrated with 50 mM Tris-HCl (pH 8.5), 8 M urea. Attention should be paid to avoid using high concentrations of reducing chemical reagents (like 100 mM of 2-ME) at the solubilization stage, which will lead to low refolding efficiency that may be caused by blocking the formation of disulfide bonds. After absorption, the I.B. was ready for refolding. For other toxins with higher expression level and more compact inclusion bodies, a solubilization buffer containing 50 mM Tris-HCl (pH 9.0), 6 M guanidine-HCl, 5 mM 2-ME was used. A simple but useful test to assess the compactness of the I.B. is to poke the I.B. after the last washing step at the bottom of the 50 ml conical centrifuge tube with a 200 μ l pipette tip, if the I.B. is compact, the pipette tip should be able to stand, and if the I.B. is loose, the pipette tip will tend to fall. This is particularly useful when selecting between urea based and guanidine hydrochloride-based solubilization buffers.

Refolding condition was optimized with a screening protocol scouting for NaCl concentration (0 or 200 mM),

L-cysteine concentration (0 to 64 mM), L-arginine-HCl (0 or 0.5 M, pH 8.8), and detergent, such as NDSB-201(0 or 0.2 M), etc. Standard refolding trial was made by diluting 200 μ l of I.B. solution into 10 ml of refolding screen solution with detailed recipe presented in Material and Method. After dilution, the solutions were left at 4 °C overnight, and concentrated each with Amicon Ultra-15 (Millipore, 3 kDa NMWL) ultrafiltration device to about 200 μ l. The retention was centrifuged at 18,000g/4 °C for 15 min and the supernatant analyzed with non-reducing SDS-PAGE. Generally, mis-paired disulfide bonds could lead to formation of intermolecular disulfide bonds and multimeric species, which are shown as a ladder pattern on non-reducing SDS-PAGE (Additional file 1: Figure S1, grey arrow), while correctly paired disulfide bonds facilitate formation of monomeric species, which are usually shown as the smallest band on non-reducing SDS-PAGE (Additional file 1: Figure S1, black arrow). The rest of the concentrated solutions were each divided into three parts and dialyzed against low ionic strength buffer with various pH values, such as 20 mM sodium acetate (NaAc, pH 5.0), 20 mM HEPES (pH 7.0, adjusted with NaOH), or 20 mM Tris-HCl (pH 8.0), using a set of home-made micro-dialysis devices [41]. Finally, the dialyzed solution was centrifuged at 18,000g/4 °C for 15 min. The supernatant was analyzed with non-reducing SDS-PAGE to assess the stability of different refolding species (i.e., monomer and multimers). For quantitative comparison between different refolding conditions, two metrics, 'monomer ratio' and 'relative monomer amount' were calculated. 'Monomer ratio' is the ratio between the band amount of monomeric species and the band amount of the total, which is the band amount sum of monomeric, dimeric, trimeric, tetrameric, etc. species, if presented. 'Relative monomer amount' (shown as 'rel. monomer amount' in Additional file 1: Figure S1) was calculated by dividing the monomer band amount from each condition with the highest monomer band amount for the same rTFP in refolding screen experiment. By simply eyeballing the gel graph, or more accurately by comparing the 'rel. monomer amount' and 'monomer ratio', the best refolding condition was selected, which gave relatively high level of monomeric species shown by high 'rel. monomer amount' and 'monomer ratio', with relatively less usage of L-cysteine and without usage of L-arginine and NDSB-201. For some rTFNs, such as rec- κ Btx, which is a dimer not interconnected with inter-chain disulfide bonds in its native form, we looked on non-reducing SDS-PAGE for conditions that produced the least amount of dimer and highest amount of monomer, as these dimers (which are interconnected with disulfide bonds) could be very hard to be separated from native dimeric rec- κ Btx (data not shown).

Preparative refolding of rTFPs

For preparative refolding, freshly solubilized I.B. was poured all in once at a volume ratio of 1:50 into a freshly prepared, ice-cold refolding solution which was stirred rapidly by a magnetic bar throughout the whole process. The refolded solution was left static over one day at 4 °C and concentrated with a compressed nitrogen-gas (or air) driven ultrafiltration device (350 ml Amicon Stirred Cell, 3 kDa NMWL membrane, Millipore). As the ultrafiltration device's volume is small compared to the total volume of the refolding solution (i.e., several liters), repeated refilling was needed where O₂ in the air was brought in, helping cysteine molecules in the solution to gradually react to form cystine and precipitated out, during which the disulfide bonds of the rTFP also formed. Considering this, we designed and crafted a 'storage and oxidation tank' and connected it in serial in between the compressed gas tank and the stirred ultrafiltration cell (Fig. 1). The 'storage and oxidation tank' not only eliminated the needs for repeated refilling of the stir cell during the ultrafiltration, but also served as a reaction vessel, making it easier in selection of compressed nitrogen gas or air or the mixture of the two, ensuring complete and homogenous oxidation formation of disulfide bonds under higher dissolved oxygen environment with the high pressure. Close to the end of the ultrafiltration process, the refolded solution was either concentrated to a very small volume of several ml ('not to dry'), or 'to dry', leaving no visible liquid on the ultrafiltration membrane, depending on the type of the rTFPs being refolded. For some toxins, like recombinant MT α (rec-MT α), Hannalgesin (rec-Hannalgesin), mouse Pate B (rec-mPateB), κ -Bungarotoxin (rec- κ Btx), α -Bungarotoxin (rec- α Btx), it is better to concentrate to 'dry', which dramatically increased the purity and quality of the final product. For other toxins we tried, such as recombinant mambalgin-1 (rec-Mambalgin-1), mouse and human Slurp1 (rec-mSlurp1 and rec-hSlurp1), concentrating to dry significantly lowered the final yield. So, trial experiments should be done at this point. The concentrated product was then re-solubilized with a low ionic strength buffer, which was pre-determined in the dialysis experiment. Normally, proteins with isoelectric point (pI) over 7 was re-solubilized in 30 to 50 ml of 20 mM NaAc (pH 5.0), while proteins with pI less than 7 was solubilized in 20 mM Tris-HCl (pH 8.0) or 20 mM HEPES (pH 7). The solution was then filtered with a 0.2 μ m filter and applied to mono S 5 50 GL or mono Q 5 50 GL column driven by a FPLC system (ÄKTA™ Purifier, GE Healthcare) based on the isoelectric point (pI) of the proteins. Bound proteins were eluted with a linear

gradient of NaCl to 1 M. The eluted peaks were again analyzed by non-reducing SDS-PAGE. Those eluted later usually contained contaminating proteins, or species inter-connected by intermolecular disulfide bonds. For those concentrated 'not to dry' but to small volume, an additional dialysis step was usually added, in which the concentrated solution was dialyzed against the low-ionic strength buffer before being subjected to cation exchange chromatography. For the proteins we tried, a single, large peak was usually seen using the mono S column, and several large peaks were seen using the mono Q column, in which the target species was usually contained in the first peak. At this stage, the refolded rTFP was pure, but for crystal growth, gel filtration was usually done with a Superdex 75 10 300 GL column (GE Healthcare), to further increase the purity of the product and to buffer-exchange to 200 mM ammonium acetate (pH 7 °C).

Native gel shift assay

5 μ g of HAP peptide [42, 43] were mixed with 5 μ g of each of the rTFPs, respectively (with molar ratio HAP:rTFP > 6), incubated at room temperature for 15 min, and run on a 15% native PAGE gel with 50 mM NaAc (pH 5.0) at 120 v/60 min/ 4 °C. For the binding assay with the nicotinic acetylcholine receptors, 5 μ g of recombinant α -Cobratoxin (rec- α CTX), recombinant Hannalgesin (rec-Hannalgesin), or α -Cobratoxin (α CTX) (Sigma-Aldrich, C6903) was mixed with 5 μ g of the recombinant extracellular domain of the α 1 subunit of muscle type nicotinic acetylcholine receptor (rec- α 1ECD) [44, 45] (molar ratio, rTFP (TFP): α 1ECD = 3), incubated on ice for 15 min and run on 12% native gel (standard discontinuous PAGE gel without SDS, 6% for top layer and 10% for bottom layer) with Tris-Glycine buffer (pH 8.3, without SDS) as the running buffer, at 120 v/90 min/4 °C. Gels were stained with Coomassie Brilliant Blue G-250 as described [46].

Labeling of rec-mPate B with fluorescence dye and visualization of binding of rec-mPate B to the mouse spermatozoa

rec-mPate B was labeled with NHS-rhodamine (Thermo Scientific) according to the manufacturer's recommended protocol. Briefly, 25 μ l of rec-mPate B (27.2 mg/ml in PBS, pH 7.4) at was mixed with 20 mM HEPES (pH 7), 4.13 μ l of 18.9 mM NHS-Rhodamine DMSO solution and incubated at room temperature for 60 min, and dialyzed exhaustively against 20 mM HEPES, 0.15 M NaCl. Mouse spermatozoa was obtained as described [47], and was mixed with 1:1000 dilution of the Rhodamine labeled

rec-mPate B, washed three times with PBS, and observed under a laser confocal fluorescence microscope.

X-ray protein crystal diffraction structural validation of rTFPs

For crystallization of rec- α Btx, rec- α Btx was mixed with HAP peptide [42] at a molar ratio of 1:1.5, incubated at room temperature for 30 min and then diluted 100 fold with 20 mM NaAc, pH 5.0 and applied to mono S column. Bound protein was eluted with linear gradient of NaCl to 1 M and the sharp peak containing the rec- α Btx-HAP complex was collected, pooled, and concentrated to about 13 mg/ml, dialyzed against 0.1 M HEPES (pH 7.0) exhaustively at 4 °C. For other rTFPs, purified toxin proteins were concentrated to 80 to 150 mg/ml with Amicon Ultra-15 and Amicon Ultra-0.5 (3 kDa NMWL) tubes. Sitting drop crystal screening was done using a robotic system (Crystal Gryphon, Art Robbins Instrument). Hanging drop method was then done manually to optimize the growth condition, by mixing equal volume of well solution and the toxin protein and incubating both at 4 °C and 18 °C. Crystals were then harvested and stored in cryo-conditions and X-ray diffraction data of rec-kBtx, rec-mambalgin 1 and rec- α Btx-HAP complex were collected either with a Rigaku MicroMaxTM-007 home X-ray source coupled with an R-AXIS IV++ image plate. For rec-MT α crystals, X-ray diffraction data was collected at Advanced Photon Source (Argonne National Laboratory, Lemont, IL). The X-ray diffraction data of rec-Hannalgesin and rec- α CTX were collected at Advanced Light Source (Lawrence Berkeley National Laboratory, Berkeley, CA). Data was processed with HKL2000 [48] or IMosflm [49], Pointless, Aimless [50], Ctruncate from the CCP4 suite [51], Molecular Replacement, structure build and refinement was done in Phenix [52] and Coot [52]. Structural visualization, alignment, and calculation were done in open source PYMOL (Version 2.0 Schrödinger, LLC).

Results

Our pipeline is applicable to a wide variety of TFPs with high yield, with 6 rTFPs being structurally validated.

Our idea is to use *E. coli* as the workhorse to produce high quality TFPs of biomedical interests. Our pipeline includes codon optimization of the encoding DNA sequence for *E. coli* expression, recombinant protein expression in *E. coli*, isolation of I.B., refolding condition screening, preparative oxidation refolding and purification, structural validation with x-ray diffraction and biochemical methods (Fig. 1). From construction of the expression vector with known protein or encoding DNA sequences, production of a purified rTFP usually took 4~5 weeks. For each of the rTFPs, non-reducing

SDS-PAGE were carried out to visualize the monomeric and multimeric species with inter-molecular disulfide bonds and to check the purity of the final product (Fig. 2). For a couple of TFPs of various origin (Additional file 6: Table S2), our pipeline was shown to be robust and successful (Fig. 2, Additional file 7: Table S3, Additional file 4: Materials).

To structurally validate the rTFPs, we screened for crystal growth for each rTFP we purified (the conditions for crystal growth are listed in Additional file 4: Materials). Most of our rTFPs' crystals were formed at high protein concentrations (Additional file 4: Materials). They were beautiful looking under the microscope with polarized light (Additional file 2 Figure S2), and diffracted x-ray quite well. Six rTFPs' structures were solved with x-ray crystal diffraction data using molecular replacement with known homologous structures. The statistics for the diffraction data were summarized in Table 1. From the structural alignment of the solved structures with their native counterparts (such as rec- α Btx-HAP complex, rec- α CTX, rec-kBtx, rec-mambalgin) or with their most homologous native counterparts (such as rec-Hannalgesin and rec-MT α , whose crystal structures were not reported, known homologous TFPs with known structure, such as α CTX and MT1, respectively, were used as the alignment counterpart). Our rTFPs are shown to be almost identical to their natural counterparts, with one or several additional amino acids at the N-terminal (a.a. sequences are shown in Additional file 4: Materials and the additional a.a. sequences are highlighted in red color), which is a unique mark for their recombinant origin) (Fig. 3).

Though it was generally considered very difficult to obtain rTFPs using either *E. coli* or other expression systems, with our pipeline, we could repeatedly obtain over one hundred mg of rec-MT α , rec-Hannalgesin, rec- α CTX and rec-mPate B, tens of milligrams of rec-mambalgin-1, Slurp1 and milligrams of rec-kBtx and rec- α Btx (Additional file 4: materials) through one round of experiment (usually finished within 4~5 weeks). The detailed expression level, final yield and refolding species analysis are summarized in Additional file 7: Table S3.

Most useful scouting conditions for refolding rTFPs are cysteine, salt concentration, and pH

The optimized refolding condition for each recombinant neurotoxin was summarized in the Additional file 4: Materials. The most critical factors are the concentration of sodium chloride and L-cysteine, and pH value. A weak basic solution, with different NaCl and L-cysteine concentration provided the essential refolding solution. Without Tris base, increasing L-cysteine concentration significantly alter the pH from neutral to acidic, leading

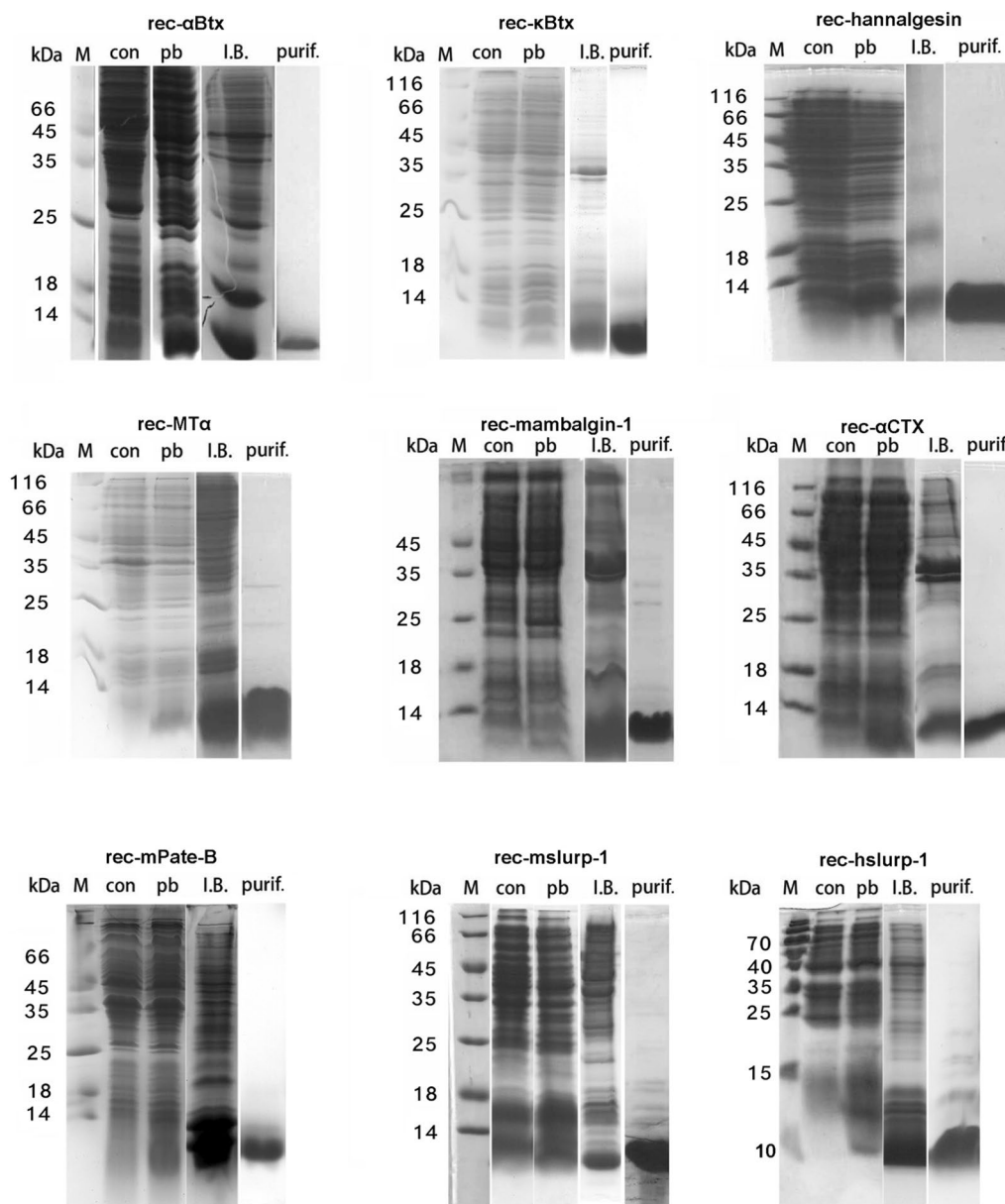


Fig. 2 SDS-PAGE analysis of rTFPs at different stages of production. *con*: control (not induced *E. coli* cells), *pb* IPTG induced *E. coli* cells, *I.B.* isolated inclusion bodies, *purif.*: purified final product (in non-reducing SDS-PAGE)

to lowered yield of monomeric species (data not shown). L-arginine [53–56] and NDSB-201 [57, 58], two known supplements which are widely used in inclusion body refolding, even though significantly increased the yield of monomeric species in the screening experiment as reflected by non-reducing SDS-PAGE (Additional file 1: Figure S1), lead to formation of a lot of precipitates in the subsequent dialysis removal of these supplements (data not shown), and thus didn't help much. What's more, L-arginine and NDSB-201 are relatively expensive and not cost-effective in large scale production. Taken

together, L-arginine and NDSB-201 are generally not recommended for refolding rTFPs, at least for those we tried. We generally only visually check the gel graph to determine the refolding condition for the scaled-up production, according to the band intensities of the monomer. For more accurate estimation, protein band corresponding to different refolding species was quantified with densitometry measurement and two metrics, 'monomer ratio' and 'relative monomer amount' were calculated, which would be helpful for selection of the optimal refolding condition. The condition selected

Table 1 Data collection and refinement statistics

| | rec-Mambalgin-1 | rec-kBtx | rec-αBtx-HAP | rec-Hannalgesin | rec-MTα | rec-αCTX | |
|----------------------|----------------------------|--------------------------|--------------------------|----------------------------|----------------------------|----------------------------|----------------------------|
| Resolution range (Å) | 24.09–2.49 (2.579–2.49) | 36.71–1.8 (1.864–1.8) | 33.35–2.4 (2.486–2.4) | 39.46–2.2 (2.279–2.2) | 29.28–1.8 (1.865–1.8) | 30.77–1.57 (1.626–1.57) | |
| Space group | P 1 21 1 | P 1 | P 1 21 1 | P 32 | P 63 | P 3 2 1 | |
| Unit cell | Length (Å) | 30.984 80.753 54.032 | 27.0899 30.9899 39.43 | 42.5295 73.1098 79.8397 | 65.5095 65.5095 164.752 | 58.555 58.555 35.395 | 73.9695 73.9695 110.841 |
| | Angle (°) | 90 90.47 90 | 70.44 75.97 71.94 | 90 103.68 90 | 90 90 120 | 90 90 120 | 90 90 120 |
| Multiplicity | | 3.9 (4.0) | 4.0 (4.0) | 7.6 (7.7) | 6.1 (6.0) | 9.3 (7.2) | 21.7 (20.6) |
| Completeness (%) | | 94.97 (93.12) | 93.40 (89.76) | 96.32 (94.14) | 99.76 (99.58) | 99.88 (99.84) | 99.83 (99.79) |
| Mean I/sigma (I) | | 20.12 (7.47) | 16.92 (8.91) | 10.58 (4.53) | 11.67 (3.38) | 53.4 (17.65) | 21.76 (3.96) |
| Wilson B-factor | | 55.19 | 13.48 | 18 | 41.94 | 21.26 | 20.76 |
| R-merge | | 0.04624 (0.1469) | 0.05006 (0.1079) | 0.1477 (0.4164) | 0.08822 (0.5762) | 0.049 (0.129) | 0.08866 (0.8743) |
| R-meas | | 0.05376 (0.1683) | 0.05787 (0.1247) | 0.1584 (0.4462) | 0.09657 (0.631) | 0.052 (0.137) | 0.09088 (0.8962) |
| CC1/2 | | 0.997 (0.978) | 0.998 (0.985) | 0.992 (0.943) | 0.996 (0.904) | | 0.999 (0.951) |
| R-work | | 0.2179 (0.3432) | 0.1658 (0.2067) | 0.2393 (0.2844) | 0.2115 (0.2708) | 0.1897 (0.3007) | 0.1736 (0.2363) |
| R-free | | 0.2670 (0.3537) | 0.1912 (0.2233) | 0.2750 (0.3626) | 0.2431 (0.3349) | 0.2202 (0.3696) | 0.1844 (0.2515) |
| macromolecules | | 1845 | 1018 | 2671 | 6094 | 532 | 2040 |
| ligands | | 15 | 0 | 96 | 25 | 28 | 39 |
| RMS (bonds) | | 0.008 | 0.007 | 0.006 | 0.007 | 0.009 | 0.012 |
| RMS (angles) | | 1.18 | 1.01 | 1.09 | 0.98 | 1 | 1.25 |
| Average B-factor | | 72.89 | 20.55 | 23.99 | 63.69 | 28.06 | 26.68 |
| macromolecules | | 73.10 | 19.43 | 23.52 | 63.95 | 26.95 | 24.44 |
| ligands | | 68.76 | | 63.94 | 72.89 | 47.62 | 51.78 |

(Values in parentheses are for the highest-resolution shell)

by eyeballing the gel gave good if not best ‘monomer ratio’, ‘relative monomer amount’ values (Additional file 1: Figure S1).

Normally, rTFPs with high isoelectric point (pI) remained soluble upon challenge with weak acidic solution (such as 20 mM NaAc, pH 5.0), while certain mammalian three finger toxin-like protein, such as Slurp1, remained soluble only in neutral and slight basic solutions, such as 20 mM HEPES (pH 7.0) and 20 mM Tris–HCl (pH 8.0).

Complete oxidation is the key to the production of high quality rTFPs

It is common to see I.B. refolding protocols in which people dissolve the I.B. with solutions containing high concentration of reducing agents (such as 100 mM β-mercaptoethanol or 2-ME). While these agents are useful in keeping the free cysteine residue in reduced form and it might not be a problem in certain cases, we found 100 mM 2-ME in I.B. solubilization buffers inevitably lead to failed refolding experiments, which was shown by the extremely low yield and formation of multimeric species [38], thus should be avoided when solubilizing the I.B. For correct disulfide bonds pairing between the cysteine residues, a classical and widely used

approach is the disulfide shuffling or mixed disulfide bond reactions, in which a predefined redox pairs such as a fixed ratio of reduced-glutathione: oxidized-glutathione, or cysteine:cystine are used [59–61]. In our pipeline, we used a simple, straightforward approach by scouting cysteine and NaCl concentration in screening of refolding conditions, and we noticed that different rTFPs had different sensitivity to cysteine concentration in the refolding experiment (Additional file 1: Figure S1). We would choose a condition with the less usage of L-cysteine that produced the highest level of monomeric species, as shown by non-reducing SDS-PAGE. In preparative refolding, we used compressed N₂ gas and/or air to drive the ultrafiltration device (Fig. 1). Because of the large volume of the diluted refolding mixture, we had to frequently reopen and refill the device, which inevitably brought air in. The O₂ in the air should help the oxidation of cysteine to form cystine, and disulfide bonds in the rTFPs. In refolding of rec-kBtx, we found N₂ gas was not as good as compressed air, which dramatically decreased the multimeric species in the final product (data not shown), and only the purified rec-kBtx from this special protocol yielded crystals. Clearly, ultrafiltration with the stirred cell is not only a physical process, but also a biochemical process in which

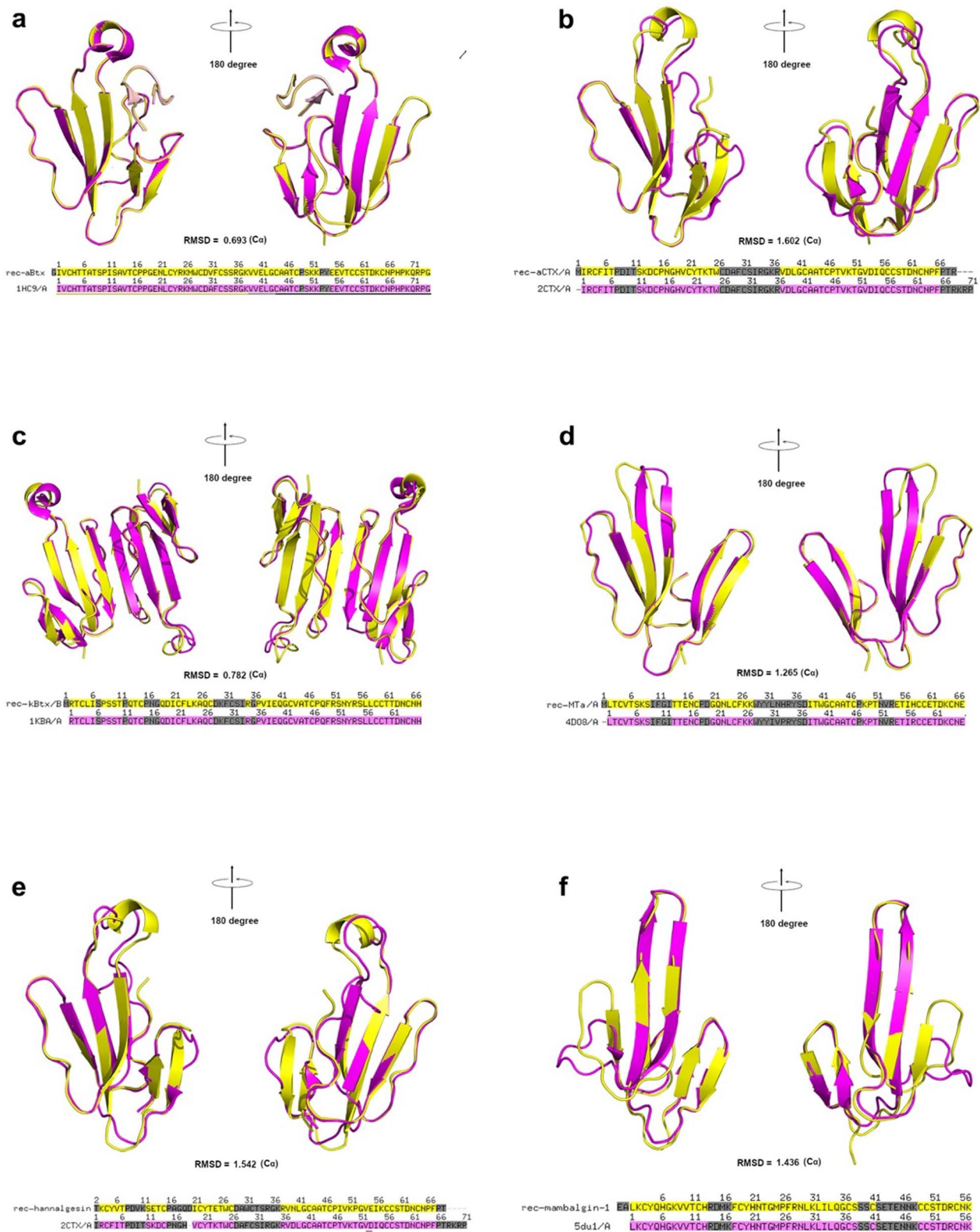


Fig. 3 Structural alignment of the crystal structure of rTFPs and their natural counterpart or most homologous natural counterpart. Yellow: rTFP, Magenta: Reported native (homologous) or synthetic counterpart. **a.** rec-aBtx-HAP vs aBtx-HAP; **b.** rec-aCTX vs aCTX; **c.** rec-kBtx vs kBtx; **d.** rec-MTa vs MT1; **e.** rec-Hannalgesin vs aCTX; **f.** rec-Mambalgin-1 vs Mambalgin-1. RMSD was calculated based on the coordinates of C_{α} of the aligned structures

the dissolved oxygen level is critical for the correct and complete formation of disulfide bonds. Considering this, we designed and crafted a 'storage and oxidation tank' (Fig. 1), which acted as an oxidation reservoir for the refolded protein mixture and gas, eliminating the needs for opening the ultrafiltration cell to refill and making it easier to adjust the ratio between air and N₂.

A typical preparative ultrafiltration procedure took about 4 weeks, during the last a few days of which a large amount of white precipitate (which turned out to be cystine, the oxidization product of cysteine) showed up in the concentrated solution, which were found to be a sign of complete oxidation, since most of our high quality

rTFPs were produced in this way. Some of our rTFPs were tested for their stability with prolonged storage at 4 °C, and were shown to be stable even after one year of storage at 4 °C, and only trace amounts of dimeric and multimeric species were found (Fig. 4a), consistent to previous report about the stability of natural TFNs [62], and proving the high quality of the rTFP from our pipeline.

Concentrate to dry is a simple and efficient step for removing incorrectly folded species

It is interesting to note this point, since we found that multimeric species, which were generally regarded

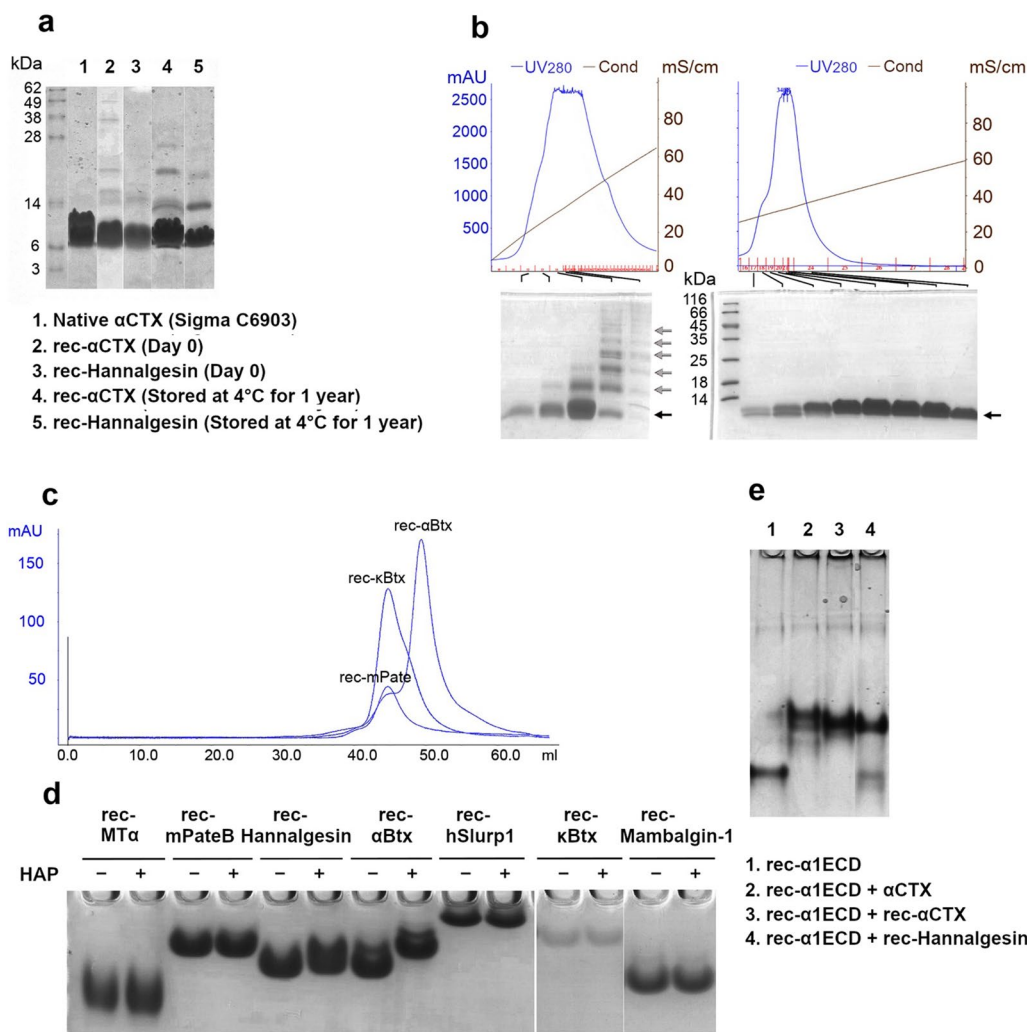


Fig. 4 Biochemical characterization of the rTFPs **a**. Stability of rec- α CTX and rec-Hannalgesin upon prolonged storage at 4 °C, analyzed with non-reducing SDS-PAGE; **b**. 'Concentrate to dry' strategy (right, with black arrow pointing to the monomeric species) efficiently removed multimeric species that were hard to be separated with the mono S column (left, in which the refolded product was not 'concentrate to dry', grey arrow); **c**. Gel filtration analysis of rec- α Btx, rec-mPate B and rec-kBtx. **d**. Native gel shift assay of various rTFPs with HAP peptide. **e**. Native gel shift assay of rec- α 1ECD with native α CTX, rec- α CTX and rec-Hannalgesin

as incorrectly folded product with wrong pairing of disulfide bonds, were inevitable as a part of refolding product and hard to be completely separated from the correctly folded species using common chromatography approaches, such as gel filtration (data not shown) and ion exchange. However, it turned out that ultrafiltration of the refolded product to dry dramatically increased the purity for some rTFPs (Fig. 4b). It is possible that incorrectly folded but soluble rTFPs tend to form insoluble aggregates only when they are concentrated to extremely high concentrations, which was achieved in the “concentrate to dry” approach. It is thus noteworthy to try two ultrafiltration strategies, “to dry, or not to dry”, which in most cases could make a great difference.

Recombinant rTFPs showed expected elution profile in gel filtration chromatography and were shown to be active in biochemical and morphological assays

We compared the elution volume of rec- α Btx, rec- κ Btx, rec-mPate B in gel filtration column (Superdex 75 10 300 GL, GE Healthcare), and found that rec- α Btx, which were known as a monomer, eluted much later than rec- κ Btx, suggesting rec- κ Btx was not a monomer, which is in accordance with earlier reports [6] that κ Btx exists in dimeric form and also with the solved crystal structures (Fig. 3c), while rec-mPate B elute at similar volume as rec- κ Btx, suggesting rec-mPate B was also a dimer (Fig. 4c). To test the binding specificities of the rTFPs, HAP peptide, a known peptide derived from the nicotinic acetylcholine receptor [42], was mixed with various rTFPs and separated on a 15% native PAGE gel at pH 5.0. HAP peptide was only able to shift rec- α Btx and only slightly shift rec-Hannalgesin, but not rec-MT α , rec-mPate B, rec- κ Btx, rec-Mambalgin-1 and rec-hSlurp1 (Fig. 4d). Also, rec- α CTX and rec-Hannalgesin was shown to bind the extracellular domain of α 1 subunit of the nicotinic acetylcholine receptor (rec- α 1ECD) [45], like the native α CTX isolated from *Naja Kaouthia* (Fig. 4e).

To test the binding activity of rec-mPate B to sperm, we labeled rec-mPate B with NHS-rhodamine and visualized the binding of rec-mPate B to spermatozoa freshly isolated from the epididymis of the mouse under the fluorescence microscope. The preliminary result suggested binding of rec-mPate B to the head and tail of mouse spermatozoa (Additional file 3: Figure S3).

Discussion

TFPs are a large collection of proteins with important functions and applications. Traditionally, such proteins were isolated from the venom of the snakes, with very few recombinantly obtained in the lab with in-depth analysis and verification. Because of their scarcity,

unique properties and applications, these proteins are very expensive (at the level of hundreds to thousands of US dollars per milligrams) and some are not commercially available. κ Btx, for example, a unique $\alpha_3\beta_2$ nicotinic acetylcholine receptor binder, is not commercially available (personal communications). Because TFPs usually contain 4 to 5 pairs of disulfide bonds, it is usually very hard to recombinantly express them, and those commercially available are mostly purified from snake venoms. Some researchers used chemical synthesis that successfully obtained these rTFPs, such as mambalgin-1 and mambalgin-2 [1, 36, 37, 63–65]. However, due to the high cost in chemical synthesis and limited yields, these successful attempts did not change the overall scenario for production of other TFPs.

With our pipeline, however, milligrams to hundreds of milligrams of rTFPs could be obtained in the lab within weeks. Through extensive biochemical assays and structural analysis, we showed that our rTFPs were almost identical to their native counterparts. Since several of our rTFPs reached milligrams to hundreds of milligrams on a single lab-scale production cycle, these rTFP could thus replace their natural counterparts, and the pipeline is worth to be exploited for production of other TFPs further, which is of general interest in the field.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12934-024-02316-1>.

Additional file 1: Figure S1. Refolding condition screening of rTNFs.

Additional file 2: Figure S2. Microscopic view of protein crystals from various rTFPs.

Additional file 3: Figure S3. Fluorescence microscopic picture showing the binding of rec-mPate B to the spermatozoa from mouse epididymis.

Additional file 4: Coding DNA sequences, a.a. sequences, protein properties, refolding conditions and key points and crystallization conditions for rTFPs.

Additional file 5: Table S1. Solutions and media table.

Additional file 6: Table S2. Properties and references of various TFPs we successfully refolded and purified.

Additional file 7: Table S3. Expression level, monomer fraction and relative amount from dilution refolding and final yield table for rTFPs.

Acknowledgements

This research used resources of the Advance Light Source, which is a DOE Office of Science User Facility under contract no. DE-AC02-05CH11231. This research used resources of the Advanced Photon Source, which is a DOE office of Science User Facility under contract no. DE-AC02-06CH11357. We thank Advanced Light Source beamline 8.2.1 and Advanced Photon Source beamline 23ID-D and 19ID-D staff members and scientists for help with data collection. We thank Dr. Aaron Wolfe for help with data collections.

Author contributions

JX conceived the idea and designed the experiments. JX and JL constructed the expression plasmids, JX did the *E. coli* fermentation, recombinant protein expression, refolding, and purification. JL did the animal experiment. JX, AL did the crystallization screen. JX harvested the crystals and collected the diffraction data. JX, XL, SL and LC solved the structure and did the structural analysis.

Funding

The research is supported in part by National Institutes of Health grants R01GM064642.

Availability of data and materials

Coordinates and structure factors have been deposited to the Protein Data Bank with accession number of 7ULB (rec-Mambalgin-1), 7ULR (rec-kBtx), 8VY8 (rec- α Btx-HAP complex), 7ULQ (rec-Hannalgesin), 7ULS (rec-MTa) and 7ULG (rec- α CTX).

Declarations

Ethics approval and consent to participate

Animal experiment was approved by the Animal Care and Use Committee of University of South China.

Consent for publication

All authors are consent for publication.

Competing interests

The Authors declare that there is no competing interests.

Received: 25 April 2022 Accepted: 10 August 2022

Published online: 12 February 2024

References

1. Diochot S, Baron A, Salinas M, Douguet D, Scarzello S, Dabert-Gay A-S, et al. Black mamba venom peptides target acid-sensing ion channels to abolish pain. *Nature*. 2012;490:552–5.
2. Pu XC, Wong PT, Gopalakrishnakone P. A novel analgesic toxin (hannalgesin) from the venom of king cobra (*Ophiophagus hannah*). *Toxicon*. 1995;33:1425–31.
3. Servent D, Fruchart-Gaillard C. Muscarinic toxins: tools for the study of the pharmacological and functional properties of muscarinic receptors. *J Neurochem*. 2009;109:1193–202.
4. Näreoja K, Kukkonen JP, Rondinelli S, Toivola DM, Meriluoto J, Näsman J. Adrenoceptor activity of muscarinic toxins identified from mamba venoms. *Br J Pharmacol*. 2011;164:538–50.
5. Chiappinelli VA. Kappa-bungarotoxin: a probe for the neuronal nicotinic receptor in the avian ciliary ganglion. *Brain Res*. 1983;277:9–22.
6. Dewan JC, Grant GA, Sacchetti JC. Crystal structure of kappa-bungarotoxin at 2.3-Å resolution. *Biochemistry*. 1994;33:13147–54.
7. Chiappinelli VA, Hue B, Mony L, Sattelle DB. Kappa-bungarotoxin blocks nicotinic transmission at an identified invertebrate central synapse. *J Exp Biol*. 1989;141:61–71.
8. Faure G, Shelukhina IV, Porowinska D, Shulepko MA, Lyukmanova EN, Dolgikh DA, et al. Interaction of three-finger proteins from snake venoms and from mammalian brain with the cys-loop receptors and their models. *Dokl Biochem Biophys*. 2016;468:193–6.
9. George AA, Bloy A, Miwa JM, Lindstrom JM, Lukas RJ, Whiteaker P. Isoform-specific mechanisms of $\alpha 3\beta 4^*$ -nicotinic acetylcholine receptor modulation by the prototoxin lynx1. *FASEB J*. 2017;31:1398–420.
10. Lyukmanova EN, Shulepko MA, Shenkarev ZO, Bychkov ML, Paramonov AS, Chugunov AO, et al. Secreted Isoform of Human Lynx1 (SLURP-2): spatial structure and pharmacology of interactions with different types of acetylcholine receptors. *Sci Rep*. 2016;6:30698.
11. Lyukmanova E, Arvaniti M, Jensen M, Shulepko M, Dolgikh D, Pinborg L, et al. Lynx1 competes with a beta 1–42 for the binding to nicotinic acetylcholine receptors. *Hoboken: Wiley*; 2017.
12. Thomsen MS, Arvaniti M, Jensen MM, Shulepko MA, Dolgikh DA, Pinborg LH, et al. Lynx1 and $\alpha\beta 1-42$ bind competitively to multiple nicotinic acetylcholine receptor subtypes. *Neurobiol Aging*. 2016;46:13–21.
13. Fischer J, Bouadjar B, Heilig R, Huber M, Lefèvre C, Jobard F, et al. Mutations in the gene encoding SLURP-1 in Mal de Meleda. *Hum Mol Genet*. 2001;10:875–80.
14. Shah K, Nasir A, Irfanullah SS, Khan S, Ahmad W. A novel homozygous mutation disrupting the initiation codon in the SLURP1 gene underlies mal de Meleda in a consanguineous family. *Clin Exp Dermatol*. 2016;41:675–9.
15. Swamynathan S, Swamynathan SK. Secreted Ly6/Urokinase-type plasminogen activator receptor-related protein-1 (SLURP1) suppresses angiogenic inflammation through a pathway that involves Src, RhoA and NF κ B. *Invest Ophthalmol Vis Sci. Assoc Res Vision Ophthalmol*. 2016;57:1447–1447.
16. Taylor JA, Bondavalli D, Monif M, Yap LM, Winship I. Mal de Meleda in Indonesia: Mutations in the SLURP1 gene appear to be ubiquitous. *Australas J Dermatol*. 2016;57:e11–3.
17. Akbar A, Prince C, Payne C, Fasham J, Ahmad W, Baple EL, et al. Novel nonsense variants in SLURP1 and DSG1 cause *Palmoplantar keratoderma* in Pakistani families. *BMC Med Genet*. 2019;20:145.
18. Kryukova EV, Egorova NS, Kudryavtsev DS, Lebedev DS, Spirova EN, Zhmak MN, et al. From synthetic fragments of endogenous three-finger proteins to potential drugs. *Front Pharmacol*. 2019;10:748.
19. Paramonov AS, Kocharovskaya MV, Tsarev AV, Kulbatskii DS, Loktyushov EV, Shulepko MA, et al. Structural diversity and dynamics of human three-finger proteins acting on nicotinic acetylcholine receptors. *Int J Mol Sci*. 2020. <https://doi.org/10.3390/ijms21197280>.
20. Levitin F, Weiss M, Hahn Y, Stern O, Papke RL, Matusik R, et al. PATE gene clusters code for multiple, secreted TFP/Ly-6/uPAR proteins that are expressed in reproductive and neuron-rich tissues and possess neuromodulatory activity. *J Biol Chem*. 2008;283:16928–39.
21. Margalit M, Yogev L, Yavetz H, Lehavi O, Hauser R, Botchan A, et al. Involvement of the prostate and testis expression (PATE)-like proteins in sperm-oocyte interaction. *Hum Reprod*. 2012;27:1238–48.
22. Rajesh A, Yenugu S. shRNA mediated ablation of prostate and testis expressed (Pate) messenger RNA results in impaired sperm function and fertility. *Andrology*. 2017;5:541–7.
23. Turunen HT, Sipilä P, Pujianto DA, Damdimopoulos AE, Björkgren I, Huhtaniemi I, et al. Members of the murine Pate family are predominantly expressed in the epididymis in a segment-specific fashion and regulated by androgens and other testicular factors. *Reprod Biol Endocrinol*. 2011;9:128.
24. Kieffer B, Driscoll PC, Campbell ID, Willis AC, van der Merwe PA, Davis SJ. Three-dimensional solution structure of the extracellular region of the complement regulatory protein CD59, a new cell-surface protein domain related to snake venom neurotoxins. *Biochemistry*. 1994;33:4471–82.
25. Huang Y, Qiao F, Abagyan R, Hazard S, Tomlinson S. Defining the CD59-C9 binding interaction. *J Biol Chem*. 2006;281:27398–404.
26. Nevo Y, Ben-Zeev B, Tabib A, Straussberg R, Anikster Y, Shorer Z, et al. CD59 deficiency is associated with chronic hemolysis and childhood relapsing immune-mediated polyneuropathy. *Blood*. 2013;121:129–35.
27. Michielsen LA, Budding K, Drop D, van de Graaf EA, Kardol-Hoefnagel T, Verhaar MC, et al. Reduced expression of membrane complement regulatory protein CD59 on leukocytes following lung transplantation. *Front Immunol*. 2017;8:2008.
28. Antil-Delbeke S, Gaillard C, Tamiya T, Corringier PJ, Changeux JP, Servent D, et al. Molecular determinants by which a long chain toxin from snake venom interacts with the neuronal $\alpha 7$ -nicotinic acetylcholine receptor. *J Biol Chem*. 2000;275:29594–601.
29. Fiordalisi JJ, Fetter CH, TenHarmsel A, Gigowski R, Chiappinelli VA, Grant GA. Synthesis and expression in *Escherichia coli* of a gene for kappa-bungarotoxin. *Biochemistry*. 1991;30:10337–43.
30. Rosenthal JA, Hsu SH, Schneider D, Gentile LN, Messier NJ, Vaslet CA, et al. Functional expression and site-directed mutagenesis of a synthetic gene for alpha-bungarotoxin. *J Biol Chem*. 1994;269:11178–85.
31. Levandoski MM, Caffery PM, Rogowski RS, Lin Y, Shi QL, Hawrot E. Recombinant expression of alpha-bungarotoxin in *Pichia pastoris* facilitates identification of mutant toxins engineered to recognize neuronal nicotinic acetylcholine receptors. *J Neurochem*. 2000;74:1279–89.

32. Krajewski JL, Dickerson IM, Potter LT. Site-directed mutagenesis of m1-toxin1: two amino acids responsible for stable toxin binding to M(1) muscarinic receptors. *Mol Pharmacol*. 2001;60:725–31.
33. Fiordalisi JJ, James PL, Zhang Y, Grant GA. Facile production of native-like kappa-bungarotoxin in yeast: an enhanced system for the production of a neuronal nicotinic acetylcholine receptor probe. *Toxicon*. 1996;34:213–24.
34. Trémeau O, Lemaire C, Drevet P, Pinkasfeld S, Ducancel F, Boulain JC, et al. Genetic engineering of snake toxins: the functional site of erabutoxin a as delineated by site-directed mutagenesis includes variant residues. *J Biol Chem*. 1995;270:9362–9.
35. Fruchart-Gaillard C, Mourier G, Blanchet G, Vera L, Gilles N, Ménez R, et al. Engineering of three-finger fold toxins creates ligands with original pharmacological profiles for muscarinic and adrenergic receptors. *PLoS ONE*. 2012;7: e39166.
36. Schroeder CI, Rash LD, Vila-Farrés X, Rosengren KJ, Mobli M, King GF, et al. Chemical synthesis, 3D structure, and ASIC binding site of the toxin mambalgin-2. *Angew Chem Int Ed Engl*. 2014;53:1017–20.
37. Pan M, He Y, Wen M, Wu F, Sun D, Li S, et al. One-pot hydrazide-based native chemical ligation for efficient chemical synthesis and structure determination of toxin Mambalgin-1. *Chem Commun*. 2014;50:5837–9.
38. Xu J, Li J, Wu X, Song C, Lin Y, Shen Y, et al. Expression and refolding of bioactive α -bungarotoxin V31 in *E. coli*. *Protein Expr Purif*. 2015;110:30–6.
39. Grote A, Hiller K, Scheer M, Münch R, Nörtemann B, Hempel DC, et al. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucl Acids Res Oxford Acad*. 2005;33:W526–31.
40. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*. 2012;9:671–5.
41. Fiala GJ, Schamel WWA, Blumenthal B. Blue native polyacrylamide gel electrophoresis (BN-PAGE) for analysis of multiprotein complexes from cellular lysates. *J Vis Exp*. 2011. <https://doi.org/10.3791/2164>.
42. Harel M, Kasher R, Nicolas A, Guss JM, Balass M, Fridkin M, et al. The binding site of acetylcholine receptor as visualized in the X-Ray structure of a complex between alpha-bungarotoxin and a mimotope peptide. *Neuron*. 2001;32:265–75.
43. Kudryavtsev DS, Tabakmakher VM, Budylin GS, Egorova NS, Efremov RG, Ivanov IA, et al. Complex approach for analysis of snake venom α -neurotoxins binding to HAP, the high-affinity peptide. *Sci Rep*. 2020;10:3861.
44. Yao Y, Wang J, Viroonchatapan N, Samson A, Chill J, Rothe E, et al. Yeast expression and NMR analysis of the extracellular domain of muscle nicotinic acetylcholine receptor alpha subunit. *J Biol Chem*. 2002;277:12613–21.
45. Dellisanti CD, Yao Y, Stroud JC, Wang Z-Z, Chen L. Crystal structure of the extracellular domain of nAChR alpha1 bound to alpha-bungarotoxin at 1.94 Å resolution. *Nat Neurosci*. 2007;10:953–62.
46. Wittig I, Schagger H. Advantages and limitations of clear-native PAGE. *Proteomics*. 2005;5:4338–46.
47. Duselis AR, Vrana PB. Harvesting sperm and artificial insemination of mice. *J Vis Exp*. 2007. <https://doi.org/10.3791/184-v>.
48. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol*. 1997;276:307–26.
49. Batty TGG, Kontogiannis L, Johnson O, Powell HR, Leslie AGW. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr*. 2011;67:271–81.
50. Evans PR, Murshudov GN. How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr*. 2013;69:1204–14.
51. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr*. 2011;67:235–42.
52. Liebschner D, Afonine PV, Baker ML, Bunkóczi G, Chen VB, Croll TI, et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol*. 2019;75:861–77.
53. Tsumoto K, Umetsu M, Kumagai I, Ejima D, Philo JS, Arakawa T. Role of arginine in protein refolding, solubilization, and purification. *Biotechnol Prog*. 2004;20:1301–8.
54. Arakawa T, Ejima D, Tsumoto K, Obeyama N, Tanaka Y, Kita Y, et al. Suppression of protein interactions by arginine: a proposed mechanism of the arginine effects. *Biophys Chem*. 2007;127:1–8.
55. Chen J, Liu Y, Wang Y, Ding H, Su Z. Different effects of L-arginine on protein refolding: suppressing aggregates of hydrophobic interaction, not covalent binding. *Biotechnol Prog*. 2008;24:1365–72.
56. Tischer A, Lilie H, Rudolph R, Lange C. L-arginine hydrochloride increases the solubility of folded and unfolded recombinant plasminogen activator rPA. *Protein Sci*. 2010;19:1783–95.
57. Wangkanont K, Forest KT, Kiessling LL. The non-detergent sulfobetaine-201 acts as a pharmacological chaperone to promote folding and crystallization of the type II TGF- β receptor extracellular domain. *Protein Expr Purif*. 2015;115:19–25.
58. Luca VC, AbiMansour J, Nelson CA, Fremont DH. Crystal structure of the Japanese encephalitis virus envelope protein. *J Virol*. 2012;86:2337–46.
59. Creighton TE. Disulfide bond formation in proteins: methods in enzymology. Cambridge: Academic Press; 1984.
60. Okumura M, Saiki M, Yamaguchi H, Hidaka Y. Acceleration of disulfide-coupled protein folding using glutathione derivatives. *FEBS J*. 2011;278:1137–44.
61. Qin M, Wang W, Thirumalai D. Protein folding guides disulfide bond formation. *Proc Natl Acad Sci USA*. 2015;112:11241–6.
62. Nirthanan S, Awal W, Niranjana NR. Snake α -neurotoxins and the nicotinic acetylcholine receptor. In: Gopalakrishnakone P, Inagaki H, Mukherjee AK, Rahmy TR, Vogel C-W, editors. *Snake Venoms*. Dordrecht: Springer; 2015. p. 1–39.
63. Mourier G, Salinas M, Kessler P, Stura EA, Leblanc M, Tepshi L, et al. Mambalgin-1 pain-relieving peptide, stepwise solid-phase synthesis, crystal structure, and functional domain for acid-sensing ion channel 1a inhibition. *J Biol Chem*. 2016;291:2616–29.
64. Sun D, Yu Y, Xue X, Pan M, Wen M, Li S, et al. Cryo-EM structure of the ASIC1a-mambalgin-1 complex reveals that the peptide toxin mambalgin-1 inhibits acid-sensing ion channels through an unusual allosteric effect. *Cell Discov*. 2018;4:27.
65. Salinas M, Kessler P, Douguet D, Sarraf D, Tonali N, Thai R, et al. Mambalgin-1 pain-relieving peptide locks the hinge between α 4 and α 5 helices to inhibit rat acid-sensing ion channel 1a. *Neuropharmacology*. 2021;185: 108453.
66. Chang CC, Lee CY. Isolation of neurotoxins from the venom of *Bungarus multicinctus* and their modes of neuromuscular blocking action. *Arch Int Pharmacodyn Ther*. 1963;144:241–57.
67. Love RA, Stroud RM. The crystal structure of alpha-bungarotoxin at 2.5 Å resolution: relation to solution structure and binding to acetylcholine receptor. *Protein Eng*. 1986;1:37–46.
68. Loring RH, Zigmund RE. Characterization of neuronal nicotinic receptors by snake venom neurotoxins. *Trends Neurosci*. 1988;11:73–8.
69. Luetjge CW, Wada K, Rogers S, Abramson SN, Tsuji K, Heinemann S, et al. Neurotoxins distinguish between different neuronal nicotinic acetylcholine receptor subunit combinations. *J Neurochem*. 1990;55:632–40.
70. Fiordalisi JJ, Al Rabiee R, Chiappinelli VA, Grant GA. Affinity of native kappa-bungarotoxin and site-directed mutants for the muscle nicotinic acetylcholine receptor. *Biochemistry*. 1994;33(12962):7.
71. Cartier GE, Yoshikami D, Gray WR, Luo S, Olivera BM, McIntosh JM. A new alpha-conotoxin which targets alpha3beta2 nicotinic acetylcholine receptors. *J Biol Chem*. 1996;271:7522–8.
72. Pu XC, Wong PTH, Gopalakrishnakone P. Hannalgesin from king cobra venom releases nitric oxide (NO) and activates nitric oxide synthase (NOS). *Toxicon*. 1996;6:622.
73. Salinas M, Besson T, Delettre Q, Diochot S, Boulakirba S, Douguet D, et al. Binding site and inhibitory mechanism of the mambalgin-2 pain-relieving peptide on acid-sensing ion channel 1a. *J Biol Chem*. 2014;289:13363–73.
74. Koivula K, Rondinelli S, Näsman J. The three-finger toxin M α is a selective α 2B-adrenoceptor antagonist. *Toxicon*. 2010;56:440–7.
75. Karlsson E, Eaker D. Isolation of the principal neurotoxins of Naja naja subspecies from the Asian mainland. *Toxicon*. 1972;10:217–25.
76. Betzel C, Lange G, Pal GP, Wilson KS, Maelicke A, Saenger W. The refined crystal structure of alpha-cobratoxin from Naja naja siamensis at 2.4-Å resolution. *J Biol Chem*. 1991;266:21530–6.
77. Luo CW, Lin HJ, Chen YH. A novel heat-labile phospholipid-binding protein, SVS VII, in mouse seminal vesicle as a sperm motility enhancer. *J Biol Chem*. 2001;276:6913–21.
78. Grønlien JH, Håkerud M, Ween H, Thorin-Hagene K, Briggs CA, Gopalakrishnan M, et al. Distinct profiles of alpha7 nAChR positive allosteric modulation revealed by structurally diverse chemotypes. *Mol Pharmacol*. 2007;72:715–24.

79. Throm VM, Männle D, Giese T, Bauer AS, Gaida MM, Kopitz J, et al. Endogenous CHRNA7-ligand SLURP1 as a potential tumor suppressor and anti-nicotinic factor in pancreatic cancer. *Oncotarget*. 2018;9:11734–51.
80. Swamynathan S, Buela K-A, Kinchington P, Lathrop KL, Misawa H, Hendricks RL, et al. Klf4 regulates the expression of Slurp1, which functions as an immunomodulatory peptide in the mouse cornea. *Invest Ophthalmol Vis Sci*. 2012;53:8433–46.
81. Upadhyay G. Emerging role of lymphocyte antigen-6 family of genes in cancer and immune cells. *Front Immunol*. 2019;10:819.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.