



HAL
open science

Ancestral protein resurrection and engineering opportunities of the mamba aminergic toxins

Guillaume Blanchet, Doria Alili, Adèle Protte, Gregory Upert, Nicolas Gilles, Livia Tepshi, Enrico A. Stura, Gilles Mourier, Denis Servent

► To cite this version:

Guillaume Blanchet, Doria Alili, Adèle Protte, Gregory Upert, Nicolas Gilles, et al.. Ancestral protein resurrection and engineering opportunities of the mamba aminergic toxins. *Scientific Reports*, 2017, 7, pp.2701. 10.1038/s41598-017-02953-0 . hal-01548128

HAL Id: hal-01548128

<https://hal.sorbonne-universite.fr/hal-01548128>

Submitted on 27 Jun 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

SCIENTIFIC REPORTS



OPEN

Ancestral protein resurrection and engineering opportunities of the mamba aminergic toxins

Guillaume Blanchet^{1,2}, Doria Alili¹, Adèle Protte¹, Gregory Upert¹, Nicolas Gilles¹, Livia Tepshi¹, Enrico A. Stura¹, Gilles Mourier¹ & Denis Servent¹

Mamba venoms contain a multiplicity of three-finger fold aminergic toxins known to interact with various α -adrenergic, muscarinic and dopaminergic receptors with different pharmacological profiles. In order to generate novel functions on this structural scaffold and to avoid the daunting task of producing and screening an overwhelming number of variants generated by a classical protein engineering strategy, we accepted the challenge of resurrecting ancestral proteins, likely to have possessed functional properties. This innovative approach that exploits molecular evolution models to efficiently guide protein engineering, has allowed us to generate a small library of six ancestral toxin (AncTx) variants and associate their pharmacological profiles to key functional substitutions. Among these variants, we identified AncTx1 as the most α_{1A} -adrenoceptor selective peptide known to date and AncTx5 as the most potent inhibitor of the three α_2 adrenoceptor subtypes. Three positions in the ρ -Da1a evolutionary pathway, positions 28, 38 and 43 have been identified as key modulators of the affinities for the α_1 and α_{2C} adrenoceptor subtypes. Here, we present a first attempt at rational engineering of the aminergic toxins, revealing an epistasis phenomenon.

Evolutionary processes of venomous animals have selected enzymes and disulfide-rich peptides in their venoms to improve their ability to subdue their prey and defend against predators. These compounds act predominantly on few, well characterized and physiologically-relevant molecular targets of the envenomed animals¹. Recruited by convergent evolution, toxins impact mainly the haemostatic, nervous and cardiovascular systems². Some toxins present in venoms, in spite of their debilitating effects, have become live-saving drugs³⁻⁵. To exert their biological activities, toxins interact mainly with ion channels, coagulation factors, nicotinic receptors, cell membrane or enzymes but only rarely with the most important class of physiological targets, i.e. the G protein-coupled receptors (GPCRs)⁶. Compared to numerous toxins that act on voltage-gated or ligand-gated ion channels, only a few toxins that interact with GPCRs have been identified. These toxins have been isolated mainly from cone snails venoms such as conopressin interacting with vasopressin receptor⁷, contulakin-G which binds to neurotensin receptor⁸, ρ -conotoxin TIA specific of the α_{1A} -adrenoceptor⁹ and the τ -conotoxin CnVA that interacts with the somatostatin sst3 receptor¹⁰. The black widow spider and the gila monster provide further examples of GPCRs interacting toxins, namely, α -latrotoxin interacting with the latrophilin receptor¹¹ and exenatide, exploited as an anti-diabetic drug¹², that targets the GLP-1 receptor, respectively. In addition, mamba venoms contain aminergic toxins that recognize various bioaminergic receptors¹³⁻¹⁹.

Aminergic toxins belong to the three-finger fold toxin (3FT) superfamily, a structural fold known to support a large diversity of biological functions²⁰. Recently, protein engineering using a cDNA display approach was applied to 3FT to create large libraries of randomized sequences. After repeated rounds of expression and screening, interleukin-6 receptor ligands²¹ and serine protease inhibitors²² were selected, highlighting the high functional versatility of this structural template. On the same scaffold, a more rational protein engineering was developed using a loop grafting strategy leading to chimeric toxins with new functional profiles on muscarinic and adrenergic receptors²³. Despite sequence identities that vary from 70 to 98% (Fig. 1A), aminergic toxins display highly variable pharmacological profiles (Fig. 1B). On the one hand, exceptional selectivity is exemplified by MT7 or MT α which interact exclusively with one receptor subtype (Fig. 1B), while on the other hand, MT3

¹CEA Institut des Sciences du Vivant Frédéric Joliot, Service d'Ingénierie Moléculaire des Protéines (SIMOPRO), Université Paris-Saclay, Gif-sur-Yvette, Paris, 91190, France. ²UFR Sciences de la Vie, Université Pierre et Marie Curie (UPMC), Paris, 75005, France. Correspondence and requests for materials should be addressed to D.S. (email: denis.servent@cea.fr)

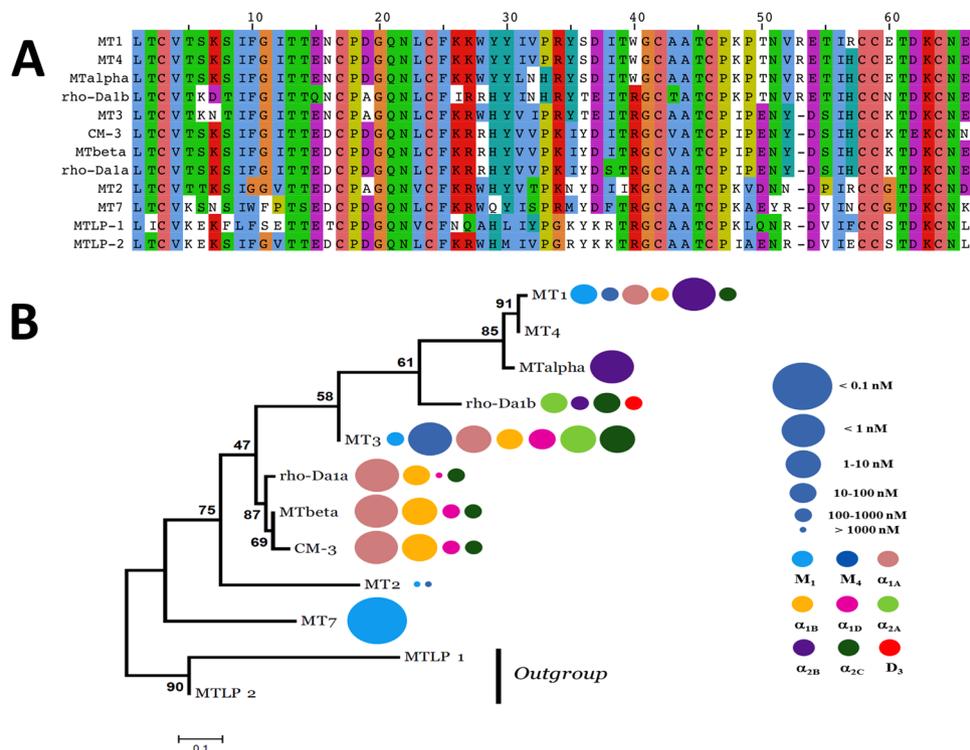


Figure 1. Sequence alignment and phylogenetic tree of aminergic toxins. **(A)** Sequence alignment of the aminergic toxin family. Color display is settled in order to emphasize amino acid conservation along all toxins. **(B)** Maximum likelihood tree and associated function of the aminergic toxins. The best evolutionary model established by MEGA5 was the Dayhoff model using a discrete Gamma distribution (+G). Bootstraps value higher than 50% are shown. MTLP-1 and 2 were used as out-group. Toxin's functions are also illustrated by circles identifying the receptor targeted (color) and the corresponding K₁ value (circle size), as described in the right legend, at the right side of each toxin name.

interacts efficiently and non-selectively with seven different aminergic receptors (Fig. 1B). These observations suggest that during the evolution of these toxins various functional properties diverged, were modified and, the least useful ones for the snake, dispensed with. It is possible that toxins with interesting activities on the bioaminergic receptor family might have existed in the past. Given the difficulty of selecting potentially interesting toxins among the vast number of variants that a traditional rational engineering approach might suggest, we decided to test an alternative semi-rational method that involves ancestral protein resurrection. By rewinding the important changes that might have occurred during evolution of a protein family, this approach might be useful to correlate sequence variations with toxin activities within this family.

In the recent years, a growing number of works relating to the use of ancestral protein resurrection have been published to address different issues like understanding protein function evolution²⁴, fold stability over time²⁵, protein solubility²⁶ or to perform protein engineering. By constructing and testing function of these ancestral proteins and comparing the results with extant proteins, this method is able to follow amino acids changes in the evolutionary history and measure their impact on their properties. This “evolutionary biochemistry”²⁷ open also a new avenue for protein engineering²⁸ by creating small but functionally rich libraries²⁹. As the first proof of concept for enzyme engineering, Gaucher and colleagues use the REAP (Reconstructing Evolutionary Adaptive Paths) to engineer new Taq polymerases with original properties^{30,31}.

Here, we use the ancestral protein resurrection methodology on the aminergic toxin group in order to pinpoint the functional substitutions which may have occurred during their evolution and use them for molecular engineering. It is the first application of this methodology on a mini-protein/peptide family. We compare the pharmacological profiles of the natural and six ancestral toxins found at different nodes of the aminergic toxin tree and describe the functional modulations between them that could be associated with key substitutions. A first attempt to engineer more potent aminergic 3FT validates these findings by obtaining the most selective α_{1A} and the most potent α₂ adrenergic toxins.

Results

Aminergic toxins pharmacology, phylogeny, ancestral states reconstruction and structure. Previously published phylogenetic analysis showed that the aminergic toxins are a functional family specifically developed in the mamba venoms and that a high sequence identity is shared among them (at least 70%, Fig. 1A)¹⁹. Based on these results, the maximum likelihood (ML) tree of the mamba aminergic toxins was reconstructed (Fig. 1B). The two MTLP 1 and 2 from *Naja kaouthia* were considered as an out group¹⁹. The ML methodology allows also the reconstruction of the sequences at each nodes of the tree, also called ancestral

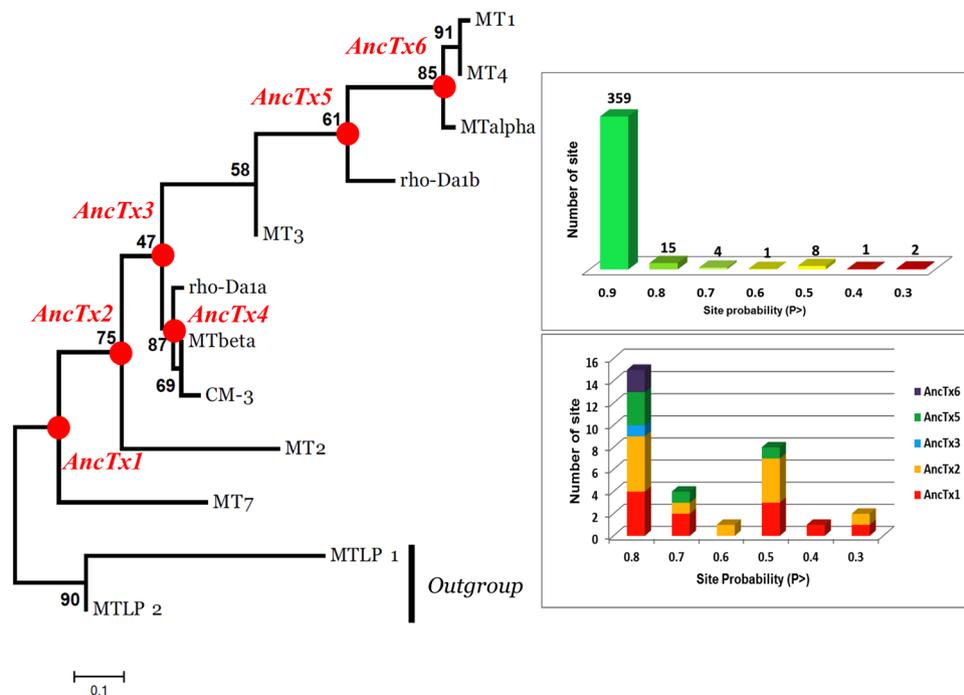


Figure 2. Maximum likelihood tree of the aminergic toxins. The best evolutionary model established by MEGA5 was the Dayhoff model using a discrete Gamma distribution (+G). Bootstraps value higher than 50% are shown. MTLP-1 and 2 were used as out-group. Six ancestral sequences are figured on the nodes of the aminergic toxins tree. Quality of the reconstruction is shown on right side upper panel. For each site, a probability P is given. In total 390 sites were reconstructed. Among them, only 31 showed a $P < 0.9$, highlighting a good quality of the reconstruction. As shown in the lower panel, a high proportion of these ambiguous sites are presented in the sequences of AncTx1 and 2. The ambiguities were resolved by chosen the residues with the best probability. Due to the high divergence of MT7 and MT2 in term of sequence and function, this choice leads AncTx1 and 2 to be closer of AncTx3 and ρ -Da1a than MT7/MT2.

sequences (Fig. 2). From this reconstruction, six ancestral toxin sequences were derived, and named from AncTx1 to AncTx6. For example, AncTx1 represents the ancestor of MT7 and AncTx2. The six sequences reconstructed are made by 390 sites in total (65×6 , Fig. 2 - upper and lower panels on the right). On these 390 sites reconstructed, 359 (92%) showed a probability $P > 0.9$, highlighting a good degree of confidence for the AncTx sequence reconstructions. Among the 31 sites with a $P < 0.9$, also called ambiguous sites, 15 had a high P value ($P > 0.8$). It is important to mention that 74% (23/31) of these ambiguous sites were from AncTx1 and 2. These ambiguities were resolved by choosing the most probable residues. Due to the high divergence of MT7 and MT2 in term of sequence and function (Fig. 1), this choice has lead AncTx1 and 2 to be closer to AncTx3 and ρ -Da1a than to MT7/MT2. More precisely, AncTx1 displayed only 7 substitutions spread over the 3 loops compared to AncTx2 (Fig. 3) but as many as 13 modifications compared with MT7. AncTx2 differs by only one substitution (G58K) compared to AncTx3 and by 16 from MT2, most of them located on finger III. A single substitution caused each transition: from AncTx3 to AncTx4 (W28R), from AncTx4 to ρ -Da1a (I38S) and from AncTx4 to MT3 (A43V) (Fig. 3).

On the other hand, 10 substitutions were needed in finger I and II to go from AncTx3 to MT3. AncTx5, at the node of ρ -Da1b, contains 10 substitutions spread over fingers II and III compared to MT3 (including the insertion of valine 52), but only 5 and 7 substitutions spread over fingers I and II compared to ρ -Da1b and AncTx6, respectively. Finally, AncTx6 was close to MT α and MT1. Only one (I31L) and three substitutions (N32V, H33P, H57R) differentiated AncTx6 from MT α , and MT1, respectively (Fig. 3).

The six toxins were synthesized by solid-phase chemical synthesis, refolded and purified by reverse phase HPLC (see materials and methods). The CD spectra of AncTxs showed pronounced maxima and minima at 195 and 215 nm, respectively, typical of β -sheet structure. The CD spectrum of AncTx6 differed slightly from the other AncTxs with additional maxima and minima at 220 nm and 230 nm, that might reflect some local structural variations, as previously observed for MT1³² (Supplementary Fig. S1).

In order to visually understand the evolution of the AncTxs, the crystallographic structure of AncTx1 (variant W28R-I38S) was solved (Fig. 4, Table S1). As it might have been expected, the structure was rigid, except at the tips of the three loops (Fig. 4A). Superimposition on other known mamba toxins, MT1 (PDB id: 4DO8²³), MT2 (PDB id: 1FF4), MT7 (PDB id: 2VLW³³) and ρ -Da1a (PDB id: 4IYE³⁴) showed strong conservation of the backbone (Fig. 4B). The RMSD from C- α for core residues (white background) shows that the backbones for all the toxins superimpose well, but the agreement deteriorates if loop residues are also added to the calculation (light blue background) (Table in Fig. 4). AncTx1W28R-I38S represents well AncTx1 as the two mutations on loop 2 do not perturb the structure, as can be verified by comparing AncTx1W28R-I38S with MT2 (Fig. 4C). Loop 1 is

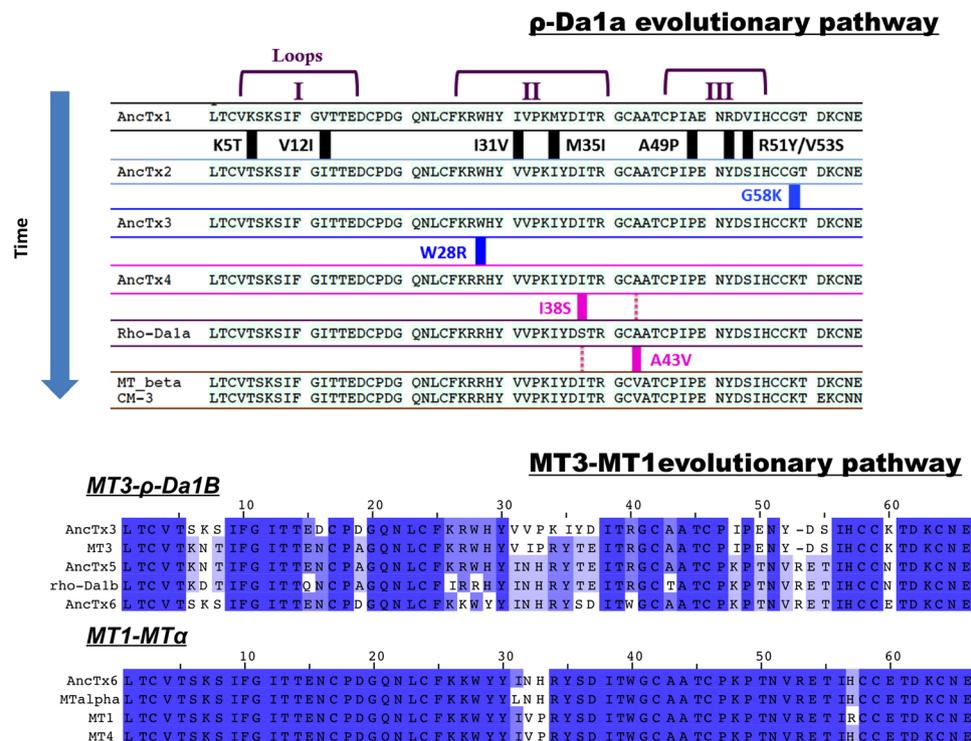


Figure 3. Amino acid substitutions in the aminergic toxins and their ancestors from the two evolutionary pathways. Sequences are following the time/tree topology of the aminergic toxins. MT7 and MT2 sequences are not shown in the ρ -Da1a pathway due to divergences in sequence and function. Amino acid substitutions are highlighted in the first pathway. MT1-MT3 evolutionary pathway is also illustrated. Color display is settled in order to emphasize amino acid conservation along the toxins. Stronger blue color indicates residue conservation among all sequences.

more susceptible to backbone variations, and both MT7 and MT2 are subjected to a large kink at the tip of this loop. However loop 1 of MT1, ρ -Da1a and AncTx1W28R-I38S remain straight (Fig. 4D).

From the first screening with 1 and 5 μ M of ancestral toxins final concentrations on various known targets of aminergic toxins, the first 4 ancestors (AncTx1 to AncTx4) display similar pharmacological properties on adrenoceptors α_{1A} , α_{1B} , α_{1D} and α_{2C} as compared to ρ -Da1a, MT β and CM-3. Despite some minor modulations in the affinity constants of ancestral and actual toxins, the targeted receptors remained the same in this phylogenetic arm. However, for toxins located in the upper part of the phylogenetic tree (MT3, AncTx5, ρ -Da1b, AncTx6, MT α , MT1), the receptor targeted and the selectivity profiles varied widely. Those toxin-receptor couples, that showed a significant radioligand displacement, were further analyzed by dose-response competition experiments in order to characterize precisely their selectivity profiles. The results are separated below along two evolutionary pathways to highlight changes in the affinity profile towards the various receptor subtypes that may have been the targets of the ancestral toxins.

ρ -Da1a evolutionary pathway. The ρ -Da1a pathway branches out from four ancestors with a primary structure similar to ρ -Da1a, MT β and CM-3. These ancestral toxins were able to interact with adrenoceptors subtypes α_{1A} , α_{1B} , α_{1D} and α_{2C} with affinity constants varying from nanomolar to micromolar range (Supplementary Fig. S2 and Fig. 5). On the α_{1A} receptor, the four reconstructed toxins showed homogeneous pK_i values, ranging from 9.62 for AncTx1 to 8.78 for AncTx3 (Table 1). Natural toxins have similar pK_i : 9.19 for ρ -Da1a and 9.43 for CM-3. Thus, the different substitutions that occurred along this evolutionary tree appear to have had a limited impact on α_{1A} subtype toxin affinity.

AncTx1 showed the lowest affinity on the α_{1B} receptor, with a pK_i value equals of 6.03. A six-fold increase in affinity characterizes the evolution from AncTx1 to AncTx2 and AncTx3 having respectively a pK_i of 6.78 and 6.89. Further gains in affinity for the α_{1B} receptor were measured for AncTx4 ($pK_i = 7.60$) and for the wild-type toxins ρ -Da1a (7.28), CM-3 (7.98) and MT β (7.99) (Table 1, Supplementary Fig. S2), suggesting that the W28R substitution between AncTx3 and AncTx4 might have been important for the affinity increase observed for this receptor subtype (Fig. 5).

A similar evolutionary pattern, as that observed on α_{1B} receptor, characterizes the pharmacological profiles of the different reconstructed toxins on the α_{2C} receptor subtype. Indeed, AncTx1 interacted with the lowest affinity on the α_{2C} receptor (pK_i of 5.53) and a seven-fold gain in affinity occurred with the evolution to AncTx2 and AncTx3 with pK_i values of 6.34 and 6.35, respectively. A minor improvement was achieved with AncTx4 that interacted with the α_{2C} subtype with a slightly higher affinity ($pK_i = 6.83$), surpassing its ancestors to reach

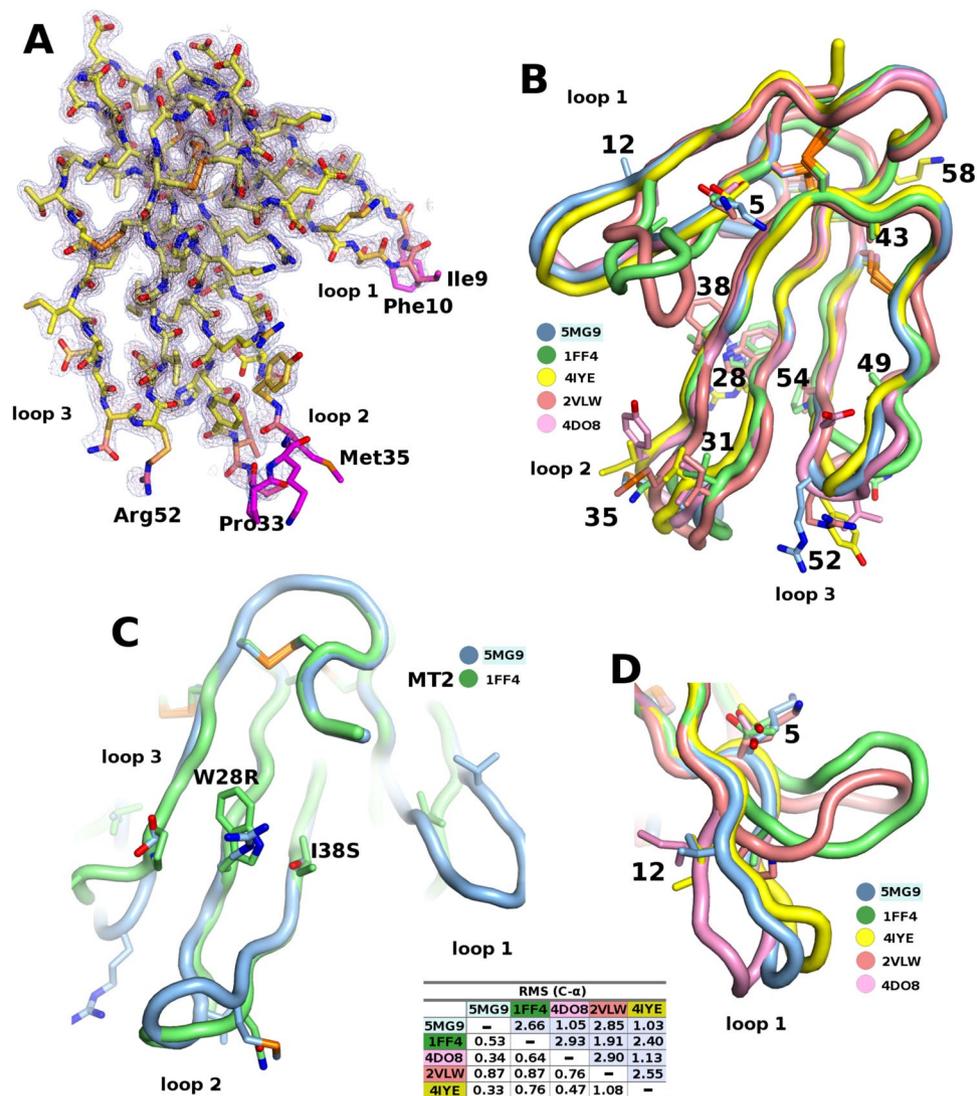


Figure 4. X-ray structure of AncTx1 (variant W28R-I38S). (A) Three-finger fold structure of AncTx1 variant. (B) superimposition of AncTx1 variant structure with other aminergic toxins from mamba, MT1 (PDB id: 4DO8), MT2 (PDB id: 1FF4), MT7 (PDB id: 2VLW) and ρ -Da1a (PDB id: 4IYE). (C) Superimposition of AncTx1 W28R-I38S and MT2 loop 2. (D) structural variations at the tip of loop 1 of aminergic toxins highlighting a large kink in MT7 and MT2. The RMSD from C- α for core residues (white background) shows that the backbones for all the toxins superimpose well, but the agreement deteriorates if loop residues are also added to the calculation (light blue background).

values comparable to actual toxins: MT β ($pK_i = 6.77$), CM-3 ($pK_i = 6.78$) and ρ -Da1a ($pK_i = 6.85$) (Table 1, Supplementary Fig. S2).

Regarding the α_{1D} subtype, the pK_i values for AncTx2 and AncTx3 (7.14 and 7.28, respectively) were higher compared to AncTx1 (6.58), AncTx4 (6.39) also compared to modern toxins with affinity constants ranging widely from 5.95 for ρ -Da1a to 6.98 for CM-3 (Table 1, Supplementary Fig. S2).

Structure-activity relationship inference: case of AncTx4. To understand the structure-activity relationship, the binding modulations that occurred between ancient and extant toxins for some receptor subtypes, needs to be analyzed. Since only a few substitutions occurred between each toxin (from AncTx2 to MT β , see Fig. 3), inferring the structure-activity relationships (SAR) becomes possible. MT β /CM-3 displayed similar pharmacological profiles on their four GPCR targets relative to ρ -Da1a, except for a slight improvement in affinity toward α_{1B} and α_{1D} subtypes by 5 and 10-fold, respectively (Table 1, Fig. 5). The double substitution S38I-A43V between ρ -Da1a and CM-3/MT β might be responsible for these differences. AncTx4 showed an intermediate sequence between CM-3/MT β and ρ -Da1a, with only one substitution (S38I or V43A) compared with these two modern toxins. AncTx4 can be positioned in term of affinity for adrenoceptors between CM-3/MT β and ρ -Da1a (Fig. 6A), suggesting that these 2 substitutions (S38I and V43A) contributed to the 5–10 fold affinity increase for the α_{1B} and α_{1D} receptors. The evolutionary pattern illustrates a slight drop in affinity between the α_{1B} and α_{1D} receptors and CM-3,

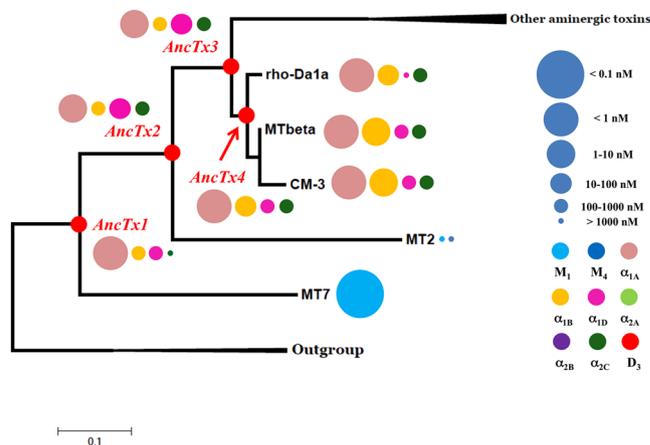


Figure 5. Maximum likelihood tree and associated function of the aminergic toxins from ρ -Da1a evolutionary pathway. Bootstraps value higher than 50% are shown. Toxin's functions are illustrated by circles identifying the receptor targeted (color) and the corresponding K_i value (circle size), as described in the right legend, at the right side of each toxin name.

Receptor subtype Toxins	α_{1A}	α_{1B}	α_{1D}	α_{2C}
AncTx1	$9,62 \pm 0,10$	$6,03 \pm 0,02$	$6,58 \pm 0,05$	$5,53 \pm 0,06$
AncTx2	$8,79 \pm 0,25$	$6,78 \pm 0,05$	$7,14 \pm 0,02$	$6,34 \pm 0,01$
AncTx3	$8,78 \pm 0,18$	$6,89 \pm 0,08$	$7,27 \pm 0,02$	$6,35 \pm 0,04$
AncTx4	$9,24 \pm 0,09$	$7,60 \pm 0,02$	$6,39 \pm 0,07$	$6,83 \pm 0,17$
ρ -Da1a ^a	9,19	7,28	5,95	6,85
MT3 ^b	9,30	7,99	6,87	6,77
CM-3 ^b	9,43	7,98	6,98	6,78
AncTx1 I38S	$9,33 \pm 0,01$	$4,75 \pm 0,22$	$6,44 \pm 0,01$	$4,90 \pm 0,19$
AncTx1 W28R-I38S	$8,59 \pm 0,09$	$6,41 \pm 0,04$	$3,71 \pm 0,11$	$4,85 \pm 0,14$

Table 1. Affinity constant ($pK_i \pm SEM$) of ρ -Da1a evolutionary pathway toxins for adrenoceptors α_{1A} , α_{1B} , α_{1D} and α_{2C} . a: [19], b: [32].

AncTx4 and ρ -Da1a, but not for the α_{1A} and α_{2C} subtypes where the pK_i remained constant (Fig. 6A). As shown on the X-ray structure of ρ -Da1a (code PDB: 4IYE, Fig. 6B), position 43 is located at the top of the toxin, at the beginning of the third loop, far away and on the opposite side to position 38 located in the middle of the loop II.

Ancestral resurrection and engineering opportunities. A SAR for the α_1 receptor subtypes was discerned from the analysis of the effect of the substitutions that differentiate AncTx3 from ρ -Da1a, which affect their respective pharmacological profiles. The substitution in AncTx4 (W28R) and those in ρ -Da1a (W28R + I38S) cause a substantial loss in affinity for α_{1D} receptor (pK_i dropping from 7.27 to 5.95) (Fig. 7). The selectivity profile of AncTx1 for the α_{1A} subtype surpasses that of all modern and ancestral aminergic toxins (Figs 5 and 7, Table 1) with a selectivity factor (the ratio between the affinity constants for the 2 most highly targeted receptors) that is 12-times better in AncTx1 compared to ρ -Da1a, currently the most selective peptide for the α_{1A} subtype³⁵. The selectivity of AncTx1 is characterized by a slightly higher affinity for α_{1A} and a much lower one for the α_{1B} and α_{2C} receptors, relative to ρ -Da1a (Fig. 7). The selectivity of AncTx1 for the α_{1A} relative to the α_{1D} subtype suffers a small setback compared to ρ -Da1a (pK_i 6.58 for AncTx1, 5.95 for ρ -Da1a). Having recognized that the substitution (I38S) in AncTx3 affects α_{1D} binding and that (W28R) increases α_{1B} and α_{2C} binding, the appropriate substitutions were integrated stepwise in AncTx1 to boost its selectivity. The selectivity profile for the variant AncTx1 I38S is characterized by pK_i values of 9.33, 4.75, 6.44 and 4.90 for α_{1A} , α_{1B} , α_{1D} and α_{2C} subtypes, respectively (Fig. 7, Table 1). These values attest to a drop in affinity of AncTx1 I38S for α_{1B} and α_{2C} compared to AncTx1, but not for the α_{1D} subtype for which the pK_i hardly changed (6.44 vs. 6.58 for AncTx1). The pK_i values for the AncTx1 double variant W28R-I38S were calculated to be 8.59, 6.41, 3.71 and 4.85 toward α_{1A} , α_{1B} , α_{1D} and α_{2C} respectively (Fig. 7, Table 1) showing that despite an impressive loss in α_{1D} affinity (3.71 vs. 6.58 for AncTx1), the selectivity of this variant suffered from a large pK_i gain toward α_{1B} (6.41 vs. 4.75 for AncTx1 I38S). While ancestral toxin resurrection gave valuable binders with novel characteristics and pointed out key substitutions that can modulate aminergic toxins binding properties, combining this knowledge together in a rational manner to refine the pharmacological characteristics, ran into problems typical of classical rational and semi-rational protein engineering projects.

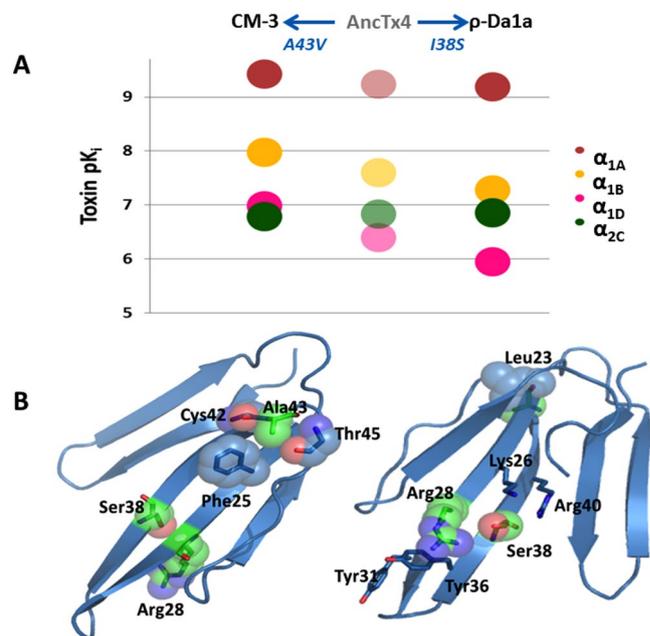


Figure 6. Affinity and structure analysis of CM-3, AncTx4 and ρ -Da1a toxins on adrenoceptors. **(A)** pK_i values are reported for CM-3, AncTx4 and ρ -Da1a on the 4 adrenoceptors α_{1A} , α_{1B} , α_{1D} and α_{2C} . **(B)** The X-ray structure of ρ -Da1a K34A (code PDB: 4IYE), the position 43 is in the top of the third loop quite far and on the opposite side compare to the substitution S38I. The lateral chain of Ala43 is surrounded by the phenyl ring of Phe25, the side chain of Leu23, and by oxygen atoms from Cys42 and Thr45 main chains. This region of the globular core is indeed compact. Hence the substitution A43V has to reorganize the packing to accommodate the bigger lateral chain of the valine.

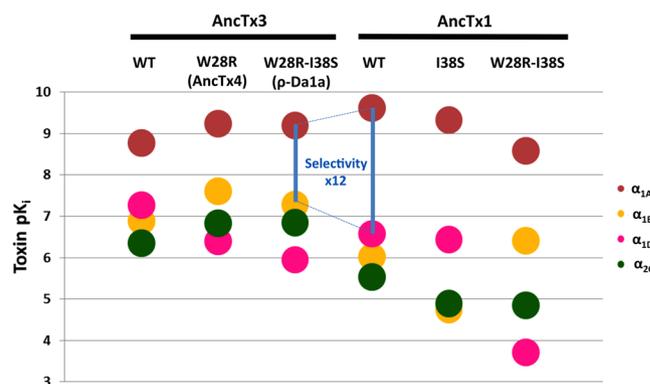


Figure 7. Affinity constants of AncTx1 and AncTx3 variants on adrenoceptors. pK_i values are reported for AncTx3, AncTx4, ρ -Da1a, AncTx1 and its 2 variants (AncTx1 I38S and W28R-I38S) on the 4 adrenoceptors α_{1A} , α_{1B} , α_{1D} and α_{2C} . The 12-times increase in selectivity for α_{1A} between ρ -Da1a and AncTx1 is shown. AncTx4 and ρ -Da1a are here represented as AncTx3 variants (W28R and W28R-I38S respectively), like I38S and W28R-I38S for AncTx1, in order to compare the effect of these substitutions in AncTx3 and AncTx1.

MT3-MT1 pathway. The second branch of the evolutionary tree comprises toxins that are more functionally divergent (compare ρ -Da1a/CM-3/MT3 to MT3 and MT1 in Fig. 1B). While MT α is exquisitely selective for the α_{2B} receptor, MT3 has the most divergent profile with an activity that spans seven of the nine receptors studied. MT3 interacts with the adrenoceptors α_{1A} , α_{1B} , α_{1D} and α_{2C} subtypes with comparable pK_i values (Fig. 8, Table 2). AncTx3 shared such an activity with notable differences with respect to α_{2C} , for which MT3 has gained an affinity that is two orders of magnitude higher (pK_i = 8.82 vs 6.35 for AncTx3). MT3 recognized three extra targets compared to AncTx3, namely the α_{2A} adrenoceptor and the muscarinic M₁ and M₄ subtypes with pK_i values ranging from 6.22 for M₁ to 9.08 for M₄ (Fig. 8). The 10 substitutions that distinguish MT3 from AncTx3, responsible for these selectivity differences, are located exclusively on fingers I and II (Fig. 3). The branching out of AncTx5 from MT3 was associated with a drastic modification of the pharmacological profile (Fig. 8). AncTx5 preserved only the interactions with the α_{2A} and α_{2C} adrenoceptors, while those with M₁, α_{1A} , α_{1B} and α_{1D} receptor subtypes were lost and the affinity for M₄ suffered a massive decrease (pK_i = 6.81 vs. 9.08 for MT3). However, AncTx5 recognizes

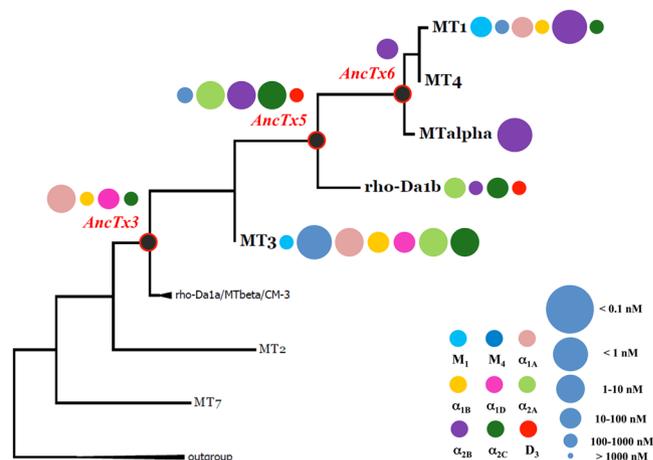


Figure 8. Maximum likelihood tree and associated function of the aminergic toxins from MT3-MT1 evolutionary pathway. Bootstraps value higher than 50% are shown. Toxin's functions are illustrated by circles identifying the receptor targeted (color) and the corresponding K_i value (circle size), as described in the right legend, at the right side of each toxin name.

Toxin Receptor subtype	MT3 ^a	AncTx5	ρ -Da1b ^a
α_{1A}	8.84	<6	<6
α_{1B}	7.68	<6	<6
α_{1D}	7.88	<6	<6
α_{2A}	8.51	8,33 \pm 0,04	7.85
α_{2B}	<6	8,11 \pm 0,07	6.66
α_{2C}	8.82	8,46 \pm 0,01	7.44
M1	6.22	<6	<6
M4	9.08	6,81 \pm 0,15	<6
D3	<6	6,81 \pm 0,09	6.12

Table 2. Affinity constant ($pK_i \pm SEM$) of MT3, ρ -Da1b and AncTx5 toxins for muscarinic (M1, M4), adrenergic (α_{1A} , α_{1B} , α_{1D} , α_{2A} , α_{2B} , α_{2C}) and dopaminergic (D3) receptors. a: [19].

the dopamine D₃ receptor, a subfamily of receptor targeted until now by only one aminergic toxin³⁶. The pK_i value of AncTx5 for this subtype was slightly higher (6.81) than for ρ -Da1b (6.12) (Table 2, Supplementary Fig. S3).

ρ -Da1b was first described for its interaction with the three α_2 adrenoceptors with pK_i values of 7.85, 6.66 and 7.44 for α_{2A} , α_{2B} and α_{2C} respectively¹⁷. Nevertheless, AncTx5 with pK_i values of 8.33, 8.11 and 8.46 on these same three subtypes was also, as for D₃, a stronger binder than ρ -Da1b (Table 2). These higher pK_i values arise from the only five substitutions that distinguish these two toxins. These include charge modifications in finger I (N7D, E15Q) and modifications at positions 26 (K to I) and 28 (W to R), already shown as key determinants for target modulation³⁶.

Due to the high number of substitutions at the other nodes (10 with an insertion at the AncTx5 node, see Figs 1A and 3), understanding the SAR was more difficult. Since finger I is identical in both MT3 and AncTx5, differences in target recognition as noted for M₄, α_{1A} , α_{1B} , α_{1D} , α_{2B} and D₃ may be derived from the changes in fingers II and III of these two proteins. As differences between AncTx3 and MT3 come from fingers I and II, we hypothesized that variations in affinity noted for M₄, α_{2A} , or α_{2C} can be mapped to changes in these fingers.

AncTx6 represents the node that leads to MT α and MT1, two toxins that recently were shown to have a completely different pharmacological profile despite a 94% sequence identity¹⁸. MT α is active on only one of the six targets of MT1, the subtype α_{2B} . Three substitutions in finger II (I31L, V32N and P33H) from MT1 to MT α are thought to be responsible for these functional differences. A single substitution marks the transition from AncTx6 to MT α . The parent shared the adrenoceptor α_{2B} as its only target (Figs 3 and 8). Considering the nature of the mutation, L31I, even a 10-fold difference in pK_i from AncTx6 to MT α (7.95 vs 9.05) was unexpected.

Discussion

In the last decades, directed evolution was largely exploited as a protein engineering strategy to mimic natural protein evolution by an iterative process combining the generation of molecular diversity by random mutagenesis and the high throughput screening selection of library members with improved functions³⁷. This approach was recently applied using cDNA display on a three-finger scaffold, in order to produce novel molecules such as interleukin-6 receptor ligands²¹ and serine protease inhibitors²². Nevertheless, in order to obtain smaller but functionally rich libraries, rational protein engineering was developed by exploiting information on protein sequence,

structure and function²⁹. We applied this strategy on aminergic toxins using a loop permutation approach and create chimeric toxins with new functional profiles on muscarinic and adrenergic receptors²³. Nevertheless, within a protein family with many functionally diverse members, a vertical strategy which reconstruct ancestral genes and study the effect of mutations on functional diversification should allow the identification of the structural determinants of protein function and the exploitation of this knowledge for protein engineering²⁴.

For the first time on animal venom toxins, we were able to apply phylogenetic analysis to resurrect ancestral proteins and obtain new and unique SAR. The aminergic toxin family is composed of closely related toxin isoforms with diverse pharmacological profiles on the biogenous amine GPCR, meaning that few substitutions can trigger important functional changes. Thus, this toxin family provided a suitable model for ancestral protein resurrection studies. Given that the phylogenetic reconstructions were sufficiently robust with very few ambiguous ancestral sites (8% under the symbolic $P < 0.9$), we retained only six ancestral toxins (AncTx) with unique and intermediate primary structures relative to extant aminergic toxins. Interestingly, most of them also showed subtle and profound modulations in their pharmacological profiles. All these changes could then be directly correlated to amino acids substitutions. We also found three key positions in the ρ -Da1a evolutionary pathway (positions 28, 38 and 43) that can modulate by varying degrees the affinities for the α_{1A} , α_{1B} , α_{1D} and α_{2C} subtypes. Based on the AncTx4 pharmacological profile, the complementary role of positions 38 and 43 in the pharmacological selectivity of these toxins was reinforced to confirm previous conclusions made for ρ -Da1a, CM-3 and MT β aminergic toxins³⁶. Given that position 43 is located in the globular core of the toxin, a direct involvement in receptor interaction is unlikely. The reason for its impact is unclear but it may involve subtle structural effects. Interestingly, this region of 3FT has already been highlighted in a directed evolution work²¹. Despite a complete “randomization” of the tips of the three loops (23 amino acids out of 61), an additional mutation at position 44 occurred spontaneously in their most represented population of interleukin 6 receptor binders²¹. Regarding position 28, for the first time it was revealed crucial in adrenoceptors binding. This position is proximal to residue 38, confirming this region as key for adrenergic receptor binding. The mechanism behind the binding modulation at position 28 needs to be clarified but it might be due to a stabilizing effect at the tip of finger II as suggested for position 38 (ref. 36). In the MT3-MT1 pathway, despite the sequence variability between ancestral and extant toxins, the important positions for the strong affinity for α_2 adrenoceptors were determined (positions 7 and 15 in finger I and 26 and 28 in finger II).

On the basis of these freshly obtained SAR, new toxins variants with desired properties were engineered. We increased the affinity of AncTx1 for the α_{1B} , α_{1D} or α_{2C} receptors with just two variants, unfortunately without achieving our primary goal to improve selectivity. Epistasis, a phenomenon already stated as important in protein evolution may have been the cause as revealed by other ancestral resurrection studies²⁷. Indeed, the functional effects of a substitution could depend on its surrounding context or sequence background. Despite close homology, the same substitutions introduced in AncTx1 or AncTx3 did not result in either a decrease nor an increase in affinity for receptors α_{1A} , α_{1B} and α_{1D} .

In this study of ancestral toxins, we have obtained new, original and potentially useful functional profiles. AncTx1 is now the most selective peptide known for the α_{1A} adrenoceptor subtypes. AncTx5 is the most potent aminergic toxin able to interact with the three α_2 adrenoceptor subtypes. The recognition mechanism and binding mode of this toxin toward the three receptor subtypes become of great interest.

Conclusion

Our study contributes to the validation of the ancestral protein resurrection methodology that uses a small but functionally rich toxin library. Most of our results come from the first evolutionary pathway related to ρ -Da1a. Nonetheless, the second pathway opens a new promising line of investigation to understand the structure-activity relationships by delineating a few important positions.

The aminergic toxins have proved to be good candidates for ancestral protein resurrection methodology as a small number of variants generated diverse pharmacological profiles. Undoubtedly, further studies will lead to methodology improvements derived from better evaluation of the sequence diversity and its link to the functional divergence before starting the resurrection project. The advantages over other protein engineering strategies are clear: a smaller number of variants with a more diverse activity profiles. Understanding the causes of the epistasis phenomenon would remove it as an obstacle and help future protein engineering work.

Methods

Materials. All tritiated radioactive ligand were purchased from PerkinElmer Life Sciences (Courtaboeuf, France). N-methylscopolamine, prazosin, yohimbine and spiperone were from Sigma-Aldrich (St Quentin-Fallavier, France). As previously described in Blanchet *et al.*¹⁹, membrane preparations were obtained from transient (COS cells) or stable (CHO cells) cell lines prepared in the laboratory.

Peptide synthesis, disulfide bond formation and protein purification. As previously described in Blanchet *et al.*³⁶, the six AncTxs and the two engineered AncTx1 variants were synthesized on a 50 μ mol scale using Fmoc chemistry on a Prelude synthesizer (Protein Technologies®) using HMPB ChemMatrix® resin functionalized with the appropriate protected amino acid (AA). Peptides were purified by reverse phase semi-preparative HPLC using a Vydac C18 column (250 \times 10 mm) with an optimized gradient (usually 40 to 60% of solvent B in 40 min (A: 0.1% TFA in H₂O, B: 60% acetonitrile and 0.1% TFA in H₂O)). The peptides were then subjected to folding in a buffer Tris-HCl 100 mM pH 8 containing 1 mM of EDTA, 1 mM of GSH and GSSG and 20% v/v of glycerol. The final reticulated peptides were purified using the same method described above and characterized by ESI MS (Bruker Esquire 3000 plus).

Circular dichroism analysis. CD spectra were recorded on a Jasco J-815 CD spectropolarimeter. As previously described in Blanchet *et al.*³⁶, measurements were routinely performed at 20 °C in 0.1 cm path length quartz cells (Hellma) with a peptide concentration of 20.10⁻⁶ M in 5 mM sodium phosphate buffer pH 7.4. Spectra were recorded in the 190 to 260 nm wavelength range. Each graphic represents the average of three spectra.

Crystallization and structure determination of AncTx1-W28R/I38S. Crystallization experiments were carried out with lyophilized toxin redissolved at 5 mg/ml in 1 M sodium acetate, pH 5.5. The crystallization trials were carried out using Cryst Chem™ sitting drop vapor diffusion plates with 1 µl drops of protein and precipitant, stored in a constant temperature incubator at 20 °C. Limited screening for crystals was carried out starting from conditions used for MT7³³. Crystals of AncTx1-W28R/I38S were obtained from 1.9 M ammonium sulfate, 4% MPD, 2% 1,4-dioxane, 2% gamma-valerolactone, 0.132 M sodium citrate, pH 5.5. Four crystals were flash frozen in a cryoloop with a cryoprotectant consisting of 80% saturated lithium sulfate, 10% dioxane and 10% gamma-valerolactone. The crystals were shipped for automatic data collection at the Massif 1 at the ESRF in Grenoble, France on beamline ID30A-1³⁸. The XDS³⁹ was used for data reduction with the script XDSME (<https://github.com/legrandp/xdsme>). The crystals belonging to the trigonal space-group P3₁21 with cell parameters a = b = 41.1 c = 67.1 Å diffracted to between 1.95 and 1.8 Å resolution. The structure was solved by molecular replacement starting from the ρ-Da1a crystal structure (PDB ID: 4IYE)³⁴ with Phaser⁴⁰. Refinement was carried out with REFMAC⁴¹ and Phenix⁴² and the electron density fitted with COOT⁴³. Statistics for AncTx1-W28R/I38S structure are reported in Table S1.

Binding experiments. The six AncTxs were first screened on the 12 previously identified aminergic receptor targets reported for the aminergic toxins (namely M₁ to M₅ muscarinic receptor subtypes, all three α₁ and α₂ adrenoceptor subtypes and the dopamine D₃ subtype) by heterologous inhibition experiment as described in Blanchet *et al.*¹⁹. In short, 5 µM of each AncTxs were incubated (in duplicate) with a fixed amount of membrane preparation and a fixed concentration of a specific radioligand in 96-wells plate with 100 µl as final volume for 20h at RT. As previously described in Blanchet *et al.*¹⁹, binding experiments for muscarinic receptor subtypes, were carried out in saline phosphate buffer (10 mM sodium phosphate, 135 mM NaCl, 2.5 mM KCl), pH 7.4, with 0.1% bovine serum albumin (BSA). [³H]-NMS was used as specific radioligand at 1 nM final concentration. The adrenergic α₁ membranes were incubated in Tris buffer (50 mM TrisHCl, 10 mM MgCl₂), pH 7.4, 0.1% BSA, with 1.6 nM of [³H]-Prazosin as radiotracer. For α₂ adrenoceptors, binding was carried out in saline HEPES buffer (50 mM HEPES, 5 mM MgCl₂, 1 mM CaCl₂), pH 7.4, 0.1% BSA and 1.7 nM of [³H]-Rauwolscine. Finally, binding assays on D₃ dopamine receptor subtype were carried out in Tris buffer (50 mM TrisHCl, 10 mM MgCl₂), pH 7.4, 0.1% BSA with 2 nM of [³H]-Methylspiperone. The affinity constants of the radioactive ligands for each membrane preparations were determined by equilibrium saturation experiments (n ≥ 2). Non-specific binding values were obtained by using N-methylscopolamine, prazosin, yohimbine and spiperone in excess (concentration up to 1000 times their K_d) in the binding assay on respectively: the muscarinic, α₁, α₂ and D₃ receptor subtype membrane preparations. Binding reactions were stopped by filtration through GF/C filter pre-soaked in 0.5% polyethyleneimine on a cell harvester (PerkinElmer, Courtaboeuf, France) and plates were dried. Ultimagold O (25 ml, PerkinElmer) was added to each well and samples were counted using a Top-Count counter (PerkinElmer, Courtaboeuf, France). After subtraction of the non-specific binding and normalization, only radiotracer displacement of 20% or more was considered as relevant and subjected to competition experiments.

AncTxs able to inhibit the radioligand at 5 µM were then subjected to dose-response competition experiments. In this experiment, various concentrations of AncTx were used in order to determine their inhibitory constant. This experiment followed exactly the same procedure described earlier. The binding data from individual experiments (n ≥ 2) were analyzed by non-linear regression analysis using Kaleidagraph 4.0 (Synergy software, Reading, PA, USA). After subtraction of the non-specific binding and normalization, a one-site inhibition curve was fitted to inhibition binding data, IC₅₀ values were converted to K_i using the Cheng-Prusoff equation⁴⁴.

Phylogenetic reconstruction and resurrection. As described previously in Blanchet *et al.*¹⁹, sequences of aminergic three-finger toxins were retrieved from the UniProt Knowledgebase. Sequence alignments were carried out using MUSCLE from the MEGA5 package⁴⁵. Alignments were visualized and inspected manually for eventual errors with the MEGA5 alignment visualizer and Jalview version 2⁴⁶ was used to edit the final colored alignment. The maximum likelihood method⁴⁷ was used to reconstruct the phylogenetic trees. The best-fit substitution model, tree reconstruction, visualization, and annotation were chosen using MEGA5. Bootstraps support values were based on 10 000 replicates. Each ancestral amino acid and its associated probability value (P) at every site of each node were retrieved from an excel file supplied by MEGA5 and used to reconstruct the most probable ancestral sequence for each node of the tree. Few ambiguous sites were found, and to minimize the number of possible sequences, the amino-acid with the highest P was every time chosen.

The position 38, located at the middle of the second loop, is quite closed to the tip of the loop II (known as a common “hot spot” for 3FTx to interact on their various targets) and may affect close and functionally important residues. Furthermore, as previously described, the S38I substitution could lead to a local stabilization of the Arg28, Arg40, Lys26 and Tyr36 lateral chains.

References

1. Casewell, N. R., Wuster, W., Vonk, F. J., Harrison, R. A. & Fry, B. G. Complex cocktails: the evolutionary novelty of venoms. *Trends Ecol. Evol.* **28**, 219–229 (2013).
2. Fry, B. G. *et al.* The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. *Annu. Rev. Genomics Hum. Genet.* **10**, 483–511 (2009).
3. King, G. F. Venoms to Drugs - Venom as a source for the development of human therapeutics. 1–306 (Royal Society of Chemistry, 2015).

4. King, G. F. Venoms as a platform for human drugs: translating toxins into therapeutics. *Expert Opin. Biol. Ther.* **11**, 1469–1484 (2011).
5. McCleary, R. J. & Kini, R. M. Non-enzymatic proteins from snake venoms: a gold mine of pharmacological tools and drug leads. *Toxicon* **62**, 56–74 (2013).
6. Maïga, A. *et al.* G protein-coupled receptors, an unexploited animal toxin targets: Exploration of green mamba venom for novel drug candidates active against adrenoceptors. *Toxicon* **59**, 487–496 (2012).
7. Dutertre, S. *et al.* Conopressin-T from *Conus tulipa* reveals an antagonist switch in vasopressin-like peptides. *J. Biol. Chem.* **283**, 7100–7108 (2008).
8. Craig, A. G. *et al.* Conatulakin-G, an O-glycosylated invertebrate neurotensin. *J. Biol. Chem.* **274**, 13752–13759 (1999).
9. Sharpe, I. A. *et al.* Two new classes of conopeptides inhibit the alpha1-adrenoceptor and noradrenaline transporter. *Nat. Neurosci.* **4**, 902–907 (2001).
10. Petrel, C. *et al.* Identification, structural and pharmacological characterization of tau-CnVA, a conopeptide that selectively interacts with somatostatin sst receptor. *Biochem. Pharmacol.* **85**, 1663–1671 (2013).
11. Sudhof, T. C. alpha-Latrotoxin and its receptors: neurexins and CIRL/latrophilins. *Annu. Rev. Neurosci.* **24**, 933–962 (2001).
12. Furman, B. L. The development of Byetta (exenatide) from the venom of the Gila monster as an anti-diabetic agent. *Toxicon* **59**, 464–471 (2012).
13. Rajagopalan, N. *et al.* Beta-cardiotoxin: a new three-finger toxin from *Ophiophagus hannah* (king cobra) venom with beta-blocker activity. *Faseb J.* **21**, 3685–3695 (2007).
14. Servent, D. *et al.* Muscarinic toxins. *Toxicon* **58**, 455–463 (2011).
15. Servent, D. & Fruchart-Gaillard, C. Muscarinic toxins: tools for the study of the pharmacological and functional properties of muscarinic receptors. *J. Neurochem.* **109**, 1193–1202 (2009).
16. Quinton, L. *et al.* Isolation and pharmacological characterization of AdTx1, a natural peptide displaying specific insurmountable antagonism of the alpha(1A)-adrenoceptor. *Br. J. Pharmacol.* **159**, 316–325 (2010).
17. Rouget, C. *et al.* Identification of a novel snake peptide toxin displaying high affinity and antagonist behaviour for the alpha2-adrenoceptor. *Br. J. Pharmacol.* **161**, 1361–1374 (2010).
18. Koivula, K., Rondinelli, S. & Nasman, J. The three-finger toxin MTalpha is a selective alpha(2B)-adrenoceptor antagonist. *Toxicon* **56**, 440–447 (2010).
19. Blanchet, G. *et al.* Polypharmacology profiles and phylogenetic analysis of three-finger toxins from mamba venom: case of aminergic toxins. *Biochimie* **103**, 109–117 (2014).
20. Kessler, P., Marchot, P., Silva, M. & Servent, D. The three-finger toxin fold: a multifunctional structural scaffold able to modulate the cholinergic functions. *J. Neurochem.* doi:10.1111/jnc.13975 (2017).
21. Naimuddin, M. *et al.* Directed evolution of a three-finger neurotoxin by using cDNA display yields antagonists as well as agonists of interleukin-6 receptor signaling. *Mol. Brain* **4**, 2–16 (2011).
22. Cai, W. *et al.* Directed evolution of three-finger toxin to produce serine protease inhibitors. *J. Recept Signal Transduct Res* **34**, 154–161 (2014).
23. Fruchart-Gaillard, C. *et al.* Engineering of three-finger fold toxins creates ligands with original pharmacological profiles for muscarinic and adrenergic receptors. *PLoS One* **7**, e39166 (2012).
24. Harms, M. J. & Thornton, J. W. Analyzing protein structure and function using ancestral gene reconstruction. *Current opinion in structural biology* **20**, 360–366 (2010).
25. Ingles-Prieto, A. *et al.* Conservation of protein structure over four billion years. *Structure* **21**, 1690–1697 (2013).
26. Gonzalez, D. *et al.* Ancestral mutations as a tool for solubilizing proteins: The case of a hydrophobic phosphate-binding protein. *FEBS open bio* **4**, 121–127 (2014).
27. Harms, M. J. & Thornton, J. W. Evolutionary biochemistry: revealing the historical and physical causes of protein properties. *Nat. Rev. Genet.* **14**, 559–571 (2013).
28. Peisajovich, S. G. & Tawfik, D. S. Protein engineers turned evolutionists. *Nat. Methods* **4**, 991–994 (2007).
29. Lutz, S. Beyond directed evolution—semi-rational protein engineering and design. *Curr. Opin. Biotechnol.* **21**, 734–743 (2010).
30. Cole, M. F. & Gaucher, E. A. Exploiting models of molecular evolution to efficiently direct protein engineering. *J. Mol. Evol.* **72**, 193–203 (2011).
31. Chen, F. *et al.* Reconstructed evolutionary adaptive paths give polymerases accepting reversible terminators for sequencing and SNP detection. *Proc. Natl. Acad. Sci. USA* **107**, 1948–1953 (2010).
32. Mourier, G., Dutertre, S., Fruchart-Gaillard, C., Ménez, A. & Servent, D. Chemical Synthesis of MT1 and MT7 Muscarinic Toxins: Critical Role of Arg-34 in Their Interaction with M1 Muscarinic Receptor. *Mol. Pharmacol.* **63**, 26–35 (2003).
33. Fruchart-Gaillard, C. *et al.* Different interactions between MT7 toxin and the human muscarinic M1 receptor in its free and NMS-occupied states. *Mol. Pharmacol.* **74**, 1554–1563 (2008).
34. Maïga, A. *et al.* Crystallization of recombinant green mamba rho-Da1a toxin during a lyophilization procedure and its structure determination. *Acta Crystallogr. F* **69**, 704–709 (2013).
35. Maïga, A. *et al.* Orthosteric Binding of rho-Da1a, a Natural Peptide of Snake Venom Interacting Selectively with the alpha1A-Adrenoceptor. *PLoS One* **8**, e68841 (2013).
36. Blanchet, G. *et al.* New alpha-adrenergic property for synthetic MTbeta and CM-3 three-finger fold toxins from black mamba. *Toxicon* **75**, 160–167 (2013).
37. Bonsor, D. A. & Sundberg, E. J. Dissecting protein-protein interactions using directed evolution. *Biochemistry* **50**, 2394–2402 (2011).
38. Bowler, M. W. *et al.* MASSIF-1: a beamline dedicated to the fully automatic characterization and data collection from crystals of biological macromolecules. *J. Synchrotron Radiat.* **22**, 1540–1547 (2015).
39. Kabsch, W. Xds *Acta Crystallogr. D* **66**, 125–132 (2010).
40. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
41. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D* **67**, 355–367 (2011).
42. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
43. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
44. Cheng, Y.-C. & Prusoff, W. H. Relationship between the inhibition constant (Ki) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochem. Pharmacol.* **22**, 3099–3108 (1973).
45. Tamura, K. *et al.* MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731–2739 (2011).
46. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).
47. Felsenstein, J. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* **17**, 368–376 (1981).

Acknowledgements

The authors thank Dr. Manolo Gouy and Dr. Mathieu Groussin for fruitful discussions and Dr. Didier Nurizzo for X-ray data collection on the automated synchrotron beamline ID30A-1 at the ESRF in Grenoble.

Author Contributions

G.B. and D.S. conceived and designed the study. G.B., D.A., A.P., G.U., L.T., E.A.S., G.M. performed the experiments. All the authors analyzed the data. D.S., G.B., and E.A.S. wrote the main manuscript text and all authors reviewed this manuscript.

Additional Information

Supplementary information accompanies this paper at doi:[10.1038/s41598-017-02953-0](https://doi.org/10.1038/s41598-017-02953-0)

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017