Synthesis of bioactive peptides inspired by venoms

Julien Giribaldi

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Ingénierie Biomoléculaire

École doctorale Sciences Chimiques Balard

Université de Montpellier, Institut des Biomolécules Max Mousseron

Synthèse de peptides bioactifs inspirés des venins

Présentée par Julien GIRIBALDI
Le 25 Novembre 2019

Sous la direction de Sébastien DUTERTRE et Christine ENJALBAL

Devant le jury composé de

M. Nicolas INGUIMBERT, Professeur, CRIOBE, Université de Perpignan Via Domitia	Rapporteur
Mme. Agnès DELMAS, Directeur de Recherche CNRS, Centre de Biophysique Moléculaire, Orléans	Rapporteur (Président)
Mme. Karine PUGET, Directeur général de GENEPEP, St-Jean-de-Védas	Examinateur
M. Loïc QUINTON, Professeur, Mass Spectrometry Laboratory, Université de Liège	Invité
M. Sébastien DUTERTRE, Chargé de Recherche CNRS, IBMM, Université de Montpellier	Directeur
Mme. Christine ENJALBAL, Professeur, Université de Montpellier	Co-directeur

UNIVERSITÉ DE MONTPELLIER

Remerciements

Tout d'abord, je souhaiterai remercier Pr. Pascal DUMY pour m'avoir accueilli au sein de l'Institut des Biomolécules Max Mousseron qu'il dirige. Je remercie également l'Ecole Doctorale Sciences Chimiques Balard, dirigée par Dr. Jean-Jacques Vasseur, pour m'avoir accordé ce financement de thèse. Merci au Dr. Muriel Amblard de m'avoir accueilli au sein du laboratoire de chimie thérapeutique de la faculté de pharmacie bien que je n'y sois pas directement rattaché.

Merci aux membres du jury : Dr. Muriel Amblard, Pr. Nicolas Inguimbert, Dr. Agnès Delmas, Dr. Karine Puget et Pr. Loïc Quinton d'avoir accepté de juger ce travail.

Un grand merci à mon directeur de thèse Dr. Sébastien Dutertre, qui a su par sa sympathie, rendre ces trois années agréables et dénuées de stress négatif, mais également me mettre en confiance et me guider tout au long du doctorat. Merci à ma co-directrice de thèse Pr. Christine Enjalbal d'avoir permis l'évolution du sujet de thèse initial vers une orientation différente.

Merci à nos collaborateurs (de longue date pour Sébastien) sans qui ce travail n'aurait pas été possible : Dr. Norelle L. Daly et Dr. David Wilson pour la partie RMN et Pr. Annette Nicke pour l'électrophysiologie.

Merci au Pr. Loïc Quinton et au Dr. Fernanda Amorim pour m'avoir accueilli au laboratoire de spectrométrie de masse à L'université de Liège et pour m'avoir formé sur le logiciel PEAKS Studio.

J'adresse mes sincères remerciements au Dr. Guillaume Laconde pour m'avoir accueilli avec bonne humeur et formé à la synthèse peptidique et à l'HPLC préparative dans les débuts de ma thèse.

Merci au Dr. Jean-Alain Fehrentz plus connu sous le nom de JAF, pour ses précieux conseils à la paillasse et sa constante bonne volonté pour aider les jeunes (en tout cas ceux qui ne l'embêtent pas).

Merci à Kawthar Belkacemi stagiaire de M1, d'avoir travaillé avec beaucoup d'intérêt et de passion, malgré que la chimie ne fasse pas partie de sa formation initiale.

Merci à Pierre Sanchez pour les analyses LC-MS et pour m'avoir formé sur l'appareillage. Merci au Dr. Gilles Valette pour les analyses MALDI et à Guillaume Cazals pour son aide précieuse sur la partie LC-MS.

Merci au Dr. Pascal Verdie pour m'avoir permis de m'installer dans le bureau normalement destiné à ses étudiants, sans lequel j'aurais certainement effectué une bonne partie de ma thèse sur un bout de 50 cm de paillasse sèche.

Merci au Dr. Luc Brunel et au Dr. Lubomir Vezenkov pour les discussions scientifiques et particulièrement au Dr. Matthieu Simon passionné par la recherche et toujours très intéressé à échanger scientifiquement et grâce à qui je peux faire des supers animations sur chimera.

Un grand merci à Manon Maurel qui m'a supporté pendant ces trois années, dans ce bureau délabré rendu chaleureux par nos éclats de rires. Merci pour son soutient et sa bonne humeur (presque constante, dans le cas contraire rdv devant la prep) qui ont rendu les journées difficiles beaucoup plus supportables.

Merci à Julie Martin pour les petites pauses discussions détentes dans le bureau et pour les nombreuses soirées (à l'époque où elle sortait encore). Merci à Titouan Montheil à qui j'ai usurpé l'identité pendant presque 6 mois pour entrer sur le parking de pharma et évidement pour les parties de squash et de padel endiablées, enfin quand il n'avait pas un empêchement… (537 parties prévues vs. 3 réellement jouées). Merci à Laurine Valot pour m'avoir rencardé sur toutes les procédures administratives de la thèse, grâce à qui j'ai pu tout faire correctement (enfin je crois). Merci à Bénédicte Drean pour sa jovialité et son humour.

Un remerciement particulier à toute la bande du CAS : Rayane Ghoteimi, Mathieu Noël, Rostom Ahmed-Belkacem, Justine Mansot et Sébastien Diot pour tous ces repas animés de nos rires, qui constituaient une pause bienvenue et toujours agréable.

Merci aux personnes du deuxième étage du bâtiment 17 : Francois Quintin, Michel Bonnard, Yves Yeboue, Elodie Fernandez, Mylène Sejalon pour leur accueil chaleureux lorsque j'étais dans les parages.

Merci au Dr. Vittoria Modica pour m'avoir renseigné et aidé pour la rédaction de mon projet postdoctoral et à Rosanna Mary ma « binome de galère » comme elle aime bien dire, qui m'a accompagné pendant ces 3 ans de thèse.

Un grand merci à Lori Gonnet pour sa bonne humeur et son humour, qui m'a soutenu, accompagné et s'est adapté mon rythme de travail inversant complètement le jour et la nuit, pendant la période de la rédaction, la rendant ainsi beaucoup plus supportable.

Merci à mes amis d'enfance que je n'ai pas trop eu le temps de voir mais qui m'ont toujours soutenu : Daniel Yasri, Baptiste Lala, Vincent Huc et Amory Panné.

Pour finir, un gros merci aux membres de ma famille : ma mère, mon père, mes frères et mon grand-père qui m'ont soutenu et supporté pendant ces trois années, avec une pensée particulière pour ma grand-mère.

Sommaire

Liste des abréviations

5'-NUC : 5'-nucleotidase
acac : acetylacetone
Acm : Acetamidomethyl
ACN : Acetonitrile
ADNc/cDNA : Acide Désoxyribonucléique cyclique/cyclic Deoxyribonucleic Acid
AMPc/cAMP : Adénosine Monophosphate cyclique/ cyclic Adenosine Monophosphate
Boc : *tert*-butyloxycarbonyl
BOP : (Benzotriazol-1-yloxy)tris(dimethylamino)phosphonium hexafluorophosphate
BRET : Bioluminescence Resonance Energy Transfer
CAA : Common Amino Acids
Ca_v : Canaux calciques dépendant du voltage
CB_1 : cannabinoid receptor type 1
cGMP : cyclic Guanosine Monophosphate
CID : Collision Induced Dissociation
CRISP : Cysteine Rich Secretary Protein
C-ter : C-terminale
CTL : C-Type Lectin
DCM : Dichloromethane
DIC : *N,N'*-Diisopropylcarbodiimide
dICK : double Inhibitor Cystine Knot
DIPEA : *N*,*N*-Diisopropylethylamine
Dis : Disintegrin
DMF : *N,N*-Dimethylformamide
DMSO : Dimethyl sulfoxide
DQF-COSY : two-dimensional Double Quantum Filter Homonuclear Correlation Spectroscopy
DTP : 2,2'-Dithiodipyridine
ECD : Electron Capture Dissociation
Eq : equivalent
ESI : Electrospray Ionization
ESI-MS : Electrospray Ionization Mass Spectrometry
ETB : Endothéline B
ETD : Electron Transfer Dissociation
FDR : False Discovery Rate
FLIPR : Fluorescent Imaging Plate Reader
Fm : 9-Fluorenylmethyl
Fmoc : Fluorenylmethyloxycarbonyl
FRET : Fluorescence Resonance Energy Transfer
FT-MS : Fourier-transform Mass Spectrometry
GHSR : Growth Hormone Secretagogue Receptor
GLP1R : Glucagon-like Peptide 1 Receptor
GnHCl : Guanidine hydrochloride
GPCR/RCPG : G protein-coupled receptor/récepteurs couplés aux protéines G
GSH : Glutathione
GSSG : Oxidized Glutathione
GTP: Guanosine Triphosphate
HATU : 1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate
HBTU : *N,N,N',N'*-Tetramethyl-*O*-(1*H*-benzotriazol-1-yl)uronium hexafluorophosphate
HCD : Higher energy Collision Dissociation
HF : Hydrofluoric acid

HMM : Hidden Markov Model
HOAt : 1-Hydroxy-7-azabenzotriazole
HOBt : 1-Hydroxybenzotriazole
hOTR : human Oxytocin Receptor
HPLC : High-pressure Liquid Chromatography
HSQC : Heteronuclear Single Quantum Correlation
hV1aR : human Vasopressin-1a Receptor
hV1bR : human Vasopressin-1b Receptor
hV2R : human Vasopressin2 Receptor
HYAL : Hyaluronidase
ICK : Inhibitor Cystine Knot
ICR-MS : Ion-Cyclotron-Resonance Mass Spectrometry
IP : Inositol Phosphate
IT : Ion Trap
K_{ATP} : potassium channels ATP-dependent
K_v : Canaux potassiques dépendant du voltage
LAAO : L-Amino Acid Oxidase
LC/MS : Liquid Chromatography coupled to Mass Spectrometry
LEAP : Liver-expressed Antimicrobial Peptide
mAChRs : muscarinic Acetylcholine Receptors
MALDI : Matrix-assisted Laser Desorption/Ionization
Meb : 4-Methylbenzyl
MeOH : Methanol
Mob : 4-Methoxybenzyl
MOR : μ-opioid Receptor
MPAA : 4-mercaptophenylacetic acid
MS/MS : Tandem Mass Spectrometry
MTs : Muscarinic Toxins
nAChRs : nicotinic Acetylcholine Receptors
NCL : Native Chemical Ligation
NEP : Neutral Endopeptidase
NGF : Nerve Growth Factor
NGS : Next Generation Sequencing
NMR/RMN : Nuclear Magnetic Resonance/Résonance Magnétique Nucléaire
nNOSs : neuronal Nitric Oxide Synthetases
NO : Nitric Oxide
NOESY : two-dimensional Nuclear Overhauser Effect Spectroscopy
NOSs : Nitric Oxide Synthetases
NPys : 3-Nitro-2-pyridylsulfenyl
N-ter : N-terminale
Pbf : 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl
PDE : Phosphodiesterase
Pfp : Pentafluorophenyl
PFP : Peptide Fragmentation Pathway
PKR1 : Prokineticin Receptor 1
PLA2 : Phospholipase A2
PL-B : Phospholipase B
PSM : Peptide Spectrum Matching
PSSM : Position-specific Scoring Matrix
PTM : Post-translational Modification
RLAP : Renin-Like Aspartic Protease
RP-HPLC : Reversed-phase High-pressure Liquid Chromatography

RT : Reverse Transcriptase
SAR : Structure-activity Relationship
ShK : Stichodactyla helianthus
SPPS : Solid-phase Peptide Synthesis
S*t*Bu : *tert*-Butylsulfenyl
SVMP : Snake Venom Metalloprotease
SVMPi : Snake Venom Metalloprotease inhibitor
SVSP : Snake Venom Serine Protease
TA : Température Ambiante
***t*Bu** : *tert*-Butyl
TFA : Trifluoroacetic acid
TIS : Triisopropylsilane
Tmob : 2,4,6-Trimethoxybenzyl
TMP : Trimethyl phosphate
TOF : Time Of Flight
TPM : Transcripts Per Million Kilobase
Tris : Tris(hydroxymethyl)aminomethane
Trt : Triphenylmethyl (trityl)
UV : Ultraviolet
Vasp : *Vipera aspis*
Vbb : *Vipera berus berus*
VEGF : Vascular Endothelial Growth Factor
Xan : 9H-Xanthem-9-yl
ZF oxy/isoR : Zebrafish oxytocin/isotocin Receptor
ZF V1a1R : Zebrafish Vasopressin-a1 Receptor
ZF V1a2R : Zebrafish Vasopressin-a2 Receptor
ZF V2R : Zebrafish Vasopressin2 Receptor
α_1-AR : α_1 Adrenergic Receptor
α-LTx : α-latrotoxine

Productions scientifiques

Les résultats obtenus lors de ces travaux de thèse ont permis la production de 5 articles scientifiques dans des journaux internationaux :

- El Hamdaoui, Y.; Wu, X.; Clark, R.J.; **Giribaldi, J**.; Anangi, R.; Craik, D.J.; King, G.F.; Dutertre, S.; Kaas, Q.; Herzig, V.; et al. Periplasmic Expression of 4/7 α-Conotoxin TxIA Analogs in E. coli Favors Ribbon Isomer Formation – Suggestion of a Binding Mode at the α7 nAChR. ***Front. Pharmacol***. 2019, *10*, 577. [Article page 149].

- M'Kadmi, C.; Cabral, A.; Barrile, F.; **Giribaldi, J**.; Cantel, S.; Damian, M.; Mary, S.; Denoyelle, S.; Dutertre, S.; Péraldi-Roux, S.; et al. N-terminal Liver-expressed antimicrobial peptide 2 (LEAP2) region exhibits inverse agonist activity toward the ghrelin receptor. ***Journal of Medicinal Chemistry*** 2018, *62*, 965–973. [Article page 179].

- **Giribaldi, J**.; Wilson, D.; Nicke, A.; El Hamdaoui, Y.; Laconde, G.; Faucherre, A.; Moha Ou Maati, H.; Daly, N.; Enjalbal, C.; Dutertre, S. Synthesis, Structure and Biological Activity of CIA and CIB, Two α-Conotoxins from the Predation-Evoked Venom of *Conus catus*. ***Toxins*** 2018, *10*, 222. **Couverture de *Toxins*, Volume 10, Issue 6 (Juin 2018).** [Article page 93].

- Drukewitz, S.H.; Fuhrmann, N.; Undheim, E.A.B.; Blanke, A.; **Giribaldi, J**.; Mary, R.; Laconde, G.; Dutertre, S.; von Reumont, B.M. A Dipteran's Novel Sucker Punch: Evolution of Arthropod Atypical Venom with a Neurotoxic Component in Robber Flies (Asilidae, Diptera). ***Toxins*** 2018, *10*, 29. [Article page 213].

- **Giribaldi, J**.; Dutertre, S. α-Conotoxins to explore the molecular, physiological and pathophysiological functions of neuronal nicotinic acetylcholine receptors. ***Neuroscience Letters*** 2018, 679, 24–34. [Article page 77].

Les résultats obtenus ont également pu être présentés lors de communications orales et posters dans des congrès internationaux :

- Communications orales

-21st meeting of the French Group of Peptides and Proteins (GFPP); 12/05/2019 – 16/05/2019; Amboise, France

Backbone cyclization turns muscle nAChR α-conotoxin CIA into highly potent neuronal nAChR ligand.

-25th meeting of the French Society for the Study of Toxins (SFET); 28/11/2018 – 29/11/2018; Paris, France

Backbone cyclization of native α-conotoxin CIA dramatically increases affinity at neuronal nicotinic receptors.

-24th meeting of the French Society for the Study of Toxins (SFET); 7/12/2017 – 8/12/2017; Paris, France

Synthesis, structure and biological activity of CIA and CIB, two new alpha-conotoxins from Conus catus.

- Présentation de poster

-Congress of Mass Spectrometry, Metabolomic and Proteomic Analysis (SMMAP); 2/10/2017 – 5/10/2017; Paris, France

Integrated proteomic and transcriptomic approaches towards the high-throughput discovery of new bioactive molecules in venoms.

Avant-propos

Jusqu'au début des années 2000, les industries pharmaceutiques se sont concentrées sur le développement de petites molécules thérapeutiques tandis que les peptides étaient laissés de côté en raison de leur faible biodisponibilité. Au cours de la dernière décennie, le concept selon lequel un médicament peut ne pas être disponible par voie orale est devenu de plus en plus accepté et a contribué à un regain d'intérêt pour les peptides comme candidats médicaments potentiels. Bien que leur poids moléculaire supérieur à celui de petites molécules puisse être un obstacle au passage à travers les barrières biologiques, les peptides sont également capables d'effectuer un nombre plus important d'interactions moléculaires avec leur cible, ce qui les rend plus puissants et plus sélectifs que les petites molécules. De plus, la similarité des peptides avec les molécules endogènes leur confère une faible toxicité avec moins d'effets secondaires que les produits pharmaceutiques traditionnels composés de petites molécules. Ainsi, les peptides sont devenus le médicament de choix pour le traitement de nombreuses maladies, comme le diabète, le VIH, l'hépatite, le cancer, etc., et les médecins et les patients acceptent de plus en plus les médicaments à base de peptides. Plus important encore, leur forte affinité et leur sélectivité pour leurs récepteurs en font des outils de recherche inestimables qui aident à disséquer les fonctions physiologiques de nombreux récepteurs humains et à élucider les mécanismes biologiques sous-jacents des maladies. La nature nous met à disposition de riches sources de peptides bioactifs. En effet, de plus en plus de peptides entrant en développement clinique proviennent d'animaux, de plantes ou encore de bactéries. Au cours de millions d'années d'évolution, les venins d'animaux ont été finement ajustés pour cibler de manière puissante, rapide et sélective les récepteurs fonctionnels clés grâce à des peptides. Un nombre croissant de peptides dérivés de venins sont en phase d'essai clinique, certains d'entre eux sont déjà approuvés par la FDA. Parmi les peptides dérivés de venins utilisés en médecine, citons l'Eptifibatide, un médicament antiplaquettaire développé à partir de l'echistatin provenant d'une vipère, le Ziconotide, un puissant analgésique identifié dans le venin d'un cône ou encore l'Exenatide, un agoniste du récepteur du glucagon-like peptide 1 isolé de la salive du monstre de Gila et utilisé pour le traitement du diabète de type 2. Les venins d'animaux se composent d'une diversité très complexe de peptides évolués pour paralyser efficacement les proies ou pour se défendre contre un large éventail d'ennemis. La conservation et la ressemblance des récepteurs de la proie/du prédateur avec les récepteurs humains, fait des peptides de venins une source unique d'éléments de base pour le design d'outils pharmacologiques et de composés thérapeutiques. Le potentiel des venins comme source naturelle de composés bioactifs est indéniable, de plus la richesse est telle qu'il est estimé que moins de 1% des peptides de venins ont été caractérisés pharmacologiquement. La marge de progression de

l'exploitation des venins reste donc conséquente et de plus en plus de nouvelles toxines bioactives seront caractérisées.

Les objectifs majeurs de ce sujet de thèse étaient la découverte, la synthèse et la caractérisation de nouveaux peptides isolés à partir de venins animaux. En effet, les quantités extraites sont rarement suffisantes pour une caractérisation pharmacologique et structurale complète. Ainsi, la synthèse par voie chimique de ces nouvelles toxines potentiellement bioactives devait permettre leur caractérisation structurale grace à la collaboration avec les Drs Norelle Daly et David Wilson du Centre for Molecular Therapeutics à l'université James Cook (Cairns, Australie) et leur caractérisation biologique avec la collaboration du Dr Annette Nicke du Walther Straub Institute of Pharmacology and Toxicology à l'université de Munich (Munich, Allemagne). L'objectif final étant la découverte de nouveaux outils pharmacologiques ou nouvelles molécules à potentiel thérapeutique.

Les venins ayant évolués pour la défense et pour faciliter la prédation, les toxines qui les composent ont souvent pour but d'incapaciter la proie ou le prédateur, notamment en ciblant des fonctions physiologiques clés telles que l'hémostase, la jonction neuromusculaire ou la perception de la douleur. Bien que la plupart des toxines typiques modulent les canaux ioniques assurant la transmission de l'information nerveuse et comportent généralement plusieurs ponts disulfures, nous avons d'abord commencé par étudier le potentiel des toxines linéaires et à 1 pont sur des cibles plus rarement explorées tels que les récepteurs couplés aux protéines G (chapitre 1). Cependant, la majorité des toxines étant composées de 2 ou 3 ponts disulfures, nous avons ensuite opté pour la mise en place d'une stratégie robuste de synthèse de peptides à deux ponts. Ainsi, la synthèse de plusieurs α-conotoxines (chapitre 2) a permis leur caractérisation structurale par RMN et pharmacologique sur les récepteurs nicotiniques. Une extension de ce travail sur une séquence prometteuse ayant une activité duale sur les récepteurs musculaire et neuronaux a été effectuée par cyclisation du squelette peptidique. Enfin, ce travail a été étendu à la synthèse de toxines plus longue ciblant d'autres canaux ioniques et comportant 3 ponts disulfures (chapitre 3). Pour finir, la synthèse de nouvelles toxines découvertes dans des banques de données issues de divers transcriptomes de glandes à venin a été initiée, et nous avons réalisé l'analyse protéo-transcriptomique d'un venin de serpent pour l'identification de nouvelles toxines (chapitre 4).

1. Introduction

a. Les peptides thérapeutiques

Au cours du dernier siècle, les petites molécules dérivées de produits naturels constituaient la principale source d'inspiration pour la plupart des programmes de découverte de candidats médicaments basés soit sur le criblage haut débit ou sur des processus rationnels de design [1]. L'aube du nouveau millénaire a été marqué par un changement de cap dans l'industrie pharmaceutique, motivé par des normes de sécurité sur les produits thérapeutiques plus strictes imposées par les autorités de réglementation [2,3] et par l'émergence de nouvelles cibles thérapeutiques pertinentes [4] telles que les interactions protéine-protéine [5,6]. En outre, les progrès des technologies « omiques » et des stratégies de conception de médicaments assistée par ordinateur [7–9], ainsi que les progrès de l'expression des protéines recombinantes et le développement d'une synthèse peptidique plus efficace et économique, ont contribué à modifier le marché pharmaceutique actuel. Ainsi, ces dernières années le secteur des produits biologiques et thérapeutiques à base de peptides s'est progressivement développé aux dépens de celui des petites molécules médicaments, et les taux de réussite de mise sur le marché sont environ deux fois plus élevés que ceux des petites molécules médicaments [10–12].

Comme beaucoup de messagers moléculaires sont des peptides ou des formes dérivées, l'utilisation des peptides pour une application thérapeutique permet d'imiter fidèlement les voies de signalisations naturelles. En effet, beaucoup de peptides thérapeutiques sont des hormones de synthèse ou des substituts administrés pour traiter les cas où les niveaux endogènes sont inadéquats [13]. A titre d'illustration, un exemple typique est l'isolation et la première utilisation thérapeutique de l'insuline, identifiée comme le principe actif des tissus pancréatiques et utilisée pour le traitement des personnes diabétiques souffrant d'une déficience hormonale en insuline [14]. Plus tard dans les années 1930, l'isolation de l'hormone corticotrope à partir des glandes pituitaires et son injection intramusculaire chez les patients présentant une sévère arthrite rhumatoïde a conduit à une diminution des symptômes [15]. Dans la première moitié du vingtième siècle, les peptides isolés de sources naturelles représentaient déjà des médicaments pouvant sauver des vies humaines (Tableau 1), cependant les quantités extraites restent très faible et dépendent largement de la quantité disponible en ressource biologique.

Dans les années 1950 quand les techniques de séquençage et la synthèse chimique ont commencé à émerger et à se répandre, de petits peptides synthétisés par voie chimique comme l'ocytocine [16] et la vasopressine [17] sont entrés en tests cliniques (Tableau 1). Cependant, l'enthousiasme concernant l'usage thérapeutique des peptides a été considérablement tempéré à cause de leur faible

biodisponibilité. En effet, l'administration par voie orale est considérée comme la méthode d'administration la plus attractive car elle permet l'autosuffisance sans avoir recours à une prise en charge externe du patient surtout pour les pathologies nécessitant un traitement chronique. Malgré les grandes avancées dans la découverte de candidats médicaments et la synthèse peptidique leur faible biodisponibilité due à la rupture des liaisons amides par les enzymes digestives ainsi que leur polarité élevée et leur haut poids moléculaire, limitent leur perméabilité intestinale (**Error! Reference source not found.**). De plus, Les peptides sont généralement éliminés par les reins quelques minutes après avoir pénétré le système sanguin. Par exemple, l'un des peptides les plus courants, l'insuline, a une demi-vie dans la circulation sanguine de seulement 4 à 6 minutes. Par conséquent les peptides thérapeutiques sont administrés par injection : sous-cutanée, intraveineuse, intramusculaire ou encore intrathécale [13].

Tableau 1 : Source ou nature chimique des premiers peptides thérapeutiques.

Peptide	Source	Tests cliniques	Description de la séquence
Insuline	Isolée des pancréas de bovins et de chiens	Années 1920	Naturelle
Hormone corticotrope	Isolée des glandes pituitaires de bovins et de chien	Années 1950	Naturelle
Calcitonine	Isolée des corps ultimobranchiaux de saumon	1971	Naturelle
Ocytocine	Synthétique	1962	Naturelle
Vasopressine	Synthétique	1962	Naturelle
Octreotide	Analogue synthétique de la somatostatine	1988	Octapeptide cyclique analogue de la somatostatine-14
Gonadoreline	Analogue synthétique de la gonadoreline	1984	Nonapeptide analogue du decapeptide gonadoreline

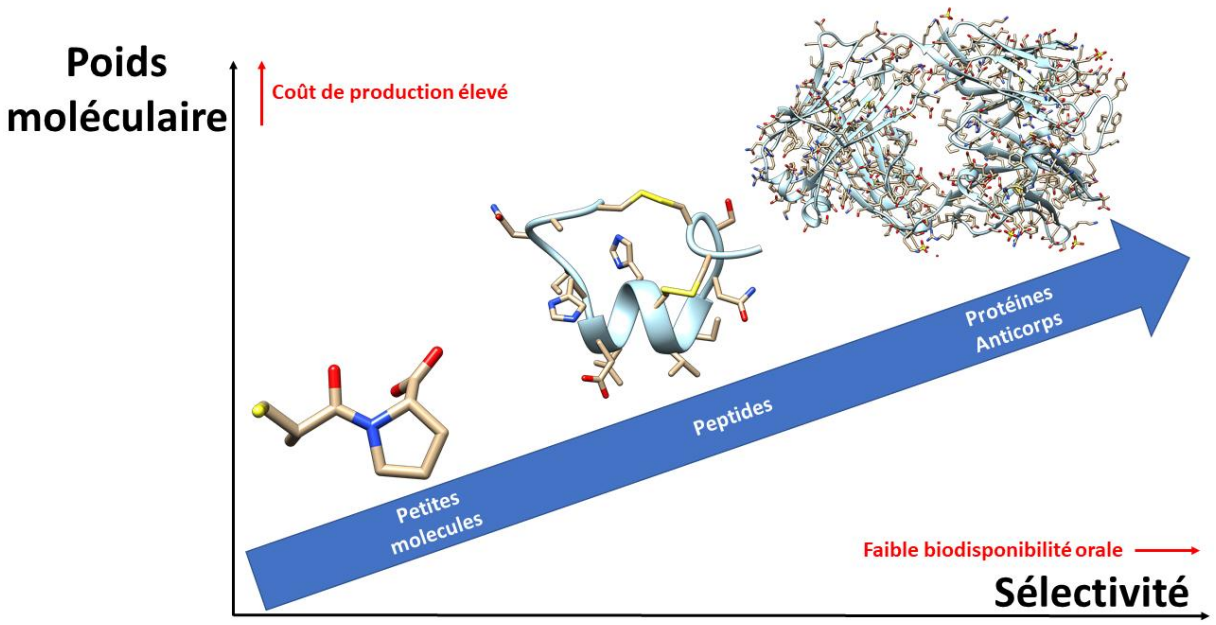


Figure 1 : Caractéristiques des petites molécules, peptides et protéines/anticorps.

Par conséquent, pendant les années 1980-2000, les industries pharmaceutiques se sont concentrées sur le développement de petites molécules thérapeutiques et les peptides étaient négligés à cause de leur faible biodisponibilité. Récemment, l'idée qu'un médicament ne doit pas forcément être disponible par voie orale est devenue de plus en plus acceptée et a contribué à un renouveau pour l'intérêt des peptides comme potentiels candidats médicaments [18]. Bien que leur masse moléculaire élevée puisse être problématique pour le passage des barrières biologiques, les peptides du fait de leur richesse structurale sont capables de générer un grand nombre d'interactions avec leur cible pharmacologique les rendant plus puissants et plus sélectifs que les petites molécules. De plus, la similarité des peptides avec les molécules endogènes assure une moindre toxicité, et provoque moins d'effets secondaires comparés aux petites molécules thérapeutiques [19]. Les médecins et les patients étant de plus en plus enclins à accepter des traitements à base de peptides, ceux-ci sont devenus un traitement de choix pour de nombreuses maladies telles que le diabète, HIV, l'hépatite, le cancer etc. Par-dessus tout, leur haute affinité et sélectivité pour leurs récepteurs cibles en font des outils pharmacologiques inestimables pour l'étude et la compréhension des fonctions physiologiques de nombreux récepteurs humains impliqués dans les mécanismes biologiques des maladies [20].

Une stratégie d'intérêt pour limiter la dégradation enzymatique *in vivo* des peptides est de s'inspirer des peptides riches en ponts disulfures, en effet ceux-ci leur confèrent une stabilité accrue comparée aux peptides linéaires, envers la digestion enzymatique. Il existe de nombreuses sources de peptides riches en ponts disulfures, tels que les plantes, les bactéries ou encore les animaux, et nous avons choisi de nous focaliser sur les venins d'animaux.

En effet, les caractéristiques des peptides longtemps perçues comme des handicaps ne sont plus problématiques : par exemple, l'injection est considérée comme une voie acceptable pour certaines indications, en partie due au développement de peptides à action prolongée et à de nouvelles formulations réduisant considérablement la fréquence des injections. Par exemple, malgré l'existence d'une grande variété de classes de médicaments disponibles par voie orale pour le traitement du diabète de type 2, le marché des peptides injectables agonistes du GLP-1 n'a cessé de grandir depuis l'approbation de l'exenatide en 2005 [13]. Au début des années 1990, le criblage haut débit de venins de lézards Nord américains (« monstres de Gila ») a conduit à la découverte de deux peptides très similaires exendin-3 et exendin-4, ne différant que de deux acides aminés l'un de l'autre (Figure 2). Les deux peptides augmentent les niveaux de AMPc (adénosine monophosphate cyclique) et la production d'insuline mais l'exendin-3, due à son manque de sélectivité interagit avec deux classes de récepteurs ce qui a pour effet l'augmentation de la production d'amylase [21–23]. Par conséquent, l'exendin-4 commercialisé sous l'appellation exenatide a servi de base pour le développement d'agonistes à action prolongée du GLP-1 (Tableau 2). Ceux-ci peuvent être injectés une fois par semaine au lieu de deux fois par jour (exenatide) ou une fois par jour (liraglutide, lixisénatide) [24]. La première injection à durée d'action d'une semaine a été effectuée par relargage contrôlé de l'exenatide [25] mais récemment deux autres analogues ont été commercialisés : l'albiglutide et le dulaglutine et encore d'autres sont en cours développement (semaglutide) [24,26]. L'intérêt pour les agonistes du peptidiques du GLP-1 augmente rapidement surtout depuis la réduction des conséquences cardiovasculaires et rénales récemment reportée lors des essais cliniques du liraglutide et du semaglutide sur les patients avec des antécédents de maladies cardiovasculaires [27,28]. Dans certains cas comme celui du teriparatide, une version tronquée de l'hormone parathyroïde présentant un mécanisme d'action bien distinct des molécules existantes sur le marché justifie une injection quotidienne pour lutter contre l'ostéoporose. En effet, celui-ci s'avère être supérieur aux bisphosphonates concernant la diminution du risque de fracture et l'augmentation de la densité minérale osseuse de la femme ménopausée [29–31].

Tableau 2 : Séquence primaire des peptides agonistes du GLP-1. L'astérisque indique un acide aspartique gamma-palmitoylé. La lettre X indique un acide 2-Aminoisobutyrique. Le symbole † une lysine acétylée avec un acide stéarique.

Nom	Séquence primaire
Exenatide	HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS
Liraglutide	HAEGTFTSDVSSYLEGQAAKE*EFIAWLVRGRG
Lixisénatide	HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPSKKKKKK
Albiglutide	Protéine de 72970.0 Da
Dulaglutine	Protéine de 59669.81 Da
Semaglutide	HXEGTFTSDVSSYLEGQAAK†EFIAWLVRGRG

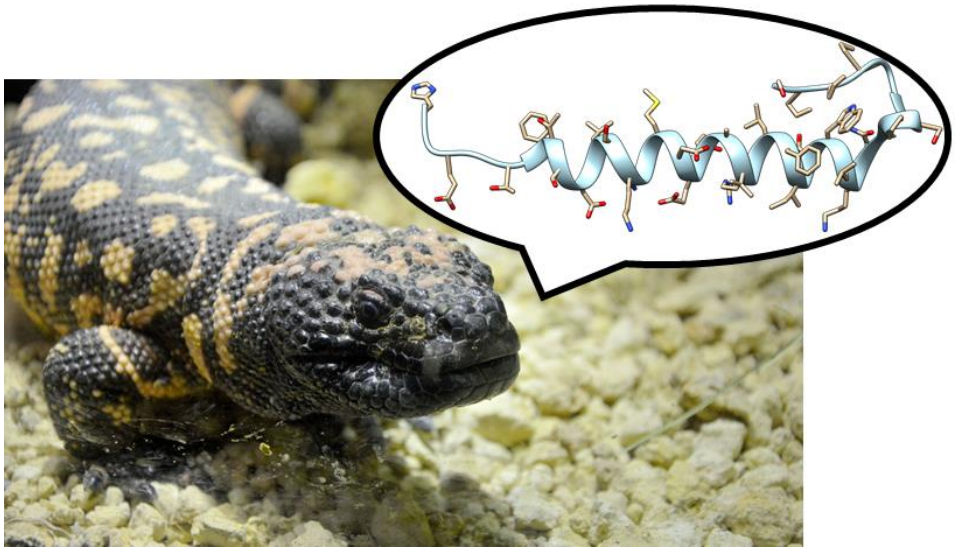


Figure 2 : Structure de l'exendin-4 (exenatide) identifié dans la salive du monstre de Gila.

De plus, les candidats médicaments peptidiques sont conçus contre des cibles moléculaire dépassant le domaine des petites molécules qui ne peuvent seulement cibler qu'une fraction des cibles moléculaire thérapeutiquement pertinentes traditionnellement dénommé « druggable targets » [32]. Ainsi des peptides empêchant les interactions protéines-protéines (longtemps classifiés comme « undruggable targets ») [33,34], ciblant les récepteurs tyrosine kinases [35] et inhibant des cibles intracellulaires [36] ont fait leur entrée en phase de test cliniques et sur le marché du médicament.

Actuellement, plus de 60 peptides médicaments ont été approuvé aux Etats-Unis, en Europe et au Japon ; plus de 150 sont en développement clinique actif et 260 autres ont été testés dans des essais cliniques humains. Le nombre de peptides entrant en développement clinique a progressivement augmenté entre 1980 et 2010, la moyenne sur cinq ans s'établissant à plus de 22 peptides par an en 2011. Le nombre cumulé de peptides approuvés a également augmenté progressivement, avec 13 approbations de peptides ayant eu lieu à partir du début de 2010 jusqu'à ce jour. Au cours de la

décennie actuelle, les candidats au développement sont répartis plus équitablement dans diverses plages de longueur allant jusqu'à 40 acides aminés (Figure 3), ce qui suggère que le nombre de résidus n'est plus une limite sérieuse pour le développement d'un médicament peptidique. Rendant ainsi encore plus pertinente l'investigation des peptides naturels, généralement plus longs que les analogues peptidiques synthétiques, comme sources de composés à potentiel thérapeutique. Il existe de nombreuses sources de peptides naturels biologiquement actifs, tels que les plantes, les bactéries ou encore les animaux, et nous avons choisi de nous focaliser sur les venins d'animaux.

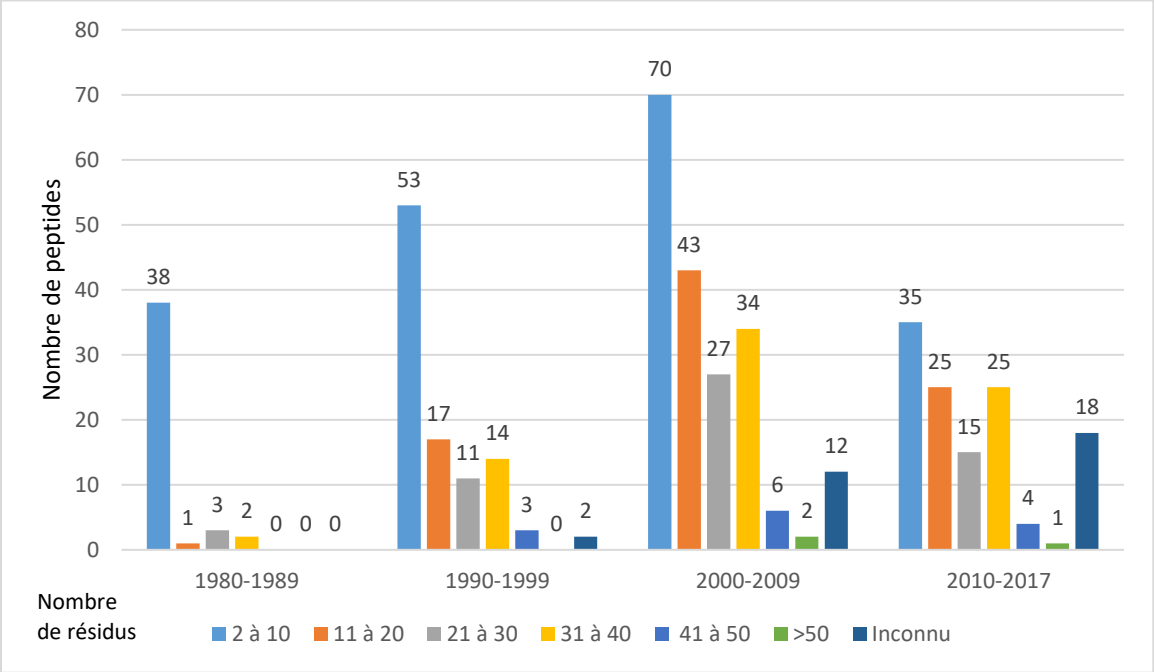


Figure 3 : Longueur des peptides entrant en développement clinique, groupés par décennie [13].

b. Stratégies de découverte de peptides bioactifs dans les venins

Les animaux ont développé des venins pour leur protection et la prédation. La diversité des animaux employant cette stratégie de survie englobe presque tous les phylums du règne animal, y compris les invertébrés tels que les annélides (ver de feu, glycera), les cnidaires (anémones de mer, méduses et coraux), les échinodermes (oursins et étoiles de mer), les mollusques (cônes et pieuvres), arthropodes (araignées, fourmis, centipèdes, abeilles, guêpes, scorpions, moustiques et tiques) et les vertébrés (poissons, grenouilles, serpents, lézards, oiseaux et mammifères) (Figure 4). Le venin est inoculé au prédateur ou à la proie via un dard, des crochets, des cellules, des harpons etc. La composition du venin varie d'un animal à l'autre mais la plupart des venins sont constitués d'un mélange hétérogène de sels inorganiques, de petites molécules à faible poids moléculaire, de peptides et d'enzymes. Ce mélange offre à l'animal une approche à plusieurs facettes pour immobiliser et / ou tuer la proie ou le

prédateur [37,38]. Depuis la nuit des temps, l'espèce humaine a intégré l'idée que les créatures venimeuses sont extrêmement dangereuses et potentiellement fatales pour les victimes mordues ou piquées. Nos ancêtres ont vite pris conscience du potentiel des venins et ont développé des armes mortelles à base de projectiles trempés dans les venins de différents animaux. Un des premiers usages thérapeutiques reporté des venins a été décrit par Appian, historien romain d'origine grecque, en 37 avant J.C, lorsque Mithradates a été blessé à la cuisse par une épée et, alors qu'il était sur le point de mourir, son médecin scythe lui a administré une petite quantité de venin de vipère des steppes (*Vipera ursinii*) aux propriétes anticoagulantes pour arrêter le saignement abondant et lui sauver la vie [39]. Les venins de serpents, araignées et scorpions ont été utilisés pendant des millénaires dans un grand nombre de remèdes traditionnels et médicaments pour traiter une variété de maladies comme l'arthrite, le cancer, et les problèmes gastro intestinaux pour en citer quelques uns. La plupart de ces remèdes traditionnels utilise de petites doses de venin pour atteindre leurs objectifs thérapeutiques. Ce n'est qu'à la fin du XXe siècle que la médecine moderne a adopté une approche plus systématique et plus rigoureuse pour utiliser les venins comme agents thérapeutiques [40].

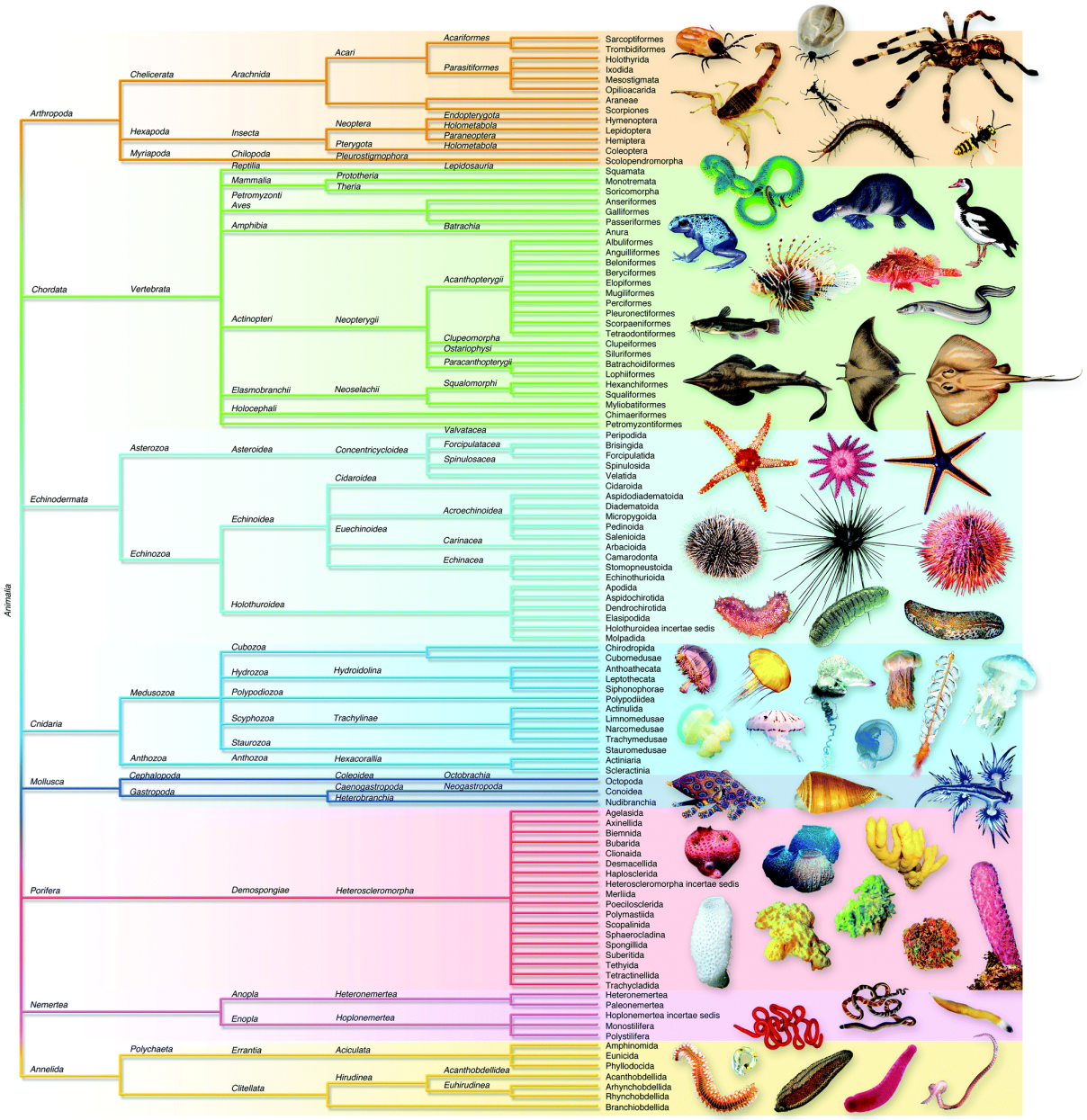


Figure 4: Arbre phylogénétique schématique du monde venimeux [41].

Au cours de millions années d'évolution, les venins ont été finement ajustés à leurs cibles. En effets ils interagissent de façon puissante, sélective et rapide avec les récepteurs assurant les fonctions clés dans l'organisme de leurs proies ou de leurs prédateurs. La conservation et la similarité de séquence des récepteurs de la proie ou du prédateur avec les récepteurs humains expliquent que les venins soient source de composés bioactifs servant de base au développement d'outils potentiellement thérapeutiques [37,38,42]. En premier lieu, la forte affinité et sélectivité de ces toxines pour leur récepteurs cibles en font de formidables outils de recherche pour l'étude des fonctions physiologiques de nombreux récepteurs humains et l'élucidation des mécanismes biologiques sous-jacent aux maladies [20].

Actuellement sur le marché 9 médicaments dérivés ou non de venins de serpents, sangsues, cônes, monstres de Gila etc. ont été approuvé pour traiter une large de gamme de maladies comme l'hypertension, l'infarctus, la douleur, le diabète etc. (Tableau 3). Cependant la diversité naturelle des venins et des espèces venimeuses est telle que nous avons à ce jour simplement effleuré le potentiel offert par les molécules bioactives présentes dans les venins. En effet, l'avènement des techniques dites « omiques » (détaillé plus tard) a permis le séquençage de nombreuses espèces venimeuse et la découverte de nombreuses autres toxines bioactives en phase de développement clinique actif comme les conotoxines α-Vc1.1, χ-MrIA, contulakin-G, conantokin-G pour le traitement de la douleur, la chlorotoxin pour l'imagerie tumorale et la toxine ShK pour les maladies auto-immunes pour n'en citer que quelques-unes.

Tableau 3 : Médicaments approuvés dérivés des venins.

Nom	Source	Description de la séquence	Cible pharmacologique	Traitement
Captopril	Crotale (*Bothrops jaracaca*)	Dérivé	Enzyme de conversion de l'angiotensine	Hypertension
Eptifibatide	Vipère des pyramides (*Echis carinatus*)	Dérivé	Récepteur αIIbβ3	Infarctus du myocarde
Tirofiban	Crotale (*Sistrusus miliarius barbouiri*)	Dérivé	Récepteur αIIbβ3	Infarctus du myocarde
Lepirudin	Sangsue (*Hirudo medicinalis*)	Dérivé	Thrombine	Thrombopénie induite par l'héparine
Bivalirudin	Sangsue (*Hirudo medicinalis*)	Dérivé	Thrombine	Thrombopénie induite par l'héparine
Ziconotide	Cône (*Conus magus*)	Naturelle	Canaux calciques de type N dépendant du voltage ($Ca_v2.2$)	Douleur non soulagée par la morphine
Exenatide	Monstre de Gila (*Heloderma suspectum*)	Naturelle	Récepteur du glucagon-like peptide-1 (GLP1R)	Diabète de type 2
Linaclotide	Bactérie (*Escherichia coli*)	Dérivé	Récepteur 2C du gros intestin	Syndrome de Bowel
Plecanatide	Bactérie (*Escherichia coli*)	Dérivé	Récepteur 2C du gros intestin	Syndrome de Bowel

La plupart des programmes de découverte de candidats médicaments commence par une forme de criblage d'une librairie de composés chimiques, dont la plupart contiennent aussi des produits naturels. Le raisonnement sous-jacent étant que la nature a développé une véritable abondance de

composés offrant des possibilités quasiment infinies de découverte de nouveaux candidats médicaments. Comme mentionné précédemment, la plupart des venins sont composés d'un mélange complexe constitué de peptides, protéines et enzymes. Etant donné les avancées technologiques du siècle dernier il est devenu de plus en plus aisé d'analyser les venins et de séparer et caractériser les protéines et les peptides les composant individuellement. Dans les années 80 les méthodes classiques de chromatographie ont été remplacées par des méthodes de chromatographies à haute pression (HPLC) pour obtenir des fractions pures qui étaient ensuite séquencées par dégradation d'Edman. Récemment, les développements parallèles en spectrométrie de masse [43], en séquençage nouvelle génération [44] et en criblage haut débit miniaturisé [45] ont largement contribué au développement de la recherche sur les venins. De par leur nature non ciblée les approches transcriptomiques et protéomiques fournissent une vue d'ensemble sur les composants des venins. L'intégration de ces approches omiques pour l'étude des venins a été nommée « venomique » [46].

i. La spectrométrie de masse au service de la protéomique

La protéomique de façon générale traite de l'identification à grande échelle des produits des gènes (peptides et protéines), mais ce domaine regroupe diverses disciplines telles que les techniques séparatives, la spectrométrie de masse et les outils de bioinformatiques qui contribuent aux études protéomiques. En effet, la spectrométrie de masse est graduellement devenue la méthode de choix pour l'analyse d'échantillons complexes de protéines. Ainsi, la protéomique basée sur les données de spectrométrie de masse est une discipline rendue possible par la disponibilité de bases de données de séquences de gènes, génomes et par les avancées techniques et conceptuelles dans de nombreux domaines, notamment la découverte et le développement de méthodes d'ionisations adaptées aux protéines, comme en témoigne le prix Nobel de chimie en 2002.

En spectrométrie de masse les mesures de masses se déroulent en phase gazeuse sur des analytes ionisés. Par définition un spectromètre de masse est constitué d'une source d'ions, d'un analyseur qui mesure le rapport masse sur charge (*m/z*) des analytes ionisés et un détecteur qui enregistre le nombre d'ions à chaque valeur de *m/z*. L'ionisation par electrospray (ESI) et la désorption/ionisation laser assistée par matrice (MALDI) sont les deux techniques les plus couramment utilisées pour volatiliser et ioniser les protéines et les peptides pour l'analyse par spectrométrie de masse [47,48]. La source ESI ionise les analytes provenant d'une solution aqueuse et peut être couplée à un outil de séparation en phase liquide (chromatographie ou électrophorèse). La source MALDI ionise les analytes co-cristallisés avec une matrice via des impulsions laser sur ce dépôt solide. La spectrométrie de masse MALDI est normalement utilisée pour l'analyse d'échantillons peu complexes ne nécessitant pas de séparation préalable alors que les systèmes de chromatographie liquide couplé à la spectrométrie de masse sont

préférés pour l'analyse de mélanges complexes. Il y a deux types d'analyseur de masse actuellement utilisés en recherche protéomique : les analyseurs de moyenne/haute résolution (le piège à ions (IT) et le temps de vol (TOF)), et les analyseurs de ultra haute résolution à Transformée de Fourrier (FT-MS) tels que l'ICR (Résonance cyclotronique ionique) et l'Orbitrap. Ils sont différents en termes de performances, chacun possédant ses propres avantages et inconvénients et ils doivent être choisis de façon adéquate en fonction de la stratégie analytique et de la précision de mesure de masse souhaitée. Des expériences de spectrométrie de masse en tandem (MS/MS) sont conduites avec des activations vibrationnelles ou électroniques selon les configurations d'analyseurs utilisés afin de caractériser les structures détectées. Celles-ci consistent en l'isolation d'un ion moléculaire, sa fragmentation par collision énergétique avec un gaz (CID) ou par transfert de charge (ECD/ETD) et l'enregistrement du spectre de masse MS/MS.

En effet, il y deux types de stratégies en analyse protéomique dénommées « bottom-up » et « top-down », la première étant la plus couramment utilisée. La spectrométrie de masse des protéines intactes qui fait l'objet des études « top-down » étant beaucoup plus complexe que la spectrométrie de masse des peptides, la stratégie « bottom-up » est basée sur l'analyse de peptides protéolytiques. La stratégie « shotgun », consitant à l'analyse d'un mélange de peptides, est la plus usitée actuellement en protéomique « bottom-up ». Ainsi, une expérience typique de protéomique « shotgun » décrite en Figure 5 comprend quatre étapes : 1) l'échantillon est d'abord digéré par voie enzymatique, souvent avec la trypsine conduisant à l'obtention d'un résidu basique protoné en position C-terminale favorisant le séquençage peptidique. 2) Les peptides protéolytiques sont analysés par LC/MS/MS. L'ensemble des données spectrales obtenues constitue les informations expérimentales qui déterminent 3) la masse moléculaire des peptides (spectres MS) et 4) l'identification de portions de leurs séquences à partir des spectres MS/MS par confrontation aux données archivées (base de données protéiques de type SwissProt, NCBI, etc.) à l'aide d'outils bioinformatiques (SEQUEST, MASCOT, PEAKS, etc.).

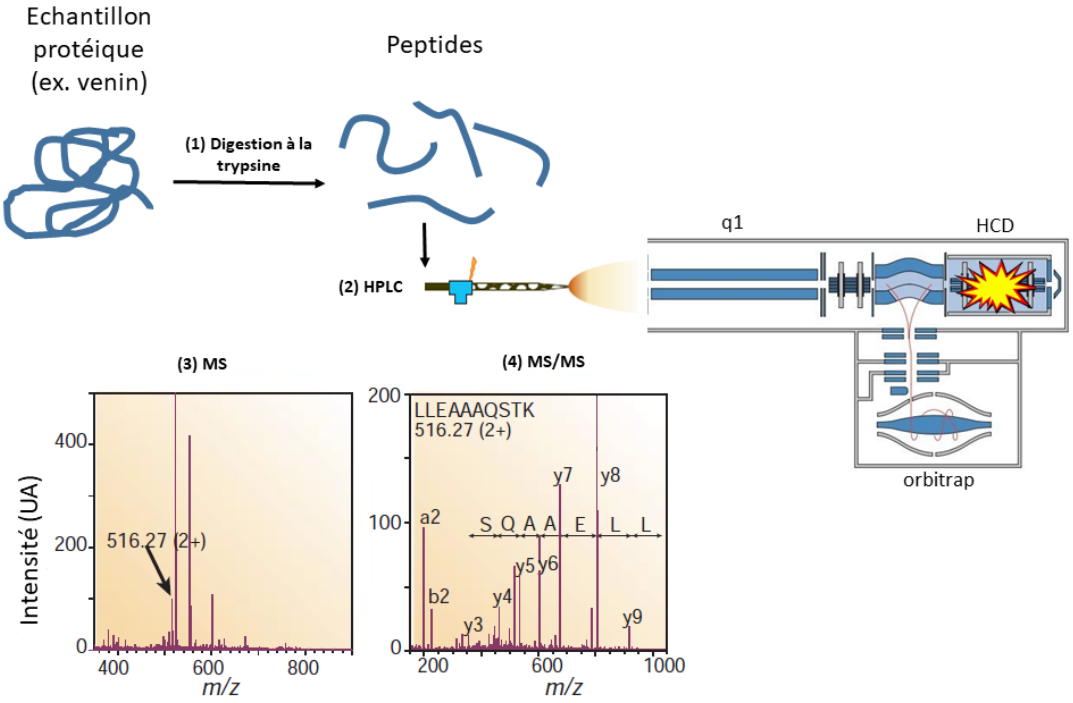


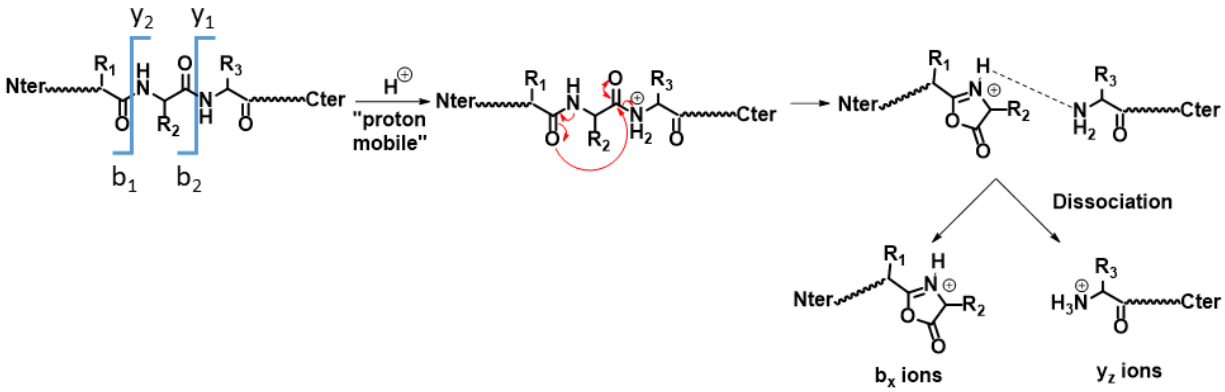
Figure 5: Représentation schématique des quatre étapes d'une expérience typique de protéomique « bottom-up » par spectrométrie de masse [49].

Les schémas de fragmentation de peptides observés en MS/MS dépendent de nombreux paramètres comme la composition en acide aminé et la taille du peptide, la méthode d'activation, l'échelle de temps de l'instrument, l'état de charge de l'ion, etc. Les ions précurseurs peptidiques dissociés dans les conditions de collision à basse énergie les plus habituelles se fragmentent le long de la chaine peptidique au niveau des liaisons amides formant des ions portant les informations structurales (appelés ions de séquences) et d'autre ions moins utiles non porteurs d'informations structurales par perte de molécules neutres comme l'eau, l'ammoniac, etc. Les ions de séquences sont les ions b et les ions y, qui contiennent les extrémités N- et C-terminale respectivement (Schéma 1) [50–52]. Les résidus d'acide aminés peuvent être déterminés à partir de la différence de masse entre deux ions de même type.

Le modèle le plus complet actuellement disponible pour décrire la manière dont les peptides protonés se dissocient est appelé « modèle du proton mobile », dans lequel le proton initialement localisé sur le site le plus basique de la molécule après activation par gain d'énergie interne, peut être transféré sur les liaisons peptidiques déclenchant ainsi la fragmentation. Ce modèle a été formalisé par Bela Paizs à la suite d'un grand nombre d'études réalisés par Wysocki, Harrison, Gaskell, Boyd et d'autres [53]. Les peptides protonés activés par des techniques de collision basse énergie se fragmentent principalement par des réactions dirigées par la charge. La dissociation des peptides protonés peut

être décrite comme une compétition entre plusieurs voies de fragmentations (PFP) dirigées par la charge dans un schéma de réactions complexes où les ions fragments sont formés selon différentes probabilités. Il existe de nombreuses PFPs, cependant seule la voie b_x-y_z donnant des ions de séquences permettant de déterminer la composition en acide aminé qui est nécessaire à la compréhension du manuscrit sera détaillée ici (Schéma 1).

Schéma 1 : Voie de fragmentation bx-yz des peptides protonés.



Il semblerait, présenté ainsi, que la fragmentation des peptides protonés permettrait le séquençage *de novo* de n'importe quelle séquence peptidique par interprétation des différences de masses entre les ions d'une même série (b ou y). Cependant, plusieurs problèmes s'y opposent, comme les mauvais taux de recouvrement de séquences dus à (i) l'inefficacité de fragmentation en CID pour les peptides longs et (ii) aux phénomènes de voies de fragmentations préférentielles. En effet, plus les peptides possèdent une taille importante plus ils vont être capables d'absorber une quantité importante d'énergie sans pour autant se fragmenter sur chaque liaison peptidique. Les phénomènes de voies de fragmentations préférentielles sont directement liés à la nature des acides aminés composants la chaine peptidique et ont pour effet de diminuer considérablement la quantité des autres ions de séquence y et b. L'effet le plus étudié est celui du résidu proline, qui donne lieu à d'abondants ions y contenant la proline en position N-terminale [53]. Le clivage b_x-y_z en C-ter des acides aminés tels qu'histidine, glutamine, asparagine, lysine, et arginine par attaque nucléophile de la chaine latérale est également favorisé [53]. La séquestration de protons notamment pour les séquences contenant un acide aspartique donne lieu préférentiellement à des fragmentations par élimination de charge conduisant à des ions b contenant l'acide aspartique en position C-terminale [54]. Enfin le dernier obstacle étant les acides aminés isobariques leucine et isoleucine. Bien que des méthodes de fragmentation à haute énergie telle que l'isolation d'ions z produit par dissociation induite par transfert d'électron (ETD) et leur subséquente fragmentation par collision à haute énergie

(HCD) produisant des ions w différents pour la leucine et l'isoleucine existent [55], ces méthodes en sont encore au stade de la recherche et ne sont pas intégrées en routine dans les analyses protéomiques.

Pour faciliter le séquençage des protéines d'un échantillon par spectrométrie de masse, celui-ci est donc assisté par des bases de données. Il existe deux types d'approches pour intégrer les bases de données dans l'identification de séquence par spectrométrie de masse : la première consiste à faire correspondre les spectres MS/MS expérimentaux à des spectres théoriques générés par digestion *in silico* des protéines et des traductions d'acide nucléiques des bases de données existantes (Uniprot, NCBI etc.). La deuxième approche consiste à effectuer une premiere étape d'identification *de novo* (sans l'appui de la base de données) afin de sélectionner des candidats potentiels pour chaque spectres MS/MS et ensuite de corriger (ou non) les candidats *de novo* avec l'utilisation de la base de données, cette approche est notamment employée par le logiciel PEAKS Studio. L'avantage de l'approche numéro 1 qui est utilisée par le logiciel Mascot par exemple est que l'identification ne requiert pas de base de données personnelles en revanche les protéines identifiées vont forcément être des protéines déjà référencées. L'approche numéro 2 permet quant à elle l'utilisation d'une base de données personnelle, ce qui est un réel avantage pour la découverte de séquences encore inconnues. Cette particularité constitue la raison pour laquelle nous l'avons utilisée, en plus de permettre la découverte de modifications post traductionnelles (PTMs) et d'autres modifications (substitutions, délétions, insertions etc.). Ainsi, dans l'approche intégrée en venomique, la transcriptomique va servir de base de données à l'analyse protéomique, on parle alors de protéo-transcriptomique (Figure 6).

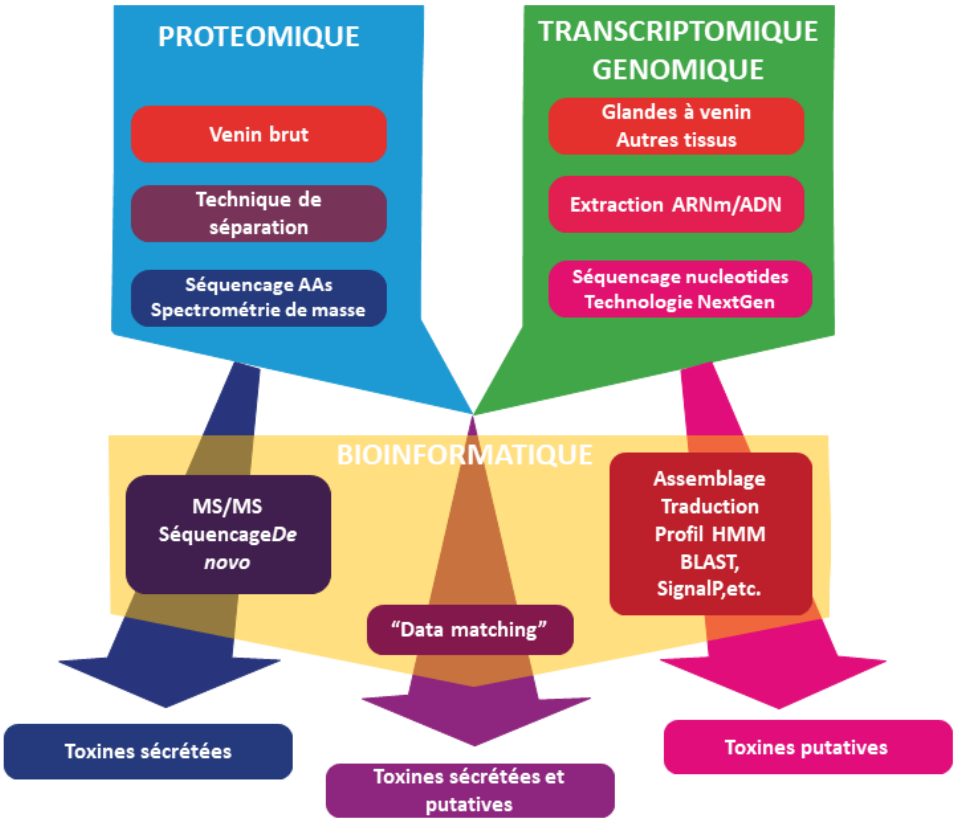


Figure 6 : Flux de travail schématique des études venomiques [41].

ii. La transcriptomique au service de la protéomique

Les premiers projets de séquençage d'ADN ont été réalisés sur un instrument de première génération basé sur la technologie Sanger. Cette technique, encore utilisée dans des études récentes, est basée sur des didésoxynucléotides, qui sont incorporés à la place des désoxyribonucléotides naturels et stoppent la progression de l'extension de la chaîne d'ADN. Les fragments sont ensuite séparés après séparation sur gels 2D et analysés [56,57]. Bien qu'extrêmement précise, largement validée et caractérisée par un faible taux d'erreur, cette approche est extrêmement coûteuse et laborieuse. Le séquençage de seconde génération ou de nouvelle génération (NGS) diffère de celui de Sanger car il n'utilise pas de nucléotides modifiés. Cette technologie repose tout d'abord sur la mesure de la synthèse du pyrophosphate : le pyrophosphate est converti en ATP par l'ATP sulfurylase, puis utilisé comme substrat pour la luciférase, produisant un rayonnement lumineux proportionnel à la quantité de pyrophosphate générée. Cette réaction a été exploitée pour le développement du premier séquenceur à haut débit : 454 GS FLX (Roche). Ce succès a été suivi par l'apparition de séquenceurs de nouvelle génération qui ont considérablement réduit les coûts, augmenté le débit et l'attractivité du séquençage de l'ADN [56]. Cela a abouti à l'émergence d'ambitieux projets de séquençage de génome entier, révolutionnant les domaines de la médecine, de la biologie, de l'écologie et de la taxonomie, avec un impact crucial sur la venomique. En dépit de ces avancées, l'absence de génomes séquencés

chez les animaux venimeux, ainsi que l'évolution extrêmement rapide de l'expression génétique des glandes à venin, demeure un défi majeur pour la venomique. De plus, l'approche classique du clonage de gènes, reposant sur des amorces (primers) spécifiques, bien qu'utile pour l'identification de toxines définies, ne convient pas au profilage global de séquences exprimées dans la glande à venin. L'avènement et la rentabilité de la technologie NGS actuelle ont ainsi offert aux scientifiques impliqués dans la venomique une possibilité inégalée de séquencer la totalité des ARNm exprimés dans la glande à venin, ce qui a été réalisé pour la première fois il y a dix ans. Cette technique, connue sous le nom de RNA-Seq, permet d'obtenir un profil transcriptome complet de la glande à venin des espèces étudiées avec un investissement en temps et en coûts raisonnables [58,59]. Dorénavant, des instruments tels que ceux developpés par Illumina (HiSeq 2500 Illumina) et Thermo Fisher Scientific (PGM) permettent le séquençage des transcriptomes des glandes à venin de façon routinière et sont de plus en plus décrits dans les publications sur la venomique [60,61]. La première étape cruciale après l'extraction de l'ARNm est la construction d'une bibliothèque d'ADNc (ADN cyclique). L'ADNc simple brin est synthétisé par la transcriptase inverse (RT) et ensuite amplifié par PCR. Une étape de fragmentation de la bibliothèque d'ADNc est nécessaire pour obtenir la longueur optimale pour le séquençage et essentielle pour obtenir des séquences de bonne qualité avec une bonne couverture [62]. La qualité et la fiabilité des séquences altérées par les lectures de séquençage courtes et le taux d'erreur relativement élevé (par rapport à la méthode Sanger) dans la détermination des bases caractérisent les NGS. Des délétions et des substitutions de nucléotides sont en particulier relativement fréquentes [63]. Ces faiblesses sont en grande partie compensées par l'utilisation de la bioinformatique, qui permet un assemblage et une analyse fiables des données. Premièrement, le filtrage et l'assemblage en contigs (séquence génomique continue et ordonnée générée par l'assemblage des clones d'une bibliothèque génomique), réalisés avec des logiciels dédiés tels que MIRA ou Trinity, nécessitent des paramètres bien définis et des critères de validation rigoureux, afin d'éviter les erreurs d'assemblage et la génération d'artefacts. Les contigs sont ensuite annotés et classés à l'aide d'algorithmes dédiés, suivant leur fonction : composants cellulaires, processus biologiques ou toxines présumées [64–66]. Jusqu'à récemment, des outils tels que BLAST [67] étaient le standard de référence pour aligner structurellement et grouper des séquences. Cependant, de nos jours, l'identification et l'annotation putatives des toxines sont fréquemment effectuées à l'aide de méthodes basées sur des profils, telles que les modèles de Markov cachés (HMM) ou la « position-specific scoring matrix » (PSSM). Ces méthodologies permettent l'identification de superfamilles géniques très divergentes et/ou non décrites [68–70]. Des algorithmes tels que SignalP ou SpiderP, spécifiquement conçus pour les toxines d'araignées, sont appliqués afin de déterminer les peptides signaux et les séquences matures (Figure 7) [71,72]. La RNA-seq génère des séquences de tous les peptides et les protéines exprimés des glandes à venin, y compris les peptides minoritaires ou les variants peu communs. Les publications décrivant

les données transcriptomiques et leur analyse sont de plus en plus fréquentes dans le domaine de la venomique, mais l'approche intégrative, couplant les études protéomiques et transcriptomiques, reste relativement nouvelle. Des logiciels dédiés au couplage des données protéomiques et transcriptomiques, tels que PEAKS, ProteinPilot ou Tandem [73,74] sont aujourd'hui largement utilisés, permettant une prévision et une annotation des PTMs fiables, une détermination des scores et une visualisation des résultats [68,75–79]. Cependant, l'identification des toxines, facilitée par ces stratégies combinant protéomique, génomique et transcriptomique, a posé un défi particulier : la mise au point d'un procédé rationnel pour nommer et classer les toxines et les toxines putatives afin d'éviter les informations redondantes, incomplètes ou incorrectes accumulées dans les publications ou les bases de données. Ce sujet a été abordé dans plusieurs publications, décrivant une nomenclature systématique pour répertorier de manière rationnelle les peptides et protéines nouvellement séquencés en fonction de leur source (espèce animale) et de leur activité (cibles moléculaires) [80–83]. Brièvement, la première lettre (grecque) indique la cible pharmacologique de la toxine. La ou les lettres suivantes (romain, première lettre en majuscule) dérivent du nom de l'espèce à partir duquel la toxine a été isolée. Cette lettre (ou ces lettres en cas d'ambiguïté) est suivie d'un chiffre romain, qui fournit des informations sur la répartition des cystéines. Enfin, une lettre majuscule indique l'ordre de découverte de la toxine dans cette catégorie. Ces « lignes directrices » aident les chercheurs à interroger les bases de données pour les annotations, les fonctions et l'origine des séquences.

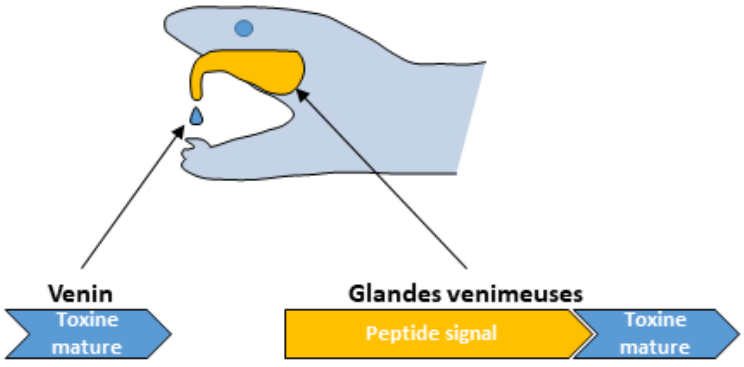


Figure 7 : Différence entre peptide précurseur et peptide mature.

c. La Synthèse Peptidique en Phase Solide (SPPS), un puissant outil pour la production de peptides

La synthèse de peptide peut être accomplie par voie recombinante c'est-à-dire en modifiant le code génétique d'une bactérie par exemple pour le programmer à la production du produit désiré. Cette méthode présente l'avantage de proposer un accès rapide à des peptides longs (>50 résidus). Cependant elle présente aussi de sévères limitations [84], en effet l'incorporation des PTMs (fréquemment observées dans les peptides venins) et des acides aminés non naturels constitue un réel défi et les difficultés d'isolation du produit désiré avec un haut degré de pureté demeurent. Beaucoup de laboratoires ont opté pour l'alternative de la synthèse chimique, cette approche permet de s'affranchir de la contamination indésirable par les protéines (hôte) souvent rencontrée dans la production de protéines par voie recombinante et donne aux chercheurs plus de contrôle et de liberté quant aux modifications. La synthèse chimique est rapide, hautement automatisée et adaptée à la montée en échelle de production, fournissant des quantités significatives de peptides natifs et modifiés de haute pureté. Elle permet l'incorporation d'acides aminés non naturels, d'une large gamme de PTM, de groupements pour l'imagerie et de modifications structurelles telles que la cyclisation du squelette peptidique ou le remplacement des ponts disulfures améliorant la stabilité contre la dégradation enzymatique.

Le travail pionnier de Bruce Merrifield [85] qui a introduit la synthèse peptidique en phase solide (SPPS) a dramatiquement changé la stratégie de synthèse des peptides en simplifiant les étapes de lavage et de purification associés à la synthèse en solution. De plus, l'implémentation de la méthode dans l'instrumentation robotisée a permis le développement de la synthèse peptidique automatisée qui est dorénavant la méthode de choix pour la production de peptides thérapeutiques bien que la synthèse en solution soit toujours utilisée pour de la production à grande échelle de petits peptides.

Le principe est simple (Figure 8) : le peptide est ancré sur une bille de résine par l'intermédiaire d'un bras clivable (linker), cet ensemble pouvant être considéré comme le groupement protecteur de la fonction acide C-terminale par analogie avec la synthèse en solution. Cette matrice polymérique est insoluble dans les solvants et réactifs utilisés lors de la synthèse. Elle est aussi importante que le solvant pour la synthèse en solution, et est classiquement constituée de polystyrène (PS) réticulé avec un autre composé chimique qui va moduler les propriétés du polystyrène. En effet les matrices polymériques hydrophobes constitué de PS réticulé à 1% de divinylbenzène vont avoir des bonnes propriétés de gonflement et donc une bonne diffusion des réactifs dans les solvants organiques aprotiques peu polaire (DCM, Toluène, THF). En revanche cela va avoir pour effet d'entraîner l'agrégation des chaines peptidiques et donc de rendre la synthèse de peptide plus « longs » (15 résidus) difficile. Pour la synthèse de ces derniers il est recommandé d'utiliser des matrices polymériques réticulées à 20 ou

40% par du polyethyleneglycol (PEG) ou même composé entièrement de PEG, celles-ci étant plus hydrophiles cela va permettre d'espacer (par rapport à une matrice hydrophobe) les chaines peptidiques et donc de limiter les phénomènes d'agrégation. Un cycle de synthèse peptidique comprend trois étapes : 1) ancrage du premier acide aminé C-terminal sur la résine. 2) assemblage de la séquence peptidique de façon linéaire de l'extrémité C-terminale vers N-terminale par cycle répété de N^{α} déprotection et réaction de couplage d'acide aminé. Le groupement amine α est protégé par une protection temporaire (généralement un dérivé uréthane) qui est facilement déprotégé en conditions douces préservant l'intégrité du peptide et réduisant le taux de racémisation. Les groupements fonctionnels des chaines latérales quant à eux sont protégés par des groupement protecteurs permanents pendant les étapes d'élongation de la chaine peptidique 3) Lorsque la séquence désirée a été assemblée le peptide est détaché de la résine et les protections des chaines latérales sont éliminées lors de l'étape de clivage [86].

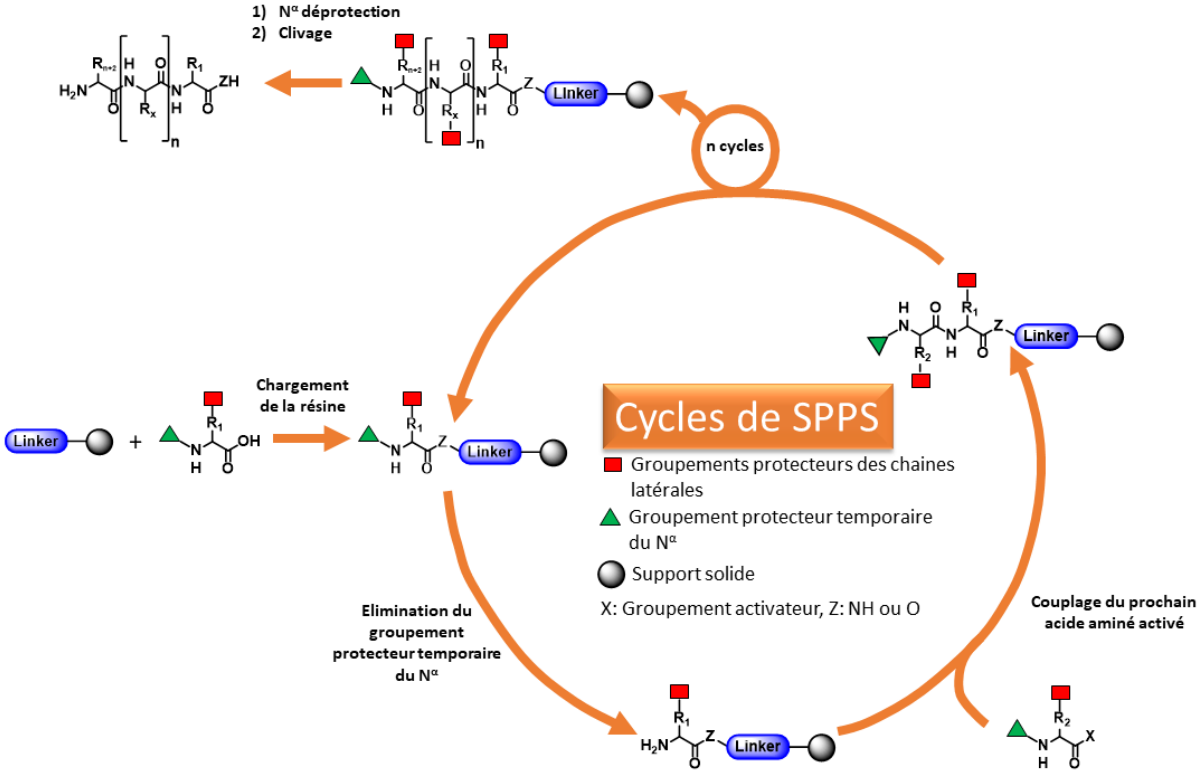


Figure 8 : Principe général de la SPPS.

Deux stratégies de SPPS existent : la Boc/Bzl et la Fmoc/*t*Bu. La stratégie Boc/Bzl utilise le groupement Boc (*tert*-butyloxycarbonyl) comme protection temporaire labile en présence d'acide trifluoroacétique (TFA) et l'étape de clivage se fait en présence d'acide fluorhydrique (HF). Bien que cette méthode permette la synthèse efficace de peptides longs voire de petites protéines, l'utilisation contraignante et dangereuse de HF font que cette technique a été progressivement supplantée par la stratégie Fmoc/*t*Bu moins dangereuse et plus facile à mettre en œuvre. En effet cette technique utilise le

groupement Fmoc (Fluorenylmethyloxycarbonyl) comme protection temporaire labile en condition basique et l'étape de clivage se fait en présence de TFA [87]. L'avantage de l'utilisation de conditions orthogonales pour le clivage et pour les déprotection N-α permet l'utilisation de conditions plus douces pour la séparation du peptide de la résine. La majorité des peptides synthétiques sont préparés par Fmoc-SPPS, le succès de cette stratégie est dû à son adoption rapide par les non-chimistes, comme les biologistes réalisant qu'ils pouvaient préparer rapidement des peptides adaptés à la production d'anticorps en utilisant des machines peu coûteuses et éviter l'utilisation de HF anhydre [88]. La Fmoc-SPPS a été plus facile à automatiser car il n'y a nul besoin d'acide trifluoroacétique corrosif lors des cycles de synthèses et parce que la déprotection du groupement Fmoc libère un groupement fluorene avec de fortes propriétés absorption ultra-violet (UV) fournissant un bon indicateur de la réussite de la synthèse [89]. Pour les chimistes, la stratégie Fmoc a permis de surpasser certaines conditions limitantes de la chimie Boc. En effet les conditions de déprotection sont compatibles avec les peptides modifiés comme les peptides phosphorylés, glycosylés et pour les chimiothèques de peptides [90]. Il y a eu des avancées considérables dans la longueur des peptides synthétisés. C'est en partie dû à l'amélioration de la pureté des acides aminés (« building blocks ») Fmoc qui sont commercialement disponibles avec une pureté >99% déterminée par chromatographie liquide de phase inverse à haute pression (RP-HPLC). Cependant, cela ne garantit pas l'absence de sels et le contenu peptidique en poids net est rarement indiqué. Malgré certains avantages de la stratégie Fmoc sur la stratégie Boc (listé précédemment), la stratégie Boc est plus adaptée à la synthèse de séquences difficiles notamment grâce à l'alternance des conditions acides (déprotection du N^{α}) et basiques (couplage) qui permet de briser l'agrégation des chaines peptidiques. Le succès de la stratégie Fmoc est donc aussi dû à l'application des pseudoprolines [91,92] et des protections de la chaine peptidique [93] permettant de s'affranchir des problèmes d'agrégations des séquences difficiles. Bien que la limite arbitraire de 50 acides aminés soit souvent citée dans les publications comme la longueur maximale de synthèse en routine, des séquences parfois bien plus courtes peuvent s'avérer extrêmement problématiques, et le succès d'une synthèse dépend plus de la séquence que de la taille du peptide.

d. Synthèse des peptides riches en pont disulfures

Les toxines peptidiques issues de venins sont souvent des peptides riches en pont disulfures ce qui leur confère une forte rigidité conformationelle et donc une stabilité et activité accrues [94]. La réussite d'une synthèse de ce type de peptide dépend de deux facteurs principaux : la suppression ou minimisation de la racémisation des résidus cystéines pendant l'élongation du peptide et l'obtention de l'appariement désiré des cystéines sous forme d'un pont disulfure. En effet bien souvent un mauvais appariement conduit à une structure tridimensionnelle différente (Figure 9) qui peut entraîner une activité réduite voire l'obtention d'un peptide totalement inactif [95–97] (bien qu'il existe des contre

exemples). Concernant le premier point, divers groupements protecteurs de la fonction thiol en position β accompagnés de procédures précises pour l'incorporation du résidu cystéine ont été étudiés et appliqués [98]. D'après la littérature les conditions de couplages pour l'incorporation des cystéines protégées (Tableau 4) (Xan, Tmob, Trt et Acm) avec un minimum de racémisation (<1% par étape) en Fmoc-SPPS sont : 1) BOP (ou HBTU ou HATU)/HOBt (ou HOAt)/TMP (4/4/4) sans préactivation dans DCM/DMF (1/1) 2) DIC/HOBt (ou HOAt) (4/4) avec 5 min de préactivation et 3) l'utilisation d'ester de pentafluorophenyl (Pfp) préformé dans DCM/DMF (1/1). Aux vues de cette étude bibliographique, les conditions retenues et utilisées pour les synthèses décrites dans ce manuscrit, alliant un minimum de racémisation et la disponibilité des réactifs présents au laboratoire sont les suivantes : HATU/HOBt/DIPEA (5/5/5) sans préactivation dans DCM/DMF (1/1). Le risque de racémisation du résidu cystéine est particulièrement présent lorsque la cystéine C-terminale est ancrée *via* une liaison ester aux résines hydroxymethyles car l'ancrage requiert des méthodes d'activation fortes qui favorise la racémisation [99]. Dans le cas d'une cystéine C-terminale est ancrée *via* une liaison ester il est recommandé d'utiliser une résine qui ne requiert pas d'activation comme la résine chloro-2-chlorotrityle [100] ou des méthodes d'activation utilisant une quantité catalytique de DMAP (4-Dimethylaminopyridine) [99]. Concernant le deuxième point le repliement (folding) oxydatif sous controle thermodynamique est la stratégie générale pour la synthèse de peptides riches en ponts disulfures [101–104]. Cette stratégie simple en comparaison à la stratégie régiosélective, consiste à oxyder sous contrôle thermodynamique les résidus cystéines. Cependant le folding est un processus compliqué impliquant des réactions d'oxydation (formation de S-S), de réduction (rupture de S-S) et d'isomérisation ou d'échange de pont disulfure (réarrangement de S-S, scrambling) [105].

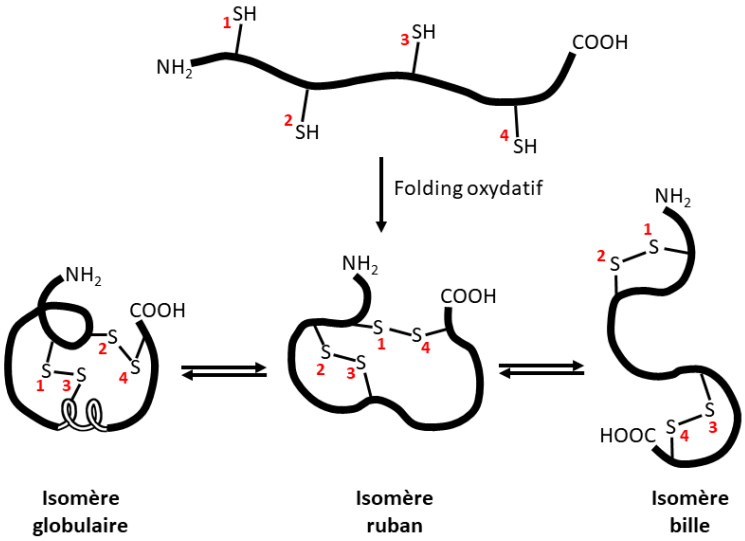
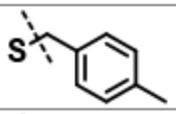
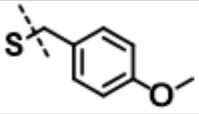
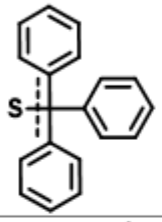
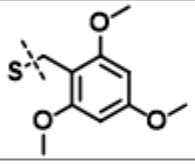
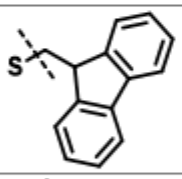
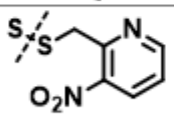
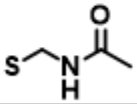
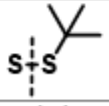
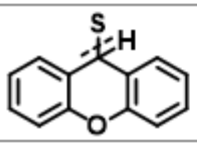


Figure 9 : Exemple des différents isomères de folding obtenu pour un peptide à deux ponts disulfures [82].

En fonction du nombre de ponts disulfures présents le procédé de folding augmente rapidement en complexité en raison du nombre croissant d'isomères possibles $(2n)!/(2^n n!)$ avec n étant le nombre de ponts disulfures formés. Par conséquent 2, 3 ou 4 ponts disulfures représentent 3, 15 ou 105 isomères possibles [106]. La réussite du folding dépend principalement de plusieurs facteurs : les informations de structure intrinsèquement encodées dans la séquence en acides aminés, la réactivité des groupements thiolates de la séquence, la stabilité thermodynamique du produit final et les conditions de folding utilisées [94,107]. En effet le peptide même complètement réduit (pas de S-S) de par sa séquence en acide aminés peut, en formant des interactions faibles intramoléculaires, se préstructurer de façon favorable au folding [105]. La stabilité thermodynamique du produit d'arrivée va faire évoluer au cours du temps les « mauvais » isomères de folding vers le « bon » isomère de folding par échange de ponts disulfures [105]. Un autre aspect important à prendre en compte également dans le folding de produits naturels est que le folding *in vivo* contrairement au folding par voie chimique peut-être assisté par des chaperonnes ou par les propeptides N- et C-terminaux pouvant jouer le rôle de chaperonnes intramoléculaires [108]. Ainsi la conformation active du peptide n'est pas forcément la plus stable thermodynamiquement et l'obtention de la conformation active via le folding en conditions oxydantes par voie chimique peut s'avérer difficile. Des conditions typiques de folding oxydatif sous contrôle thermodynamique sont : NH_4OAC (0.33 M), GnHCl (0.5 M), GSH/GSSG/peptide (100/10/1), pH 7.8, 4 °C [82]. Cependant il existe une myriade de conditions de folding alternatives, et il n'y a pas de règles générales pour une séquence donnée (certaines conditions peuvent mener au bon folding alors que d'autres non). Pour s'affranchir de ce problème, les chimistes ont développé divers groupements protecteurs et approches sophistiquées pour l'appariement régioselectif des ponts disulfures. L'approche régiosélective permet la formation dirigée des ponts disulfures par paire de cystéines. De façon génerale la stratégie est basée sur une déprotection sélective et l'oxydation 2 à 2 des résidus cystéines. Comme plusieurs ponts disulfures sont formés étapes par étapes, des conditions de réaction qui empêche (ou limite) la rupture ou le scrambling des liaisons disulfures déjà existantes sur la molécule, sont nécessaires. Par conséquent l'exposition à des conditions alcalines, des thiols ou autres nucléophiles et des temps de réaction trop longs qui pourraient conduire à l'échange des ponts disulfures sont à éviter. Les groupements protecteurs de la cystéine peuvent être regroupés en 4 catégories : labiles en présence de base, labiles en présence d'acide, labiles en présence d'ions métalliques et labiles en présence d'agents réducteurs (Tableau 4). La versatilité du groupement Acm utilisé en combinaison avec des groupements labiles en conditions acide en font l'un des groupements les plus utilisé en Fmoc- ou Boc-SPPS [109]. En effet celui-ci peut en présence de sels de mercure pour obtenir des groupements thiols ou, comme c'est souvent le cas pour la synthèse de peptides riches en ponts disulfures, être clivé de façon oxydante par l'iode ou le trifluoroacetate de thallium pour former une liaison disulfure avec un autre groupement Acm [109,110]. Les conditions de réactions pour le

clivage [111] et les étapes de déprotection doivent être choisies précautionneusement pour éviter les réactions secondaires telles que l'oxydation des résidus méthionines, histidines, tryptophanes ou tyrosines ainsi que le déplacement S→O du groupement Acm pour les peptides contenant des résidus serines et thréonines [112]. La formation régiosélective de ponts disulfures peut être effectuée soit pendant que le peptide est encore attaché au support solide, soit en solution après l'étape de clivage, chacune des deux approches à ses propres avantages et inconvénients.

Tableau 4 : Les différents groupements protecteurs de la cystéine classiquement utilisés. La ligne hachurée indique le site de clivage.

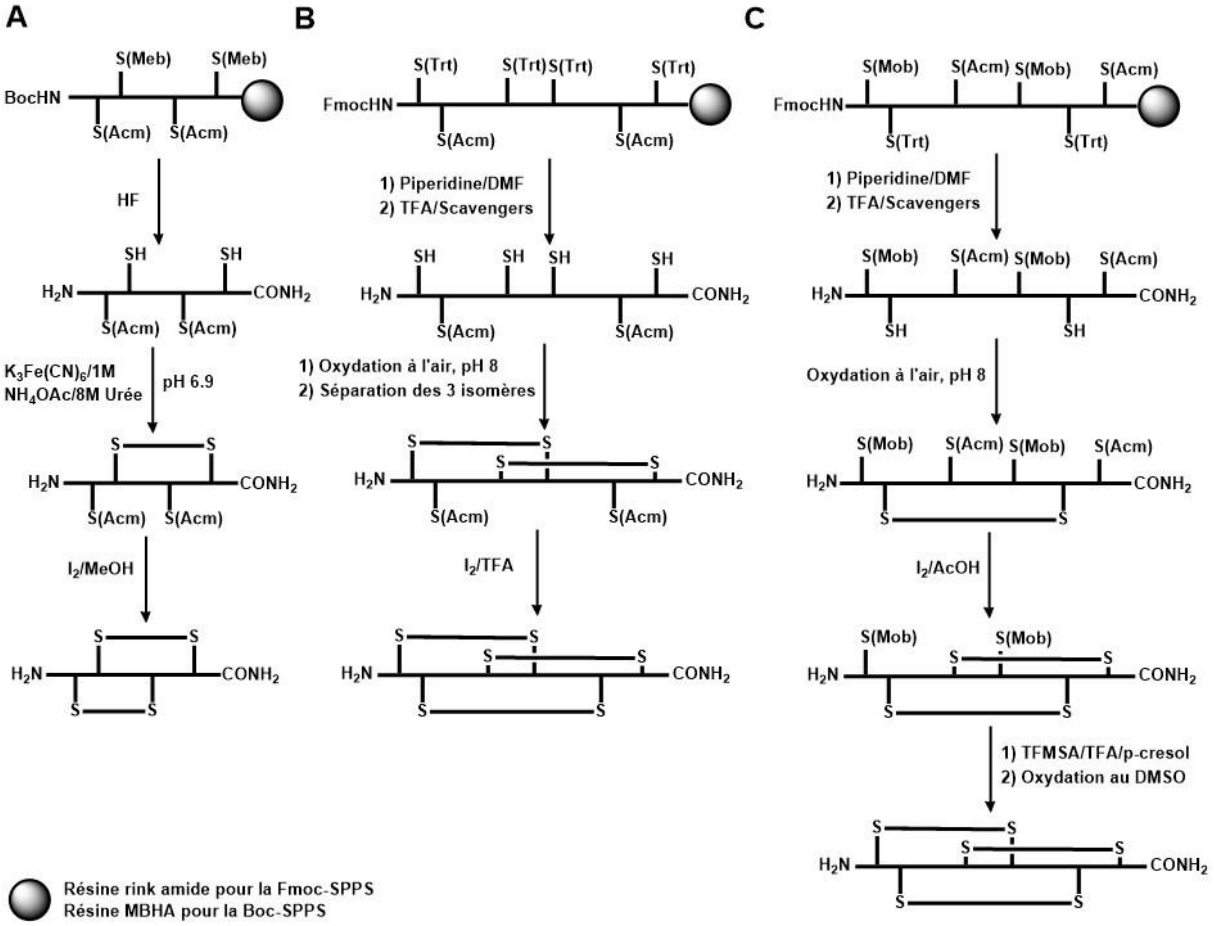
Chimie	Groupement protecteur	Structure	Stabilité	Conditions d'élimination
Synthèse génerale				
Boc (Fmoc)	4-Methylbenzyl (S-Meb)		Base TFA	HF 5%DMSO/TFA 60°C
Boc (Fmoc)	4-Methoxybenzyl (S-Mob)		Base TFA	HF TFMSA
Fmoc	Triphenylmethyl (S-Trt)		Base	1% TFA Tl(III) I_2
Fmoc	2,4,6-Trimethoxybenzyl (S-Tmob)		Base	7% TFA/scavengers
Synthèse orthogonale				
Boc	9-Fluorenylmethyl (S-Fm)		TFA HF	Base
Boc	3-Nitro-2-pyridylsulfenyl (S-Npys)		TFA HF	Agents réducteurs Thiols
Boc **Fmoc**	Acetamidomethyl (S-Acm)		Base TFA HF	I_2 Hg(II), Ag(II),Tl (II)
Boc **Fmoc**	*tert*-Butylsulfenyl (S-*S*tBu)		TFA HF (partiel)	Agents réducteurs Thiols
Fmoc **(Boc)**	*tert*-Butylmercapto (S-*t*Bu)		TFA Base	5% DMSO/TFA 25°C HF 20°C
Fmoc	9H-Xanthem-9-yl (S-Xan)		Base	1% TFA/scavengers I_2 Tl(III)

i. Les approches en solution

La stratégie la plus commune utilise le groupement acetamidométhyle (Acm) protecteur de la fonction thiol. Typiquement le peptide est clivé du support solide ainsi que toutes les protections des chaines

latérales à l'exception du groupement Acm. Les résidus cystéines libres sont oxydés en conditions douces pour former le premier pont disulfure et le second pont est généralement formé par traitement avec l'iode. Cette approche a été utilisée avec succès pour la synthèse de l'α-conotoxine GI et de ses isomères de folding par chimie Boc [113] (Schéma 2A) et a été appliquée par la suite à de nombreuses autres synthèses tant en chimie Boc qu'en chimie Fmoc [114] (Schéma 2). La stratégie de protection par le groupement Acm a pu aussi être appliquée à la synthèse de toxines à trois ponts disulfures en utilisant une stratégie semi sélective pour la synthèse de la conotoxine ω-MVIID (Schéma 2B) [115]. L'oxydation simultanée des quatre résidus cystéines ont conduit à la formation majoritaire d'un seul isomère qui a pu être isolé et engagé dans une deuxième étape de déprotection/oxydation concomitante en présence d'iode. Dans une approche plus sélective le groupement Mob a été introduit en plus du groupement Acm et du groupement Trt pour permettre la formation consécutive des trois ponts disulfures dans la synthèse en chimie Fmoc de la ω-MVIIA (Schéma 2C) [82], plus connue sous le nom de ziconotide et qui constitue à ce jour la seule conotoxine ayant reçu l'autorisation de mise sur le marché.

Schéma 2 : Exemples de stratégie régiosélective de folding en solution. (A) Stratégie de synthèse de la conotoxine α-GI [113]. (B) Stratégie de synthèse de la conotoxine ω-MVIID [115]. (C) Stratégie de synthèse de la conotoxine ω-MVIIA [82].



La synthèse « one-pot » de la conotoxine α-SI a été effectuée en utilisant une combinaison des groupements protecteurs Meb et *t*Bu pour laquelle la différence de stabilité des groupements protecteurs selon la température en condition d'oxydation DMSO/TFA/anisole a été exploitée [116]. Les résidus cystéines protégés avec le groupement *t*Bu sont rapidement déprotégés et convertis en pont disulfure à température ambiante (TA) alors les cystéines protégées Meb restent intactes. Le chauffage de la solution mène à la déprotection/oxydation pour former le deuxième pont disulfure (Schéma 3A) [116,117]. Cette approche a également été combinée à l'approche Acm/Trt pour aboutir à la synthèse sélective d'un dimère de la conotoxine α-SI composé de 4 ponts disulfures (Schéma 4) [118]. Cette approche d'oxydation double « one-pot » est le premier exemple utilisant 4 groupements protecteurs distincts pour la production d'analogues de toxines. Malgré l'utilité d'une telle procédure, elle reste limitée par la nature des acides aminés composant la séquence particulièrement les résidus tryptophanes et méthionines connus pour subir une oxydation irréversible s'ils ne sont pas protégés

dans ces conditions [119]. Comme cela a pu être mis en évidence lors de la synthèse de l'α-ImI en utilisant une procédure d'oxydation « one-pot » où aucune quantité détectable de produit n'a pu être récupérée [117]. De façon alternative, le groupement S*t*Bu a été utilisé en combinaison avec le groupement Acm pour la synthèse de l'α-conotoxine GI, le premier pont disulfure est formé après clivage en condition réductrice du groupement S*t*Bu par oxydation à l'air, s'ensuit une une oxydation à l'iode permettant la formation du deuxième pont (Schéma 3B) [120]. La différence de stabilité à l'iode du groupement Acm et *t*Bu a été explorée pour la synthèse de l'α-conotoxine GI où le premier pont disulfure a été obtenu par oxydation directe des groupements Acm suivi d'une oxydation des groupements *t*Bu via la procédure chlorosilane/sulfoxide (Schéma 3C) [121]. Bien que le premier pont ait été obtenu avec succès le traitement avec le système chlorosilane/sulfoxide a conduit à l'obtention d'un isomère non natif et donc non désiré, mais cependant l'isomère natif a pu être régénéré par échange de pont disulfure (shuffling) en condition tampon Tris en présence de cystéine.

Schéma 3 : (A) Stratégie sélective pour l'oxydation en solution de l'α-SI et α-GI en utilisant une stratégie "one-pot" [116,117]. (B) Folding stratégie de l'α-GI utilisant le couple S*t*Bu /Acm [120]. (C) Stratégie sélective non efficace pour la synthèse de l'α-GI [121].

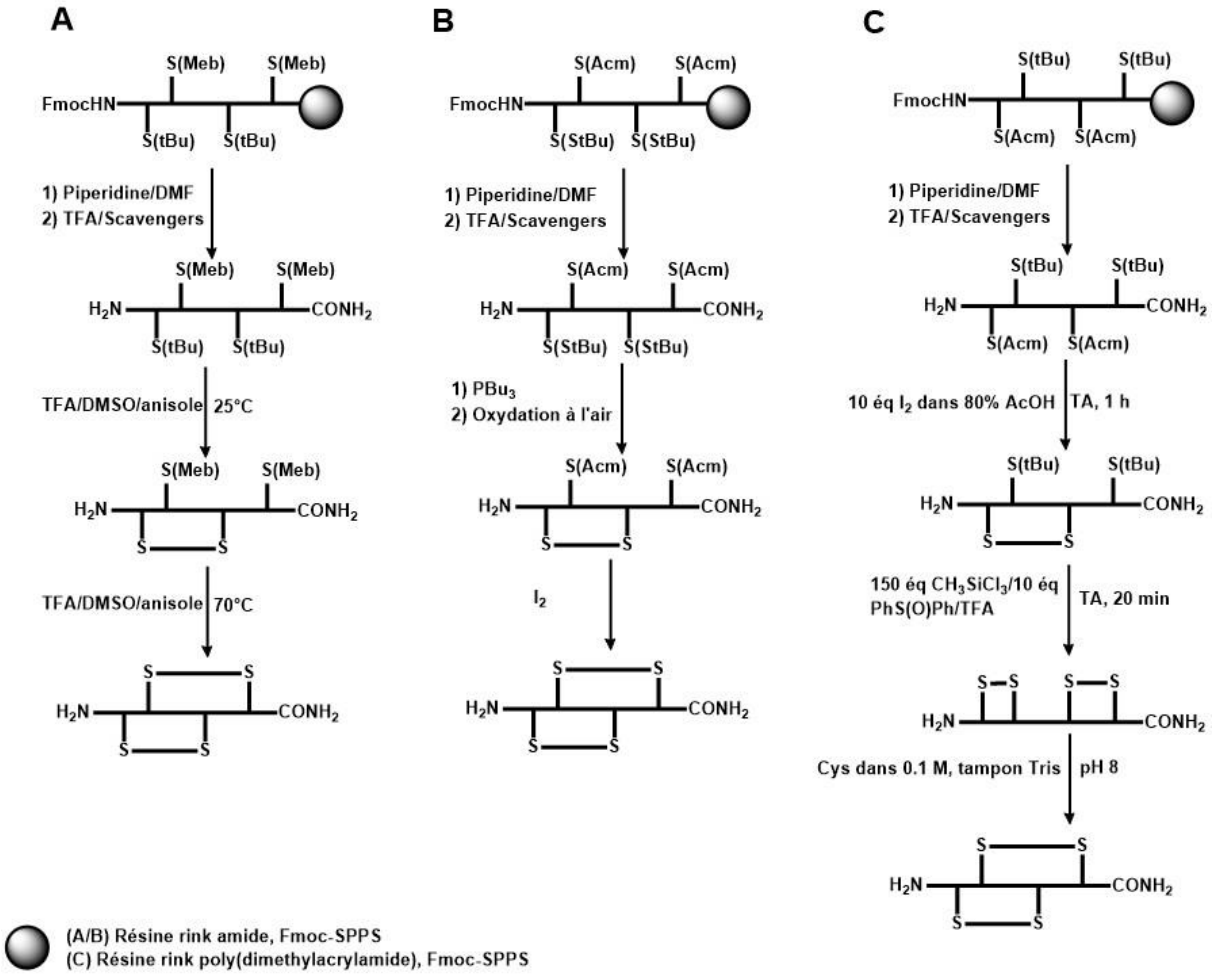
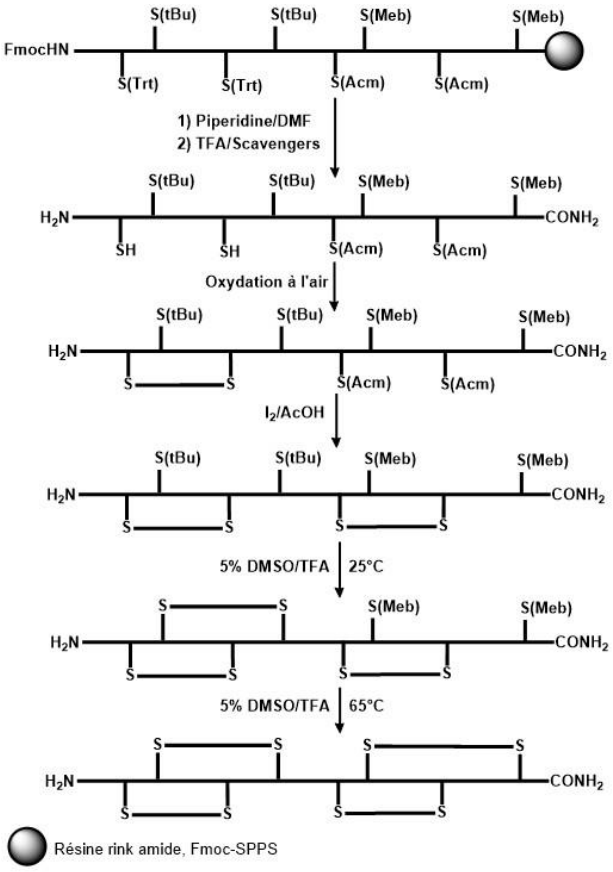


Schéma 4 : Stratégie de synthèse du dimère de α-SI [118].

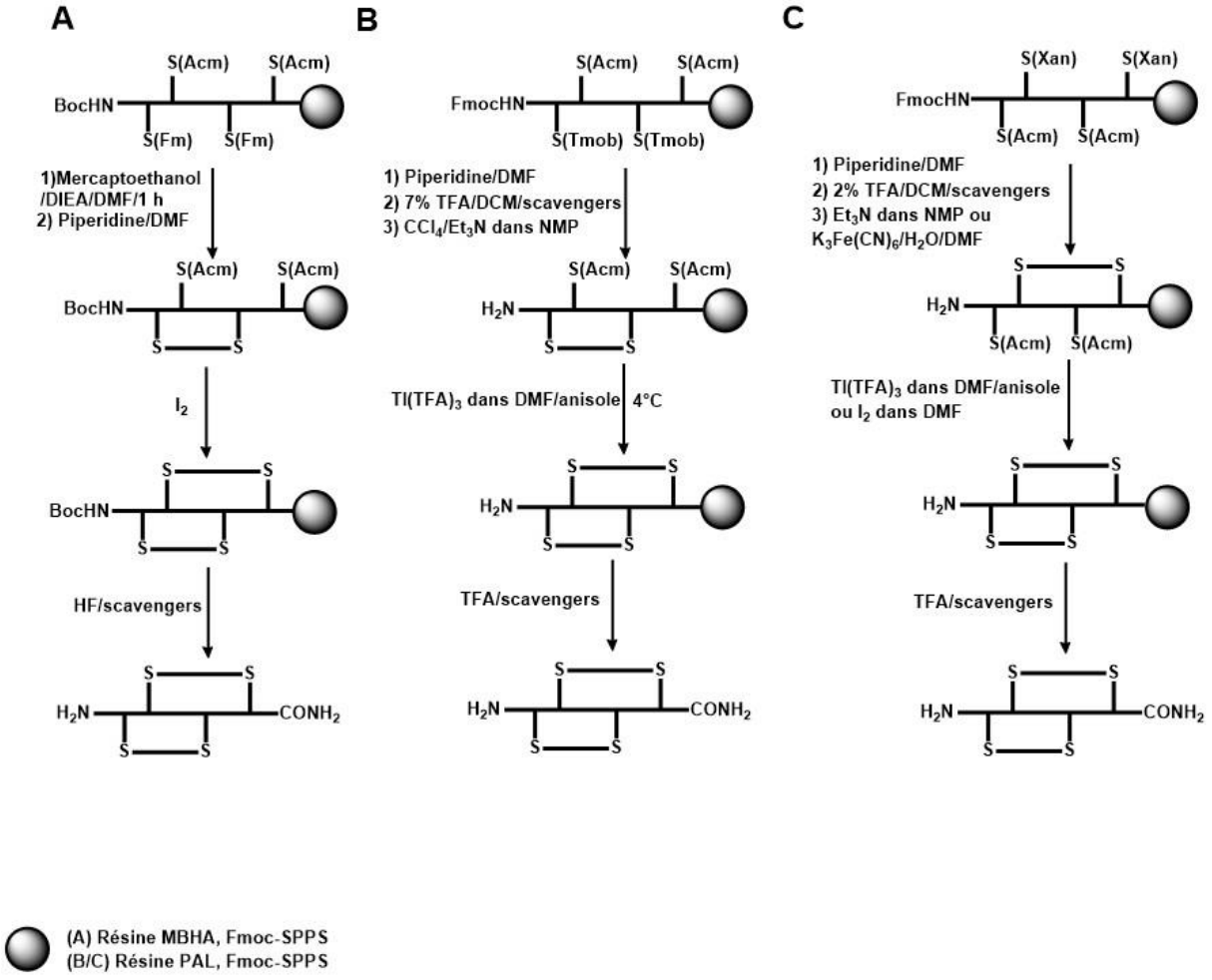


ii. Les approches sur résines

Le principe des stratégies sur résine consiste à former un ou plusieurs ponts disulfures alors que le peptide est toujours ancré au support solide. Les approches sur résine sont plus pratiques sur le plan opérationnel, évitent les étapes de purification fastidieuses et préviennent un grand excès de solvants d'oxydation en raison de l'effet de pseudodilution, qui simule une dilution élevée, favorisant ainsi la formation de liaisons disulfures de façon intramoléculaire. Dans la perspective de créer des librairies de peptides riches en ponts disulfures pour des études de relation structure-activité (SAR) ou de découverte de candidats médicaments, la formation de ponts disulfures sur résine devient pertinente. Les limitations actuelles des approches sur résine sont souvent un faible taux de récupération du produit souhaité dû à une polymérisation non voulue causée par la formation de ponts intermoléculaires (au lieu d'intramoléculaire) ou des interactions entre les ponts disulfures et la résine [122], à cela on peut ajouter le manque de conditions de folding uniformes ou de stratégies pour contrôler la formation multiple de ponts disulfures. Par exemple, dans la synthèse sur résine de l'α-SI, la formation de la boucle la plus petite avant la plus grande boucle était importante pour obtenir l'isomère souhaité, tandis qu'en solution, quel que soit l'ordre d'oxydation la synthèse était réussie [122]. En outre, la formation de ponts disulfures sur support solide peut souvent être lente en raison de l'encombrement stérique causé par les groupements protecteurs de la chaîne latérale, et les

rendements de récupération sont souvent inférieurs à ceux des méthodes en solution [122,123]. La formation de ponts disulfures sur résine assistée par micro-ondes peut améliorer les rendements de folding. Le Schéma 3 décrit trois exemples de stratégies orthogonales d'oxydation sur résine qui ont été appliquées sur des conotoxines, dans tous les cas le folding en solution était plus efficace en termes de rendement et de taux de récupération de l'isomère désiré. La combinaison du groupement Fm labile en condition basique avec le groupement Acm a été utilisée en chimie Boc pour la synthèse de l'α-GI (Schéma 5A) [82]. Dans cette approche, le groupement Fm a été retiré et oxydé sur résine à l'aide de la pipéridine et la formation du deuxième est réalisée par déprotection/oxydation des fonctions Acm. Bien que des études antérieures aient suggéré que les liaisons disulfures sont instables dans des conditions très acides, une sélection minutieuse de piégeurs (scavengers) pour le clivage de HF a permis l'obtention de l'isomère intact. Le laboratoire de Barany (United states, Minneapolis) a étudié différentes méthodologies orthogonales pour la synthèse de l'α-SI par chimie Fmoc (Schéma 5B, C) [122,124]. La méthodologie initiale consistait à déprotéger les groupements Tmob sans perte de peptide sur la résine puis de former ensuite le pont entre les cystéines libres par oxydation avec CCl_4-Et_3N dans NMP, avec la deuxième paire de cystéine toujours protégée par le groupement Acm (Schéma 5B). L'oxydation des groupements Acm pour former le deuxième pont a été réalisée avec du $Tl(TFA)_3$ dans du DMF avec de l'anisole en tant que scavenger. Les rendements globaux de la conotoxine monomère atteignaient 14% avec présence d'oligomères supplémentaires retenus sur le support solide. Alternativement, la combinaison orthogonale des groupements Xan et Acm peut être utilisée, ce qui a été illustré dans la synthèse sélective sur résine (et en solution) des trois isomères possibles de l'α-SI (Schéma 5C). S'il est certes possible d'améliorer les rendements de la formation des ponts disulfures sur résine, il ne faut pas oublier que dans une optique de conception de librairie de molécules, de faibles rendements peuvent être sacrifiés pour une efficacité de criblage plus élevée. Une fois qu'une tête de série (« hit ») a été isolé et caractérisé, le folding en solution peut être la meilleure option pour un meilleur scale-up et une optimisation plus poussée ou des études de SAR.

Schéma 5: Exemples de stratégie orthogonale d'oxydation sur résine. (A) Synthèse de l'α-GI par Boc-SPPS [82]. (B, C) Deux stratégies différentes pour la synthèse de l'α-SI en Fmoc-SPPS [122,124].



Un nombre conséquent de conditions existent pour la synthèse régiosélective de ponts disulfures, le travail de cette thèse s'inscrit dans le cadre de la synthèse de produits naturels et dérivés en quantité suffisante pour conduire les tests biologiques et déterminer la structure tridimensionnelle, c'est-à-dire généralement moins d'une dizaine de milligrammes. Ce n'est pas un travail d'optimisation de synthèse, par conséquent les méthodologies employées sont celles qui ont permis l'obtention du produit compte tenu de la disponibilité des réactifs présents au laboratoire, et certaines d'entre elles peuvent être améliorées ou modifiées.

2. Chapitre 1 : Synthèse de toxines linéaires et/ou pauvres en ponts disulfures ciblant les récepteurs couplés aux protéines G

a. Introduction

La famille des récepteurs couplés aux protéines G (RCPGs) est l'une des plus grandes superfamilles de gènes avec environ 370 membres réagissant aux ligands endogènes tels que les hormones ou les neurotransmetteurs et un nombre à peu près égal de récepteurs sensibles aux stimuli externes [125]. Ces récepteurs doivent leur appellation à leur capacité de couplage, une fois activés, à des protéines hétérotrimériques liant le GTP (guanosine triphosphate), aussi appelées protéines G. Ces protéines transmettent le signal provenant du récepteur à différents effecteurs intracellulaires permettant la génération d'une réponse cellulaire appropriée (Figure 10). Un certain nombre de récepteurs de cette superfamille sont des cibles bien reconnues pour le traitement médical de diverses maladies, alors que pour de nombreux autres, leur implication dans les pathologies est encore à déterminer. Un problème général associé à la recherche et au développement de médicaments ciblant les RCPGs est la spécificité insuffisante des ligands disponibles pour différencier les différents sous-types de récepteurs étroitement homologues. En effet, les RCPGs partagent tous la même topologie avec 7 hélices transmembranaires, une partie N-terminal extracellulaire et un C-ter intracellulaire, et au sein d'une même famille pharmacologique, la diversité des différents sous types existant représente un véritable challenge pour le développement de médicaments et de ligands sélectifs de certains sous-types. Malgré les efforts constants de la recherche académique et privée, cette tâche reste extrêmement difficile, surtout pour les petites molécules. Dans ce contexte, les peptides de venin pourraient apporter une contribution significative au développement de médicaments plus spécifiques. Bien que les toxines ciblent majoritairement les canaux ioniques, certaines d'entre elles ciblent les RCPGs et sont brièvement détaillées ci-après.

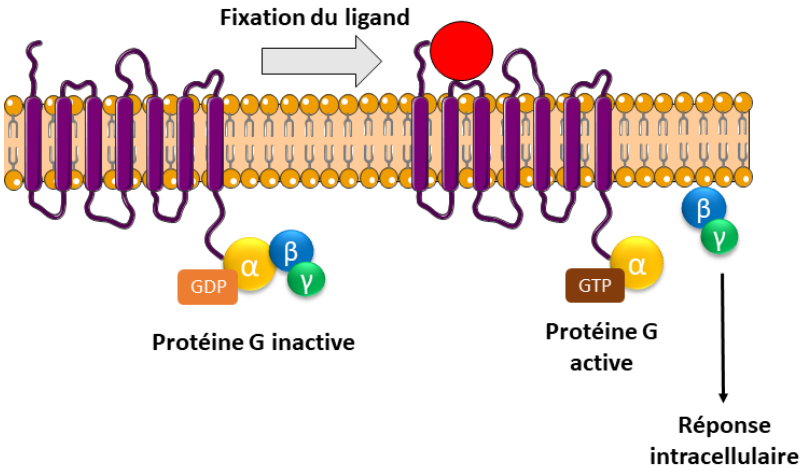


Figure 10 : Représentation schématique et fonctionnement des RCPGs.

- Les toxines à trois doigts (three-finger toxins, Figure 11)

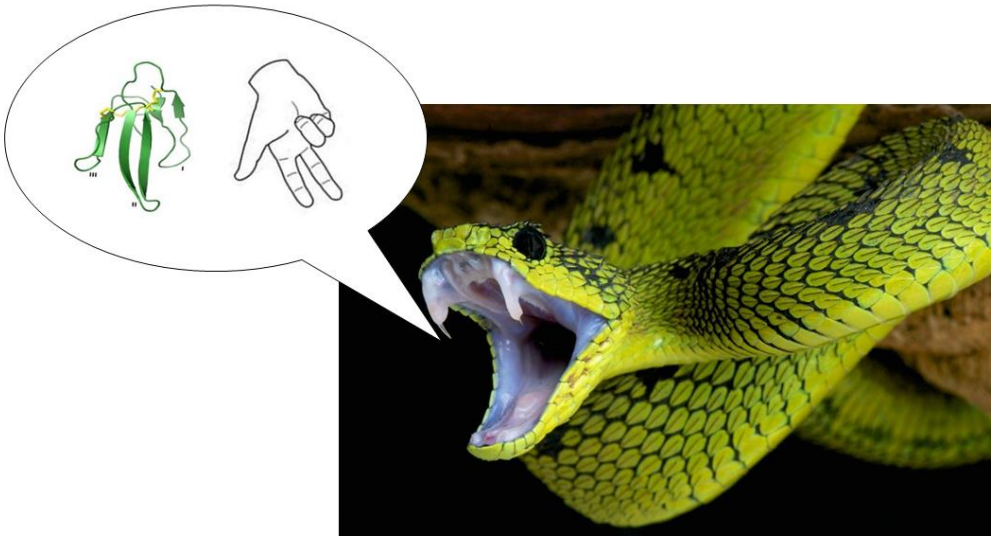


Figure 11 : Topologie des toxines à trois doigts (three-finger toxins).

La famille des toxines à trois doigts est une des familles les plus largement représentées et les mieux caractérisées des venins de serpents. Ces toxines présentes dans les venins d'elapidae (cobras, mambas), d'hydrophiinae et de colubridae [126] sont des peptides non enzymatiques composés de 60 à 75 résidus. La plupart de ces toxines à trois doigts ciblent les canaux ioniques (nAChRs et canaux calciques) cependant, la famille des toxines muscariniques (MTs) isolées des mambas, montre des activités sélectives pour le blocage des différents sous types des récepteurs muscariniques à l'acétylcholine (mAChRs) (Tableau 5). Les mAChRs sont les homologues métabotropes (RCPGs) des nAChRs, ils sont composés de 5 sous types différents M_1-M_5 [127]. Malgré la grande quantité de ligands muscariniques découverts dans les plantes et d'autres développés synthétiquement, les ligands vraiment sélectifs d'un sous type en particulier sont rares. Les toxines à trois doigts isolées des venins

de mambas sont parmi les ligands les plus sélectifs qui existent. Les premières MTs (MT1 et MT2) ont été isolées il y a 20 ans de cela à partir des venins de mamba vert d'Afrique de l'est (*Dendroaspis angusticeps)*. D'autres ont continué à être découvertes et sont regroupées dans le Tableau 5. Plus récemment des toxines du récepteur adrénergique ont aussi été identifiées comme appartenant à la famille des toxines à trois doigts (Tableau 5) [128–130]. D'autres toxines à trois doigts sont suspectées d'agir sur les RCPGs telles que les cardiotoxines ressemblant fortement à la chaine courte des neurotoxines mais généralement plus basiques et hydrophobiques, ainsi que les toxines à trois doigts dites non conventionnelles caractérisées par leur faible toxicité et un cinquième pont disulfure additionnel dans la boucle numéro 1 [126,131].

Tableau 5 : Profils de sélectivité des toxines trois doigts et affinités apparentes déterminées par déplacement de ligands radiomarqués. † Pour certaines toxines les valeurs d'IC_{50} sont plus appropriées à cause de la nature non compétitive des interactions. M_1-M_5 représente les 5 sous types des récepteurs muscariniques à l'acétylcholine et les sous types du récepteur adrénergique sont indiqués par la lettre α.

Toxine	Sélectivité (affinité nM)	Refs
MT1	M_1 (50) > M_4 (70-130)	[132,133]
MT2	M_1 (400) > M_4 (1200)	[132,133]
MT3	M_4 (1-2) >> M_1 (100)	[134–136]
MT4	M_1 (60) > M_4 (80)	[137]
MT5	M_1 (200) > M_4 (500)	[137]
MT6	M_4 (4) >> M_1 (200)	[137]
MT7	M_1 (0.5-1.5†)	[138–140]
MTα	α_{2B} (3†) >> α_{2A}, α_{2C} (>1000†)	[128]
ρ-Da1a	α_{1A} (0.35) >> α_{1B}, α_{1D} (300–400)	[129]
ρ-Da1b	α_{2A} (14) > α_{2B}, α_{2C} (40–70)	[130]

- Les phospholipases A2 (PLA2s)

Les venins de serpent sont une riche source de PLA2, enzymes qui hydrolysent les glycérophospholipides, les acides gras libres et lysophospholipides. Elles sont suspectées de jouer un rôle dans l'initiation de la digestion de la proie. Cependant, certaines montrent des activités neurotoxiques comme la vipoxin de la Vipère de Russell (*Vipera ruselli)* qui inhibe le récepteur à l'amine biogénique [141] avec une forte affinité (3-30 nM). Une autre toxine également de *Vipera ruselli* nommée β-RTX inhibe les récepteurs α_2 adrénergiques [142]. De plus, un inhibiteur muscarinique présentant une activité PLA2 purifié du venin de cobra indonésien (*Naja naja sputatrix*) a montré une sélectivité parmi les différents sous types de mAChRs [143,144] (M_5>M_1>M_2=M_3=M_4).

- L'exemple de l'exenatide

Au début des années 1990, le criblage haut débit de venins de monstres de Gila a conduit à la découverte de l'exendin-4 (Figure 12), aujourd'hui commercialisé sous l'appellation exenatide pour le traitement du diabète de type 2 et a servi de base pour le développement d'agoniste à action prolongée du GLP-1.

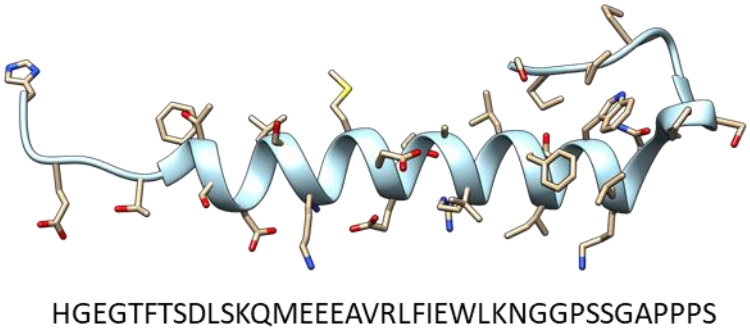


Figure 12 : Représentation tridimensionelle et séquence primaire de l'exenatide.

- Les sarafotoxines

Les sarafotoxines-a, -b et -c ont été isolées du venin de vipère *Atractaspis engaddensis*, un serpent très toxique d'Afrique et du Moyen-Orient [145]. L'action de ces peptides a entraîné une vasoconstriction prononcée chez le rat et présentait une forte homologie avec les endothélines (Figure 13) qui sont de puissants vasoconstricteurs chez l'homme [146]. La sarafotoxine-c (ou S6c) est un agoniste sélectif du récepteur de endothéline B (ETB) [147]. Elle constitue l'agoniste du récepteur de l'ETB le plus utilisé dans les modèles expérimentaux pour la détermination des réponses via le récepteur de l'ETB. Il a été démontré que la sarafotoxine-b possédait une activité inhibitrice au niveau des métalloprotéinases

matricielles (MMP) et, grâce à des travaux d'optimisation de séquence, la toxine était convertie en un inhibiteur sélectif des MMP1 et 9 [148].

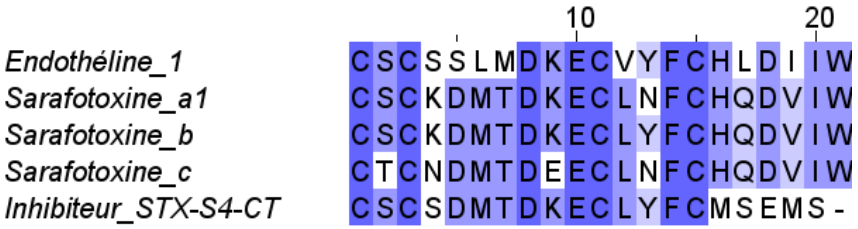


Figure 13 : Alignement montrant la similarité entre l'endothéline 1, les sarafotoxines et l'inhibiteur STX-S4-CT.

- La toxine intestinale de Mamba (MIT1) et Bv8

La toxine MIT1 (Mamba Intestinal Toxin) isolé de l'intenstin du mamba *Dendroaspis polylepis* est un peptide de 81 acides aminés de long isolé du mamba noir [149]. Un peptide de structure semblable, Bv8, a été isolé de la peau de la grenouille *Bombina variegata* en 1999 [150]. Cela a conduit à l'identification des procinéticines (PKs), PK1, et PK2, ainsi que leurs RCPGs associés, récepteur à la procinéticines 1 (PKR1) et PKR2 [151–153]. Les PKs sont impliqués dans divers cadres physiologiques, notamment le développement des organes de reproduction, la neurogenèse, le rythme circadien et la consommation de nourriture, la motilité gastro-intestinale et la perception de la douleur [154]. Bv8 et MIT1 se lient et activent les récepteurs de la PKR avec une puissance élevée, mais ils ne différencient généralement pas les différents sous-types de récepteurs [152,155].

- Les ligands des récepteurs latrophilines

La α-latrotoxine (α-LTx) composant du venin de la veuve noire a été utilisée pendant plusieurs décennies pour étudier le mécanisme de l'exocytose neuronale [156,157]. L'α-LTx s'associe également à plusieurs récepteurs de la membrane présynaptique. L'un d'entre eux s'est révélé être un récepteur de l'α-LTx indépendant du Ca^{2+}, et nommé latrophiline [158,159]. La latrophiline présente les caractéristiques structurelles d'un RCPG et est l'un des membres fondateurs d'un groupe de RCPG collectivement appelés RCPG d'adhésion [160]. On en sait peu sur la fonction de ces récepteurs et une grande partie des homologues identifiés sont encore orphelins. Le récepteur de la latrophiline, par le biais du stimulus de α-LTx, se couple aux protéines G et induit des stimuli excitateurs dans les neurones ainsi que dans les cellules neuroendocrines [161,162].

- Les peptides de cônes

Un peptide nommé ρ-TIA antagonisant les récepteurs α_1 adrénergiques (α_1-AR) a été identifié dans le venin de *Conus tulipa* bien que celui-ci présente une structure typique des α-conotoxines ciblant les nAChRs [163]. L'antagonisme fonctionnel α-AR du ρ-TIA a été démontré aux niveaux tissulaire et cellulaire par sa capacité à inhiber les augmentations des taux de Ca^{2+} intracellulaire et la contractilité musculaire [164]. La ρ-TIA se lie aux trois sous-types de α_1-AR, mais montre une affinité dix fois supérieure pour le sous type α_{1B}-AR avec une valeur d'IC_{50} de 2 nM [164,165].

Les hormones hypophysaires arginine vasopressine et ocytocine sont des nonapeptides agissant par l'intermédiaire des récepteurs à la vasopressine et à l'ocytocine. Des analogues de ces peptides, appelés conopressins, ont également été trouvés dans les venins de cônes (détaillé ultérieurement dans la section c de ce chapitre) [166,167].

Un glycopeptide 16 acides aminés appelé contulakine-G, a été isolé de *Conus geographus* et présentait une similarité de séquence avec les neurotensines des vertébrés et il s'est avéré se lier avec des affinités submicromolaires à différents récepteurs de la neurotensine, avec une faible sélectivité pour certains sous-types de récepteurs [168]. Cependant, dans les tests fonctionnels mesurant la production de phosphate d'inositol médiée par le récepteur, la contulakine-G a montré une activité sélective sur récepteur de la neurotensine 1 préférentiellement au récepteur 2. La neurotensine étant étroitement impliquée dans les voies de la douleur, la contulakin-G est actuellement en phase de test clinique comme analgésique.

Il semble que certains animaux utilisent des peptides assez simples pour imiter les homologues hormonaux de leurs proies et modifient ainsi considérablement l'activité des récepteurs lors de l'envenimation. On peut supposer que davantage de peptides analogues d'hormones seront identifiés par différentes procédures de criblage et que ceux-ci pourraient aider à découvrir les différents rôles des nombreux sous-types de récepteurs existant.

b. Synthèse d'analogues de toxines d'araignées

Les araignées utilisent principalement leur venin pour paralyser leurs proies, il n'est donc pas surprenant que ces venins contiennent une abondance de peptides qui modulent l'activité des canaux ioniques et des récepteurs neuronaux. Cependant, comme nous avons pu le voir précédemment, peu d'études ont été réalisées sur les peptides de venins d'araignées ciblant les RCPGs. La toxine d'araignée (*Phoneutria nigriventer*) PnTx2-6 a été décrite comme bloquant les canaux sodiques et potentiel traitement pour la dysfonction érectile et la perception de la douleur. Récemment, un analogue

linéaire nommée PnPP-19 a montré une activité analgésique et la cible a été décrite comme étant le récépteur μ-opioïde. Dans l'idée de déterminer son pharmacophore et de mieux identifier ses cibles pharmacologiques, nous avons effectué la synthèse du PnPP-19 et de 2 analogues tronqués supplémentaires afin de tester leurs activités biologiques sur le récepteur μ-opioïde en collaboration avec l'équipe du Dr. Sébastien Granier de l'IGF à Montpellier. Cette étude dans laquelle j'ai effectué la synthèse et la caractérisation par spectrométrie de masse, est détaillée ci-après et présentée sous la forme d'une publication, non publiable en l'état.

1. Introduction

Matavel *et al.* [169] cloned and sequenced the encoding precursor of the PnTx2-6 toxin, which has been shown to reduce the Na^+ current amplitude. PnTx2-6 has also been reported to increase relaxation *in vivo* in rat [170,171]. Thereafter, Matavel *et al.* [172] confirmed the sequence and showed that PnTx2-6 markedly delayed the inactivation kinetics of neuronal-type sodium channels. Consequently, decrease in voltage gated Na^+ channels have been proposed as a possible mechanism explaining the action on erectile function [172,173]. Moreover, they proposed a structural model of the toxin and identified a likely binding region proposed on the basis of the different affinities between PnTx2-5 and PnTx2-6 related to their sequence differences. This likely binding region have been used to design a synthetic analogue PnPP-19 [173]. Surprisingly, this synthetic analogue had no effect on sodium channels or heart, however, it has been shown to potentiate erection through nitric oxide (NO)/ cyclic guanosine monophosphate (cGMP) pathway [173]. In addition, PnPP-19 exhibited reduced toxicity compared to dose-dependent side effects observed with PnTx2-6 toxin [173] such as intense vascular congestion in kidney, liver, lungs and myocardium, a discrete brain edema and prolonged pain [174].

Since PnTx2-6 induces hyperalgesia, Freitas *et al.* [175] investigated the role of PnPP-19 in the nociceptive pathway and unexpectedly found that PnPP-19 induces antinociception in rat. They have demonstrated inhibition of the antinociceptive effect of PnPP-19 in the presence of selective inhibitors of cannabinoid receptor type 1 (CB_1), μ- and δ-opioid receptors. Moreover, they evidenced that PnPP-19 is a neutral endopeptidase (NEP) substrate and it is cleaved only after a long incubation time suggesting that it is competing with endogenous opioids thereby increasing their levels and causing the antinociceptive response. Consequently, they proposed that PnPP-19 induced antinociception through NEP inhibition and activation of CB_1, μ- and δ opioid receptors.

Considering that PnPP-19 potentiates erectile function by activating nitrergic system and that NO is an antinociceptive molecule, *Freitas et al.* [176] investigated the role of the activation of this pathway in

PnPP-19 induced antinociception. By using selective inhibitors of nitric oxide synthetases (NOSs), neuronal nitric oxide synthetases (nNOSs), guanylyl cyclase and potassium channels ATP-dependent (K_{ATP}) they hypothesized that the peripheral antinociceptive effect results from NO-cGMP-K_{ATP} pathway. Moreover, they reported tissue nitrite concentration increase after PnPP-19 administration. Finally, *Freitas et al.* [177] demonstrated that PnPP-19 selectively activates μ-opioid receptors inducing indirectly inhibition of calcium channel without β-arrestin 2 recruitment.

To further delineate the pharmacophore responsible for the analgesic activity, we synthesized PnPP-19 and two additional truncated peptides and tested their activity on μ-opioid receptors.

2. Methods

2.1 Abbreviations

Acm, acetamidomethyl; ACN, acetonitrile; Boc, *tert*-butoxycarbonyl; DCM, Dichloromethane; DIPEA, diisopropylethylamine; DMF, *N,N'*-dimethylformamide; ESI-MS, electrospray ionization mass spectrometry; Fmoc, fluorenylmethoxycarbonyl; HATU, 1[*Bis*(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate; LC/MS, liquid chromatography/mass spectrometry; MeOH, methanol; Pbf, pentamethyl-dihydrobenzofuran-5-sulfonyl; RP-HPLC, reversed phase high performance liquid chromatography; SPPS, solid phase peptide synthesis; t-Bu, *tert*-butyl;TFA, trifluoroacetic acid; TIS, triisopropylsilane; Trt, trityl; UV, ultra-violet.

2.2 Chemical synthesis

DMF, DIPEA, ACN, TIS, TFA, piperidine and all others reagents were obtained from Sigma-Aldrich (Saint-Louis, MI, USA) or Merck (Darmstadt, Germany) and were used as supplied. Fmoc (L) amino acid derivatives and HATU were purchased from Iris Biotech (Marktredwitz, Germany). Amphispheres™ 40 RAM (0.4 mmol/g) were purchased from Agilent Technologies (Les Ulis, France). The following side-chain protecting groups were used: Trt for Asn and Gln; *t*Bu for Ser, Tyr and Glu; Boc for Lys and Trp; and Pbf for Arg. Peptides were manually synthesized using the Fmoc-based solid-phase peptide synthesis technique on a VWR (Radnor, PA, USA) microplate shaker. All Fmoc amino acids and HATU were dissolved in DMF to reach 0.5 M. Fmoc deprotection was carried out with piperidine in DMF (1/2 v/v) twice for 3 min. Subsequent amino acids were coupled onto 0.1 mmol of resin twice for 10 min using an amino acid/HATU/DIPEA ratio of 5:5:10 relative to resin loading. DMF was used for resin washing between deprotection and coupling steps. After chain assembly was complete, the terminal Fmoc group was removed and the resin washed with DMF and DCM. For PnPP-19 synthesis, acetylation was carried out twice for 15 min with 50% acetic anhydride in DCM. Side-chains deprotection and cleavage from the resin was carried out by adding 10 mL of TFA/TIS/H_2O (95/2.5/2.5 v/v/v) and stirring the mixture for 2.5 h at room temperature. After the resin was removed by filtration and washed three

times with dichloromethane. Dichloromethane and TFA were removed under vacuum then cold diethyl ether was added to precipitate the peptide. Residues were taken up in 20% ACN acidified with 0.1% TFA and loaded onto preparative RP-HPLC and pure fractions were combined. The combined pure fractions were freeze-dried and their purity were confirmed by LC/ESI-MS.

2.3 Mass spectrometry

Solvents used for LC/MS were of HPLC grade.

Intermediate products were characterized using a LC/MS system consisting of a Waters (Milford, OH, USA) Alliance 2695 HPLC, coupled to a Waters Micromass ZQ spectrometer (electrospray ionization in positive mode (ESI+) fitted with a quadrupole mass analyzer). All the analyses were carried out using a Chromolith (Fontenay sous Bois, France) HighResolution RP-18e (4.6 x 25 mm, 15 nm–1.15 µm particle size, flow rate 3.0 mL/min) column. A flow rate of 3 mL/min and a gradient of 0–100% B over 2.5 min for routine analyses and 0-30% B over 30 min for quality control of pure products were used. Eluent A was water/0.1% HCO_2H and eluent B consisted of acetonitrile/0.1% HCO_2H. UV detection was performed at 214 nm. Electrospray mass spectra were acquired at a solvent flow rate of 200 µL/min. Nitrogen was used for both the nebulizing and drying gas. The data were obtained in a scan mode ranging from 100 to 1000 m/z or 250 to 1500 m/z to in 0.7 s intervals.

2.4 Preparative RP-HPLC

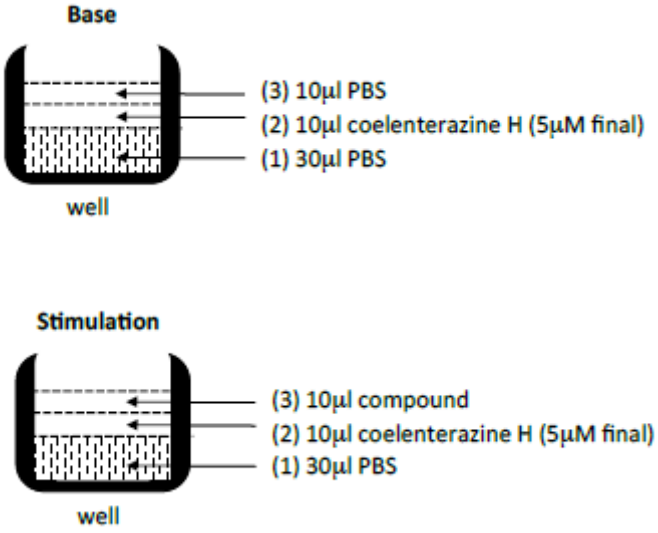
Preparative RP-HPLC was run on a Gilson PLC 2250 Purification system (Villiers le Bel, France) instrument using a preparative column (Waters DeltaPak C18 Radial-Pak Cartridge, 100 Å, 40 × 100 mm, 15 µm particle size, flow rate 50.0 mL/min). Buffer A was 0.1% TFA in water, and buffer B was 0.1% TFA in acetonitrile.

2.5 *Cell culture and transfection*

HEK293 cells were grown in DMEM supplemented with 10% FBS (without antibiotics) at 37°C, 5% CO_2 Transient transfection was performed using electroporation in a volume of 150µl with 1µg FLAG-SNAP-MOR, 1µg Gi1/oA-Rluc8, 1µg FLAG-β2, 2µg Venus-γ2 plasmids and 5 millions of HEK293 cells in electroporation buffer (50mM K_2HPO_4, 20mM CH_3COOK and 20mM KOH, pH 7.4). After electroporation (250 V, 0.5mF, Bio–Rad Gene Pulser electroporator; Bio-Rad Laboratories, Hercules, CA), cells were resuspended in 5ml DMEM supplemented with 10% fetal bovine serum and seeded for 24h in a white 96-well plate (pretreated with Poly-L-Ornithine 1X) at a density of 100,000 cells per well.

2.6 BRET assay

24 hours after transfection, cells were washed with PBS Ca^{2+}/Mg^{2+} and reagents were added as described below:



BRET between Rluc8 and Venus was measured after the addition of the Rluc8 substrate coelenterazine H (5µM). BRET readings were collected using a Mithras 2 plate reader. The BRET signal was calculated by the ratio of emission of Venus/YFP (535nm) to Rluc8 (480nm): mBRET = {(Ratio 535/480) assay - (Ratio 535/480)Rluc8 alone)} X 1000.

3. Results

Most endogenous ligands of µ-opoid receptors share all the same N-terminal motif Y-[P/G]-[G/F]-F (Figure 1). Considering the highly conserved tyrosine residue at position 1 in endogenous ligands of µ-opioid receptors we decided to generate truncated forms of PnPP-19 sequence at tyrosine residues, namely analogues PnPP-14 and PnPP-8.

Met-enkephalin	YGGFM
Leu-enkephalin	YGGFL
Beta-endorphin	YGGFMTSEKSQTPLVTLFKNAIIKNAYKKGE
Dynorphin-A	YGFFLRRIRPKLK
Endomorphin-2	YPFF
PnPP-19	GERRQYFWIAWYKLANSKK
PnPP-14	YFWIAWYKLANSKK
PnPP-8	YKLANSKK

Figure 1: Alignment of synthesized PnPPs peptides and endogenous ligands of µ-opioid receptors.

3.1 Chemical synthesis

Peptides have been synthesized using Fmoc-SPPS chemistry. According to its described sequence PnPP-19 [175] has been prepared on rink amide resin and the N-ter has been acetylated. Considering that PnPP-8 and PnPP-14 have been designed as truncated sequences they also have been synthesized on rink amide resin but not acetylated since endogenous ligands are not. After successful assembly of the peptides, they were purified by preparative RP-HPLC purification and their homogeneity was assessed by analytical RP-HPLC and MS (Figure 2).

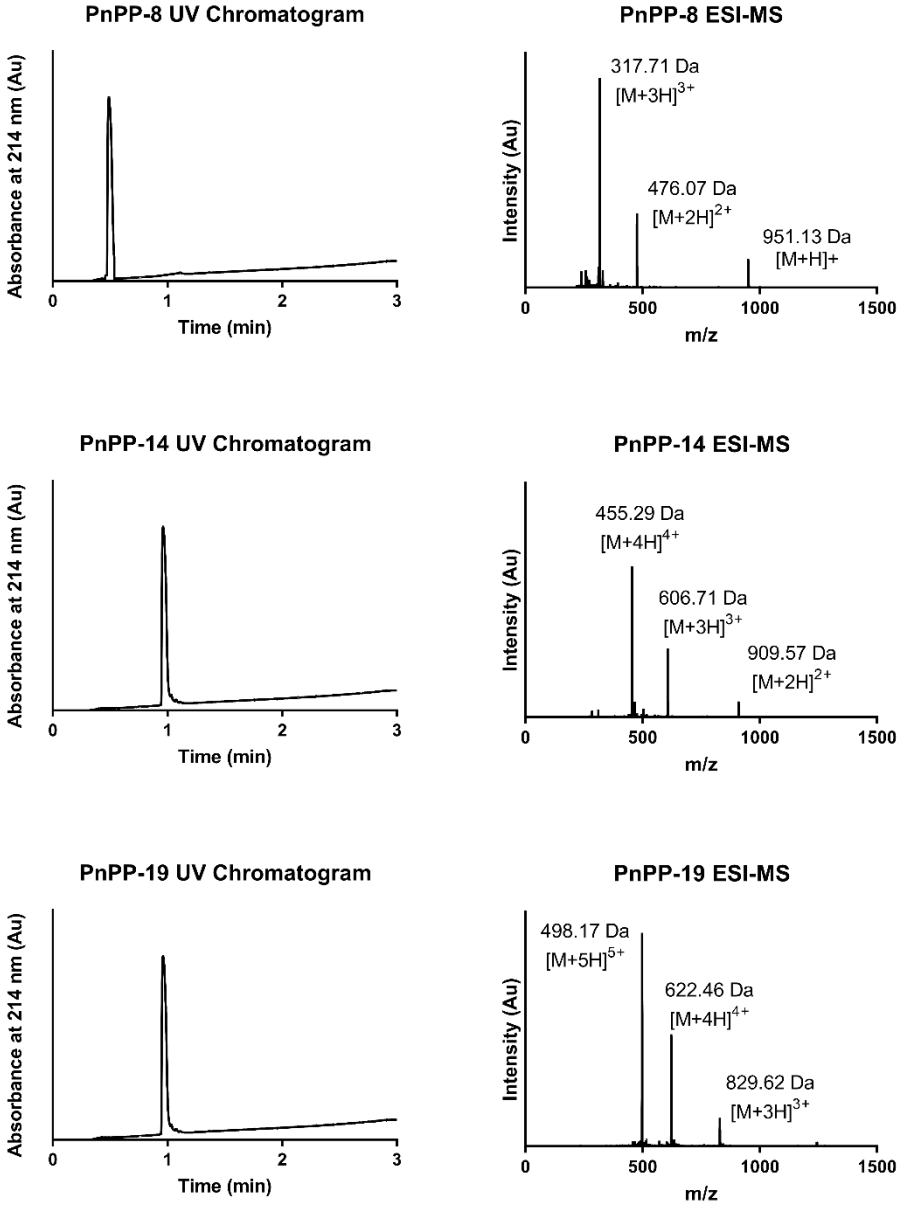


Figure 2: RP-HPLC/ESI-MS analyses of the synthesized PnPP peptides. ACN gradient from 0 to 100% over 2.5 min.

3.2 BRET assays on μ-opioid receptors

We investigated the signaling profile of the activated μ-opioid receptor expressed in HEK293 cells towards G proteins. The three synthetic peptides were tested on two GPCRs. G heterotrimeric complex dissociation is monitored by BRET (Bioluminescence Resonance Energy Transfer). The Gα protein is fused to Renilla Luciferase8 (donor) and gamma-2 to Venus (acceptor). Activation of MOR by an agonist promotes the dissociation/conformational modification of the Gαβγ protein complex resulting in the BRET signal decay.

3.2.1 Determination of MOR coupling to Gi1/oA upon peptides stimulation.

For Gi1/oA dissociation assays, DAMGO agonist was used as positive control (Figure 3) which is an enkephalin analogue displaying the following sequence: H-Tyr-D-Ala-Gly-N-MePhe-Gly-ol. Naloxone is an antagonist of the μ-opioid receptor.

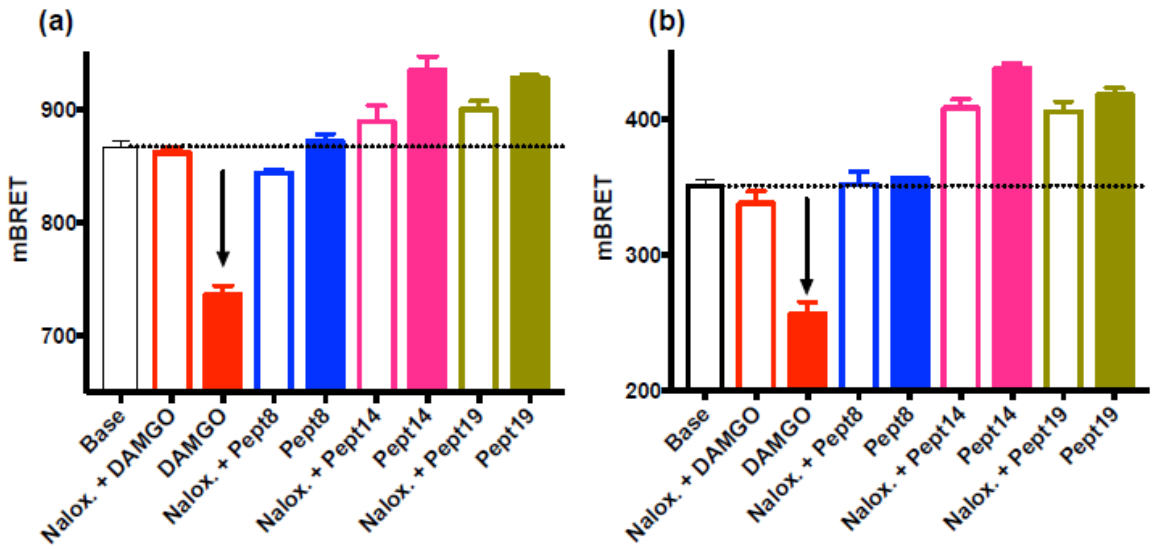


Figure 3. Stimulation of μ-opioid receptor with DAMGO and synthetic peptides (agonist activity). HEK293 cells expressing μ-opioid receptor with the G protein complex Gi1-Rluc8 (a) GoA-Rluc8 (b), FLAG-β2 and Venus-γ2 were stimulated with 10μM DAMGO, PnPP-8, PnPP14 and PnPP-19 +/- an excess of Naloxone at 100μM. Data are means +/- SD of triplicates of a representative experiment.

None of the synthetic peptides were able to induce the dissociation of the Gαβγ protein complex (no agonist activity). Surprisingly, upon stimulation with PnPP-14 and PnPP-19, the BRET signal measured between Gi1/oA-Rluc8 and Venus-γ2 was increased, possibly meaning i) an increased number of G protein complex formation, ii) a conformational change of the G protein complex promoting the energy transfer between Rluc8 and Venus.

The same experiment was then conducted in presence of DAMGO at EC80 +/- peptides in order to determine if peptides were able to block DAMGO activity on μ-opioid receptor (Figure 4).

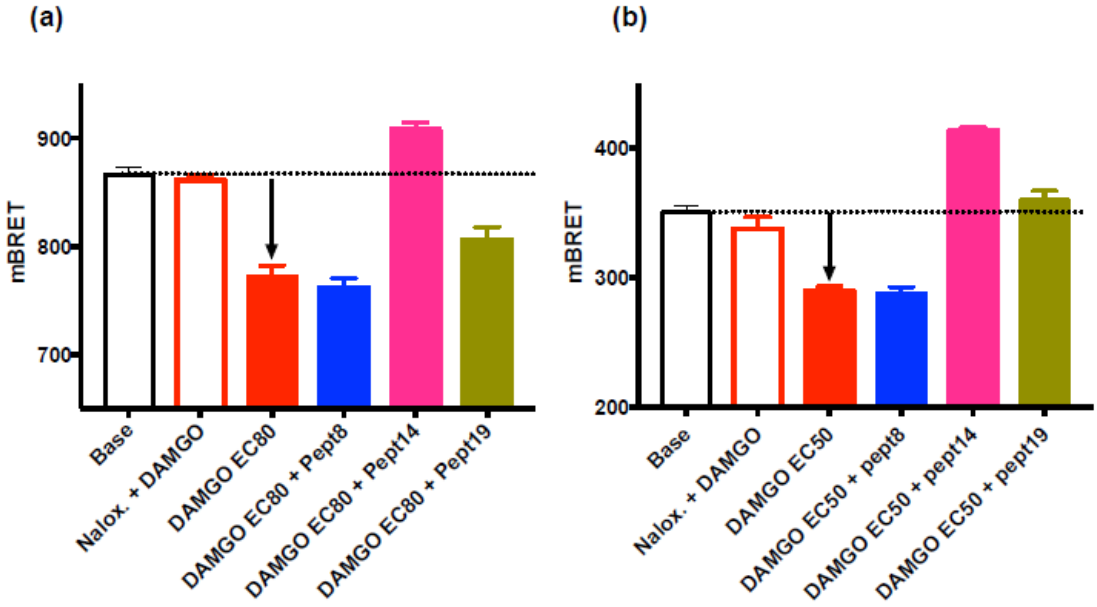


Figure 4. Stimulation of μ-opioid receptor with DAMGO at EC80 +/- peptides (antagonist activity). HEK293 cells expressing μ-opioid receptor with the G protein complex Gi1-Rluc8 (a) GoA-Rluc8 (b), FLAG-β2 and Venus-γ2 were stimulated with 100nM DAMGO (EC80) +/- 10μM PnPP-8, PnPP14 and PnPP-19. Data are means +/- SD of triplicates of a representative experiment.

PnPP-14 and PnPP-19 were able to block DAMGO activity on Gi1/oA protein complex, with PnPP-14 being the most effective. Interestingly, the signal obtained with PnPP-14 is higher than the base signal.

3.2.2 Effects of peptides on other GPCRs.

To determine the selectivity of the effect seen on opoid receptors, the same experiments were reproduced using urotensin receptor activated by UII agonist (Figure 5).

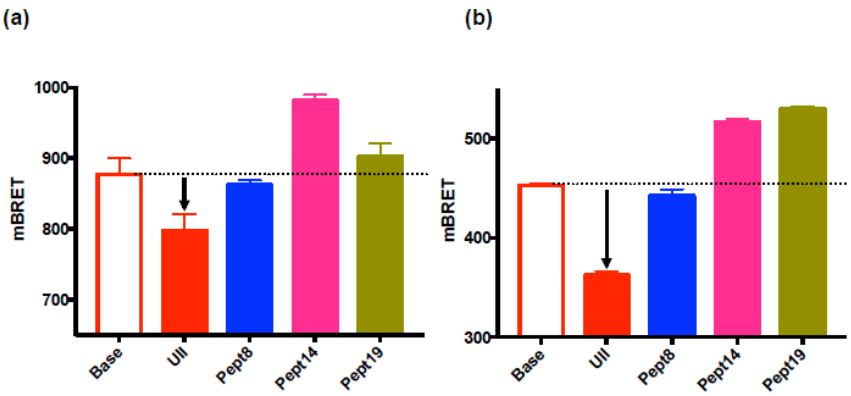


Figure 5. Stimulation of urotensin receptor with UII and the synthetic peptides. HEK293 cells expressing UT with the G protein complex Gi1-Rluc8 (a) GoARluc8 (b), FLAG-β2 and Venus-γ2 were stimulated with 10µM UII, PnPP-8, PnPP14 and PnPP-19. Data are means +/- SD of triplicates of a representative experiment.

Remarkably, the same effects were observed on UT and µ-opioid receptor with PnPP-14 and PnPP-19.

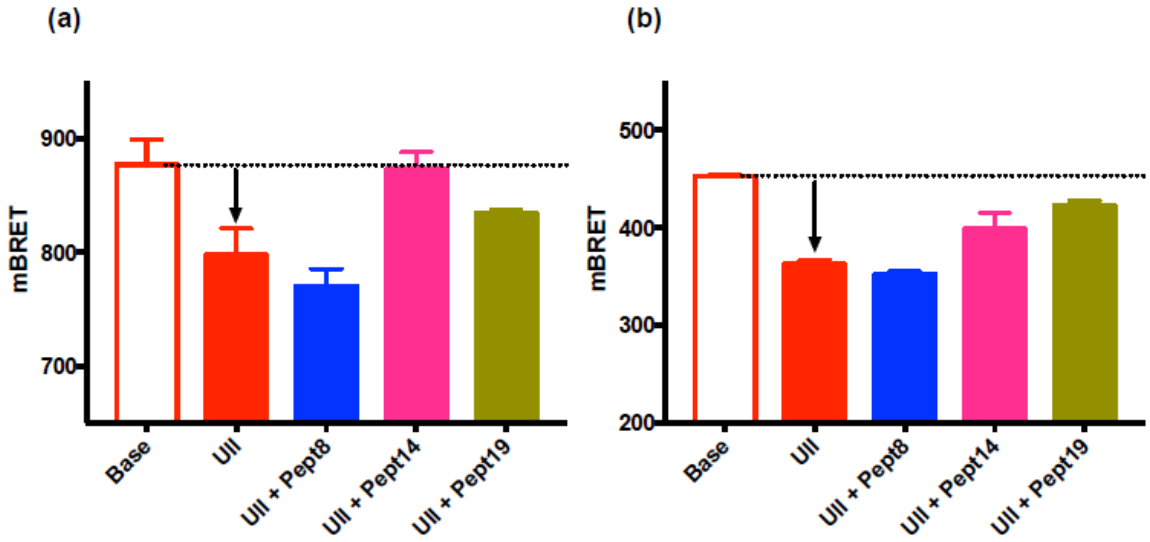


Figure 6. Stimulation of urotensin receptor with UII +/- peptides. HEK293 cells expressing ST-UT with the G protein complex Gi1-Rluc8 (a) GoA-Rluc8 (b), FLAG-β2 and Venus-γ2 were stimulated with 10µM UII +/- 10µM PnPP-8, PnPP14 and PnPP-19. Data are means +/- SD of triplicates of a representative experiment.

Given that the same effects were observed on urotensin receptor for PnPP-14 and PnPP-19 peptides, we can conclude that these peptides do not have a specific activity on µ-opioid receptor and may act via a yet unknown mechanism.

4. Discussion

PnPP-19, a linear peptide derived from the spider toxin PnTx2-6, was recently described to produce antinociception in rat, likely through activation of µ-opioid receptors. However, we found that neither PnPP-19 nor its truncated analogues activated the µ-opiod receptor. Unexpectedly, they seem to induce the recruitment of additional GoA/Gi1 proteins, therefore increasing the number of associated complexes and the BRET values. Intrestingly, PnPP-14 is more active than PnPP-19 and PnPP-8 completely lost the activity, suggesting that the required pharmacophore is located between residues 8 and 14. Another surprising result concerns the apparent lack of selectivity of this effect. Indeed, the

same effects were obtained using the urotensin receptor, another GPCR, clearly indicating that the observed effect is not selective for the μ-opiod receptor. Whereas it is difficult and preliminary at this stage to propose a mechanism of action, one plausible explanation is that PnPP-14 and PnPP-19 induce the recruitment of an increasing number of GoA/Gi1 proteins via modulation of another receptor, probably an ion channels considering the literature. In order to determine if this effect is selective for GoA/Gi1 proteins, further assays are currently conducted on others G protein subtypes. Another exciting hypothesis would be that PnPP-14 and PnPP-19 are able to cross the cell membrane as cell-penetrating peptides do, and somehow directly potentiate the association of G proteins therefore increasing the BRET values. Indeed cell-penetrating peptides contain a high relative abundance of positively charged amino acids such as lysine or arginine or have sequences that exhibit an alternating pattern of polar/charged amino acids and non-polar, hydrophobic amino acids. Considering the high content of positively charged lysine as well as hydrophobic residues (Leu, Trp, Tyr and Phe) in the PnPPs sequences, the hypothesis that PnPPs peptides are able to cross the membrane is reasonably conceivable and is worth investigating.

c. Synthèse de nouvelles conopressines

Les venins d'araignées étant habituellement pauvre en toxines ciblant les RCPGs, nous avons orienté notre exploration vers les toxines de cônes, dont certaines d'entre elles sont déjà bien répertoriées dans la littérature et leurs activités sur les GPCRs avérées. Une famille en particulier, les conopressines, regroupe des analogues naturels de l'hormone vasopressine.

Contrairement aux autres toxines précédemment explicitées dans l'introduction de ce chapitre, et aux vues de la petite taille des peptides de cônes, leurs venins constituent sans aucun doute la source la plus attractive de toxines ciblant les RCPGs. Pour cette raison nous avons interrogé nos transcriptomes de cônes à la recherche de nouvelles séquences de conopressines. Deux d'entre-elles ont été identifiées dans le transcriptome de *Conus miliaris* et sélectionnées pour leur originalité. Cette étude dans laquelle j'ai effectué la recherche transcriptomique, la synthèse et la caractérisation par spectrométrie de masse est présentée ci-après sous la forme d'une publication en préparation. Seules les études structurales (RMN) manquent encore aux parties discussion et résultats.

Synthesis, pharmacological and structural characterization of novel conopressins from *Conus miliaris*

Julien Giribaldi [1], Lotten Ragnarsson [2], Edward Evans [3], Tom Pujante [1], Christine Enjalbal [1], David Wilson [3], Norelle L. Daly [3], Richard J. Lewis [2] and Sebastien Dutertre [1,*]

[1] *Institut des Biomolécules Max Mousseron, UMR 5247, Université de Montpellier-CNRS, 34095 Montpellier, France*
[2] *Institute for Molecular Bioscience, The University of Queensland, St Lucia, Queensland 4072, Australia*
[3] *Centre for Biodiscovery and Molecular Development of Therapeutics, Australian Institute of Tropical Health and Medicine, James Cook University, Cairns, QLD 4878, Australia*

1. Introduction

Cone snail venoms represent a unique source of bioactive peptides, yet it is estimated that less than 1% of all conopeptides have been pharmacologically characterized so far. Nevertheless, the therapeutic potential is undeniable with one peptide approved by FDA for the treatment of chronic pain [178–180] and several others that are in clinical trials. Furthermore, these natural peptides have become indispensable tools to decipher ion channel subtypes and other membrane receptors [20,181]. Conopeptides are divided in two mains groups, based on the number of disulfide bonds, the disulfide-poor conopeptides (0 or 1 disulfide bond) and the disulfide-rich conotoxins (2 or more disulfide bonds) [182]. Disulfide-poor conopeptides are minor components of cone snail venom subdivided into many subclasses such as contulakins, conantokins or conopressins that target several G protein-coupled receptors (GPCRs) [182]. Since they are not as abundant as conotoxins targeting ion channels they have been less studied, but recent venomics analyses of venom glands reveal that disulfide-poor conopeptides remain an untapped resource of novel bioactives.

Almost thirty years ago, the first vasopressin-oxytocin-related contoxins Lys-conopressin-G and Arg-conopressin-S were isolated from *Conus geographus* and *Conus striatus* venoms, respectively [167]. Intracerebral injection into mice induced similar effects compared to neurohypophysal hormone injection, and therefore, it was hypothesized that conopressin-G/S might act as agonists of the same class of GPCR in the brain. While only one subtype of oxytocin receptor (OTR) has been identified, three different vasopressin receptor subtypes have been characterized that possess different pharmacological and G protein coupling properties [183]. V2 receptors (V2Rs) are coupled to adenylyl cyclase and their activation produces an antidiuretic effect on the kidneys. V1aR and V1bR, as well as OTR are coupled to phospholipase C but produce different biological effects upon activation: The V1aR regulates blood pressure and vasoconstriction, whereas the V1bR is responsible for the corticotrophin

release from the pituitary gland. Compounds targeting the OTR are clinically used to stimulates the contraction of uterine and mammary myocytes [184,185]. Interestingly, unlike vasopressin and oxytocin, conopressins G and S display an additional positive charge at position 4 (Figure 1A) which is only found in two other endogenous vasopressin analogues cephalotocin (*Octopus vulgaris*) and annetocin (*Eisenia foetida*) [167,186]. Lys-conopressin-G was later isolated from other non-venomous snail species and was therefore proposed as an endogenous invertebrate vasopressin analog. An underlying question was to determine whether conopressins are endogenous or venom peptides in cone snails, since a role in prey capture has never been demonstrated. However, given that more conopressins have been identified in different venomous cone snail species, it clearly suggests they are venoms peptides with a role in envenomation [187]. Among them, the unique γ-conopressin-vil purified from *Conus villepinii* venom, which displays a carboxyglutamate residue (Figure 1A) conferring the capacity of the peptide to change its conformation in the presence of calcium ions [188]. Another unique example is conopressin-T (Figure 1A) isolated from *Conus tulipa*, which, unlike vasopressin-oxytocin related peptides, is a selective antagonist of V1a receptors [186]. Interestingly, position 9 was described as an antagonist switch and replacement of glycine at this position by a valine residue turn oxytocin and Arg-Vasopressin (AVP) from full agonists to full antagonists. The L7P-conopressin-T mutant displays an increased affinity for the V1a receptors, whereas activity at the V1b and V2 receptors remain unchanged, suggesting a favorable V1aR selective conformational change induced by the proline residue [186].

Docking studies on a three-dimensional model of the V1a receptor revealed that Arg-vasopressin binds into a 15–20 Å deep cleft defined by the transmembrane helices of the receptor. Residues located in this region interacting with agonist ligands are highly conserved in all the vasopressin and oxytocin receptors suggesting that the agonist-binding pocket is common to all the different subtypes of this receptor family [186]. Interestingly, mutations of residues from the receptor agonist binding site do not affect antagonist activity, suggesting a different binding mode of antagonist ligands [186]. Indeed, the different binding sites of various antagonist ligands are formed by transmembrane helices 1, 2 and 7, whereas the agonist binding site is mainly made by the three extracellular domains of the oxytocin receptor, as evidenced by Postina et al [189]. Based on the bovine rhodopsin structure, three-dimensional molecular models of the V1aR and V1bR complexed with vasopressin suggested that four receptor's key residues $E^{1.35}$, $D^{2.65}$, $V^{4.61}$ and $P^{5.35}$ fine tune the binding of vasopressin and related peptide agonists to both receptor subtypes [190]. These predictions have been validated by receptor mutants and should enable the identification of further new V1a and V1b receptor selective agonists [190].

In this study, we report on the synthesis, the pharmacological characterization and structure of two new conopressin-related sequences identified in the *Conus miliaris* transcriptome and displaying an unusual negatively charged amino acid in position 8. All peptides were characterized on both human and zebrafish receptors using conopressin-G, vasotocin, oxytocin and vasopressin as controls. Together with their NMR structures and considering their unique sequences, this study provides valuable structure-function information that might be useful to guide the design of new vasopressin receptor selective ligands.

2. Materials and methods

2.1 Abbreviations

Acm, acetamidomethyl; ACN, acetonitrile; Boc, *tert*-butoxycarbonyl; DCM, Dichloromethane; DIPEA, diisopropylethylamine; DMF, *N,N'*-dimethylformamide; DTP, 2,2'-Dithiopyridine; ESI-MS, electrospray ionization mass spectrometry; Fmoc, fluorenylmethoxycarbonyl; HATU, 1[*Bis*(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate; LC/MS, liquid chromatography/mass spectrometry; MeOH, methanol; nAChR, nicotinic acetylcholine receptor; NMR, nuclear magnetic resonance; Pbf, pentamethyl-dihydrobenzofuran-5-sulfonyl; RP-HPLC, reversed phase high performance liquid chromatography; SPPS, solid phase peptide synthesis; t-Bu, *tert*-butyl; TFA, trifluoroacetic acid; TIS, triisopropylsilane; Tris, 2-Amino-2-(hydroxymethyl)propane-1,3-diol;Trt, trityl; UV, ultra-violet.

2.2 Chemical synthesis

DMF, DIPEA, ACN, TIS, TFA, piperidine and all others reagents were obtained from Sigma-Aldrich (Saint-Louis, MI, USA) or Merck (Darmstadt, Germany) and were used as supplied. Fmoc (L) amino acid derivatives and HATU were purchased from Iris Biotech (Marktredwitz, Germany). AmphiSpheres™ 20 HMP resin (0.6 mmol/g) and Amphispheres™ 40 RAM (0.4 mmol/g) were purchased from Agilent Technologies (Les Ulis, France). The following side-chain protecting groups were used: Trt for Asn and Cys; *t*Bu for Ser, Tyr and Glu; Boc for Lys; and Pbf for Arg. Peptides were manually synthesized by using the Fmoc-based solid-phase peptide synthesis technique on a VWR (Radnor, PA, USA) microplate shaker. All Fmoc amino acids and HATU were dissolved in DMF to reach 0.5 M. For acid peptides the first residue was anchored on 20 HMP resin using the method described by *Grandas et al*. [99]. The resin was washed with DMF, DCM, MeOH, and DMF. Fmoc deprotection was carried out with piperidine in DMF (1/2 v/v) twice for 3 min. Subsequent amino acids were coupled onto 0.1 mmol of resin twice for 10 min using an amino acid/HATU/DIPEA ratio of 5:5:10 relative to resin loading. DMF was used for resin washing between deprotection and coupling steps. After chain assembly was complete, the terminal Fmoc group was removed and the resin washed with DMF and DCM. Side-chain deprotection and cleavage from the resin was carried out by adding 10 mL of TFA/TIS/H_2O (95/2.5/2.5

v/v/v) and stirring the mixture for 2.5 h at room temperature. After the resin was removed by filtration and washed three times with dichloromethane. Dichloromethane and TFA are removed under vacuum then cold diethyl ether was added to precipitate the peptide. The disulfide bridge is formed between the free cysteine residues by dissolving the peptide at 0.2 mM in 50 mM Tris-HCl buffer adjusted to pH 8 and adding dropwise 7 equivalents of DTP at 10 mM in MeOH. When reaction was complete, the reaction mixture was acidified to pH 3 and loaded onto preparative RP-HPLC and pure fractions were combined. The combined pure fractions were freeze-dried and their purity were confirmed by LC/ESI-MS.

2.3 Mass spectrometry

Solvents used for LC/MS were of HPLC grade.

Intermediate products were characterized using a LC/MS system consisting of a Waters (Milford, OH, USA) Alliance 2695 HPLC, coupled to a Waters Micromass ZQ spectrometer (electrospray ionization in positive mode (ESI+) fitted with a quadrupole mass analyzer). All the analyses were carried out using a Chromolith (Fontenay sous Bois, France) HighResolution RP-18e (4.6 x 25 mm, 15 nm–1.15 µm particle size, flow rate 3.0 mL/min) column. A flow rate of 3 mL/min and a gradient of 0–100% B over 2.5 min for routine analyses and 0-30% B over 30 min for quality control of pure products were used. Eluent A was water/0.1% HCO_2H and eluent B consisted of acetonitrile/0.1% HCO_2H. UV detection was performed at 214 nm. Electrospray mass spectra were acquired at a solvent flow rate of 200 µL/min. Nitrogen was used for both the nebulizing and drying gas. The data were obtained in a scan mode ranging from 100 to 1000 m/z or 250 to 1500 m/z to in 0.7 s intervals.

Folded peptides were characterized using a Synapt G2-S high-definition MS system (Waters, Corp., Milford, MA, United States) equipped with an ESI source and an hybrid QTof mass analyzer configuration. Chromatographic separation was carried out at a flow rate of 0.4 ml/min on a Acquity H-Class ultrahigh performance liquid chromatography (UPLC) system (Waters, Corp., Milford, MA, United States), equipped with a Kinetex C18 100Å column (100 mm x 2.1 mm, 2.6 mm particle size) from Phenomenex (France). The mobile phase consisted of water (solvent A) and ACN (solvent B) with both phases acidified by 0.1% (v/v) formic acid. Mass spectra were acquired in the positive ionization mode.

2.4 Preparative RP-HPLC

Preparative RP-HPLC was run on a Gilson PLC 2250 Purification system (Villiers le Bel, France) instrument using a preparative column (Waters DeltaPak C18 Radial-Pak Cartridge, 100 Å, 40 × 100

mm, 15 µm particle size, flow rate 50.0 mL/min). Buffer A was 0.1% TFA in water, and buffer B was 0.1% TFA in acetonitrile.

2.5 Cell culture method and transient expression of human and zebra fish oxytocin and arginine vasopressin receptors

The human oxytocin receptor (hOTR), the human arginine vasopressin receptor V1a (hV1aR), the human arginine vasopressin receptor V1b (hV1bR) and the human arginine vasopressin receptor V2 (hV2R) complementary DNAs (cDNAs) were obtained from OriGene Technologies. The corresponding arginine vasopressin zebra fish (ZF) receptors; arginine vasopressin receptor 1Ab (V1a1, The National Center for Biotechnology Information (NCBI) accession number NP_001284605.1), arginine vasopressin receptor 1Aa (V1a2, NCBI accession number NP_001288043.1), arginine vasopressin receptor 2 (V2, NCBI accession number NP_001103595.1) and the oxytocin receptor (OTR, NCBI accession number NP_001186299.1), were synthetically synthesized by GenScript.

COS-1 cells (American Type Culture Collection (ATCC)) grown in Dulbecco's modified Eagle's medium (DMEM) and 5% fetal bovine serum (FBS) were transiently transfected with plasmid DNA encoding the hOTR, hV1aR, hV1bR, hV2, ZF V1a1, ZF V1a2, ZF V2 and ZF OTR using FuGENE HD in a 1:3 ratio of DNA and FuGENE, following the manufacturer's protocol.

2.6 FLIPR assay measuring intracellular Ca^{2+} responses

24 h post-transfection, transiently transfected COS-1 cells were seeded at a density of 15,000 cells/well in black-walled imaging plates (Corning, Sigma Aldrich) and maintained for another 24 h at 37°C in a 5% humidified CO_2 incubator. Assay measuring ligand-induced Ca^{2+} responses was performed 48 h post-transfection. On the day of the assay, cells were loaded with the Calcium 4 No-wash dye (Molecular Devices) by diluting the lyophilized dye in physiological salt solution (PSS: 140 mM NaCl, 11.5 mM glucose, 5.9 mM KCl, 1.4 mM $MgCl_2$, 1.2 mM NaH_2PO_4, 5 mM $NaHCO_3$, 1.8 mM $CaCl_2$, 10 mM HEPES, pH 7.4), and incubated for 30 min at 37°C in a 5% humidified CO_2 incubator. Intracellular Ca^{2+} responses were measured in response to ligands in a Fluorometric Imaging Plate Reader (FLIPR) (Molecular Devices) using a cooled CCD camera with excitation at 470–495 nM and emission at 515–575 nM. Camera gain and intensity were adjusted for each plate to yield a minimum of 1000 arbitrary fluorescence units (AFU) baseline fluorescence. Prior to addition of control agonists or conopressins, 10 baseline fluorescence readings were taken, followed by fluorescent readings every second for 300 s. Concentration-response curves were established by plotting Delta F/F_0 values, where F_0 is the base-line level of fluorescence and Delta

F is the change in fluorescence from the baseline level, against agonist concentration using Prism (GraphPad Software). The conopressins were also tested for antagonist activity at the human and ZF receptors. After addition of 10 µM conopressin peptide, cells were incubated for 10 minutes before stimulating the receptors with an EC_{90} concentration of agonists (oxytocin for the hOTR (0.5 µM), vasopressin for hV1a and hV1b (1 µM), vasotocin for ZF V1a1, V1a2, V2 and OTR (0.1 µM)). Changes in fluorescence responses were assessed for 10 s to set the baseline, then 600 s after addition of antagonist and for a further 300 s after addition of agonists, using the FLIPR as previously described.

2.7 LANCE Ultra cAMP assays

Assays measuring cAMP accumulation were performed 48 hours after transfection following the manufacturer's instructions (LANCE Ultra cAMP kit, PerkinElmer). To test for agonist activity at the hV2R, increasing concentrations of control agonists or conopressins (10 pM to 100 µM) were added to 500 transfected cells in stimulation buffer in a white 384-well plate (OptiPlate, PerkinElmer Life Sciences). When testing the conopressins for antagonist activity, conopressins (100 µM) were added in the presence of an EC_{90} concentration of vasopressin (0.1 nM agonist) to the cells as previously described. The plates were incubated for 30 min at room temperature. Cells were then lysed by the addition of the europium (Eu) chelate-labeled cAMP tracer and the cAMP-specific monoclonal antibodies labeled with the ULight dye, diluted in cAMP detection buffer (LANCE Ultra cAMP kit, PerkinElmer), followed by incubation for 1 hour at room temperature. The emission signals were measured at 615 and 665 nm after excitation at 340 nm using a Tecan microplate reader (Tecan).

3. Results

Two new conopressin-related sequences were identified in the venomous gland transcriptome of *Conus miliaris*. Both had the characteristic mature nonapeptide containing one disulfide bond forming a cyclic sequence of 6 amino acids plus three exocyclic residues. Surprisingly, both sequences had an aspartic acid instead of the highly conserved basic residue at position 8 (Figure 1A). Considering this unusual substitution is present in the functionally critical exocyclic region, it was of interest to investigate the pharmacological and structural properties of these novel conopressins.

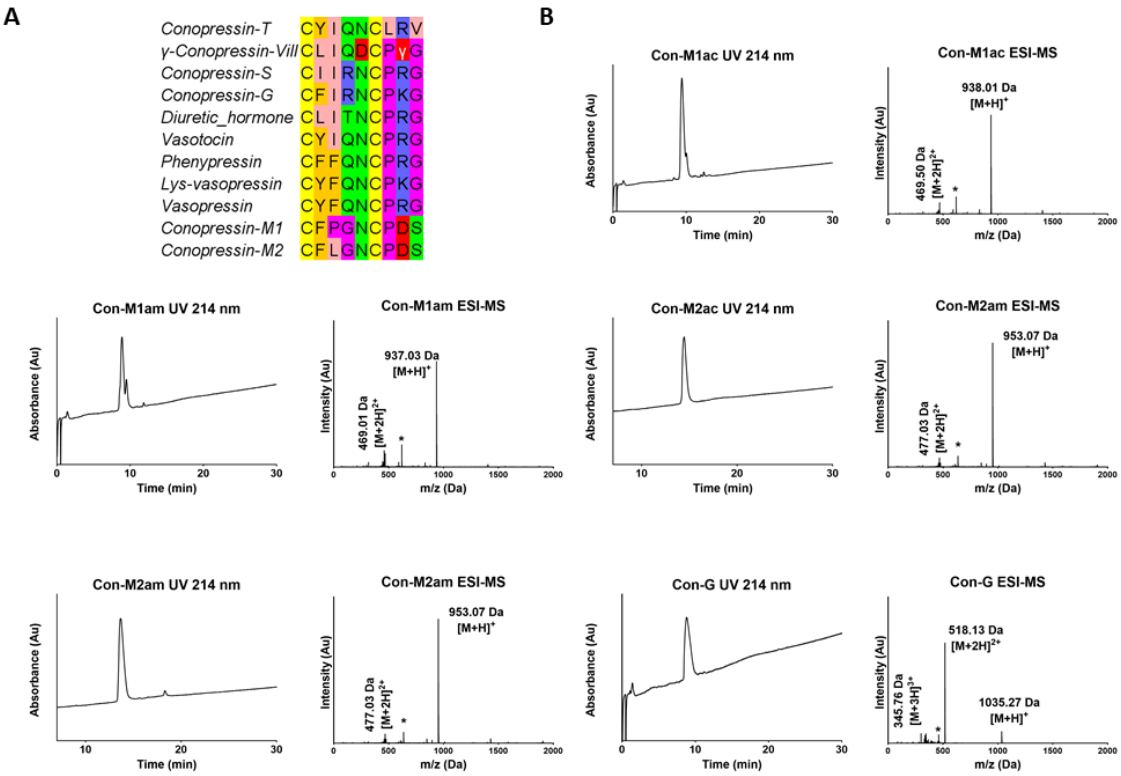


Figure 1. Alignment of conopressin-related sequences and RP-HPLC/ESI-MS analyses of the synthesized conopressins (A) Alignment of conopressin-related sequences. All the sequences display an amidated C-terminal. Conopressin-M1 and M2 with γ-Conopressin-vill are the only ones that display a negatively charged amino acid at position 8. Interestingly, Conopressin-M1 also displays an unusual proline residue at position 3. Highly conserved glycine residue at position 9 is replaced by a serine residue in Conopressin-M1 and M2. (B) RP-HPLC/ESI-MS analyses of the synthesized conopressins. ACN gradient from 0 to 30% over 30 min. The ac ending indicates an acidic C-ter and the am ending indicates a C-ter amide. For Con-M1 the two peaks display the same mass thus, the dynamic conformational exchange leading to the splitting of UV chromatogram peak is probably caused by the two proline residues inducing cis-trans isomerization [191–193]. The asterisk * on ESI-MS insets indicates an ion resulting from in source fragmentation of the proline residue [194].

3.1. Chemical synthesis

All conopressins discovered so far have an amidated C-terminal, yet the precursor of conopressin-M1/M2 sequences do not display the usual Gly_{37},Lys_{38},Arg_{39} (Figure 2) motif which is a typical enzymatic recognition site where the glycine residue at position 37 is enzymatically converted to a C-terminal amide group [195]. Therefore, without MS evidence for one or the other we

synthesized both C-terminal amide (Con-M1/M2am) and acid (Con-M1/M2ac) versions of the peptides. Con-G was also synthesized as its yet three-dimensional structure and pharmacological characterization are incomplete. After RP-HPLC purification of the folded peptide, their homogeneity was assessed by analytical RP-HPLC and MS (Figure 1B). Pure peptide yields were nearly two-fold better for the amidated peptides (around 30% compared to acid peptides around 17%), which can be attributed to the better stability of the amide *vs*. ester bond. Con-G yield was up to 70%, however, it is difficult to conclude if this difference comes from a better folding yield or a better synthesis yield since linear products have not been isolated and purified prior to oxidation.

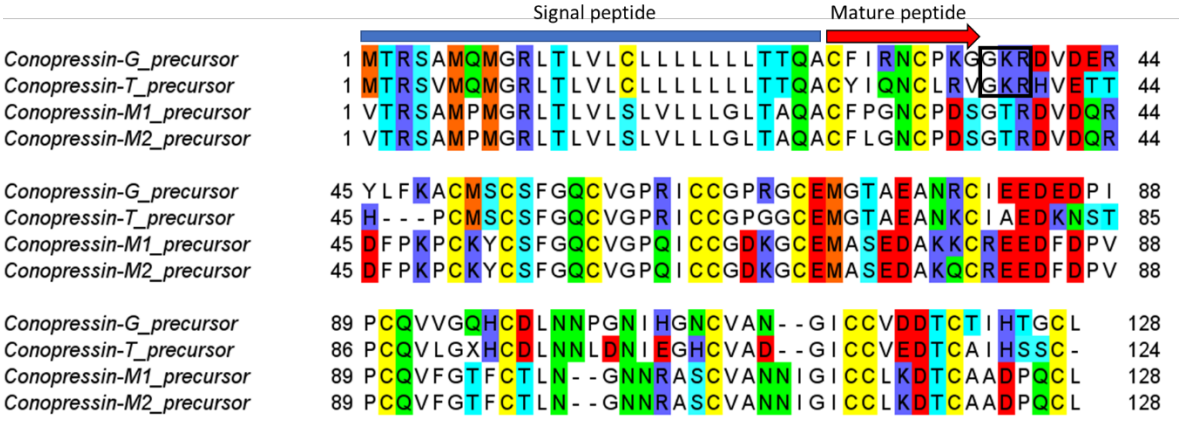


Figure 2. Alignement of conopressin precursors. Enzyme recognition site is black boxed. Conopressin-M1 and -M2 precursors have been identified in our transcriptome of *C. miliaris* based on sequence homology.

3.2. Pharmacological characterization on human and zebrafish receptors

We performed fluorescent imaging plate reader (FLIPR) Ca^{2+} mobilization assay and a second-messenger cyclic adenosine monophosphate (cAMP) assay (for V2R) to determine the biological activity of the four new conopressins and Con-G. Considering a structure-activity relationship (SAR) study in a drug design perspective we investigated the agonist and antagonist activity of the conopressins on human oxytocin receptor (hOTR), human vasopressin-1a receptor (hV1aR), human vasopressin-1b receptor (hV1bR) and human vasopressin2 receptor (hV2R). However, since *Conus geographus* is a piscivorous cone snail feeding on fish, it was of interest to investigate conopressin activities on zebrafish (*Danio rerio*) receptors e.g. zebrafish vasopressin1-a1 receptor (ZF V1a1R), zebrafish vasopressin1-a2 receptor (ZF V1a2R), zebrafish vasopressin2 receptor (ZF V2R) and zebrafish oxytocin/isotocin receptor (ZF oxy/isoR) (Figure 3 and 4). Oxytocin, vasopressin and vasotocin were used as reference compounds. We can delimitate roughly three order of potency: (i) reference compounds displaying EC_{50} values in the picomolar/nanomolar range, (ii)

Con-G exhibiting potencies in the hundred nanomolar range and *miliaris* conopressins that mostly show no activity except at the ZF V2R in the micromolar range. Unfortunately, there is no activity below the micromolar range for the *miliaris* conopressin peptides, suggesting the negative influence of the conserved residue substitutions which will be further discussed. Interestingly, Con-G is more active on zebrafish (*Danio rerio)* receptors than on its human counterpart receptors, which suggest a role of Con-G in the envenomation process. Con-G acts as a partial agonist at hOTR, ZF V1a2R and ZF oxy/isoR with respectively 28%, 69% and 62% response of control but as a full agonist at other receptors. Con-M1am is the only *miliaris* peptide showing a weak response at hV1b and hV1a with respectively 26% and 58% response of control. Endogenous mammalian peptides (oxytocin and vasopressin) are more active than vasotocin non mammalian peptide at all human receptors except for the OTR. Consistent from an evolutionary point of view vasotocin was more active than endogenous mammalian peptides at all zebrafish (*Danio rerio*) receptors. The EC_{50} values are represented in Table 1. There was no antagonist activity detected for any of the synthesized conopressin up to 10 µM (data not shown).

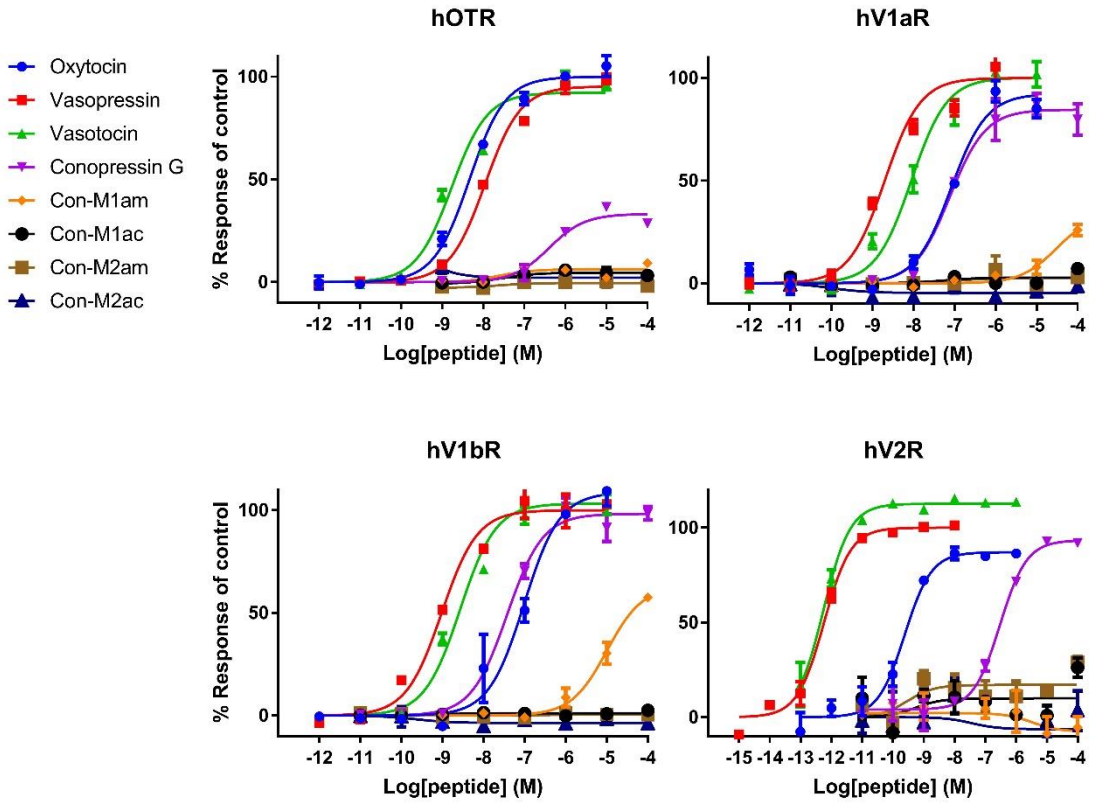


Figure 3. Representative concentration-response curves of all tested compounds against human oxytocin-vasopressin related receptors. Each point represents the mean of measurements from one experiment performed in triplicate. Errors bars represent S.E.M.

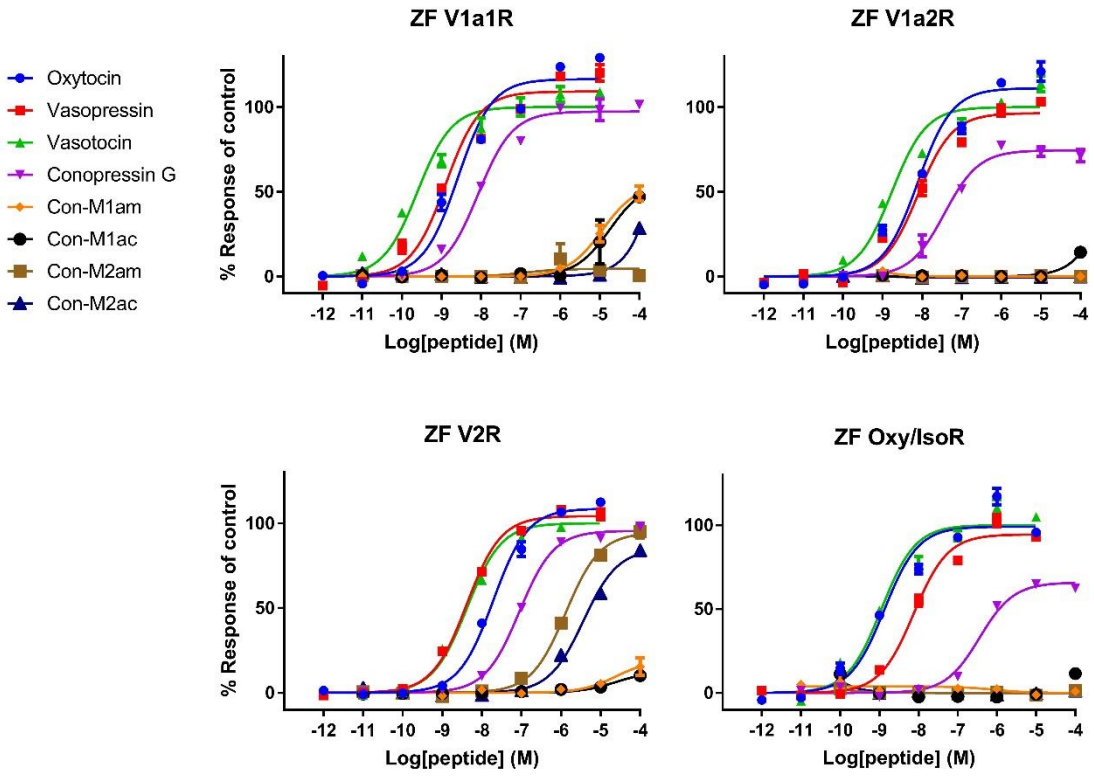


Figure 4. Representative concentration-response curves of all tested compounds against *Danio rerio* (zebrafish) oxytocin-vasopressin related receptors. Each point represents the mean of measurements from one experiment performed in triplicate. Errors bars represent S.E.M.

4. Discussion

Considering the unusual sequence deviation of *miliaris* conopressins compared to oxytocin/vasopressin related peptides i.e. proline residue at position 3 of Con-M1 instead of usual hydrophobic residue, glycine at position 4 rather than glutamine/arginine, negatively charged aspartic acid at position 8 in place of conserved basic residue and the lack of terminal glycine residue (Figure 1A) it was of interest to investigate their structure and pharmacological characterization. Since it has been demonstrated that a little change in the sequence such as glycine substitution by a valine residue at position 9 switched oxytocin and arginine-vasopressin (AVP) from agonist to antagonist [186], we also investigated the antagonist activity of synthesized peptides unfortunately, none of them shown antagonist activity up to 10 µM. Overall, this study provides valuable structure-activity relationship information for the design of vasopressin receptor selective ligands.

Chemical synthesis generated peptides with good purity and in sufficient yield to perform pharmacological and structural characterization. Interestingly, Con-M1 did not resolve as a sharp UV

peak but instead two partially separated peaks displaying the same mass. One plausible explanation is that the additional proline residue at position 3 of Con-M1 peptides is causing dynamic conformational exchanges between cis-trans proline. The two confomers might have different three dimensionnal structures conferring them different physico-chemical properties. This conformational heterogeneity probably due to cis/trans isomerization has been described in other conotoxins, but at this stage there is no data supporting this hypothesis [191–193]. Therefore, it is not possible to conclude whether only one or both of the two conformations is/are responsible for the activity, or whether each conformation has the same level of activity [196].

By taking vasotocin as a reference and comparing the EC_{50} successively with oxytocin, vasopressin and Con-G, we can extract the following information (Table 1) : (i) absence of a basic residue at position 8 decreases the potency towards all tested receptors especially at the hV2R, (ii) an aromatic residue at position 3 improves the selectivity for hV1aR and hV1bR by decreasing the potency at all other receptors and especially ZF oxy/isoR , (iii) replacing the higly conserved glutamine residue at position 4 by a basic residue reduces the potency for all tested receptors, drastically at hV2R but also significantly at ZF oxy/isoR and hOTR. For the sequence comparison, when considering their chemical properties, tyrosine and arginine residues were considered chemically equivalent to phenylalanine and lysine residues, respectively. We assumed that the EC_{50} values differences were not due conservative substitutions but rather significantly chemically different amino acid substitutions. Based on the X-ray structure of bovine rhodopsin (PDB 1F88), Rodrigo et al. [190] built three-dimensional molecular models of the complexes between AVP and the two receptor subtypes V1a and V1b. Their predictions have been confirmed by receptor mutational studies and identified four key residues that finely tune the binding of vasopressin and related peptide agonists to both receptor subtypes. $Glu^{1.35}$ and $Asp^{2.65}$ (Asp112 of hV1aR and Asp95 of hV1bR, Figure 5) residues are described as key anchoring residues to the Arg^{8}, which is evidenced by higher EC_{50} values of oxytocin, and similarly, the *miliaris* conopressins do not display a basic residue in position 8 (Table 1). Surprisingly, vasopressin and vasotocin are almost 10000-fold more potent on hV2R than ZF V2R, which is not the case for oxytocin, suggesting a crucial role of the basic residue in position 8 to enable a tight interaction with hV2R but not with ZF V2R (Table 1). Rodrigo et al. also identified $Val^{4.61}$ and $Pro^{5.35}$ (Val169 and Pro196 of hV1bR, Figure 5) (as a hydrophobic subsite specific of V1bR according to its high affinity with d[Cha^{4}]AVP [165,190]. The latter probably partly explains why vasopressin, which has a hydrophobic Phe residue at position 3, displays a slightly better affinity at hV1b compared to vasotocin (Table 1). A previous study from Mouillac et al. has shown that hydrophobic parts of the peptide hormone is accommodated by a hydrophobic pocket lying deep in the 7-TM domain delineated between TMIII and TMVI of V1aR [184,190], confirming the slightly better affinity at hV1a compared to vasotocin (Table 1). A tight hydrogen bond network also

greatly contributes to crucial interactions with V1aR and V1bR. Indeed, conserved Gln residues located at the rim of the cavity H-bonds side chains of Gln4 and Asn5 of the peptide hormone [190]. As a result, substitution of vasotocin glutamine residue at position 4 by arginine residue in Con-G leads to a loss of potency at hV1aR and hV1bR (Table 1). Moreover, SAR studies show that a glycine carboxamide moiety is required for biological activity by enabling H-bonds between AVP/OT and Gln214, Gln218 of V1aR (Gln104 and Gln 108 of hV1aR, Figure 5) [184,197] providing another justification for the lack of activity of *miliaris* conopressins where the terminal glycine residue is replaced by a serine residue (Table 1). Therefore, the low activity on all tested receptors of Con-M1 and Con-M2 can be explained by (i) the substitution of the Gln4 by a glycine residue, (ii) the absence of a basic residue in position 8 as well as (iii) the missing glycine residue in position 9. Interestingly, Con-G is more active on zebrafish (Danio rerio) receptors than on its human counterpart receptors, which suggest a role of Con-G in the envenomation process. Consistent with the feeding behavior of C. *geographus* and C. *miliaris*, potencies displayed by Con-G are in the hundred nanomolar range with best potency at ZF V2R (10 nM). However, C. *miliaris* conopressins mostly show no activity except at the ZF V2R in the micromolar range. Indeed C. *miliaris* is a worm hunting cone snail and consequently, the interspecies receptor molecular differences (worm/fish) is most likely the cause of the C. *miliaris* conopressins' inactivity at zebrafish/human receptors. We can hypothesize that C. *miliaris* conopressins were likely designed to target worm key function receptors.

Table 1. Mean EC_{50} (nM) values of all tested peptides on all receptors (3 independent experiments). Amino acid substitutions relative to vasotocin inducing pharmacological properties differences are highlighted. Up arrow or down arrows are relative to vasotocin mean EC_{50} values. Selectivity values are calculated according to the highest EC_{50} value. N.D means not determined because of high value > 100 µM.

Name	Sequence	hOTR	hV1aR	hV1bR	hV2R	ZF V1aR	ZF V1a2R	ZF V2R	ZF oxy/isoR	Selectivity
Vasotocin	CYIQNCPRG*	1.62	15.92	4.26	0.00055	0.41	2.77	5.07	0.85	h2VR (x28945.5)>ZF V1a1R (x38.8)>ZF oxy (x18.7)>hOTR(x9.8)>ZF V1a2R (x5.7)>hV1bR (x3.7)>ZF V2R (x3.1)>hV1aR
Oxytocin	CYIQNCPLG*	4.57↗	89.26↗	84.92↗	0.24↗	3.68↗	8.02↗	15.98↗	1.74↗	h2VR (x371.9)> ZF oxy (x51.3)> ZF V1a1R (x24.3)>hOTR(x19.5)>ZF V1a2R (x11.1)> ZF V2R (x5.6)>hV1bR (x1.1)> hV1aR
Vasopressin	CYFQNCPRG*	8.86↗	3.33↘	1.30↘	0.00058=	1.28↗	8.24↗	4.97=	7.42↗	h2VR (x15275.9)>ZF V1a1R =hV1bR(6.9)>hV1aR (x2.7)>ZF V2R (x1.8)>ZF oxy (x1.2)>ZF V1a2R (x1.1)>hOTR
Con-G	CFIRNCPKG*	455.66↗	123.78↗	51.92↗	299.2↗	10.61↗	44.06↗	61.05↗	353.73↗	ZF V1a1R (x42)>hV1bR (x10.1)>ZF V1a1R (x8.6)> ZF V2R (x7.3)>hV1a2R (x3.6)>hV2R (x1.5)>ZF oxy (x1.3)>hOTR
Con-M1ac	CFPGNCPDS	N.D	N.D	N.D	N.D	116 950	N.D	N.D	N.D	
Con-M1am	CFPGNCPDS*	N.D	N.D	38 194	N.D	13 614	N.D	N.D	N.D	
Con-M2ac	CFLGNCPDS	N.D	N.D	N.D	N.D	N.D	N.D	3 656	N.D	
Con-M2am	CFLGNCPDS*	N.D	N.D	N.D	N.D	N.D	N.D	1 722	N.D	

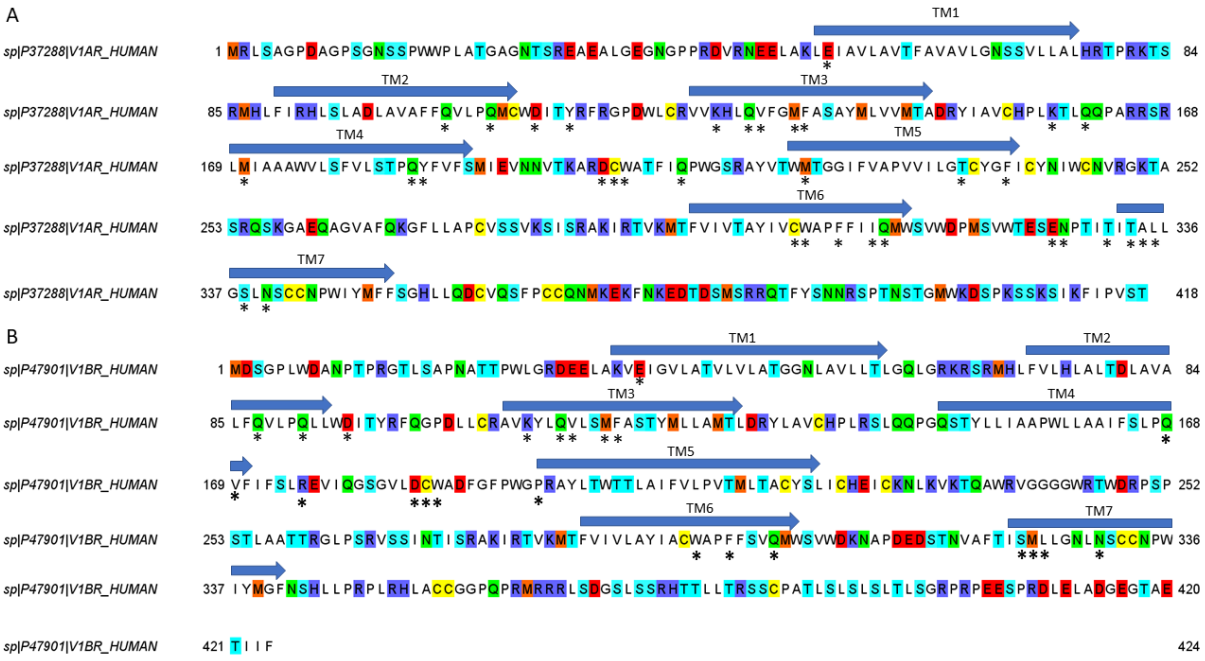


Figure 5. Primary amino acid sequences of (A) human V1A receptor and (B) human V1B receptor. Asterisks indicate amino-acid residues that have been suggested to participate and to be important in receptor–ligand interaction. Arrows indicate the seven putative transmembrane domains (TM 1–7).

Conclusions et perspectives

Des résultats intéressants ont été obtenus concernant les peptides pauvres en ponts disulfures ciblant les RCPGs. Bien que faiblement actives sur les récepteurs testés, les conopressines de *C. miliaris* ont fourni des premiers résultats de relation structure-activité sur les récepteurs de *Danio rerio* (zebrafish) qui pourront être approfondis par la suite avec des études incluant de la modélisation et du docking. Contrairement, aux données publiées dans la bibliographie, les peptides PnPPs n'ont pas montré d'activité agoniste sur le récepteur μ-opioïde mais semble affecter de la même manière le récepteur à l'urotensine, un autre RCPG utilisé en contrôle. Il serait intéressant d'étudier si les peptides sont capables de traverser la membrane cellulaire et d'agir directement sur la protéine G ou alors s'ils agissent sur un autre récepteur qui après une cascade de signalisation influe sur la protéine G.

La majorité des toxines comportant plusieurs ponts disulfures, la suite de ce travail à été la mise en place d'une stratégie permettant la formation de toxines à deux ponts disulfures.

3. Chapitre 2 : Synthèse de toxines à 2 ponts disulfures

a. Introduction

Pour poursuivre notre exploration des toxines de cônes et dans l'optique d'élaborer une stratégie de synthèse régiosélective de toxines à 2 ponts disulfures, nous nous sommes intéressés aux α-conotoxines. Ces toxines ont la particularité de bloquer les récepteurs nicotiniques à l'acétylcholine. Les α-conotoxines sont majoritairement représentées par des séquences relativement courtes (12-20 résidus) comportant deux ponts disulfures, néanmoins leur forte sélectivité et affinité pour leur cible en font de formidables outils pharmacologiques comme décrit dans la revue suivante [20].

α-Conotoxins to explore the molecular, physiological and pathophysiological functions of neuronal nicotinic acetylcholine receptors

Julien Giribaldi and Sébastien Dutertre*

Institut des Biomolécules Max Mousseron, UMR 5247, Université Montpellier - CNRS, Place Eugène Bataillon, 34095 Montpellier Cedex 5, France

Abstract

The vast diversity of neuronal nicotinic acetylcholine subunits expressed in the central and peripheral nervous systems, as well as in non-neuronal tissues, constitutes a formidable challenge for researchers and clinicians to decipher the role of particular subtypes, including complex subunit associations, in physiological and pathophysiological functions. Many natural products target the nAChRs, but there is no richer source of nicotinic ligands than the venom of predatory gastropods known as cone snails. Indeed, every single species of cone snail was shown to produce at least one type of such α-conotoxins. These tiny peptides (10-25 amino acids), constrained by disulfide bridges, proved to be unvaluable tools to investigate the structure and function of nAChRs, some of them having also therapeutic potential. In this review, we provide a recent update on the pharmacology and subtype specificity of several major α-conotoxins.

1. General introduction about α-conotoxins

Cone snails belong to the large Conoidean group of venomous marine gastropods, comprising Turridae, Terebridae and Conidae mollusks. These predatory snails display complex and specific feeding behavior, whether they prey on worms (vermivorous), mollusks (molluscivorous) or fish (piscivorous). They use fast-acting and often paralyzing venoms that is usually injected into their prey or predator through a hypodermic needle-like modified radula tooth [198]. Cone snail venoms contain a mixture of many diverse compounds, including small molecules, peptides and enzymes, yet peptidic toxins called conotoxins largely dominate these cocktails [199]. Typically, each cone snail venom contains 50-200 major conotoxins, which are mostly small disulfide rich peptides (between 10-40 residues and 2-4 disufilde bonds) that synergize together to induce the rapid immobilization of preys [187]. Disulfide bridges provide exceptional structural stability, enabling a tight interaction with their molecular targets, mainly ion channels, transporters and GPCRs [82]. Conotoxins are classified after i) the signal sequence of their precursors, which defines the gene superfamily (A-Q, S, T, V, Y) ii) their cysteine frameworks, which describes the cysteine pattern (I-XXVI) iii) the pharmacology, which is determined

by the pharmacological target (α, γ, δ, ε, ι, κ, μ, ρ, σ, τ, χ, ω) [82]. Among all the different classes of conotoxins, α-conotoxins are the most studied and the most pharmacologically characterized [200,201]. α-Conotoxins target the nicotinic acetylcholine receptors (nAChRs), which are well known targets for other potent animal toxins, such as snake neurotoxins [202]. Whereas a toxin that blocks the muscle type nAChR is an obvious weapon to neutralize preys, there are surprisingly more α-conotoxins that target neuronal nAChRs identified, suggesting that cone snail venoms work in a much more subtle way than simply muscle paralysis. Most of these α-conotoxins act as competitive antagonists of nAChRs, however there are few examples where α-conotoxins show allosteric modulation or even agonist action in certain circumstances. Historically, the bioassays were performed on rat or human nAChRs, although it is expected that conotoxins show a greater potency for their prey's membrane receptors. In this review, we will focus on recent advances on α-conotoxins targeting neuronal nAChRs and their use as pharmacological tools to decipher the function of these receptors in normal and pathological states.

2. Classification and primary structure of α-Conotoxins

α-conotoxins usually belong to the A superfamily and present a type I cysteine framework (CC-C-C, ≈75%) or eventually a type IV (CC-C-C-C-C, ≈10%). However conotoxins targeting nAChRs were found outside the A superfamily (≈20%), including nine other superfamilies (B, D, J, L, M, O1, S, T, and a yet unnamed family) with at least seven more cysteine frameworks. The rules established for naming conotoxins are based on what suggested McIntosh *et al*. Olivera *et al*. and Gray *et al*. [203–206]. The first greek letter specify the conotoxin's pharmacological target, a majuscule letter can be added to indicate the superfamily. The next one or two letters stand for the species from which the toxin was discovered. This letter is followed by a roman number that determines the cysteine framework and finally, a capital letter designates the order of the discovery within that category [206]. However, this nomenclature is not always respected for every conotoxins, mainly for historical reasons. In addition, for the conotoxins discovered by sequencing techniques (e.g. cloning or RNAseq), the cysteine framework and order of the discovery of the clone is noted by Arabic numbers separated by a dot [207]. The classic α-conotoxin contains 12-20 residues with four cysteines connected to form the "globular fold" (Cys_1-Cys_3/Cys_2-Cys_4). They can be further classified based on the number of amino acids in between the cysteine residues (loops). Interestingly, in most case the size and the nature of these cysteine loops seem to predispose the specificity of the α-conotoxin for α-homomeric (3/4), muscle-type (3/5), or neuronal (4/4, 4/6, and 4/7) nAChR subtypes[208].

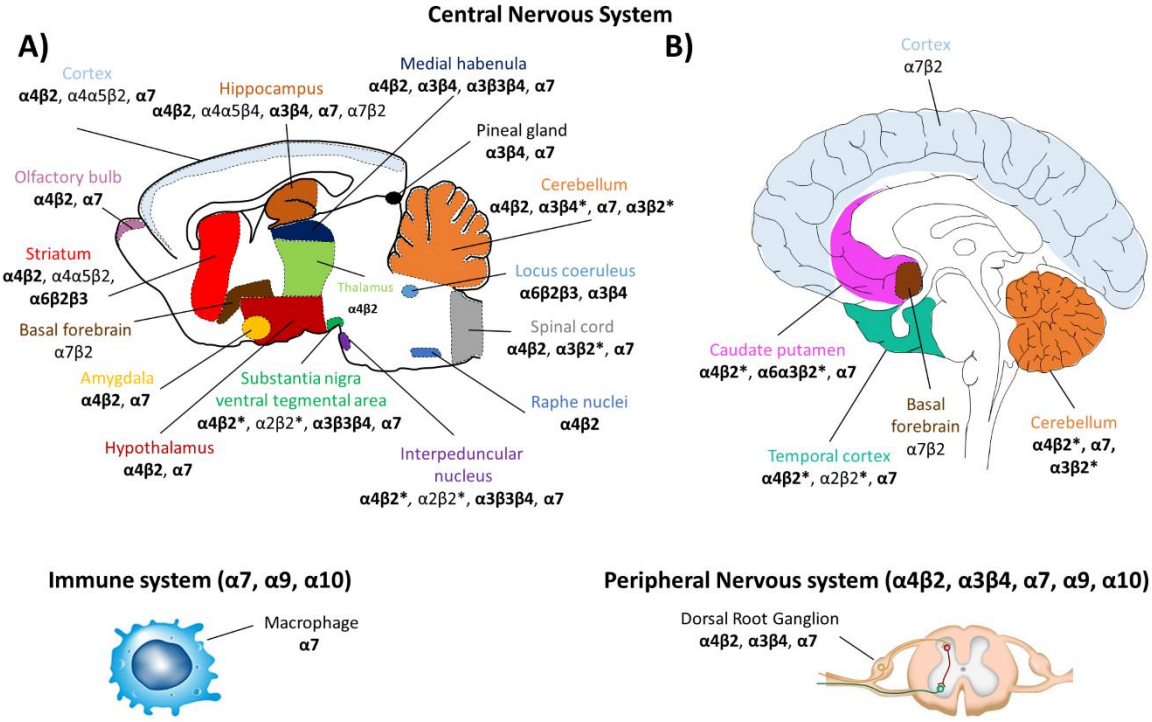


Fig. 1. Distribution of neuronal nAChR subunits in CNS (top panel): A) subtypes identified to date in rodent brain B) subtypes identified to date in human brain, Immune system (left panel) and PNS (right panel). Conotoxin-targeted subunit combinations are shown in bold. Top panel A) was adapted from Gotti *et al.* [209] and B) from Zoli *et al.* [210].

3. α-conotoxins as probes for α7 nAChRs

Homomeric α7 neuronal nAChR has been intensely studied since its original discovery [211]. In the central nervous system, this subtype is mainly distributed in the hippocampus and the cerebral cortex, in regions associated with learning and memory mechanisms [212] (Fig. 1). Major human pathologies such as epilepsy, myasthenic syndromes, schizophrenia, Parkinson's and Alzheimer's diseases result from a dysfunction of the α7 nAChRs [213,214]. The α-conotoxin's binding sites are localized at the interfaces between identical α7 subunits. Like the snake neurotoxin α-bungarotoxin, which has been used to determine the precise location of α7 nAChRs at cellular and subcellular levels [215], the high selectivity of some α-conotoxins for α7 nAChRs can provide unique opportunities for deciphering the complex pharmacology of these receptors. One unusual example is the [A10L]PnIA which has been described by Hogg *et al.* as acting as an antagonist on the wt-α7 but as an agonist on the α7-L247T mutant [216]. One plausible hypothesis is that [A10L]PnIA mutant has the capacity to stabilize the desensitized form of the receptor, which is not conducting for the wt-α7 but conducting for the α7-L247T mutant. Since many channelopathies are due to a mutation that modifies the biophysical properties of the channel, such α-conotoxins that can act differently on mutated neuronal nAChRs

(agonist versus antagonist) may have potential as therapeutics[217]. Ellison *et al.* reported α-conotoxin ImII [218] to be structurally closely related to previously described α-conotoxin ImI [219] although they both block α7 with similar affinity (respectively 571 nM and 595 nM). However, in contrary to α-conotoxin ImI, α-conotoxin ImII does not compete with radiolabeled α-bungarotoxin (a classical competitive inhibitor of the α7 nAChR). This suggests that the binding site of ImII in α7 is different from that of α-bungarotoxin (at the interfaces between identical α7 subunits) and ImI. Indeed, two mutations (Q117S and N111S) within the orthosteric ligand binding site of the α7 nAChR were shown to strongly affect antagonism potency of α-ImI but not α-ImII [220]. Therefore, α-conotoxin ImII represents a unique probe for targeting a novel antagonist-binding site, or microsite, on the α7 nAChR. α-conotoxin ImI has been used as a template molecule for the design of synthetic analogues with a significant higher binding affinity and antagonist activity at the α7 nAChR [221]. For instance, [P6/5-(R)-Phenyl]ImI exhibited a K_i value of 0.19 µM at the α7/5-HT_{3A} (*vs.* 1.2 µM for the wild type) and an IC_{50} value of 0.7 µM (*vs.* 2.6 µM for the wild type) at the α7. The authors demonstrated, by studying the binding site with the AChBP, that there were additional hydrophobic interactions at proline-modified residue in position 6. Armishaw *et al.* used combinatorial chemistry to generate many synthetic variants of α-conotoxin ImI and three analogs containing the Nva9–Dmt10–His11 (Dmt: 2,6-dimethyltyrosine), Leu9–Aph10–Abu11 (Abu: α-aminobutyric acid), and Nva9–Dmt10–Trp11 combinations exhibited ~12-, 14- and 10-fold increases in α7 nAChR inhibition respectively, when compared to wild type ImI [222]. α-conotoxin ArIB was synthesized directly from a cDNA library of *Conus arenatus* and its synthetic variant [V11L,V16D]ArIB was 800- to >1000-fold more potent against α7 than all other nAChR [223] (table 1). This synthetic mutant has been found to reduce both the lactate dehydrogenase level in tumor cells and the inflammatory leukocyte infiltration in tumor tissue in mice[224]. In addition, ^{125}I-α-conotoxin [V11L,V16D]ArIB was proposed and validated as an alternative to radiolabeled α-bungarotoxin as pharmacological tool [225].^{125}I-α-conotoxin [V11L,V16D]ArIB is a highly α7-selective conotoxin confirmed by autoradiography using α7-null mutant tissue. Because the ^{125}I-α-conotoxin [V11L,V16D]ArIB conotoxin has better selectivity for the α7 subtype than α-bungarotoxin (wich also interact α1* and α9α10 nAChRs), there is less unspecific binding, hence it may represent a more reliable tool for binding experiments. Surprisingly, Jin *et al.* reported that conotoxin MrIC is a near full agonist at endogenous human α7 nAChRs in SH-SY5Y cells in the presence of a positive allosteric modulator, with no activity at endogenous α3β2 and α3β4 nAChRs in SH-SY5Y cells [226]. However, it should be noted that this agonist activity could not be confirmed on heterologously expressed α7 nAChRs in *Xenopus* oocytes. Indeed, MrIC acted as a classic antagonist at human α7 nAChRs heterologously expressed in *Xenopus* oocytes, indicating that important functional differences of yet unknown origin exist between neuronal and oocyte expressed α7 nAChRs[21]. Understanding the structure-activity and mode of nAChR activation by MrIC may lead to

the development of novel α7 nAChR modulators with potential for the treatment of a range of neurological disorders.

Table 1: Most selective α-conotoxins for each subtype and their selectivity against others subtypes. Amino acid residues potentially important for the selectivity (or activity) are shown in bold characters. #: amidated C-terminus, O: hydroxyproline residue, γ: carboxyglutamic acid, δ: citrulline, Y*: 3-I-Tyr.

α-conotoxin	Amino acid sequence	Subtype	Affinity (nM)	Selectivity against others subtypes	Refs
[V11L,V16D]ArIB	DECCSNPACR**L**NNPH**D**CRRR	α7	1.09	α6/α3β2β3 (x760) α3β2 (>10 000) others subtypes (>10 000)	[225]
LvIA	RGCCSHPACNV**D**HPEIC#	α3β2	8.67	α6/α3β2β3 (x13) α6/α3β4 (x14) α3β4 (x17) α7 (x346) α2β4 (x1790) others subtypes (>10 000)	[228]
[S9A]TxID	GCCS**HPV**C**A**A**M**S**P**IC#	α3β4	3.9	α6/α3β4 (x46) others subtypes (>10 000)	[229]
[γ4E,V18N]GID	IRDECCSNP**A**CR**V**NNOH**N**C	α4β2	1.85	α3β2 (>10 000)	[230]
[P6O,T5A]BuIA	GCCS**AO**PCAVLYC#	α6/α3β4	58.1	α3β4 (x21) α2β4 (x86) other subtypes (>10 000)	[231]
[S4A,E11A,L15A]MII	GCC**A**NPVCHL**A**HSN**A**C#	α6/α3β2β3	1.2	α3β2 (x83) α3β4 (x317) others subtypes (>10 000)	[232]
[H9A,L15A]MII	GCCSNPVC**A**LEHSN**A**C#		2.4	α3β2 (x2020)	[233]
TxIB	GCCSDPPCRNKHPDLC#		28.4	others subtypes (>10 000)	[234]
[A7V,S9H,V10A,N11R,E14A]PeIA	GCCSHP**VC**H**AR**HP**A**LC#		2.16	α3β2 (>10 000)	[235]
GVIIIB	SGSTCTCFTSTNCQGSCECLSPP GCYC	α9α10	9.8	others subtypes (>10 000)	[236]
RgIA4	SNNGIRQRGCSCTCPGT# GCC**T**DPRC**δY*Q**CY**		1.5	others subtypes (>10 000]	[237]

4. α-conotoxins as probes for α3β2 nAChRs

Cardiovascular functions are modulated by autonomic ganglia that contain neuronal nAChRs, including α3 subunits (α3*). The α3* containing receptors expressed by the nociceptive cells in the dorsal root ganglia are involved in regulation of pain awareness [238]. In the brain, the medial habenula expresses high levels of α3* nAChRs [239] (Fig. 1). Importantly, fear, anxiety, response to stress and nicotine addiction are all linked to the medial habenula [240,241]. Consequently, the habenula's α3* nAChRs are potential pharmacological targets to treat these disorders. When the cholinergic signalling pathway in the medial habenula of mice is compromised and blocked, nicotine withdraw from β4 containing nAChRs [242]. Thus, α-conotoxins with antagonist activity towards the α3* nAChRs delivered to medial habenula could lead to potential drugs against nicotinic addiction, whereas those delivered to the nociceptive cells in the dorsal roots may induce an analgesic effect. However there is some controversy concerning the molecular target responsible for the anti allodynic effect of α-conotoxins since both α9α10 nAChRs and GABAB receptors/N-type calcium channels were proposed

by different groups to be the correct entity [243,244]. Given that α3* nAChRs display high sequence homology to the α6* nAChRs, it is important to test the selectivity of α-conotoxins towards this subtype. The most convincing examples of α-conotoxins that can distinguish between α3* and α6* subunits are the α-conotoxin LvIA and OmIA. LvIA shows a selectivity ratio of 13 and 14 (in favor of α3β2) for the α6/α3β2β3 and the α6/α3β4 respectively [228] (table 1). Luo *et al.* suggested that the negatively charged Asp11 residue in LvIA allows for the discrimination between the different electrostatic properties between α3* and α6* subtypes [228]. Two years later, Zhu *et al.* used LvIA produced by recombinant expression in *Escherichia Coli* to evaluate its analgesic properties in a mouse hot-plate test model of pain [245]. They showed that rLvIA significantly increased the reaction time (latency) to the thermal stimulus at 30, 90, 120 min after treatment. On the other hand, α-conotoxin OmIA blocks α7 nAChRs subtype but also α3β2 and α6/α3β2β3 with IC_{50} values of 27.1 nM, 11 nM and 201 nM, respectively, which corresponds to a selectivity ratio α3β2 *vs.* α6/α3β2β3 of 18. The molecular interactions that favor selectivity for the α3* subunit instead of α6* remain undetermined, however, Dutertre *et al.* have shown that the size of the hydrophobic side chain of the residue in position 10 contributes to the interaction at the complementary subunit (β2 vs. α7 subunit)[246]. Overall, OmIA could be used as a structural template to produce ligands with improved α3β2 nAChRs selectivity, potentially leading to substantial therapeutic compounds. Kompella *et al.* used α-conotoxin RegIIA and its more potent analog [N11A, N12A] to probe the pharmacological properties of rat and human α3β2 nAChRs [247]. Interestingly, a 70-fold lower potency was observed at human α3β2 compared to rat α3β2. Receptor mutagenesis and molecular dynamics studies revealed that this dramatic difference of selectivity was due to a single amino acid change: a glutamic acid residue at position 198 on the rat α3 instead of a proline in the human subunit. α-conotoxin MII has also been used through alanine scan in order to determine the important residues for the interaction with α3β2 nAChRs [248]. It was found that N5A, P6A, and H12A mutations resulted in a >2700-fold, 700-fold, and ~2700-fold loss in potency respectively. The weak potency of [P6A]MII mutant could be explained by the loss of secondary structure initially induced by the proline residue but the two others mutants displayed a structure identical to MII, indicating that it was the nature of sidechains that was responsible for the altered activity. In particular, histidine and asparagine at position 12 and 5, respectively, are key residues for the interaction. Protonated Histidine and possibly Asparagine could interact with negatively charged β2E61 and β2D169 [248] (Fig. 2). Both Napier *et al.* and Klimis *et al.* reported an anti allodynic action of α-conotoxin MII [249,250]. Indeed, α-conotoxin MII, when administrated peripherally in allodynic rats, exhibited a dose dependent increase of mechanical paw withdrawal threshold (PWT) [250]. Furthermore, intrathecal administration into neuropathic rats reduced mechanical allodynia for up to 6 hours without unwanted side effects [249]. Since α-conotoxin MII do not interact with α9α10 nAChRs

or $GABA_B$ receptors/N-type calcium channels, these studies indicate that α3* targeting conotoxins are potentially valuable for the treatment of neuropathic pain [249,250].

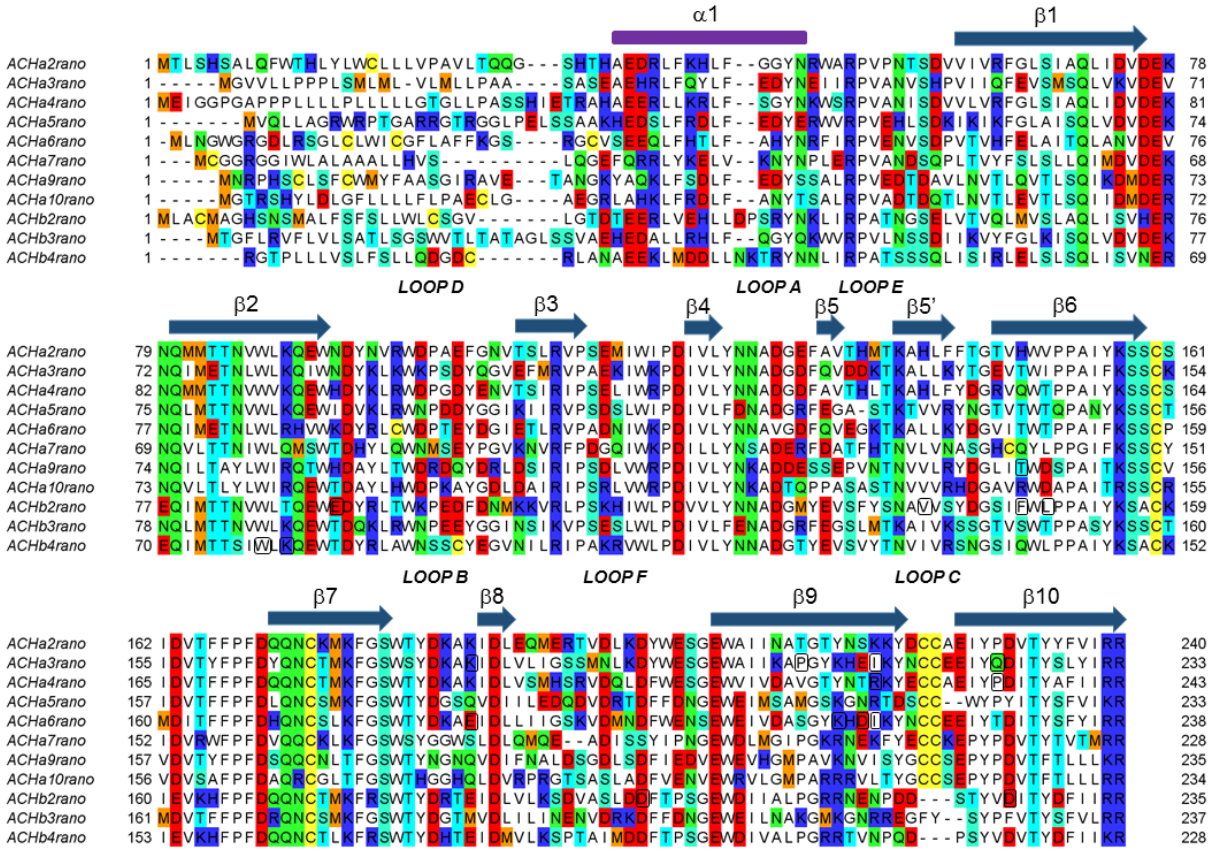


Fig 2. Alignment of neuronal nAChR subunit sequences (extracellular binding domain). Residues involved in the affinity and/or selectivity of conotoxins are boxed in black.

α-conotoxin MII was also used to highlight the involvement of α3β2/α6* nAChRs in the regulation of metabolic processes and cell-cell interactions related to carcinogenesis and tumor-associated inflammation [224]. Accordingly, MII injections led to a significant reduction in Ehrlich tumor growth. Dutertre *et al.* used an Ls-AChBP binding assay to screen *Conus textile* venom and isolated α-conotoxin TxIA, which was highly potent on α3β2 nAChRs (IC_{50} 3.6 nM) [166]. Co-crystallization of [A10L]TxIA with Ac-AChBP showed a critical salt bridge between Arg5 and an acid aspartic residue from Ac-AChBP. The highly similar sequence PnIA loses its activity on α3β2 and Ls-AChBP when this arginine in position 5 is mutated, suggesting that the mode of interaction between AChBP and nAChR is conserved. Dutertre *et al.* also used nAChR structural models based on AChBP to determine the role of the β2 subunit in the binding of three α4/7-conotoxins, namely MII, PnIA and GID [246]. It appears that all three conotoxins target a small binding cleft in the β2 subunit, and point mutations in this region identified two important residues, namely β2L119 and β2V109 (Fig. 2). There is a direct hydrophobic interaction between the conotoxin's residue at position 10 and a conserved proline residue in the receptor's cleft

that locks the ligand, which is consistent with docking simulations. In addition, Everhart *et al.* used the preferred selectivity of α-conotoxin PnIA for α3β2 over α2β2 subtype to delineate the molecular basis for this selectivity [251]. Three key residues Pro182, Ile188 and Glu198 of the amino-terminal extracellular domain of the α3 subunit have been identified by directed mutagenesis (Fig. 2). They are located on the loop C (β strands 9 and 10) of α3, with isoleucine 188 positioned the nearest to the ACh binding pocket. Recent co-crystallization experiments of α-conotoxin GIC with Ac-AChBP highlighted two key conotoxin residues His5 and Gln13 [252]. His5 interacted with Tyr91, Tyr186 and Tyr193 of the Ac-AChBP, whereas Gln13 was nested in a pocket formed by residues Arg57, Val106, Thr108, Ser112 and Met114. The importance of these two residues has been demonstrated experimentally using analogues of GIC and their ability to interact with Ac-AChBP and α3β2 nAChR. It appears that the >600-fold selectivity for α3β2 over α3β4 is mainly due to the steric clashes between Gln13 residue and Arg115 residue of the human β4 subunit, which is replaced with threonine in Ac-AChBP and serine in the human β2 subunit. These studies show the importance of co-crystallization and docking experiments with AChBPs (and nAChR models) for characterizing the interactions in order to design α3β2 selective ligands. By using molecular docking and point mutation, Luo *et al.* could explain the selectivity of α-conotoxin LtIA for α3β2 *vs.* α3β4 nAChRs [253]. They show that LtIA use a shallow binding site that is different from the binding site of α-conotoxin MII mainly because LtIA lacks the typical Ser4-Xaa-Pro6 pattern of others neuronal active α-4/7 conotoxins. This study suggests that even conserved parts of conotoxins of the same subfamily can be altered to modulate their selectivity or probe new binding site.

5. α-conotoxins as probes for α3β4 nAChRs

The α3β4* nAChRs are involved in noradrenaline release in the hippocampus [254]. The pineal gland expresses high level of α3β4 nAChRs[255]. α3β4 and α3β2β4 have also been found in the cerebellum, the retina, the interpeduncular nucleus and the medial habenula [256,257] (Fig. 1). The α3β3β4 combinaison is selectively present in the habenulo-interpeduncular pathway and can modulate the reinforcing effect of addictive drugs and the anxiolytic effect of nicotine [257]. Furthermore, the α3β4 nAChR subtype has been linked to several pathophysiological disease conditions such as lung cancer, nicotine addiction and drug abuse [213,258,259]. However, because very few specific probes of α3β4 nAChRs exist, it is difficult to determine their pathological and pathophysiological roles. The first identified α3β4 selective α-conotoxin was AuIB, and despite its low affinity (IC_{50} 750 nM) it has been widely used to probe this particular subtype [260]. Dutton *et al.* described that the ribbon isomer of α-conotoxin AuIB is less structured (the helical nature of native AuIB is lost) but shows ten times greater potency on dissociated neurons of the rat parasympathetic ganglia [261]. However, it appears that for the AuIB ribbon the maximum inhibition is only 50% because it is α/β stoichiometry dependent, which

is not the case for native globular AuIB [95]. The binding between AuIB ribbon and low sensitivity α3β4 (10α:1β) is reversible by increasing the agonist concentration. On the contrary, the block of high sensitivity α3β4 (1α:10β) by native globular AuIB was insurmountable. Docking studies suggests that AuIB ribbon bind to the orthosteric binding site while the native AuIB bind to both receptors (high and low sensitivity α3β4) outside of the ACh binding pocket in a non-competitive way [95]. These two studies show that disulfide isomers are potentially valuable tools to further determine ligand-receptor interactions and for the design of novel conopeptides with different inhibitory mechanism of action. Napier *et al.* show that intrathecal administration of α-conotoxins MII, Vc1.1 and AuIB into neuropathic rats reduced mechanical allodynia for up to 6 hours without significant side effect [249]. Despite the fact that AuIB was the least potent among these three conotoxins, they made the hypothesis that because of their weak activities on α9α10, $GABA_B$ receptor, N-type calcium channels, the anti-allodynic action comes from interactions with α3* nAChRs. Grishin *et al.* performed an alanine scan of AuIB and mutagenesis of α3β4 nAChR to identify molecular determinants responsible for the interaction [262]. Two mutants lost their activities for α3β4 nAChRs, including [P6A]AuIB which showed low affinity mainly due to a loss of secondary structure. Phenylalanine residue in position 9 seems to play a major role in binding. Indeed, molecular dynamic and point mutation revealed that residues Trp59 and Lys61 form a hydrophobic binding pocket that interacts with phenylalanine in position 9 of AuIB in the loop D of β4 subunit (Fig. 2). For another α-conotoxin, namely RegIIA, molecular dynamic simulations of Cuny *et al.* suggest that the key interaction with the β subunit are different [263]. From these simulations, β4 loop D residue Lys59 and loop E residue Arg113 appear to form hydrogen bonds with RegIIA residues in loop 2 explaining why RegIIA is more α3β4 selective over α3β2, whereas the side chains of β2-T59 and β2-S113 were not long enough to form such interactions (Fig. 2). However, Franco *et al.* in 2012 found that the selectivity of RgIIA is slightly in favor of α3β2 subtype [264]. Intriguingly, a synthetic mutant of RegIIA in which Asn11 and Asn12 of loop 2 were replaced by alanine exhibited a 3-fold increase in selectivity for α3β4 over α3β2 nAChR subtype [265]. Another α4/6-conotoxin, TxID, has been recently discovered and it is much more potent (IC_{50} 12.5 nM) at α3β4 nAChRs than α-conotoxin AuIB, but it also targets closely related α6/α3β4 nAChRs with an IC_{50} 7.5 fold lower [266]. Taken in consideration that the first loop of TxID is also present in others conotoxins, its selectivity likely comes from its unique second loop. It appears that Pro13 residue plays a key structural role and that an isoleucine residue in position 14 is responsible for its higher potency compared to AuIB. Very recently, Wu *et al.* redesigned α-conotoxin TxID to produce the most selective ligand for α3β4 over α6/α3β4 to date [229] (table 1). Indeed [S9A]TxID mutant exhibits a 46-fold greater potency at α3β4 over α6/α3β4 with no loss of activity at α3β4 compared to the native TxID. Molecular dynamic simulations revealed that serine residue in position 9 forms a hydrogen bond with β4K81 in α6β4 but not in α3β4 (Fig. 2). Remarkably, combinatorial chemistry based on α-conotoxin BuIA lead to the

discovery of a synthetic mutant TP-2212-59 with low nanomolar inhibition constant (IC_{50} 2.3 nM) at α3β4 subtype but no inhibition at α7 or α3β2 nAChRs, although Chang *et al.* did not test this peptide on α6β4 subtype [267].

6. α-conotoxins as probes for α4β2 nAChRs

The neuronal nAChRs α4β2 subtype is the most abundant in the human brain and particularly in the cortex, striatum, superior colliculus, lateral geniculate nucleus and cerebellum [215,256,268] (Fig. 1). It plays a key role towards the effectiveness of synaptic connection by regulating the release of other neurotransmitters [209,269,270]. The α4β2 nAChR also plays an important role in nicotine addiction and cognitive performance [209,269–271] which makes it a valuable pharmacological target for improved cognition, smoking cessation, analgesia, and possibly for a wide range of neurological disorders such as Alzheimer's and Parkinson's disease, depression, and attention deficit disorders [272–274]. Only a few conotoxins were shown to interact with α4β2 nAChR. Millard *et al.* used α-conotoxin GID alanine scan to determine the important residues for interactions with the α4β2 subtype [275]. It appears that the Pro9 residue is a structural determinant for the activity on the three subtypes α7, α3β2, α4β2 and that the charge state provided by Arg2 and Asp3 of the unusual N-terminal tail is also crucial. Interestingly, very few mutants retain activity for the α4β2 subtype, suggesting that interaction with this subtype is very specific. In addition to the negative charge induced by the carboxyglutamic acid in positon 4, the positively charged C-terminal amidation also interferes with α4β2 binding. Furthermore, Nicke *et al.* suggest that Arg12 residue in loop II is important for the activity at α4β2 [276]. Because α3β2-F117A and α4β2-F117A mutants both increase the potency of α-conotoxin GID, Dutertre *et al*. suggest that α-conotoxins MII and GID have the same binding mode within the binding pocket, with the β2 subunit having a major role whereas the α subunit is weakly involved [246] (Fig. 2). Molecular dynamic simulation reveals that the binding pathway of [γ4E]GID to α4β2 differs compared to α7, with α4-R185 causing a steric clash with Arg2 of GID [277]. Based on this information, three mutants were proposed that would theoretically improve the selectivity in favor of α4β2 over α7 subtype. Electrostatic interaction between Arg4 of [E4R]GID and β2-196D is particularly promising while [γ4E,R12F]GID and [γ4E,V13R]GID destabilizes the α4β2 subtype less than the α7 subtype (Fig. 2). Unfortunately, none of these putatively α4β2 selective mutants could be validated by Banerjee *et al*. [230]. Other mutants, including [γ4E,A10S]GID and [γ4E,V13I]GID, were found more selective for α4β2 over α3β2, while [γ4E,V18N]GID exhibited no measurable inhibitory activity for the α3β2 subtype, yet it retained the same inhibitory activity as wild type GID towards the α4β2 subtype (table 1). Computational homology modeling highlighted three residues Asp3, Arg12 and Asn14 as crucial for the interaction with the α4β2 subtype. Interestingly, all α-conotoxins targeting the α4β2 subtype also target the α3β2 with greater potency, except for the α-conotoxin TxVC that exhibits a 3-

fold selectivity in favor of α4β2 over α3β2 nAChRs [278]. This selectivity maybe due to the presence of Ile6 and Leu14 as hydrophobic pharmacophore. Beissner *et al.* showed by point mutagenesis that the lower affinity of α-conotoxin MII, TxIA and [A10L]TxIA for α4β2 subtype is mainly due to α4-R185 and α4-P195 [279], residues that are replaced by Ile and Gln in α3β2 (Fig. 2). Overall, more selective and potent probes are required to study the physiological role of this subtype.

7. α-conotoxins as probes for α6* nAChRs

Until now, it was assumed that α6* nAChRs were mainly present in the catecholaminergic nuclei of the central nervous system (Fig. 1). However, recent data indicate that the α6 subunit is widely expressed in the visual pathways and is also present in peripheral tissues [280–282]. Although not very widespread in the brain, the α6* subtype is abundant in midbrain dopaminergic regions which are related to pleasure, reward and mood control [283–286]. Yang *et al.* proposed that α6* nAChRs may assume critical roles in nicotinic reward and in the regulation of mood by nicotine [287]. Quik *et al.* suggested that striatal α6 nAChRs mediate dopamine release, which makes them potential valuable target for neurodegenerative disorders linked to the nigrostriatal system such as Parkinson's disease [288]. As mentioned previously, the α6 subunit is structurally closely related to α3 nAChRs, consequently conotoxins that distinguish between α6* and α3* subtypes were very seldom identified. α-conotoxin BuIA binds the extracellular rat α6β2 *vs.* α4β2 interface with a very significant 60 000-fold selectivity. Kim *et al.* identified by mutation that the region 184-207 contains molecular determinants responsible for this selectivity and more precisely residues α6-Lys185, α6-Thr187 and α6-Ile 188 [289] (Fig. 2). Azam *et al.* proposed a synthetic mutant [P6O,T5A]BuIA, which exhibits a 108-fold selectivity for α6/α3β4 over α6/α3β2β3 and the authors used it to demonstrate that α6β4* nAChRs modulate norepinephrine release in hippocampus but not dopamine release in striatum [231] (table 1). Hone *et al*. also used this synthetic variant to indicate that the predominant heteromeric nAChR expressed by human adrenal chromaffin is constituted of β4 subunit [290] and to determine that α6β4* are among the subtype expressed by dorsal root ganglia [291]. Unlike MII conotoxin that discriminate poorly between α3β2 and α6β2*, α-conotoxin PIA shows a 75-fold lower IC_{50} for α6/α3β2β3 nAChRs and a prolonged k_{off} relative to α3β2 [292]. Off-rate kinetics was dependent on the β subunit suggesting that α-conotoxin PIA possesses binding determinants on the β subunit, since the presence of a β2 subunit versus β4 subunit led to lower IC_{50}. Since the first loop is identical to MII, the reduced α3 subunit affinity must be caused by the N-terminal tail or by differences in its second loop. α-Conotoxin MII has been widely engineered to develop synthetic mutant that were more selective for α6*. In 2004 McIntosh *et al.* proposed a synthetic variant [H9A,L15A]MII with 29-fold higher potency towards α6β4 versus α3β4 subtype, 590-fold higher potency towards α6/α3β2 versus α3β2 subtype and 2020-fold higher potency towards α6/α3β2β3 versus α3β2 subtype [233] (table 1). Four years later Azam *et al.* also develop a

synthetic mutant [S4A,E11A,L15A]MII with 1000-fold selectivity for α6/α3β2β3 over α3β2 and non α6 subunit [232] (table 1). They determined by mutagenesis that Glu152, Asp184 and Thr195 of the α6 subunit were crucial for the activity. More recently, in 2011, based on the observation that the N-terminal tail sequence RDP of α-conotoxin PIA could be responsible for its selectivity for α6 subtypes and that MII-E11 likely binds in the proximity of a highly negatively charged area of the α6 chain, two variants [E11R]MII and RDP-[E11R]MII were synthetized to produce 31-fold and 14-fold lower affinity, and 21-fold and 3.5-fold lower potency for α3β2* over α6β2*, respectively [293]. Head to tail cyclisation of α-conotoxin MII has also been employed by Clark *et al.* to confer resistance to proteolytic degradation without loss of activity [294]. Radiolabeled α-conotoxin MII using ^{125}I has been widely employed to understand the pathophysiological roles of α3/α6 nAChRs in native tissues. Thus, Ray *et al.* demonstrated that the deficit in α6/α3 nAChRs observed in patient with dementia may in part be related to the loss of dopaminergic neurons [295]. It appears that dopaminergic neurons expressed on substantia nigra-ventral tegmental area are nAChR β3 dependent, which in turn, have a physiologically significant role in dopaminergic neurotransmission [296]. Indeed, administration of dopaminergic neurotoxin induced damage to the nigrostriatal system and autoradiographic analysis showed that α-conotoxin MII ^{125}I binding sites in the basal ganglia were all (or mostly) on nigrostriatal dopaminergic neurons [297]. By using ^{125}I-α-conotoxin MII Whiteaker *et al.* identified a novel nAChR population in mouse brain, which was less cytisine sensitive and consequently assumed to be composed of α3*/α6* subunit [298]. Recently described α-conotoxin TxIB display remarkable selectivity towards the α6/α3β2β3 combination with no activity at any other nAChRs subtype [234] (table 1). Although conotoxin TxIB is unlikely to cross the blood-brain barrier, it could be used as a lead for designing molecule targeting α6β2* nAChRs. Hone *et al.* substituted specific amino acids of PeIA with those of MII and PnIA that are known to be critical for the activity and the mutant [S9H,V10A,E14N]PeIA showed a 291-fold selectivity for α6/α3β2β3 over α6/α3β4 nAChRs subunit [281]. This property should allow the toxin to be used across species and prove particularly useful for distinguishing α6β2* from α6β4* in tissues in which these receptors are potentially co-expressed. This synthetic mutant is also 1300-fold more potent at blocking α3β2 than α3β4 nAChRs and may be useful for discriminating between these two subtypes in neurons such as those found in intracardiac ganglia [299] and superior cervical ganglia[300,301], which mostly express a mixed population of α3-containing nAChRs. The same group used positional scanning mutagenesis of α-conotoxin PeIA in combination with mutagenesis of the α6 and α3 subunits to probe the molecular determinants for the interaction [235]. Asn11, which interacts with the α6 and α3 subunit, when replaced with a positively charged amino acid becomes unfavorable for interaction with the α3-Lys152 but not for the interaction with the α6-Glu152 (Fig. 2). This led to the design of [A7V,S9H,V10A,N11R,E14A]PeIA which is 15 000-fold more potent on α6/α3β2β3 than α3β2 nAChRs subtype (table 1).

8. α-conotoxins as probes for α9α10 nAChRs

The α9α10 nAChR subtype is constituted of two α9 and three α10 subunits [302] and has a very limited tissue distribution with no known CNS expression or peripheral nervous system protein expression [303–306] (Fig. 1) but nonetheless has a critical role in the cochlea, auditory system [306–308], immune responses [309,310], pain[243,311] and in cancer therapy [312]. The emergence of analgesic conotoxins targeting nAChR α9 raised the question of the role of this subtype in pain [313]. The involvement of α9 nAChRs in pain is controversial since Mohammadi *et al*. suggested that inhibition of this receptor alone cannot account for the analgesic effects of conotoxin. However, α9 nAChRs may represent a valid target to treat long-term mechanical hyperalgesia, although the precise mechanism remains to be determined. Very recently Romero *et al.* confirmed the expression of α9α10 nAChRs subtypes in immune cells which possibly explain their role in analgesia and inflammation [237]. α-conotoxin PeIA has been used to distinguish between α7 and α9α10 nAChRs subtypes from inner hair cells and confirmed that the efferent fibers-hair cell synapse is most likely mediated by α9α10 nAChRs [314]. However, these results should be interpreted with caution given that conotoxin PeIA also has a nanomolar affinity for the α3β2 subtype. One very α9α10 selective conotoxin is αS-conotoxin GVIIIB, which exhibits low nanomolar potency towards α9α10 and no affinity for others nAChRs subtype [236] (table 1). Christensen *et al.* determined by functional competition studies the importance of the (+) face of the α10 rather than the (+) face of the α9 nAChR subunit as critical for the interaction [236]. Despite its remarkable selectivity as a pharmacological tool, the major disadvantage of αS-conotoxin GVIIIB is its ten cysteine residues, which make its chemical synthesis very challenging. Recently, a dimeric αD-conotoxin GeXXA showed an unusual mode of inhibition of α9α10 nAChRs [315]. Indeed the two C-terminal domains bind to the α9α10 interface at the top surface of the nAChR causing an allosteric perturbation of the receptor. Later, Yang *et al.* investigated the influence of the N-terminal domain and it appears that it is responsible for the inhibitory activity [316]. However the very N-terminus of the N-terminal domain does not directly interact with nAChRs, but after optimization of this N-terminal domain, a fully active single-chain cyclic molecule based on 4 Arg residues was found to interact with nAChRs. This demonstrates that the N-terminal domain of αD-GeXXA acts like a lid covering the pore, which is different from the mode of action of other allosteric ligands. The discovery of this unusual mode of inhibition opens new perspective for the designing of drugs with original mode of action. Another example of a conotoxin that uses a different binding site is α-conotoxin GeXIVA [317]. Interestingly, all three isomers are active, with bead and ribbon folds being nearly equipotent, and the globular fold being less potent. Both bead and ribbon isomers have been used to treat mechanical hyperalgesia in the rat chronic constriction injury model of neuropathic pain [317,318]. α-conotoxin RgIA is one of the most studied α-conotoxin for its analgesic properties. First, Ellison *et al.*

established that the C-terminal half of RgIA was responsible for its preferential inhibition of α9α10 nAChRs [319]. However, it was later determined that Asp5, Pro6, and Arg7 in loop 1 were also critical for the activity on α9α10 and α7 nAChRs, and that Arg9 in loop 2 was crucial for the selectivity towards α9α10 [320]. Azam *et al.* showed that RgIA potency for the rat isoform of the α9α10 subtype is 100-fold greater than for the human isoform, due to a single amino acid substitution (Thr56 located in the (-) complementary binding face instead of Ile56 on human α9α10) [321] (Fig. 2). Point mutagenesis and molecular modeling based on α9α10 extracellular domain crystal structure clearly indicated that the (+) face is located on α10 subunit and the (-) complementary face on the α9 rather than the contrary [322]. The analgesic properties of α-conotoxin RgIA make it an attractive potential therapeutic compound, and consequently chemists attempted to improve its *in vivo* stability. Halai *et al.* found that head to tail cyclisation improved its proteolytic cleavage resistance and its $GABA_B$ potency[323]. Chhabra *et al.* have replaced disulfide bridges with dicarba bridge analogues to make them less sensitive to *in vivo* reduction [324]. Interestingly [3,12]-dicarba RgIA retained α9α10 inhibitory activity but loss its activity on N-type calcium channel. It should be noticed that a similar behavior is observed for α-conotoxin Vc1.1 in which the [2,8]-dicarba Vc1.1 isomer retains activity at $GABA_B$ G protein-coupled receptors, whereas the isomeric [3,16]-dicarba Vc1.1 peptide retains activity at the α9α10 nAChR subtype. Intramuscular injection of α-conotoxin RgIA and Vc1.1 in rodent reduced the chronic-constriction injury (CCI) induced mechanical hypersensitivity and the oxaliplatin-dependant hypersensitivity to mechanical and thermal stimuli [325,326]. In L4-L5 dorsal root ganglia, RgIA prevented morphometric change and reduced the inflammatory infiltrate [326,327]. In the spinal cord RgIA prevented both CCI-induced and oxaliplatin induced activation of microglia and astrocytes [326,327]. Buildup of macrophages and T-cells was also reduced, and these results demonstrate that α9α10 is a critical mediator of peripheral nerve injury-induced immune cell buildup and mechanical hypersensitivity [325]. The molecular basis for the analgesic activity of RgIA is still controverted, with Callaghan *et al.* assuming that RgIA and Vc1.1 modulate native N-type calcium channel by acting as agonist on G-protein coupled $GABA_B$ receptor, potentially mediating their analgesic action [244]. However Romero *et al.* designed RgIA4, which exhibits high potency for both rα9α10 and hα9α10 and is at least 1000 fold more selective for α9α10 *versus* other subtypes including opioid and $GABA_B$ receptors [237] (table 1). Daily dose of RgIA4 prevented chemotherapy-induced neuropathic pain in rat, which demonstrate that α9α10 are basis for the efficacy of RgIA4 and that α9 are a critical target for the prevention of chronic cancer chemotherapy-induced neuropathic pain [237,328]. Using α-conotoxin RgIA, Colomer *et al.* also revealed a novel functional roles for α9α10 receptors in the stimulus-secretion coupling of the medulla [329]. They showed that α9-containing nAChRs contribute to excitatory postsynaptic current and that they dominantly contribute to ACh-induced current in cold stressed rat. α-conotoxin Vc1.1 has a similar mode of action compared to RgIA, and positional scanning

revealed that Asp5, Arg7, Asp11 and Ile15 are key residues for the interaction with α9α10 receptors. Based on this result, Halai *et al.* synthesized five variants that were more potent than native Vc1.1 towards rα9α10 and hα9α10 [330]. By using computational modeling and electrophysiology experiments, Yu *et al.* demonstrated that Vc1.1 binds in the α10/α9 pocket (and not α9/α10) and that a single hydrogen bond formed with residue in position 59 of the pocket is responsible for the specificity to rat *versus* human α9α10 nAChRs [331] (Fig. 2). The first evidence of *in vivo* analgesic properties came with the injection of Vc1.1 in rats suppressing the vascular response to unmyelinated sensory nerve C-fiber activation, such responses being involved in pain transmission [332]. Later Satkunanathan *et al.* showed that Vc1.1 could inhibit pain response associated with peripheral neuropathy in rats *in vivo* and that it could accelerate the functional recovery of injured neurons[311]. Rationally, Clark *et al*. attempted to apply head to tail cyclization to Vc1.1, which surprisingly led to an orally active compound to treat neuropathic pain [333]. However, head to tail cyclization usually lead to better potency towards $GABA_B$ receptors, which, as previously mentioned, may also be the molecular basis of its analgesic action.

Perspectives

This review aims to give a recent update about the utilization of α-conotoxins to probe nAChRs in order to better understand their physiological and pathophysiological roles. Indeed, nAChRs are implicated in various diseases : α7, α4β2 and α6* are involved in Alzheimer and Parkinson diseases, but they also play a role in more complex mechanisms such as learning, memory, epilepsy, schizophrenia, depression, attention deficit, pleasure reward and mood control. α3* and α9α10 are both implicated in pain awareness but remarkably α3* nAChRs are also involved in nicotine addiction, fear, anxiety and response to stress while α9α10 play a role in immune response and cancer biology. Overall, α-conotoxins are not only natural peptides with high affinity for nAChRs, but thanks to their unique selectivity towards particular subunit combinations, they also hold great potential as drug candidate, in particular for the treatment of pain and addiction to nicotine. However, the antagonist mode of action of most α-conotoxins prevents their applications for the treatment of neurodegenerative diseases, where agonist action is required. In the future, significant efforts should be engaged into the discovery and design of new α-conotoxins able to probe nAChRs for which pharmacological tools are still lacking, such as α2*, α5* and the recently described α7β2 subtype [334–340], whose functions and localization are still poorly known.

b. Synthèse de toxines de *Conus catus*

i. Synthèse des toxines CIA et CIB

Malgré la déconvolution de son venin récemment décrite par *Himaya et al.* [341] et une composition riche en α-conotoxines, aucune toxine de *Conus catus* n'a encore été pharmacologiquement caractérisée comme ligand avéré des récepteurs nicotiniques. A partir de cette constatation, nous avons donc synthétisé deux α-conotoxines CIA et CIB choisies en raison de leur identification dans le venin injecté de *Conus catus* [341], de leur composition en acides aminés et de l'arrangement des résidus cystéines présentant une homologie avec d'autres conotoxines déjà caractérisées comme active sur les nAChRs.

Cette étude est présentée sous la forme d'une publication parue [196] dans laquelle j'ai effectué la synthèse des deux toxines CIA et CIB, la caractérisation par spectrométrie de masse et les tests biologiques *in vivo* sur les poissons zèbres (*Danio rerio*).

Synthesis, Structure and Biological Activity of CIA and CIB, Two α-Conotoxins from the Predation-Evoked Venom of *Conus catus*

Julien Giribaldi [1], David Wilson [2], Annette Nicke [3], Yamina El Hamdaoui [3,†], Guillaume Laconde [1], Adèle Faucherre [4], Hamid Moha Ou Maati [4], Norelle L. Daly [2], Christine Enjalbal [1] and Sébastien Dutertre [1,*]

[1] *Institut des Biomolécules Max Mousseron, UMR 5247, Université de Montpellier—CNRS, 34000 Montpellier, France; julien.giribaldi@umontpellier.fr (J.G.); guillaume.laconde@umontpellier.fr (G.L.); christine.enjalbal@umontpellier.fr (C.E.)*

[2] *Centre for Biodiscovery and Molecular Development of Therapeutics, Australian Institute of Tropical Health and Medicine, James Cook University, Cairns, QLD 4878, Australia; david.wilson4@jcu.edu.au (D.W.); norelle.daly@jcu.edu.au (N.L.D.)*

[3] *Walther Straub Institute for Pharmacology and Toxicology, Ludwig-Maximilians Universität, Nußbaumstraße 26, 80336 Munich, Germany; annette.nicke@lrz.uni-muenchen.de (A.N.); elhamdao@uni-mainz.de (Y.E.H.)*

[4] *Département de Physiologie, Institut de Génomique Fonctionnelle, CNRS/INSERM UMR 5203, Université de Montpellier, 34000 Montpellier, France; Adele.Faucherre@igf.cnrs.fr (A.F.); Hamid.Moha@igf.cnrs.fr (H.M.O.M.);*

* *Correspondence: sebastien.dutertre@umontpellier.fr; Tel.: +33-4671-43809*

† *Present address: Institute of Pharmacy and Biochemistry, Johannes Gutenberg-University Mainz, Staudinger Weg 5, 55128 Mainz, Germany*

Abstract: Cone snails produce a fast-acting and *often* paralyzing venom that is usually injected into their prey or predator through a hypodermic needle-like modified radula tooth. Many diverse compounds are found in their venom including small molecules, peptides and enzymes. However, peptidic toxins called conotoxins (10–40 residues and 2–4 disulfide bonds) largely dominate these cocktails. These disulfide rich toxins are very valuable pharmacological tools for investigating the function of ions channels, G-protein coupled receptors, transporters and enzymes. Here, we report on the synthesis, structure determination and biological activities of two α-conotoxins, CIA and CIB, found in the predatory venom of the piscivorous species *Conus catus*. CIA is a typical 3/5 α-conotoxin that blocks the rat muscle type nAChR with an IC_{50} of 5.7 nM. Interestingly, CIA also inhibits the neuronal rat nAChR subtype α3β2 with an IC_{50} of 2.06 µM. CIB is a 4/7 α-conotoxin that blocks rat neuronal nAChR subtypes, including α3β2 (IC_{50} = 128.9 nM) and α7 (IC_{50} = 1.51 µM). High resolution NMR structures revealed typical α-conotoxin folds for both peptides. We also investigated the in vivo effects of these toxins on fish, since both peptides were identified in the predatory venom of *C. catus*. Consistent with their pharmacology, CIA was highly paralytic to zebrafish (ED_{50} = 110 µg/kg), whereas CIB did not affect the mobility of the fish. In conclusion, CIA likely participates in prey capture through muscle paralysis, while the putative ecological role of CIB remains to be elucidated.

Keywords: conotoxins; Conus catus; electrophysiology; in vivo; nicotinic receptors; structure; synthesis

Key Contribution: Two new alpha-conotoxins identified in a previous study from the injected venom of C. *catus* have been synthesized and structurally characterized by NMR. CIB display a typical α-4/7

conotoxin sequence that blocks rat neuronal nAChR subtypes, including α3β2 and α7. CIA is a typical 3/5 α-conotoxin that blocks the rat muscle type nAChR and was highly paralytic to zebrafish.

1. Introduction

Cone snails are predatory marine mollusks comprising more than 800 different species [342]. They capture prey using a venom gland that produces a fast-acting paralyzing venom injected through a hypodermic needle-like radula tooth [198]. Conotoxins, small (10–40 residues) and highly constrained peptides (2 to 4 disulfide bridges), are the main components of cone snail venom, which also contains small molecules and enzymes [199]. Conotoxins are classified into various gene superfamilies based on their conserved signal sequence. Further classification is based on their cysteine framework and their target receptor [343]. Conotoxins often have highly specific and selective biological activity, and many of them proved to be very valuable pharmacological tools or even drug leads [344,345]. Indeed their high selectivity and affinity combined with their small size make these toxins good candidates for the design of therapeutic peptides or peptide mimetics [346,347]. One particular class of conotoxins, the α-conotoxins, acts as antagonists of the nicotinic acetylcholine receptors (nAChRs), a diverse family of ligand-gated ion channels formed by the pentameric assembly of homologous subunits [348].

Both neuronal and muscle type nAChR associated channels open in response to binding of the neurotransmitter acetylcholine [349], and antagonism of nAChR may be of interest for the treatment of pain, nicotine addiction or epilepsy [350,351], whereas agonistic action is desired for treating neurological disorders. The numerous neuronal nAChRs subtypes are involved in a wide variety of biological mechanisms [209,254,352,353] in the central nervous system [209,254] but also in the peripheral nervous system [352] and the immune system [353]. In the central nervous system, they are involved in neurological disorders such as Parkinson's disease, Alzheimer's disease, epilepsy and schizophrenia [213,214], but also in more complex mechanisms such as learning, memory or mood control [212]. Whereas some neuronal subtypes are already validated targets, several combinations of neuronal nAChR subunits remain orphan of selective ligands and therefore more efforts need to be directed at finding or designing novel α-conotoxins with tailored pharmacological profiles.

Historically, α-conotoxins have proved to be remarkable probes for characterizing nAChRs subtypes and establishing their physiological/pathophysiological roles [20,209]. The first α-conotoxins to be characterized were isolated from *Conus geographus* venom gland extracts [354]. Their potent paralytic action at the neuromuscular junction prompted the interpretation of their ecological role as being essential for prey capture [205]. However, recent findings demonstrated that in the case of *C. geographus*, α-conotoxins are injected massively for defense purposes, not for prey capture [198,355]. In order to determine the ecological role of α-conotoxins in other fish-hunting species, we report on

the synthesis, structure determination and biological activities of two new α-conotoxins CIA and CIB discovered in the venom gland transcriptome and confirmed by proteomic analysis to be present in the injected predatory venom of the piscivorous species *Conus catus* [341]. Both CIA and CIB were investigated at the functional and structural level, as well as tested for in vivo effect on fish.

4. Methods

2.1. Abbreviations

Acm, acetamidomethyl; ACN, acetonitrile; Boc, *tert*-butoxycarbonyle; CHCA, α-Cyano-4-hydroxycinnamic acid; DCM, Dichloromethane; DIPEA, diisopropylethylamine; DMF, *N*,*N*'-dimethylformamide; DTP, 2,2'-Dithiopyridine; eq, equivalent; ESI-MS, electrospray ionization mass spectrometry; Fmoc, fluorenylmethoxycarbonyl; HATU, 1 [*Bis*(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate; LC/MS, liquid chromatography/mass spectrometry; MALDI, Matrix Assisted Laser Desorption/Ionization; MeOH, methanol; nAChR, nicotinic acetylcholine receptor; NMR, Nuclear Magnetic Resonance; Pbf, pentamethyl-dihydrobenzofuran-5-sulfonyl; RP-HPLC, reversed phase high performance liquid chromatography; SPPS, solid phase peptide synthesis; t-Bu, *tert*-butyl; TFA, trifluoroacetic acid; TIS, triisopropylsilane; Tris, 2-Amino-2-(hydroxymethyl)propane-1,3-diol; Trt, trityl; UV, ultra-violet.

2.2. Chemical Synthesis

DMF, DIPEA, ACN, TIS, TFA, piperidine and all others reagents were obtained from Sigma-Aldrich (Saint-Louis, MI, USA) or Merck (Darmstadt, Germany) and were used as supplied. Fmoc (L) amino acid derivatives and HATU were purchased from Iris Biotech (Marktredwitz, Germany). AmphiSpheres 40 RAM resin (0.37 mmol/g 75–150 µm) was purchased from Agilent Technologies (Les Ulis, France). The side chain protecting groups for amino acids are t-Bu for Asp, Glu, Ser, Thr and Tyr; Trt for $Cys_{3,16}$ of CIB and $Cys_{5,15}$ of CIA, Acm for $Cys_{2,8}$ of CIB and $Cys_{4,9}$ of CIA; Trt for Gln; Pbf for Arg; Boc for Lys and Trp. CIA and CIB were manually synthesized by using the Fmoc-based solid-phase peptide synthesis technique on a VWR (Radnor, PA, USA) microplate shaker. All Fmoc amino acids and HATU were dissolved in DMF to reach 0.5 M. Chain elongation was performed step by step using 0.1 mmol of AmphiSpheres 40 RAM resin. Fmoc deprotection was performed with 20% piperidine in DMF two times, each for 1 min at room temperature, then the resin is washed three times with DMF. Each Fmoc-protected amino acid (5 eq) was coupled in the presence of HATU (5 eq) and *N*,*N*-diisopropylethylamine (DIPEA, 10 eq) in DMF at room temperature for two min. After completion of coupling reaction, the resin was sequentially washed two times with DMF. Cleavage of peptide from the resin and removal of side-chain protecting groups were carried out using 10 mL of a solution containing TFA/TIS/H_2O (95:2.5:2.5, *v*/*v*/*v*) for 15 min at 60 °C. After the resin was removed by filtration and washed three times with dichloromethane. Dichloromethane and TFA are removed under vacuum then cold diethyl ether was

added to precipitate the peptide. Crude peptides were purified by preparative RP-HPLC, and their purity were confirmed by LC/ESI-MS. A twostep oxidation procedure was then carried out. Trt groups were removed during the cleavage step while the Acm protective groups are resistant to cleavage conditions. The first disulfide bridge is formed from free cysteines with the use of DTP and the second disulfide bridge is formed by concomitant deprotection and oxidation of the Acm groups [356,357].

2.3. Mass Spectrometry

Solvents used for LC/MS were of HPLC grade. The LC/MS system consisted of a Waters (Milford, OH, USA) Alliance 2695 HPLC, coupled to a Waters Micromass ZQ spectrometer (electrospray ionization mode, ESI+). All the analyses were carried out using a Chromolith (Fontenay sous Bois, France) HighResolution RP-18e (4.6 × 25 mm, 15 nm–1.15 µm particle size, flow rate 3.0 mL/min) column. A flow rate of 3 mL/min and a gradient of 0–100% B over 2.5 min were used. Eluent A: water/0.1% HCO_2H; eluent B: acetonitrile/0.1% HCO_2H. UV detection was performed at 214 nm. Electrospray mass spectra were acquired at a solvent flow rate of 200 µL/min. Nitrogen was used for both the nebulizing and drying gas. The data were obtained in a scan mode ranging from 100 to 1000 *m*/*z* or 250 to 1500 *m*/*z* to in 0.7 s intervals. MALDI mass spectrometry analyses were performed on an Ultraflex III instrument from Bruker Daltonics (Champs sur Marne, France). Each sample was analyzed from CHCA matrix deposit in positive-ion mode.

2.4. Preparative RP-HPLC

Preparative RP-HPLC was run on a Gilson PLC 2250 Purification system (Villiers le Bel, France) instrument using a preparative column (Waters DeltaPak C18 Radial-Pak Cartridge, 100 Å, 40 × 100 mm, 15 µm particle size, flow rate 50.0 mL/min). Buffer A was 0.1% TFA in water, and buffer B was 0.1% TFA in acetonitrile.

2.5. Electrophysiology Measurements

cDNAs of rat nAChR subunits used in this study were provided by J. Patrick (Baylor College of Medicine, Houston, TX, USA) and subcloned into the oocyte expression vector pNKS2. cRNA was synthesized with the SP6 mMessage mMachine Kit (Ambion, Austin, TX, USA) and adjusted to a concentration of 0.5 mg/mL. nAChR subunit RNAs were mixed in the ratios 1:1 (α3/β2), 5:1 (α4/β2), and 2:1:1:1 (α1/β1/δ/γ). *Xenopus laevis* (Nasco International, Fort Atkinson, WI, USA) oocytes were injected with 50 nl aliquots of cRNA (Nanoject automatic oocyte injector, Drummond Scientific, Broomall, PA). Antagonist dose-response curves were measured as described previously (Dutertre et al., 2005) in ND96 (96 mM NaCl, 2 mM KCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$, and 5 mM HEPES at pH 7.4). In brief, current responses to acetylcholine were measured 1–6 days after cRNA injection and recorded at −70 mV using a Turbo Tec 05X Amplifier (NPI Electronic, Tamm, Germany) and Cell Works software. A standard concentration of 100 µM ACh for α3β2, α4β2 nAChRs and $(\alpha1)_2\delta\gamma\beta$ muscle nAChR and a standard

concentration of 100 µM nicotine for the α7 nAChR was used to keep the data comparable to previous studies. A fast and reproducible solution exchange (<300 ms) was achieved with a 50-µL funnel-shaped oocyte chamber combined with a fast vertical solution flow fed through a custom-made manifold mounted immediately above the oocyte. Agonist pulses were applied for 2 s at 4-min intervals. Following one minute of perfusion directly after the agonist application, peptides were applied in a static bath for 3 min. IC_{50} values were calculated from a nonlinear fit of the Hill equation to the data (Prism version 4.0; GraphPad Software, Inc., San Diego, CA). Data are presented as means ± S.E. from at least three oocytes.

2.6. NMR Spectroscopy

Lyophilized synthetic peptides were resuspended to a final concentration of ~2.5 mM in 90%H_2O:10%D_2O. 2D ^{1}H-^{1}H TOCSY, ^{1}H-^{1}H NOESY, ^{1}H-^{1}H DQF-COSY, ^{1}H-^{15}N HSQC, and ^{1}H-^{13}C HSQC spectra were acquired at 290 K using a 600 MHz AVANCE III NMR spectrometer (Bruker, Karlsruhe, Germany) equipped with a cryogenically cooled probe. All spectra were recorded with an interscan delay of 1 s. NOESY spectra were acquired with mixing times of 200–250 ms, and TOCSY spectra were acquired with isotropic mixing periods of 80 ms. Two-dimensional spectra were collected over 4096 data points in the f2 dimension and 512 increments in the f1 dimension over a spectral width of 12 ppm. Standard Bruker pulse sequences were used with an excitation sculpting scheme for solvent suppression. Slowly exchanging amide protons were detected by acquiring a series of one-dimensional and TOCSY spectra over a 24 h period, following dissolution of the peptides in D_2O. The two-dimensional NOESY spectra of CIA and CIB were automatically assigned and an ensemble of structures calculated using the program CYANA [358]. Torsion-angle restraints from TALOS+ were used in the structure calculations. Distance restraints between the beta carbons and sulfur atoms of the cysteine residues that are disulfide bonded were included in the structure calculations. Restraints were included between Cys I-Cys II and Cys II-Cys IV. The final structures were visualized using Pymol.

2.7. In vivo Bioassay

Zebrafish were maintained under standardized conditions and experiments were conducted in accordance with the European Communities council directive 2010/63. Briefly, CIA and CIB were diluted in milli-Q water and 5 µL of incremental doses were injected intramuscularly into adult zebrafish with a 10 µL Neuros Syringe from Hamilton (Bonaduz, Switzerland). Each dose was repeated three times on three different fishes to determine error bars. The onset of paralysis was measured over a maximum observation time of 10 min. Paralysis was considered total when the fish went on its back. We performed negative control experiments according to the same protocol by injecting milli-Q water instead of toxins.

3. Results

3.1. Chemical Synthesis

Two α-conotoxins, CIA and CIB (Figure 1), were found to be relatively abundant in the predatory venom of the piscivorous species *Conus catus* [341]. To determine the biological activity, structure and ecological role of these conotoxins, both peptides were manually synthesized using Fmoc-based solid-phase peptide synthesis (see experimental procedures). Assuming the canonical globular disulfide bond connectivity for α-conotoxins (C_1–C_3; C_2–C_4), linear peptides were folded using a regioselective protection (Acm group) of cysteine residues C_1–C_3. According to the number of residues within the loops, CIA and CIB are 3/5 and 4/7 α-conotoxins and we expected that they would target muscle type and neuronal type nAChRs, respectively [192,359–361].

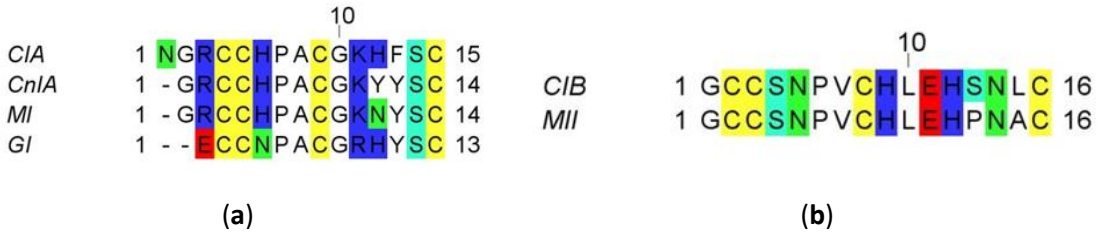


Figure 1. Sequences of α-conotoxins. (**a**) Alignment of CIA α-conotoxin with other closely related 3/5 α-conotoxins; (**b**) Alignment of 4/7 α-conotoxin CIB with MII.

As part of this two-step folding strategy, the standard procedure for the formation of the first disulfide bond from the two non-protected cysteine residues (C_2–C_4) typically consists of stirring the linear peptide in an aqueous basic buffer [356]. However, when the kinetics of disulfide bond formation are too slow, it is common to add potassium hexacyanoferrate or glutathione and oxidized glutathione [191] to facilitate the bridge formation [113,362]. Formation of the first disulfide bridge in an aqueous basic buffer (without any additives) took 24 h to obtain 60% conversion (based on UV chromatogram) and 40 h to obtain 50% conversion for CIA and CIB respectively. Alternatively, we found that DTP (2,2'-Dithiopyridine) was very effective in greatly accelerating the disulfide bridge formation. Surprisingly, this reagent is not widely used, despite its effectiveness being demonstrated by Maruyama et al. [357] 20 years ago. We used this method on several other conotoxins and to our experience, if used in a diluted medium, an excess of DTP induces the quasi-total formation of the intramolecular disulfide bridge in less than 10 min (Figure 2). The second disulfide bond is formed between C_1-C_3 after removal of the Acm protecting groups.

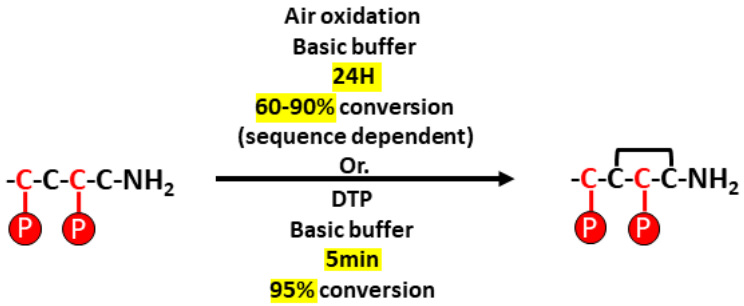
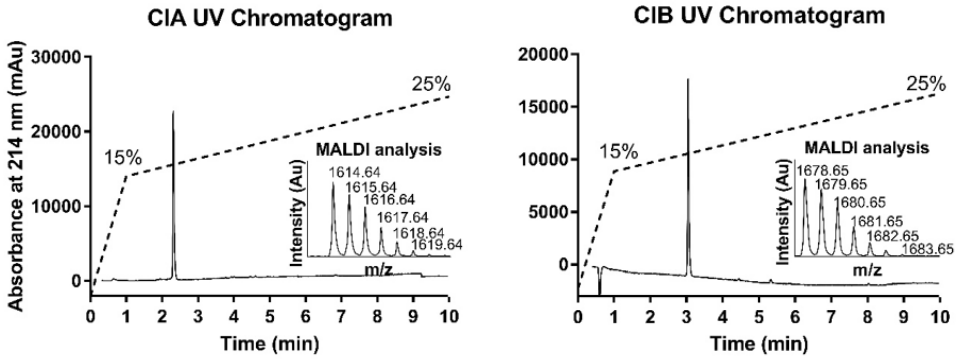


Figure 2. Comparison of disulfide bond formation between free cysteine residues by using air oxidation or DTP. P stands for usual cysteine lateral chain protective groups.

The homogeneity of folded CIA and CIB was assessed by analytical RP-HPLC and MS (Figure 3). MALDI-MS(+) confirmed a monoisotopic mass of 1614.63 Da (calculated 1614.64 Da for $[M + H]^+$) for CIA and 1678.65 Da (calculated 1678.64 Da for $[M + H]^+$) for CIB. Overall, 7.7 mg and 3.4 mg of pure CIA and CIB were obtained, respectively.



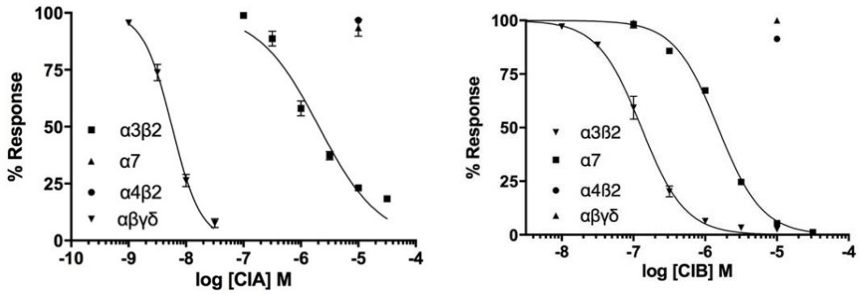
(**a**) (**b**)

Figure 3. HPLC and MS analyses of synthetic CIA and CIB. (**a**) Synthetic folded CIA UV chromatogram at 214 nm and mass spectrometry MALDI analysis; (**b**) Synthetic folded CIB UV chromatogram at 214 nm and mass spectrometry MALDI analysis. Dashed line indicates the acetonitrile gradient.

3.2. Electrophysiology

Next, the biological activity of CIA and CIB was investigated using a two-electrode voltage clamp analysis on three neuronal nAChRs subtypes (α3β2, α7, α4β2) and the muscle type $(\alpha1)_2\delta\gamma\beta1$ nAChR from rat expressed in *Xenopus laevis* oocytes. As expected, the 3/5 α-conotoxin CIA potently blocks the muscle type nAChR with an IC_{50} of 5.7 nM (Figure 4a). Surprisingly, CIA also inhibits the neuronal nAChR subtype α3β2, although with >350-fold lower affinity (IC_{50} of 2.06 µM), whereas no activity was detected on the two other subtypes at concentration up to 10 µM. In contrast, CIB is a 4/7 α-conotoxin that blocks the neuronal nAChR α3β2 subtype with an IC_{50} of 128.9 nM and α7 subtype with an IC_{50} of

1.51 µM (Figure 4b). No activity was detected on the muscle and α4β2 subtypes at concentrations up to 10 µM.



(**a**) (**b**)

Figure 4. Concentration-response analysis of α-conotoxins CIA (**a**) and CIB (**b**) on wild type and mutant rat nAChRs. The indicated subunit combinations were expressed in *Xenopus laevis* oocytes and analyzed by 2-electrode voltage clamp. Responses to 2-s pulses of 100 µM ACh (or nicotine in case of the α7 receptor) were recorded after a 3-min preincubation with the indicated toxin. Each point represents the mean of measurements from at least 3 different oocytes. Error bars represent S.E.M.

3.3. NMR Spectroscopy

High resolution NMR spectroscopy allowed us to determine the three-dimensional structures of both CIA and CIB (Figure 5). The structure statistics are given in Table 1. The NMR spectra for CIA indicate the presence of multiple conformations based on the presence of additional cross-peaks, whereas CIB indicates the presence of a single conformation. The structure determined for CIA was for the major conformation, as the minor conformation displayed weak peaks. The calculated structures are well-defined with low RMSDs for residues 5–11. However, CIB has a larger number of NOE restraints and consequently a better defined structure. CIA has a 3_{10} helix from residues 7–9 and CIB an α-helix from residues 6 to 11. CIA has a similar three-dimensional structure to the closely related peptide, CnIA. Both peptides have a 3_{10} helix involving the Pro7-Ala8-Cys10 sequence (CIA numbering). The ensembles of CIA and CnIA (PDB code 1B45) superimpose with an RMSD of 1.26 Å (calculated using MOLMOL). Similarly, CIB is structurally very similar to MII, where both peptides have a central α-helical region comprising residues 6–11 in CIB and residues 7–11 in MII. The ensembles of CIB and MII (PDB code 1MII) superimpose with an RMSD of 0.55 Å (calculated using MOLMOL).

The conformational heterogeneity observed for CIA is common to most α-3/5 conotoxins. Indeed, Favreau et al. reported similar conformational heterogeneity in solution for α-3/5 conotoxin CnIA, but this also holds true for α-3/5 GI and α-3/5 MI [192,363,364]. This heterogeneity could be due to cis to trans isomerization of a peptidic bond, but at this stage there is no data to support this hypothesis

[192]. The presence of a minor conformation has been identified by Maslennikov et al. for α-conotoxin GI [363]. These two conformations are interconvertible in solution and differ in the regions of the second loop and C terminus [363]. Therefore, it is not possible to conclude whether only one or both of the two conformations is responsible for the activity, or whether each conformation has the same level of activity.

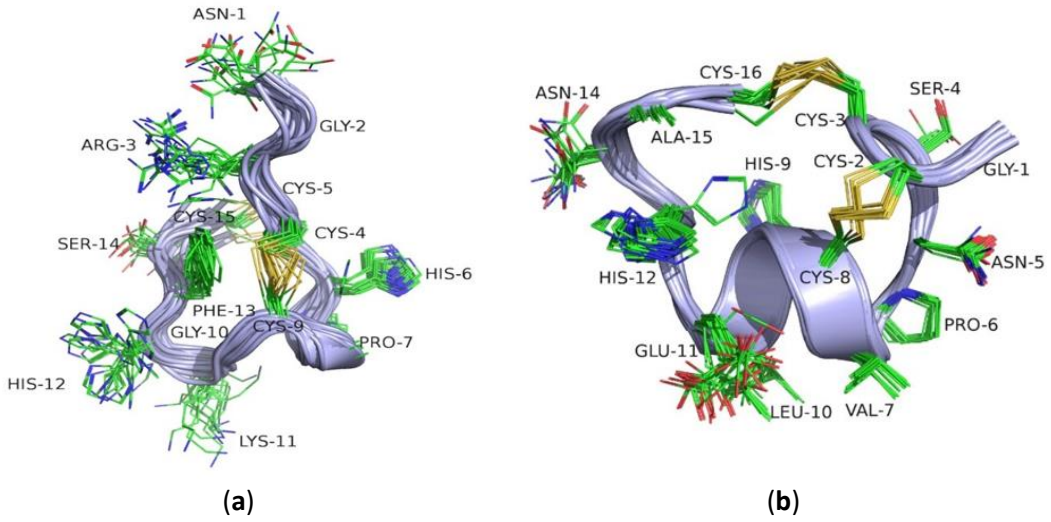


Figure 5. Three-dimensional structures of CIA and CIB. The 20 lowest energy NMR structures, superimposed over the backbone atoms for CIA (**a**) and CIB (**b**). The backbone is shown in ribbon format and the side-chains in stick format.

Table 1. Structural statistics for CIA and CIB.

	CIA	CIB
Experimental restraints		
Interproton distance restraints	69	98
Intraresidue	30	30
Sequential	28	47
Medium range (*i-j* < 5)	11	18
Long range (*i-j* ≥5)	0	3
Disulfide-bond restraints	4	4
Dihedral-angle restraints	21	22
R.m.s deviations from mean coordinate structure (Å)		
Backbone atoms	0.95 ± 0.33	0.48 ± 0.16
Backbone atoms (res 5–11)	0.08 ± 0.04	0.19 ± 0.10
All heavy atoms	1.92 ± 0.46	0.95 ± 0.22
All heavy atoms (res 5–11)	0.39 ± 0.41	0.80 ± 0.19
Ramachandran Statistics		
Clashscore, all atoms	0 ± 0	0 ± 0
% in most favoured region	85.7 ± 0	91 ± 10
MolProbity score	2.13 ± 0.14	1.68 ± 0.45

3.4. In vivo Bioassays

Since both CIA and CIB were detected in the predatory venom of the piscivorous *C. catus*, they were injected into fish (*Danio rerio*) in order to infer the possible ecological role of each conotoxin in prey capture. Intramuscular injection of α-conotoxin CIA produces a rapid flaccid paralysis of skeletal muscles, as evidenced by a loss of equilibrium of the fish, and ultimately a complete immobilization. Paralysis induced by conotoxin α-CIA was a dose-dependent effect, with an IC_{50} of 6.88 µM (Figure 6). Based on the calculated average weight of our adult zebrafish (0.5 g) and the volume injected (5 µL), CIA has an ED_{50} of 110 µg/kg. Injection of up to 1 mM of CIB, however, does not produce any noticeable effect on the locomotion of zebrafish, which is consistent with the absence of activity on muscle nAChR.

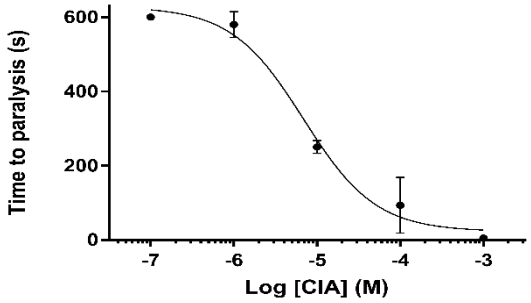


Figure 6. Paralytic effect of CIA on zebrafish. Paralysis induced by conotoxin α-CIA shows a dose-dependent effect, with an IC_{50} of 6.88 µM.

3. Discussion

Animal venoms are generally complex mixtures of hundreds or more biologically active compounds. It is assumed that each of these molecules has been selected through evolution for a specific ecological role. Whereas pharmacologically characterized conotoxins from cone snail venom duct extracts have traditionally been associated to primary roles in prey-capture due to their paralytic action, recent findings suggest other possible venom-ecology relationships [198]. Indeed, independently collected predation and defense-evoked venoms unexpectedly showed that the potent paralytic conotoxins well characterized from *C. geographus* were almost exclusively injected to defend against predators, not for prey capture. This is contrary to the consensus literature published on this topic for the past 30 years. Noteworthy, piscivorous cones of the Gastridium clade such as *C. geographus* and *C. tulipa* employ a rather unusual and unique prey capture strategy to catch fish, producing an apparent sedative effect through passive release of venom components in the surrounding water [198]. On the other hand, species in the largest clade of fish-hunting cones, namely the Pionoconus, use a “Taser-like” effect to rapidly stun fish. This effect is thought to be the result of a combination of two synergic

conotoxin types: δ-conotoxins, which inhibit the inactivation of voltage-gated sodium channels, and κ-conotoxins, which block voltage-dependent potassium channels [365]. However, the predation-evoked venom of *C. striatus* and *C. consors*, two of the largest species of Pionoconus, shows a composition completely devoid of these conotoxins, composed instead of nearly exclusively unrelated κA-conotoxins, and the occasional α- and ω-conotoxins [366,367]. Similarly in the predation-evoked venom of one of the smallest fish-hunting species, *C. catus*, Himaya et al. reported mainly the presence of κA-conotoxins, but consistently also some vertebrate-active and paralytic α-, ω-, and μ-conotoxins in significant amounts [341].

Based on this previous study [341], the sequences corresponding to two major α-conotoxins with the cysteine pattern CCX_3CX_5C and CCX_4CX_7C were synthesized and their biological activity assessed using electrophysiology and fish bioassays. The first α-conotoxin, named CIA, is a typical muscle type α-3/5 conotoxin and its sequence closely resembles those of other α-3/5 conotoxins such as MI, GI and CnIA. CIA blocks the muscle type nAChR with high affinity (~5 nM), but surprisingly, it was found to also block the α3β2 neuronal subtype of nAChR, albeit with lower potency (~2 μM). To our knowledge, CIA is the only known nAChR muscle type α-conotoxin that can also target the neuronal α3β2 subtype with significant affinity. Whereas the molecular basis for this neuronal nAChR activity remains to be investigated, conserved proline and tyrosine/phenylalanine residues (corresponding to position 7 and 13 in CIA sequence, respectively) have been shown to be crucial for the strong hydrophobic interaction between the δ subunit of the muscle type nAChR and conotoxin MI, suggesting a similar mode of action [359,368] for CIA.

The second α-conotoxin, CIB, blocks the neuronal nAChR α3β2 subtype with an IC_{50} of 128.9 nM and the α7 subtype with an IC_{50} of 1.51 μM. The sequence of CIB is most similar to α-conotoxin MII, sharing 14 out of 16 amino acids. However, its potency towards α3β2 is approximatively 29-fold lower compared to MII [369]. The two different residues that most likely account for this difference are proline in position 13 for CIB instead of a serine for MII and alanine in position 15 for CIB instead of leucine for MII. By comparing the three-dimensional structures, it appears that the proline residue in position 13 of CIB induces a kink that prevents the alanine in position 15 from filling the hydrophobic pocket in the same way the leucine does in MII (Figure 7). However, Dutertre et al. [246] using an AChBP model showed that MII possesses an hydrophobic core (Pro-6, Val-7 and Leu 10) where Pro-6 is involved in a direct interaction with β2-Leu-119 and additional hydrophobic contacts due to interactions of Val-7 and Leu-10 with β2-Val-109 and β2-Phe-117 respectively. Docking studies would be very helpful to further investigate how this difference at position 15, which seems to not be directly

involved in the binding with the receptor, can explain the significant loss of activity of CIB compared to MII.

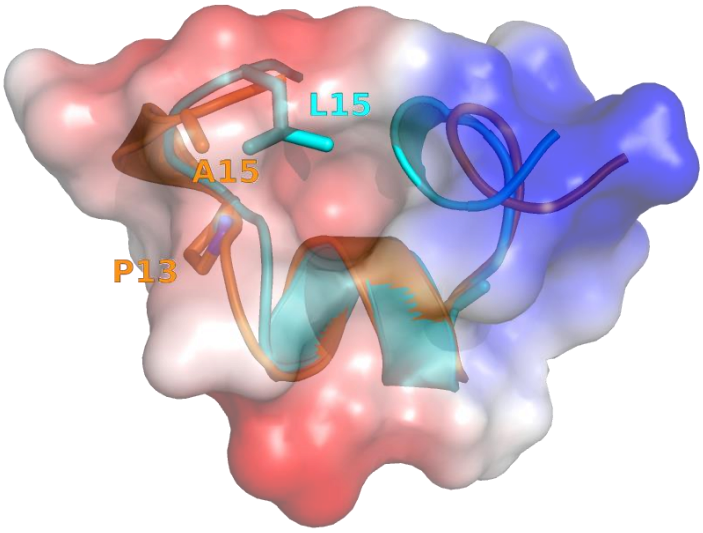


Figure 7. Charge distribution of CIB (orange) and MII (blue; PDB: 1MII), superimposed over residues 1 to 16. Charges are represented with different shades from red (negative charge states) to blue (positive charge states). A cavity appears in the molecular surface of CIB, which is partially filled by L15 in MII.

Conotoxins are traditionally tested on mammalian systems, including human and rodent receptors. However, in the case of these two α-conotoxins CIA and CIB found in the predatory venom of *C. catus*, it was of interest to evaluate their activity on their natural prey (here fish for piscivorous cone snails). Intramuscular injection of α-conotoxin CIA into fish causes flaccid paralysis of the skeletal muscles with an effective dose ED_{50} = 110 µg/kg, highlighting a potent biological effect compatible with a role in prey capture. At the highest dose (1 mM) tested, the paralysis is almost instantaneous, however this dose is unlikely to be biologically relevant. Indeed, it was demonstrated for *Conus purpurascens* that most conotoxins expressed in injected venoms range up to 3.51–121.01 µM within a single sting sample [370]. Therefore, the quasi-instantaneous tetanic paralysis provoked by *Conus catus* sting [371] is likely due to the most abundant conotoxins injected, namely the κA-conotoxins [341,372]. Indeed, Kelley et al. observed that the injection of purified κA-conotoxins is able to reproduce the biological effects of the whole injected venom on fish prey [372]. The role of α-3/5 conotoxins in prey envenomation is most likely secondary, and the resulting blockade of the neuromuscular junction may prevent the escape response of a fish prey that would recover from the “Taser-effect” of κA-conotoxins. Interestingly α-3/5 conotoxin CIA is also found in *Conus achatinus* and *Conus consors* [64,191], belonging also to the Pionoconus clade. On the other hand, injection of α-conotoxin CIB into fish did

not cause any significant change compared to the control. The role of conotoxins targeting neuronal nAChRs in prey capture remains to be elucidated, but we propose that they may interfere with sensory circuitry and the escape response of prey.

ii. Synthèse d'analogues cycliques de la CIA

Compte tenu du profil pharmacologique intéressant de l'α-conotoxine CIA précédemment décrite, nous avons effectué la cyclisation du squelette peptidique de la toxine dans le but d'améliorer ses propriétés pharmacologiques mais aussi sa stabilité métabolique. Cela était également l'occasion d'appliquer la méthodologie de formation des ponts disulfures mise en place précédemment, sur la synthèse de peptides cycliques non naturels.

Cette étude est présentée sous la forme d'un article en cours de préparation dans lequel j'ai effectué la synthèse des analogues, la caractérisation par spectrométrie de masse, les tests de stabilité dans le sérum et les tests biologiques sur les poissons zèbres (*Danio rerio*).

Backbone cyclization turns α-conotoxin CIA into a potent and stable ligand with dual muscle/neuronal-type nicotinic receptor activity

Julien Giribaldi[1], Yves Haufe[2], Edward Evans[3], Muriel Amar[4], Anna Krautloher[2], Casey Schmidt[3], Adèle Faucherre[5], Hamid Moha Ou Maati[5], Christine Enjalbal[1], Jordi Molgó[4,6], Denis Servent[4], David T. Wilson[3], Norelle L. Daly[3], Annette Nicke[2] and Sébastien Dutertre[1]

[1]Institut des Biomolécules Max Mousseron, UMR 5247, Université de Montpellier-CNRS, 34095 Montpellier,France ; [2]Walther Straub Institute of Pharmacology and Toxicology, Faculty of Medicine, LMU Munich, Nußbaumstraße 26, 80336 Munich, Germany ; [3]Centre for Molecular Therapeutics, Australian Institute of Tropical Health and Medicine, James Cook University, Cairns, QLD 4878, Australia ; [4]Commissariat à l'Energie Atomique et aux énergies Alternatives (CEA), Institut des Sciences du Vivant Frédéric Joliot, Service d'Ingénierie Moléculaire des Protéines (SIMOPRO), CEA de Saclay, UniversitéParis-Saclay, F-91191 Gif-sur-Yvette, France ; [5]Département de Physiologie, Institut de Génomique Fonctionnelle, CNRS/INSERM UMR 5203, Université de Montpellier, 34095 Montpellier, France ; [6]Institut des Neurosciences Paris-Saclay, UMR 9197 CNRS/Université Paris-Sud, 91198 Gif-sur-Yvette, France.

Abstract

Conotoxins have always been considered as promising therapeutic leads due to their small size and exquisite potency/selectivity. However, like most peptides their therapeutic utilization is severely limited by their low oral bioavailability. By cyclizing the previously described α-conotoxin CIA using a range of linkers from two to four residues, we have designed peptides that preserve the activity of the native toxin towards muscle-type nAChRs but have greatly improved resistance to proteolytic degradation and, surprisingly, display better potency for α3β2 neuronal nAChR subtypes. The most potent analogue, cCIA-3, displays up to 52-fold higher potency (IC_{50} 1.3 nM) for α3β2 compared to the native α-CIA. Structural conservation between the cyclic analogues and the native toxin was confirmed by NMR structure determination. However, the cCIA-3 analogue reveals the most well-defined structure and some structural deviations in the second loop, potentially explaining the gain of potency for α3β2 nAChRs. Consistent with their activity on muscle-type nAChRs, all the peptides were highly paralytic when injected into *Danio rerio* fish and display oral/transcutaneous activity when incubated in *Danio rerio* fish larvae. Finally, the unique dual activity of the most potent cCIA-3 analogue was used to investigate the contribution of α3β2 subtypes in TOF fade, which is usually observed in the presence of a neuromuscular blocker. Surprisingly, cCIA-3 does not induce the expected TOF fade suggesting that dual blockade of α3β2 and muscle-type nAChRs is able to prevent this phenomenon.

1. Introduction

Nature provides one of the richest sources of bioactive peptides, some of which are useful for treating human diseases. Indeed, more and more peptides entering clinical development come from animals, plants and bacteria [37,373]. Animal venom peptides are of particular interest as they have been finely

tuned to potently and selectively target receptors involved in key physiological functions and, consequently, can be used as highly potent pharmacological tools [42].

Nowhere is peptide diversity more apparent than within cone snail venoms. Cone snails are predatory marine mollusks that have evolved and radiated into more than 800 different species [374]. They have developed highly potent venoms mainly composed of small peptides called conotoxins that are frequently highly constrained by disulfide bridges (10-40 residues and 2-4 bridges), and target a wide range of key receptors for rapid prey immobilization and defense [199,375]. Disulfide bridges provide highly structured conformations, enabling a tight interaction with their molecular targets, which are mainly ion channels, but also transporters and GPCRs [20,82].

Nicotinic acetylcholine receptors (nAChRs) are ligand-gated ion channels existing as homopentamers or heteropentamers composed of numerous homologous subunits. Subunits α2-α10 and β2-β4 assemble into the diverse group of neuronal nicotinic receptors, while the muscle-type nicotinic receptor exist only as $(\alpha1)_2\beta1\delta\varepsilon$ and $(\alpha1)_2\beta\gamma\delta$ subunit combinations for the adult and fetal type, respectively [376]. nAChRs are involved in a wide range of physiological and pathophysiological processes, for example muscle contraction, pain, nicotine addiction, and neurological disorders such as Parkinson's and Alzheimer's disease [20].

Cone snail venoms are composed of various subfamilies of conotoxins with different molecular targets. Members of the α-conotoxin subfamily specifically block nAChRs and some represent unique ligands to study the involvement of the diverse nAChRs in specific biological processes [20]. α-conotoxins can be further classified according to the number of residues within the loops formed by two conserved disulfide bridges. Interestingly, in most cases, the size of the loops seems to determine their specificity for different nAChR subtypes. Thus, 3/4 α-conotoxins target α-homomeric nAChRs, 3/5 α-conotoxins target muscle-type nAChRs and 4/4, 4/6 and 4/7 α-conotoxins target different hetoromeric and/or α-homomeric neuronal nAChRs [20].

Despite great advances in peptide synthesis and drug development, the poor bioavailability of conotoxins due to amide bond breakdown by digestive enzymes, as well as their high polarity and molecular weight, limit their intestinal permeability and, consequently, their utilization as therapeutics. Thus alternative routes of application, such as intrathecal administration have been developed [13]. Head-to-tail cyclization, N-acetylation and C-amidation are widely employed strategies to improve peptide stability through prevention of degradation by aminopeptidases and carboxypeptidases [19,377]. Although N-acetylation and C-amidation are relatively easy to implement via chemistry techniques, the more challenging head-to-tail cyclization is often preferred because of the increased permeability through biological barriers provided to the cyclic analogues [378–381].

Inspired by previous work describing the successful cyclization that improved the stability of several α-conotoxins with specificity for neuronal nAChRs [294,323,333,382,383], we investigated head-to-tail cyclization of a native muscle-type α-conotoxin CIA. To the best of our knowledge, backbone cyclization of a muscle-type conotoxin has not been attempted before. Here we describe the first synthesis, pharmacological characterization, and structure determination of three cyclic analogues of α-CIA and demonstrate their increased stability and modified subtypes selectivity. In addition, we show their potential utilization as pharmacological tools to study the weakening of the muscle during repetitive nerve stimulation (fade of "train of four").

2. Methods

2.1 Abbreviations

Acm, acetamidomethyl; ACN, acetonitrile; Boc, *tert*-butoxycarbonyl; DCM, Dichloromethane; DIPEA, diisopropylethylamine; DMF, *N,N'*-dimethylformamide; DTP, 2,2'-Dithiopyridine; ESI-MS, electrospray ionization mass spectrometry; Fmoc, fluorenylmethoxycarbonyl; HATU, 1[*Bis*(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate; LC/MS, liquid chromatography/mass spectrometry; MeOH, methanol; nAChR, nicotinic acetylcholine receptor; NMR, nuclear magnetic resonance; Pbf, pentamethyl-dihydrobenzofuran-5-sulfonyl; RP-HPLC, reversed-phase high performance liquid chromatography; SPPS, solid-phase peptide synthesis; t-Bu, *tert*-butyl;TFA, trifluoroacetic acid; TIS, triisopropylsilane; Tris, 2-Amino-2-(hydroxymethyl)propane-1,3-diol;Trt, trityl; UV, ultra-violet.

2.2 Chemical synthesis

DMF, DIPEA, ACN, TIS, TFA, piperidine and all others reagents were obtained from Sigma-Aldrich (Saint-Louis, MI, USA) or Merck (Darmstadt, Germany) and were used as supplied. Fmoc (L) amino acid derivatives and HATU were purchased from Iris Biotech (Marktredwitz, Germany). PS-2-Chlorotrityl chloride resin (100-200 mesh, 1.6 mmol/g) was purchased from Iris Biotech (Marktredwitz, Germany). The following side-chain protecting groups were used: Trt for Asn, Cys and His; *t*Bu for Ser; Boc for Lys; Pbf for Arg; and Acm for Cys. Peptides were manually synthesized using the Fmoc-based solid-phase peptide synthesis technique on a VWR (Radnor, PA, USA) microplate shaker. All Fmoc amino acids and HATU were dissolved in DMF to reach 0.5 M. The first amino acid was coupled onto the resin for 6 h in a 1/1 (v/v) mix of DMF and DCM, with a 2.5-fold excess of amino acid and 5-fold excess of DIPEA followed by addition of methanol and further mixing for 15 min to cap any remaining reactive functionalities on the resin. The resin was washed with DMF, DCM, MeOH, and DMF. Fmoc deprotection was carried out with piperidine in DMF (1/2 v/v) twice for 3 min. Subsequent amino acids were coupled onto 0.1 mmol of prepared resin (determined loading value 0.73 mmol/g) twice for 10

min using an amino acid/HATU/DIPEA ratio of 5:5:10 relative to resin loading. DMF was used for resin washing between deprotection and coupling steps. After chain assembly was complete, the terminal Fmoc group was removed and the resin washed with DMF and DCM. Cleavage of the peptide from the resin without affecting the side-chain protecting group was carried out in a reaction vessel and treated ten times with 10 mL of 1% TFA in DCM (v/v) for 5 min. Eluates were collected and combined into a round-bottomed flask then DCM and TFA were removed under vacuum and cold diethyl ether added to precipitate the peptide. The crude side-chain protected peptide was dissolved in DMF at a concentration of 2 mM in a round-bottom flask. HATU was added to the solution to give a final concentration of 5 mM and mixed for 30 s. DIPEA was added to a final concentration of 10 mM, and the solution was stirred for 4 h at room temperature. DMF was removed under vacuum and residues were uptake in ACN/H_2O (1/1 v/v) and freeze dried overnight. Side-chain (except Acm) deprotection was carried out by adding 6.25 mL of TFA/TIS/H_2O (95/2.5/2.5 v/v/v) per 100 mg of crude peptide and stirring the mixture for 2.5 h at room temperature. Crude peptides were purified by preparative RP-HPLC and pure fractions were combined and freeze-dried. A two-step oxidation procedure was then carried out. The first disulfide bridge was formed between the free cysteine residues CysII-CysIV by dissolving the peptide at 0.2 mM in 50 mM Tris-HCl buffer adjusted to pH 8 and adding dropwise 7 equivalents of DTP at 10 mM in MeOH. When the reaction was complete the reaction mixture was acidified to pH 3 and loaded onto preparative RP-HPLC and pure fractions were combined. The second disulfide bridge CysI-CysIII was formed by deprotection/oxidation of the Acm protecting group directly on the combined pure fractions of the mono bridged intermediates by treatment with 20 equivalents of 10 mM iodine in H_2O/TFA/ACN (78/2/20 v/v/v). When the reaction was complete, the reaction mixture was quenched with 20 mM ascorbic acid until total discoloration of the solution, acidified and purified by preparative RP-HPLC. The combined pure fractions were freeze-dried and their purity were confirmed by LC/ESI-MS.

2.3 Mass spectrometry

Solvents used for LC/MS were of HPLC grade. The LC/MS system consisted of a Waters (Milford, OH, USA) Alliance 2695 HPLC, coupled to a Waters Micromass ZQ spectrometer (electrospray ionization in positive mode (ESI+) fitted with a quadrupole mass analyzer). All the analyses were carried out using a Chromolith (Fontenay sous Bois, France) HighResolution RP-18e (4.6 x 25 mm, 15 nm–1.15 µm particle size, flow rate 3.0 mL/min) column. A flow rate of 3 mL/min and a gradient of 0–100% B over 2.5 min for routine analyses and 0-30% B over 30 min for quality control of pure products were used. Eluent A was water/0.1% HCO_2H and eluent B consisted of acetonitrile/0.1% HCO_2H. UV detection was performed at 214 nm. Electrospray mass spectra were acquired at a solvent flow rate of 200 µL/min.

Nitrogen was used for both the nebulizing and drying gas. The data were obtained in a scan mode ranging from 100 to 1000 m/z or 250 to 1500 m/z to in 0.7 s intervals.

Folded peptides were characterized using a Synapt G2-S high-definition MS system (Waters Corp., Milford, MA) equipped with an ESI source and an hybrid QTof mass analyzer configuration. Chromatographic separation was carried out at a flow rate of 0.4 ml/min on an Acquity H-Class ultrahigh performance liquid chromatography (UPLC) system (Waters Corp., Milford, MA), equipped with a Kinetex C18 100 Å column (100 × 2.1 mm, 2.6 µm particle size) from Phenomenex (France). The mobile phase consisted of water (solvent A) and ACN (solvent B) with both phases acidified by 0.1% (v/v) formic acid. Mass spectra were acquired in the positive ionization mode.

2.4 Preparative RP-HPLC

Preparative RP-HPLC was run on a Gilson PLC 2250 Purification system (Villiers le Bel, France) instrument using a preparative column (Waters DeltaPak C18 Radial-Pak Cartridge, 100 Å, 40 x 100 mm, 15 µm particle size, flow rate 50.0 mL/min). Buffer A was 0.1% TFA in water, and buffer B was 0.1% TFA in acetonitrile. A gradient of 0-50% B over 50 min was used.

2.5 Electrophysiological Recordings

cRNA was synthesized from linearized plasmids with SP6 RNA polymerase using the mMessageMachine kit (Invitrogen, Thermo Fisher Scientific, USA). Lobes of *Xenopus laevis* oocytes were kindly provided by Prof. Luis Pardo, MPI of Experimental Medicine, Göttingen, and incubated with collagenase (1.0mg/ml; no. S1746502; Nordmark Biochemicals, Uetersen, Germany) in ND96 (96 mM-NaCl, 2 mM KCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$ and 5 mM Hepes, pH 7.4 for 1-2 h at 16 °C. Follicular cells were removed by subsequent incubation (15-20 min) in Ca^{2+}-free ND96, followed by several washing steps with ND96. Defolliculated oocytes of oogenesis stages V/VI were manually selected and injected with 50-nl aliquots of cRNA (0.5 µg/µl). Injected oocytes were kept at 16°C in filtered ND96 containing gentamicin (5 µg/ml).

Two-electrode voltage clamp recordings were performed in oocytes 1–5 days after cRNA injection at a holding potential of -70 mV. Pipettes were pulled from borosilicate glass and filled with 3 M KCl. Resistances were below 1 MΩ. Membrane currents were recorded using a Turbo Tec 05X Amplifier (npi electronic, Tamm, Germany) filtered at 200 Hz and digitized at 400 Hz. CellWorks software was used for recording. The perfusion medium was automatically switched between ND96 with or without agonist (100 µM ACh) using a custom-made magnetic valve system. A fast and reproducible solution exchange (<300 ms) for agonist application was achieved using a 50-µl funnel-shaped oocyte chamber combined with a fast solution flow (150 µl/s) fed through a custom-made manifold mounted

immediately above the oocyte. ACh pulses were applied for 2 s at 4-min intervals. After each application, the cell was superfused for 54 sec with agonist-free solution, and the flow was then stopped for 3 min. Immediately at the beginning of this interval, peptide (prepared in filtered ND96 containing 0.1% BSA m/V) was mixed from a 10-fold stock into the static bath when responses of three consecutive agonist applications differed by less than 10%. The use of BSA showed no change in toxin potency but produced more stable measurements at low toxin concentrations. ACh-evoked currents peak following peptide incubation were normalized to last ACh current peak before peptide exposure.

The analysis of the electrophysiological data was performed using GraphPad Prism version 8.0. Dose-response curves were fit to the data using the Hill equation: % response = Bottom + (Top-Bottom)/(1+10^((LogIC_{50}-X)*HillSlope)) and constraints of 100% and 0% for Top and Bottom, respectively. Dissociation curves were fit to the data with the equation % response = (response (time 0) - plateau)*exp(-K * time) + plateau.

The functional analysis of competitive binding was performed as previously described (Ellison et al., 2013). Briefly, 2s ACh pulses were applied for 1 min intervals for 5 consecutive ACh evoked peaks with the last three not differing by more than 10%. The perfusion was then stopped for 7 min with application of cCIA-3 after 1 min and/or MII after 4 min in the bath with static bath method depending on the experimental design and then continued. As a control, ND96 was applied instead of a peptide. All peak currents were normalized to the mean of the 4 ACh evoked peak currents before the peptide incubation.

All experiments were conducted with oocytes from at least two different frogs.

2.6 *NMR spectroscopy*

Lyophilized synthetic peptides (~1.5-2 mg) were resuspended in 90%H_2O:10%D_2O. 2D ^{1}H-^{1}H TOCSY, ^{1}H-^{1}H NOESY, ^{1}H-^{1}H DQF-COSY, ^{1}H-^{15}N HSQC, and ^{1}H-^{13}C HSQC spectra were acquired at 290 K using a 600 MHz AVANCE III NMR spectrometer (Bruker, Karlsruhe, Germany) equipped with a cryogenically cooled probe. All spectra were recorded with an interscan delay of 1 s. NOESY spectra were acquired with mixing times of 200-250 ms, and TOCSY spectra were acquired with isotropic mixing periods of 80 ms. Two-dimensional spectra were collected over 4096 data points in the f2 dimension and 512 increments in the f1 dimension over a spectral width of 12 ppm. Standard Bruker pulse sequences were used with an excitation sculpting scheme for solvent suppression. The two-dimensional NOESY spectra of the cCIA analogues were automatically assigned and an ensemble of structures calculated using the program CYANA [358]. Torsion-angle restraints from TALOS+ were used in the structure calculations. The final structures were visualized using Pymol (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.), MOLMOL[384] and UCSF Chimera [385].

2.7 In vivo assays zebrafish (Danio rerio)

2.7.1 Monitoring of the paralysis effect after Injection onto adult zebrafish

Zebrafish wild-type AB were maintained under standardized conditions and experiments were conducted in accordance with the European Communities council directive 2010/63, procedures were approved by Ethical Committee for Animal Experiment of Languedoc Roussillon n ° 36 (reference number: 2018040911129080 #14665 v4). Toxins were diluted in milli-Q water and 5 µL of incremental doses were injected intramuscularly into adult zebrafish with a 10 µL Neuros Syringe from Hamilton (Bonaduz, Switzerland). Each dose was repeated three times on three different fishes to determine error bars. The onset of paralysis was measured over a maximum observation time of 10 min. Paralysis was considered total when the fish went on its back. Negative control experiments were performed according to the same protocol by injecting water instead of toxins.

2.7.2 Monitoring of the paralysis effect after incubation into larvae swimming water

Experiments were conducted on 5 day-old larvae of zebrafish wild-type AB. 6 larvae per toxin were placed in a 96-well plate and a controlled volume of swimming water was added. Small volumes of toxin was added to reach the final desired concentration of 100 µM. Immediately after incubation, the plate was placed in the movement tracking chamber. The movement of larvae was video captured and quantified using the Zebrabox infrared camera setup and tracking extension of the ZebraLab software system (Viewpoint Life Sciences, Canada). The integration period for movement data was set to 30 min. Each time the animal speed goes above the small/large movement threshold, the large movement counter in incremented. Negative control experiments were performed according to the same protocol by adding water instead of toxins.

2.8 In vitro assays on isolated mouse nerve-muscle preparations

2.8.1 Animals

Adult male and female Swiss mice (Mus musculus, 2–5 months of age and 23–28 g of body weight), were purchased from Janvier Elevage (Le Genest-Saint-Isle, France). Mice were acclimatized at the CEA animal facility for at least 48 h before experiments. Live animals were treated according to the European Community guidelines for laboratory animal handling and the guidelines established by the French Council on animal care “Guide for the Care and Use of Laboratory Animals” (EEC86/609 Council Directive – Decree 2001-131). Mice were housed four- to six-wise in cages with environmental enrichment, in a room with constant temperature and a standard light cycle of 12-h light/12-h darkness and had free access to water and food. All experimental procedures on mice were approved by the Animal Ethics Committee of the CEA and by the French General Directorate for Research and

Innovation (project APAFIS#2671-2015110915123958v4) authorized to E. Benoit). Male and female mice were anesthetized by isoflurane (Aerrane®, Baxter S.A., Lessines, Belgium) inhalation before being euthanized by dislocation of the cervical vertebrae.

2.8.2 Recordings on isolated nerve-muscle preparations

In vitro assays were performed on left phrenic-nerve hemidiaphragm muscle preparations quickly removed and mounted in a silicone-lined organ bath (4 mL capacity). Preparations were bathed in a Krebs-Ringer solution of the following composition: 150 mM NaCl, 5 mM KCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$, 11 mM glucose, and 5 mM HEPES (pH 7.4), continuously superfused with pure O_2 throughout the experiment at a constant temperature of 22 °C, unless otherwise indicated. For nerve-evoked isometric twitch tension measurements, the phrenic nerve was usually stimulated with a suction microelectrode (adapted to the diameter of the nerve) with supramaximal current pulses of 0.25 ms duration, at a frequency of 0.1 Hz delivered by the isolation unit of a stimulator (S-44 Grass Instruments, West Warwick, RI, USA). The hemidiaphragm tendons (at the rib side) were tightly anchored onto the silicone-coated bath with stainless pins, and the other tendon (central medial tendon) was attached *via* an adjustable stainless-steel hook to a FT03 isometric force transducer (Grass Instruments). The resting tension was monitored for each preparation tested and adjusted with a mobile micrometer stage allowing variations of muscle length in order to obtain maximal contraction amplitude in response to motor nerve stimulation. Once maximal contraction was obtained, the resting tension was fixed, and monitored during the whole duration of the experiment. Signals from the isometric transducer were amplified, collected, and digitized with the aid of a computer equipped with an Axon™ Digidata-1550B A/D (interface board low noise acquisition systemplus hum silencer™), using the PClamp/Axoscope 10.7 version software (Axon Instruments, Molecular Devices Inc., Sunnyvale, CA, USA).

In some experiments a Train-Of-Four (TOF) stimuli was delivered to the phrenic nerve trunk at a frequency of 2 Hz for 2 s, at a train rate of 0.033 Hz. The ratio of muscle tension developed in the mouse hemidiaphragm by the fourth to the first stimulus was evaluated [T(4) /T(1)] at different peptide concentrations.

2.8.3 Statistical analysis

Data are presented as means ± standard deviations (S.D.) of n different experiments. Differences between values were tested using the parametric two-tailed Student's *t*-test (either paired samples for comparison within a single population or unpaired samples for comparison between two independent populations) or the Kolmogorov-Smirnov two-sample test. Differences were considered significant when $P < 0.05$.

2.9 Serum stability assay

Human AB serum (VWR, Fontenay-sous-Bois, France) was centrifuged at 12000 g for 10 min for the removal of the lipid component. Supernatant was taken out and incubated for 15 min at 37 °C before the assay. All peptides were tested at a final concentration of 30 µM after dilution in serum (water for negative control). The incubation time points were 1, 2, 4 and 8 h at 37 °C. Controls and test peptides were incubated in parallel at each time point. Serum proteins (30 µL) were denatured by quenching with 40 µL of 6 M urea (10 min, 4 °C), followed by the precipitation of proteins with an addition of 40 µL of 20% trichloroacetic acid (10 min, 4 °C). These solutions were then centrifuged at 12000 g for 10 min. 100 µL of supernatant was taken out at each time point. Chromatographic separation was carried out at a flow rate of 0.4 ml/min on an Acquity H-Class ultrahigh performance liquid chromatography (UPLC) system (Waters, Corp., Milford, MA, United States), equipped with a Kinetex C18 100A column (100 mm x 2.1 mm, 2.6 mm particle size) from Phenomenex (France). The mobile phase consisted of water (solvent A) and ACN (solvent B) with both phases acidified by 0.1% (v/v) formic acid and gradient was 0-80% B in 10 min. Mass spectra were acquired in the positive ionization mode. The elution time for each peptide was determined by the zero-time point. The stability at each time point was calculated as the area of the serum treated peptide peak on RP-HPLC at 214 nm as percentage of the area of the 0 h serum treated peptides. Controls were an eight-residue linear peptide incubated in serum for the positive control and incubated in water for the negative control. Each experiment was performed in triplicate.

3. Results

The previously described α-conotoxin CIA [196] from predation-evoked venom of *Conus catus* is a highly potent blocker of muscle-type nicotinic receptors. It has a typical 3/5 α-conotoxin disulfide framework and high sequence homology with other known muscle-type conotoxins. Remarkably, it also shows activity on the α3β2 neuronal subtype, which has not been observed for the 3/5 α-conotoxin family. The NMR structure of α-CIA revealed a relatively short inter-termini distance of about 10 Å that favors head-to-tail cyclization, and we investigated the effect of backbone cyclization of this peptide with amino acid linkers of different lengths (Figure 1). The objective of this modification was twofold: (i) to improve the stability of the native toxin and (ii) produce the first proteolysis-resistant and highly potent blocker of both neuronal and muscle-type nicotinic receptors. Every other characterized α-conotoxin distinguishes between neuronal and muscle-type nAChRs. As demonstrated in this work, compounds targeting both types of nicotinic receptors represent a unique pharmacological tool for biological questions involving both receptors.

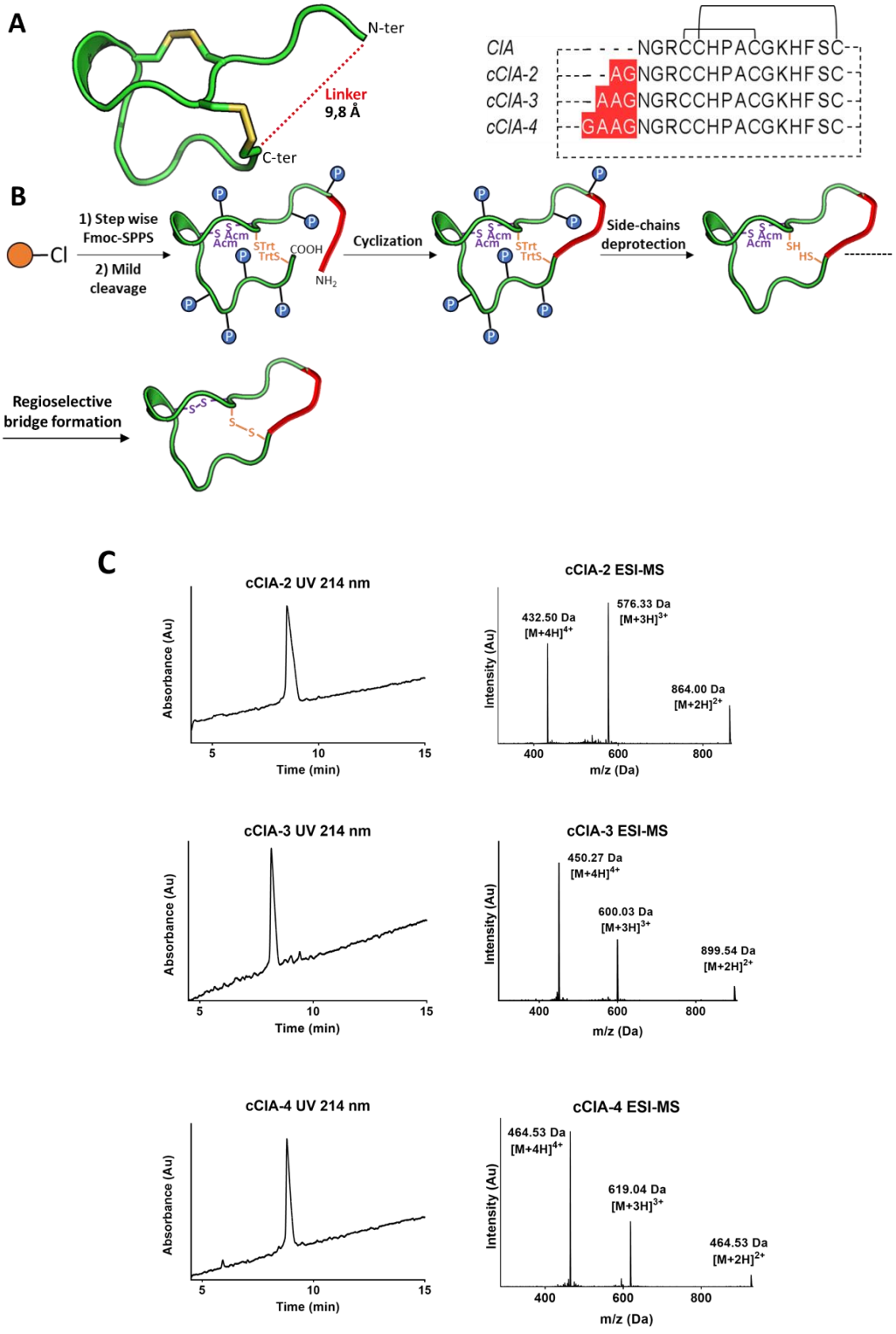


Figure 1: Synthesis strategy of CIA cyclic analogues and NMR structure representation of native CIA (A) NMR structure (backbone representation) of CIA. CIA exhibits a classical α-conotoxin fold with I-III and II-IV disulfide connectivity (represented in yellow). Based on the inter-termini spacing, amino acid linkers with a length between 10 and 19 Å were chosen. (B) Synthesis

strategy for the production of three cyclic α-CIA peptides and the respective RP-HPLC/ESI-MS analysis. The orange round circle represents a chloro-(2'-chloro) trityl resin and P represents common side chains protecting groups used in Fmoc-SPPS (Solid Phase Peptide Synthesis). (C) RP-HPLC-UV (ACN gradient from 0 to 30% in 30 min) coupled to ESI-MS analyses revealed the presence of dominant peaks of the expected masses.

3.1 Chemical synthesis

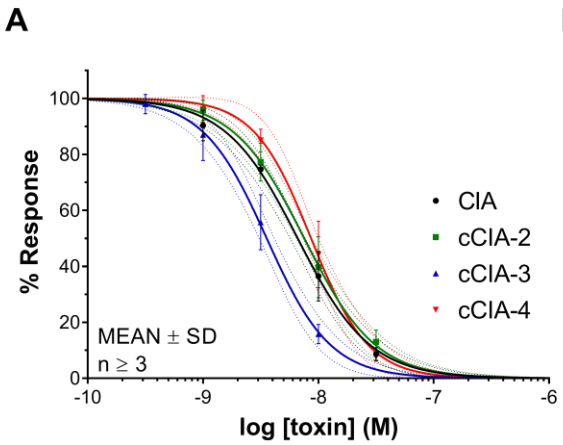
After anchoring of the C-terminal residue on a chloro-(2'-chloro) trityl resin, elongation of the peptide chain was performed manually using Fmoc-based SPPS [87,386]. Mild cleavage conditions were used to allow the separation of the peptide from the resin without affecting the side-chain protecting groups. The crude protected peptide was cyclized under coupling conditions [387]. Side-chain deprotection and subsequent purification produced the cyclic peptide. Next, a regioselective folding strategy using acetamidomethyl (Acm) protecting groups (CysI-CysIII) was employed to produce the disulfide bond connectivity of the native peptide [196]. After final purification, the homogeneity of folded cCIA peptides was assessed by analytical RP-HPLC-UV coupled to ESI-MS (Figure 1 and Figure S5).

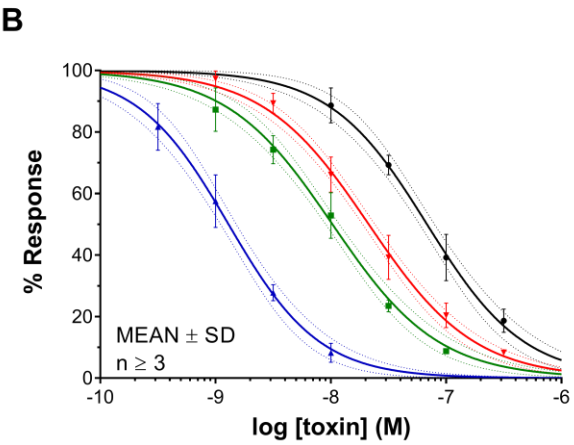
3.2 Electrophysiology and binding assays

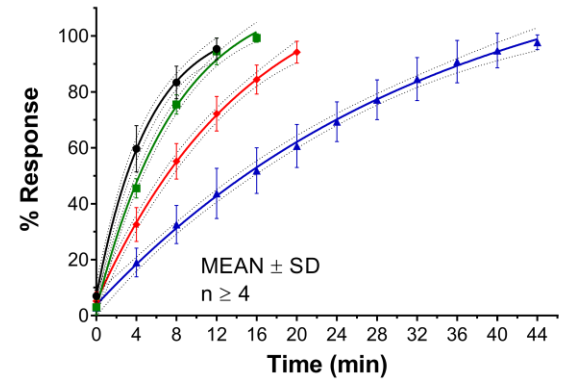
The biological activities of the cCIA peptides were investigated using the two-electrode voltage-clamp method on rat α3β2 neuronal nAChRs subtypes and fetal rat muscle-type nAChR ($(\alpha1)_2\beta\gamma\delta$) expressed in *Xenopus laevis* oocytes. All three cyclic analogues retained the low nanomolar (IC_{50} 4-9 nM) activity of the native toxin (IC_{50} 6.62 nM*) at the muscle-type nAChR. Surprisingly, cyclization significantly improved the potency of CIA (IC_{50} 68.2 nM) at the α3β2 subtype with an up to 52-fold increase in potency for the most active analogue cCIA-3 (IC_{50} 1.30 nM, Figure 2B). Thus, α-cCIA-3 is the first highly potent cyclic blocker described to display a low nanomolar IC_{50} value on both [1]neuronal subtype α3β2 and muscle-type nAChRs. As the potency increase was accompanied by a reduced recovery of the receptor from the block, we next measured the dissociation of the toxin from both muscle-type and α3β2 nAChR subtypes. Despite the similar potency on the muscle-type nAChR of cCIA analogues, they displayed significantly slower off-rates. The cCIA-3 peptide had the smallest dissociation constant (Koff) and a wash-out time of ≈30 min for 80% receptor activity recovery (Figures 2C, 2D and S1) at the muscle-type nAChRs. In contrast, complete dissociation from neuronal α3β2 subtypes occurred within minutes and could not be determined with the

* Please note that a different IC_{50} value was previously determined [196], the reason for this could not be determined, but could be due to different toxin stocks or experimenters.

established protocol. Therefore, the real-time receptor reactivation during agonist application (Figures S2 and S3) was analyzed and compared to provide an estimate of the dissociation of the cCIA analogues from the α3β2 receptor. Again, this revealed cCIA-3 has the slowest dissociation rate, and supports the finding that this analogue has a structure that increases affinity for both muscle and α3β2 nAChRs. Since the common notion is that binding pockets of muscle-type and neuronal nAChRs are quite different and 3/5 α-conotoxins are highly muscle selective, we considered if cCIA binds to an unorthodox binding site in the α3β2 receptor such as the αα interface. However, functional competition binding experiments with α-conotoxin MII, a potent competitive antagonist at the α3β2 subtype showed that preincubation with 100 nM cCIA-3 inhibited binding of MII and suggests competitive binding of cCIA-3 to the orthosteric α3β2 binding site (Figure S4).







D

	$(\alpha1)_2\beta\gamma\delta$		α3β2	
Toxin	Koff (min^{-1})	IC_{50} (nM)	K_d (min^{-1})	IC_{50} (nM)
CIA	**0.191** (0.144-0.245)	**6.62** (5.69-7.67)	**N.D**	**68.22** (58.86-79.13)
cCIA-2	**0.126** (0.102-0.151)	**7.54** (6.47-8.77)	**N.D**	**10.33** (9.01-11.83)
cCIA-3	**0.030** (0.020-0.039)	**3.57** (3.07-4.14)	**N.D**	**1.30** (1.11-1.51)
cCIA-4	**0.066** (0.044-0.089)	**8.64** (7.41-10.01)	**N.D**	**21.3** (18.87-24.09)

Figure 2: Cyclization increases potency and affinity of CIA analogues at muscle-type and α3β2 nAChRs. (A) Dose-response curves of CIA and the three cyclic analogues (cCIA-2, cCIA-3, and cCIA-4) on rat muscle-type (B) and neuronal α3β2 nAChRs. (C) Recovery of muscle-type nAChRs from toxin block measured in 4 min intervals *Xenopus laevis* oocytes expressing the indicated receptors were analyzed by 2-electrode voltage-clamp at -70 mV holding potential. Responses

to 2-s pulses of 100 µM ACh were recorded after 3 min preincubation with the respective toxin. Each point represents the mean of measurements from at least 3 different oocytes. Error bars represent S.D. Dotted lines in (A), (B) and (C) represent the 95% confidence bands of the fits. (D) Summary of inhibition and dissociation constants of the toxins. Numbers in brackets are 95% confidence intervals. N.D = not determined.

2.3 NMR spectroscopy

High-resolution NMR spectroscopy allowed the determination of three-dimensional structures of the CIA cyclic analogues (Figure 3A). Two-dimensional TOCSY, NOESY, COSY and HSQC spectra were collected for all three CIA cyclic analogues and the assignments made using standard approaches [388] (Table S1). Additional peaks are present in the cCIA-2 and cCIA-4 spectra, most likely representing alternative conformations, whereas cCIA-3 has primarily one set of peaks corresponding to a single conformation. Therefore, the linker length appears to impact the structural stability. The major conformations were assigned for the peptides and a superposition of the structures with native α-CIA is shown in Figure 3B, highlighting the similarity between the cyclic analogues and the native peptide. Structures were determined based on the NOE data, and angle restraints predicted from TALOS [389]. Ensembles representing the three peptides are shown in Figure 3A.

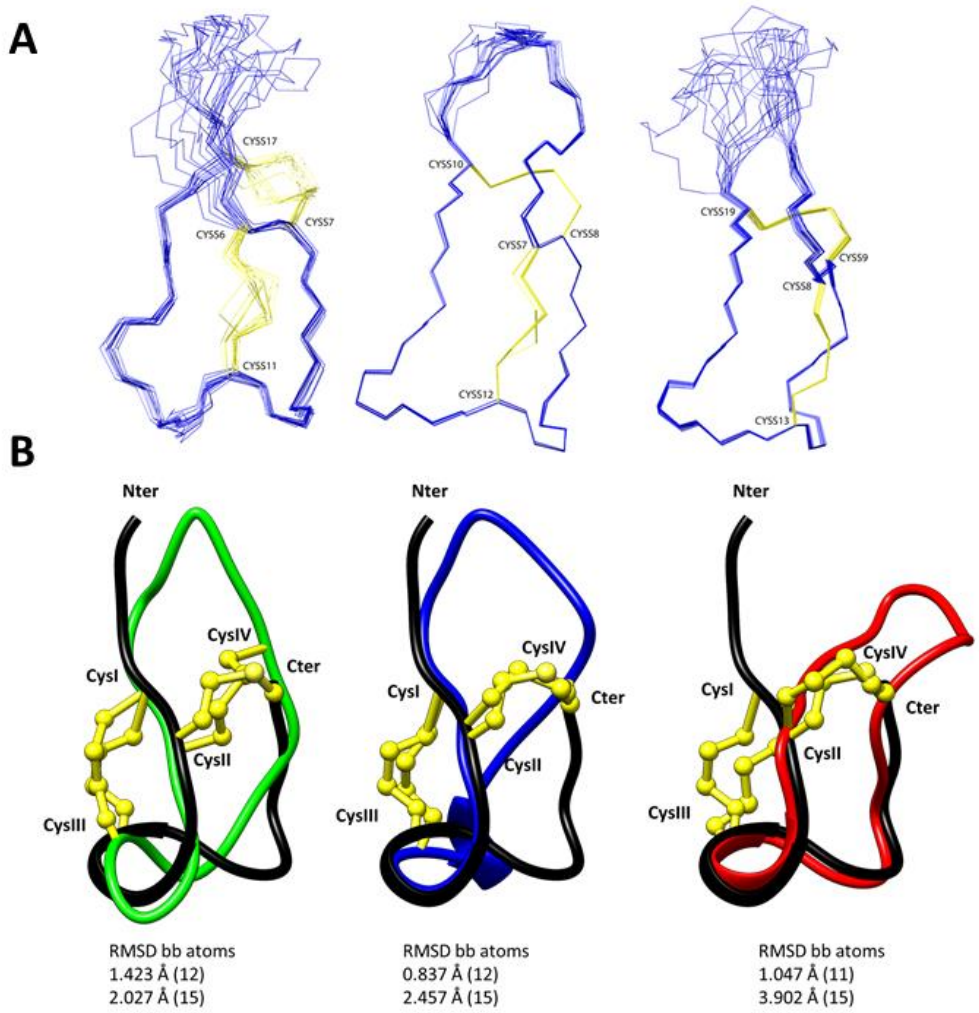


Figure 3: (A) NMR-derived structures of the cyclic CIA analogues. The 20 lowest energy structures are superimposed over the backbone atoms for residues 6-17 (cCIA-2; left), 4-18 (cCIA-3, middle) and 5-19 (cCIA-4, right). The cysteine side chains are shown in yellow. (B) Ribbon representations of cCIA-2 (left), cCIA-3 (middle) and cCIA-4 (right) superimposed over α-CIA backbone atoms (bb) for residues 3-17 (cCIA-2), 4-18 (cCIA-3) and 5-19 (cCIA-4). The disulfide bonds are shown in ball and stick format. The root-mean-square deviation (RMSD) values are given for the superposition of the lowest energy structure over bb atoms. The first value corresponds to the best fitted residues bb atoms, the number of residues is indicated between brackets and the second value corresponds to the 15 residues of α-CIA.

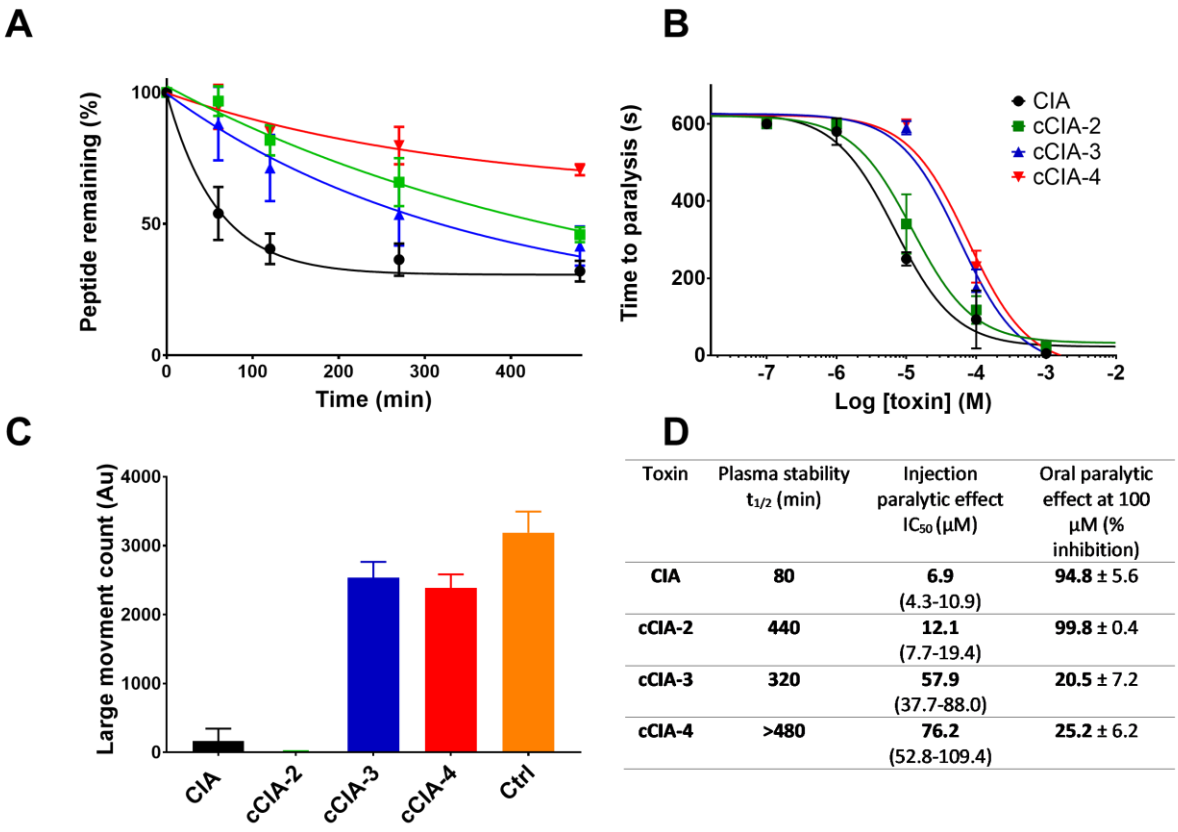
2.4 Stability assays

The stability of CIA and its cyclic analogues against enzymatic degradation was determined by incubation with human serum AB. Peptides were incubated for 8 h at 37°C and the amount of intact peptide remaining was determined by liquid chromatography/MS analysis of aliquots taken

at 0, 1, 2, 4, 8 hours post-incubation. Results revealed that especially during the first three hours the degradation kinetics of cyclic analogues were much slower in comparison to CIA (figure 4A). cCIA-4 is the most resistant cCIA analogue to proteolysis-degradation, with a serum half-life of more than 8 hours and up to 70% intact peptide remaining after 8 hours of incubation. In comparison, cCIA-2 and cCIA-3 have half-life values of 440 min and 320 min, respectively. Finally, native α-CIA is degraded the quickest, with a half-life of 80 min. CIA cyclization greatly improved the enzymatic stability and is consistent with previously published data on conotoxin cyclization.

2.5 Zebrafish (Danio rerio) in vivo assays

In order to investigate a potential oral/transcutaneous activity of CIA and its three cyclic analogues we first performed *in vivo* intramuscular injections into adult zebrafish (*Danio rerio*) in order to determine an easily visible effect that would allow us to monitor uptake of the peptide. Intramuscular injection of α-conotoxin CIA and its three cyclic analogues produces a rapid flaccid paralysis of skeletal muscle, as evidenced by a loss of an upright posture of the fish, and ultimately complete immobilization. Paralysis induced by CIA and cCIA analogues exhibited a dose-dependent effect, with IC_{50} values for CIA, cCIA-2, cCIA-3 and cCIA-4 of 6.88 µM, 12.1 µM, 57.87 µM, and 76.22 µM, respectively (Figure 4B, 4D). This evaluation of paralysis caused by muscle nAChR blockade led to investigation of the oral/transcutaneous toxin uptake/activity by tracking the movement of zebrafish larvae in 100 µM toxin for one hour using the ZebraBox tracking system (Figure 4C). Native CIA and cCIA-2 induced a strong paralysis suggesting efficient uptake of the toxins. Consistent with data from the intramuscular injection, cCIA-3 and cCIA-4 had the weakest paralysis activity.



Toxin	Plasma stability $t_{1/2}$ (min)	Injection paralytic effect IC_{50} (µM)	Oral paralytic effect at 100 µM (% inhibition)
CIA	**80**	**6.9** (4.3-10.9)	**94.8** ± 5.6
cCIA-2	**440**	**12.1** (7.7-19.4)	**99.8** ± 0.4
cCIA-3	**320**	**57.9** (37.7-88.0)	**20.5** ± 7.2
cCIA-4	**>480**	**76.2** (52.8-109.4)	**25.2** ± 6.2

Figure 4: Paralytic effect of the CIA analogues on zebrafish (*Danio rerio*) and serum stability assay. (A) Stability of CIA and cCIA analogues in human serum. Experiments were performed in triplicate with positive and negative controls. Each point represents the mean of the triplicate measurement value. Error bars represent the S.D. (B) Dose-response analysis of paralysis induced by intramuscular injection of conotoxins CIA and cCIA analogues in adult zebrafish. Experiments were performed in triplicate with a negative control. Points represent the mean of measurement performed in triplicate. Error bars represent the S.D. (C) Large movement count over one hour of *Danio rerio* larvae movement tracking after addition of the toxins into the water at a concentration of 100 µM. Each point represents the mean of measurements from 6 different larvae. Error bars represent the S.D (D) Recapitulative table, 95% confidence intervals are between brackets. ± symbol indicates the standard deviation.

2.6 Mouse muscle contraction assays

2.6.1 α-CIA and analogues block nerve-evoked muscle contraction

When applied at nanomolar concentrations to isolated phrenic-nerve hemidiaphragm muscle preparation, CIA and cCIA analogues produced a concentration- and time-dependent reduction of the isometric twitch force evoked by nerve stimulation at 0.1 Hz (Figure 5A).

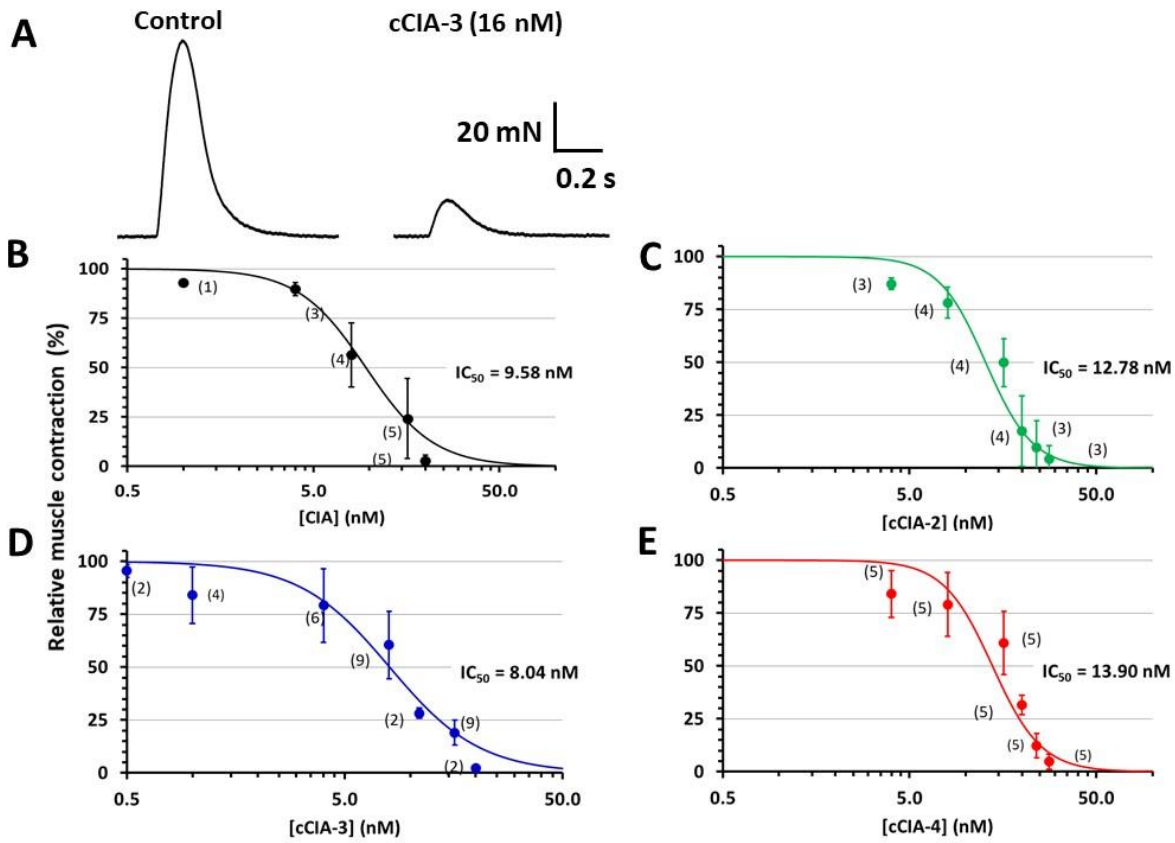


Figure 5: Typical isometric single twitch recordings upon nerve stimulation and concentration-response curves for CIA and cCIA analogues that block phrenic nerve-evoked isometric twitch force in mouse hemidiaphragm muscles. (A) Typical twitch responses upon nerve stimulation under control conditions (left trace) and during the action of 16 nM cCIA-3 (right trace). (B) Concentration-response curves and IC_{50} (nM, ± SD) values for CIA, cCIA-2 (B, C), cCIA-3 and cCIA-4 (D, E). Data points represent the mean ± S.D. of twitch responses after 20-25 min toxin exposure, relative to the respective controls. The number of nerve-muscle preparations used is indicated in parenthesis.

The rank order of potency obtained from these experimental data was cCIA-3 > CIA > cCIA-2 > cCIA-4 > αC-PrXA. For the most active peptide, cCIA-3 (IC_{50} 16 nM), the time required to block 50% of nerve evoked muscle contraction was 19.0 ± 6.2 min (n = 4). The effect of all peptides was reversible upon washing the preparations with an extracellular medium free of α-conotoxins. A 40% recovery was obtained in 32.6 ± 1.8 min (n=3) with the cCIA-3 analogue (data not shown). These results indicate that CIA and analogues are effective neuromuscular blockers, and are 2-3 fold more potent than the previously studied αC-PrXA, a highly specific and potent inhibitor of muscle-type nAChR of the neuromuscular junction [390].

It was of interest to determine if the block produced by the most potent CIA analogue, cCIA-3, could be antagonized by an increase in nerve-evoked quantal ACh. This would determine if an increase of evoked ACh release from nerve terminals could displace the cCIA-3 from its nAChR binding site. For this, we chose 3,4-diaminopyridine (3,4-DAP), which greatly increases quantal ACh by blocking voltage-gated K^+ channels in motor nerve terminals [391–393] and is an effective drug used in neurology.

As shown in Figure 6A, cCIA-3 (16 nM) blocked the peak amplitude of nerve-evoked contraction by approximately 85 %, and this effect was 80% reversed by addition of 3,4-DAP (10 µM) to the medium (Figure 6B). Interestingly, when computing twitch tension-time integrals instead of peak amplitudes the reversal was in the range of 96-98 %. This difference in values is due to the fact that nerve-evoked contractions exhibit a prolonged time course in the presence of 3,4-DAP (shown in a comparison of the control and 3,4-DAP traces in Figure 6A). Thus, it is likely that 3,4-DAP is able to displace the cCIA-3 analogue from the muscle endplate nAChR by increasing acetylcholine release from nerve terminals. These results strongly suggest that the peptide acts in a competitive manner on the muscle-type nAChR, in agreement with the general notion.

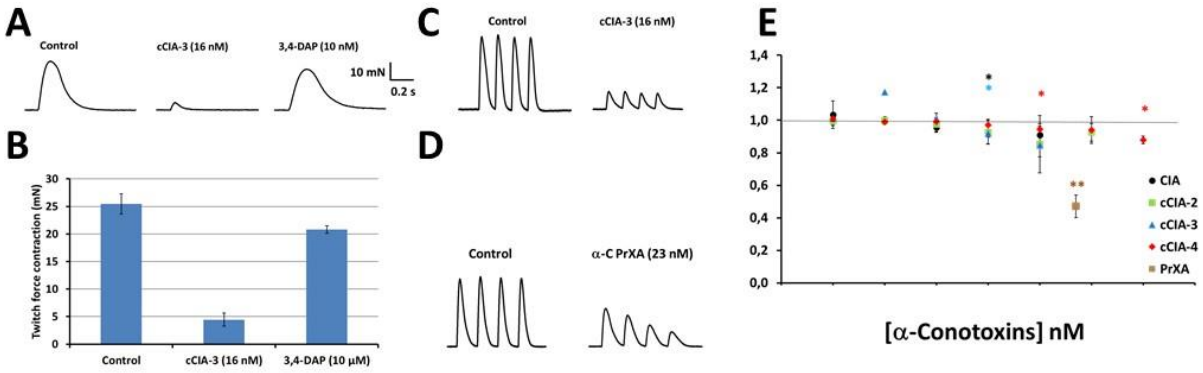


Figure 6: 3,4-DAP reverses the neuromuscular blockade produced by cCIA-3 and TOF fade ratio of the peptides. (A) Isometric twitch tension recordings evoked by nerve stimulation (0.1 Hz) on the same phrenic-nerve hemidiaphragm preparation under control conditions, during 30 min of 16 nM cCIA-3 action, and after 27 min of 10 µM 3,4-DAP action. Note in (A) the different time course of twitch recordings between the control and after 3,4-DAP-treatment. (B) Data obtained in the same preparation and expressed as means ± S.D. (C) Isometric tension developed during TOF nerve stimuli (2 Hz) under control conditions, and in the presence of cCIA-3 and (D) αC-PrXA. Note the TOF fade in the presence of αC-PrXA. (E) TOF ratio fade in the presence of α-CIA, cCIA analogues and αC-PrXA. Note that the fade is relatively small and statistically non-significant for some of the peptide concentration used.

2.6.2 Train-of-four (TOF) fade in the presence of various conotoxins

Agents that produce a non-depolarizing neuromuscular block at the neuromuscular junction are known to display a typical TOF fade, both *in vitro* and *in vivo*. The TOF fade corresponds to the T4 / T1 ratio, where T4 and T1 are the fourth and first twitch tensions in the same train. The inhibition of the presynaptic α3β2 nAChR subtype at motor nerve terminals has been suggested as an explanation for train-of-four fade seen during a non-depolarizing neuromuscular block [394,395]. Therefore, due to their original dual muscular/ α3β2 nAChRs antagonist property, it was of interest to determine whether α-CIA and the cCIA analogues were able to produce TOF fade.

Under control conditions (in the absence of peptides), as shown by the typical recordings (Figure 6C), there was no TOF fade. Interestingly, when nerve-evoked contraction was inhibited about 76 % by cCIA-3 no significant TOF fade was observed (Figure 6C). However, a marked TOF fade was observed for the highly muscle-selective αC-PrXA peptide, with only approximately 45 % neuromuscular block (Figure 6D). As shown in Figure 6E, TOF fade was calculated at different conotoxin concentrations. If a 50% decrease of TOF fade was measured with αC-PrXA (23nM), no significant effect was observed for any of the α-CIA and α-CIA analogue concentrations studied (Figure 6E).

3. Discussion

Backbone cyclization of several neuronal α-conotoxins has been previously reported to enhance stability and in some cases to improve the permeability of the cyclic analogue through biological membranes [294,323,333,382,383]. However, backbone cyclization of a muscle-type α-conotoxin has never been attempted before. Considering the unusual activity of the 3/5 α-conotoxin CIA on nAChR α3β2 subtypes, we investigated its backbone cyclization. During the cyclization process, a linker minimizing perturbations of the three-dimensional structure of a bioactive native toxin is highly desirable. Indeed, it has been shown that an inappropriate linker can distort the structure leading to a loss of bioactivity [294]. Based on the inter-termini spacing, amino acid linkers with a length between 9 and 18 Å were chosen. Clark *et al.* showed that the linker length has to be sufficiently long so as not to introduce strain into the peptide and hence alter the structure of the conotoxin [333]. Conservation of the structure between CIA and its cyclic analogues was confirmed by NMR spectroscopy. However, the linker length had an impact on how well-defined the structures are (Table S1), with cCIA-3 having the lowest RMSD and therefore the most well-defined structure, and one predominant conformation in the NMR spectra in contrast to the other two cyclic analogues.

Consistent with the NMR data, structural conservation of the cCIA analogues compared to the native CIA led to the conservation of the bioactivity towards muscle-type nAChRs at low nanomolar concentrations. However, the significant decrease in Koff values suggests stronger interaction of the

cyclic analogues within one, or both, of the two orthosteric muscle nAChR binding sites. Considering the high sequence homology of α-CIA with α-MI and α-GI, α-CIA is most likely binding at the α-δ interface [368,396]. It has been demonstrated that the ACh binding pocket is mostly composed of hydrophobic residues that interact with residues of the two conotoxin loops formed by the disulfide bridges. Although, the linker is outside of these cysteine loops, the lower Koff cCIA-3 and cCIA-4 might be due to additional hydrophobic interactions with the additional alanine residue in the linker compared to cCIA-2. Interestingly, in contrast to muscle nAChRs, dissociation rates from the neuronal α3β2 subtype were so fast that dissociation constants could not be determined with established protocols despite a strong potency increase at these subtypes. This raised the question of how exactly the native α-CIA and the cyclic analogues bind to the α3β2 subtype. Indeed, allosteric modulators usually display very fast dissociation rates. However, functional competition binding experiments suggest a competitive binding of cCIA-3 to the orthosteric $\alpha 3\beta 2$ binding site and we can reasonably extend this hypothesis to cCIA-2, -4 and native α-CIA toxin. Surprisingly, all the cyclic analogues displayed an improved affinity towards the α3β2 subtypes, with cCIA-3 being the most potent with a 52-fold increase in activity compared to native α-CIA. Favorable structural changes in the second loop and well-defined structures of cCIA-3 compared to native α-CIA (Figure 3B) might be responsible for the enhanced activity of cCIA-3 towards α3β2 nAChR subtypes.

Consistent with the previous published data on α-conotoxin cyclization [294,323,333,382,383], cyclization of α-CIA leads to an improved stability to enzymatic degradation. cCIA-4 was the most resistant to proteolysis-degradation, exhibiting a serum half-life of more than 8 hours and up to 70% remaining peptide, followed by cCIA-2, cCIA-3 and native CIA, which is degraded at least 4 times faster. Note that CIA and the cCIA analogues appear to be more readily degraded in serum than cVc1.1 and cMII, as shown in the study by Clark *et al.* [294,333]. Indeed, unlike Vc1.1 and MII conotoxins, CIA contains one arginine residue before the first cysteine residue, and one lysine residue in the second loop that can be cleaved by endopeptidases. Therefore, cutting the N-terminal part and replacing the lysine residue with an unnatural lysine isostere in cCIAs should further improve their serum stability, but potentially at the expense of the affinity.

A visible paralyzing effect of the toxins, resulting from the block of muscle-type nAChR was observed when CIA or the cCIA analogues were injected intramuscularly into zebrafish. Paralysis activity of the conotoxins could be easily monitored by movement tracking of zebrafish (*Danio rerio*) larvae, after incubation with the toxins in their swimming water. Considering the inhibition constant values obtained by intramuscular injection, we performed the assay at a concentration of 100 µM as insufficient starting material was available for higher doses. CIA and cCIAs showed an oral paralysis effect in comparison to the control, with CIA and cCIA-2 being the most potent. cCIA-3 and cCIA-4

exhibited a weaker activity, which is consistent with the intramuscular injection data. Zebrafish (*Danio rerio*) might not have the required metabolic means to completely digest the native CIA and make it completely orally inactive. However, despite being less potent than native CIA when injected, cCIA-2 still displayed a better oral/transcutaneous uptake/activity, indicating that CIA might be at least partially metabolized in contrast to cCIA-2 due to its fully cyclic nature.

Considering the unique capacity of α-CIA and the cCIA analogues to block both muscle-type and neuronal α3β2 subtype nAChRs, these peptides represent a novel pharmacological tool to study the activity of the α3β2 subtype in nerve stimulation when neuromuscular blockers are present. Consistent with the electrophysiology data, CIA and the cCIA analogues block the phrenic nerve-evoked isometric twitch force in mouse hemidiaphragm muscles in the nanomolar range, with conserved order of potency (cCIA-3>CIA>cCIA-2>cCIA-4). Increase in quantal ACh release by 3,4-DAP displaced cCIA-3 showing a competitive binding of the toxin.

Train-of-four (TOF) fade in muscle (reduction of the twitch amplitude in a train under repetitive nerve stimulation) was used to assess neuromuscular junction and it is well-known that non-depolarizing neuromuscular blocking drugs induce TOF fade [394,395,397]. The role of the pre-synaptic α3β2 nicotinic receptor in the TOF fade phenomenon has been proposed previously. One of the persevering theories is that presynaptic α3β2 nicotinic receptors would increase the release of acetylcholine via a positive feedback mechanism to maintain the contraction at the same level following repeated nerve stimulation at the level of the neuromuscular junction. Thus, the inhibition of these presynaptic receptors could be one of the causes of an attenuated release of acetylcholine leading to nerve-evoked muscle contraction fade. This hypothesis was recently challenged by using ligands with various selectivity for pre- and post-synaptic receptors [398]. For example, in *in-vivo* experiments, the TOF fade was clearly correlated with the administration of a postsynaptic muscle-type antagonist such as α-bungarotoxin or α-conotoxin GI, while DHβE, an α3β2 blocker, was shown to potentiate the TOF fade. Nevertheless, DHβE is a non-selective neuronal receptors blocker, justifying more precise examination of the role of the presynaptic α3β2 subtype. Therefore, considering their high potency and variable selectivity on both muscle-type and neuronal α3β2 nAChRs, CIA and its cyclic analogues represent unique pharmacological tools to address this question. Surprisingly, neither CIA nor the cyclic analogues induce a visible TOF fade, in contrast to the muscle-specific αC-PrXA conotoxin (Figure 6). This suggests that dual blockade of α3β2 and muscle-type nAChRs is able to prevent this phenomenon, nevertheless, their role in neuromuscular response will have to be investigated in more detail in future studies.

5. Supporting information

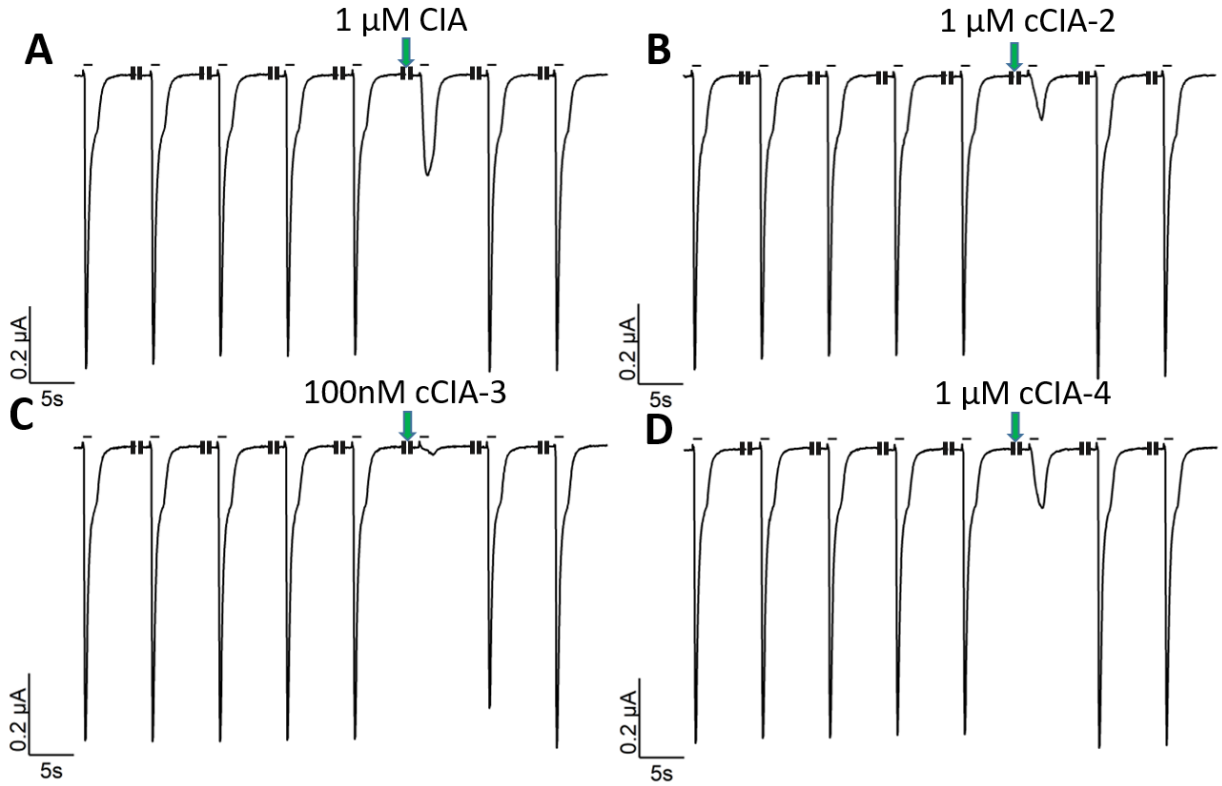


Figure S1: Recovery of ACh-induced $\alpha 3\beta 2$ currents after inhibition by CIA analogues. 2s-pulses of ACh (100 µM, black bar) were applied in 4 minute intervals (1 min under perfusion, 3 min in a static bath). Once stable current responses were obtained, toxins were applied for three minutes in the static bath (indicated by green arrows) in concentrations that produced at least 80% inhibition: (A) 1 µM native CIA, (B) 100 nM cCIA-3, (C) 1 µM cCIA-2 and (D) 1 µM cCIA-4.

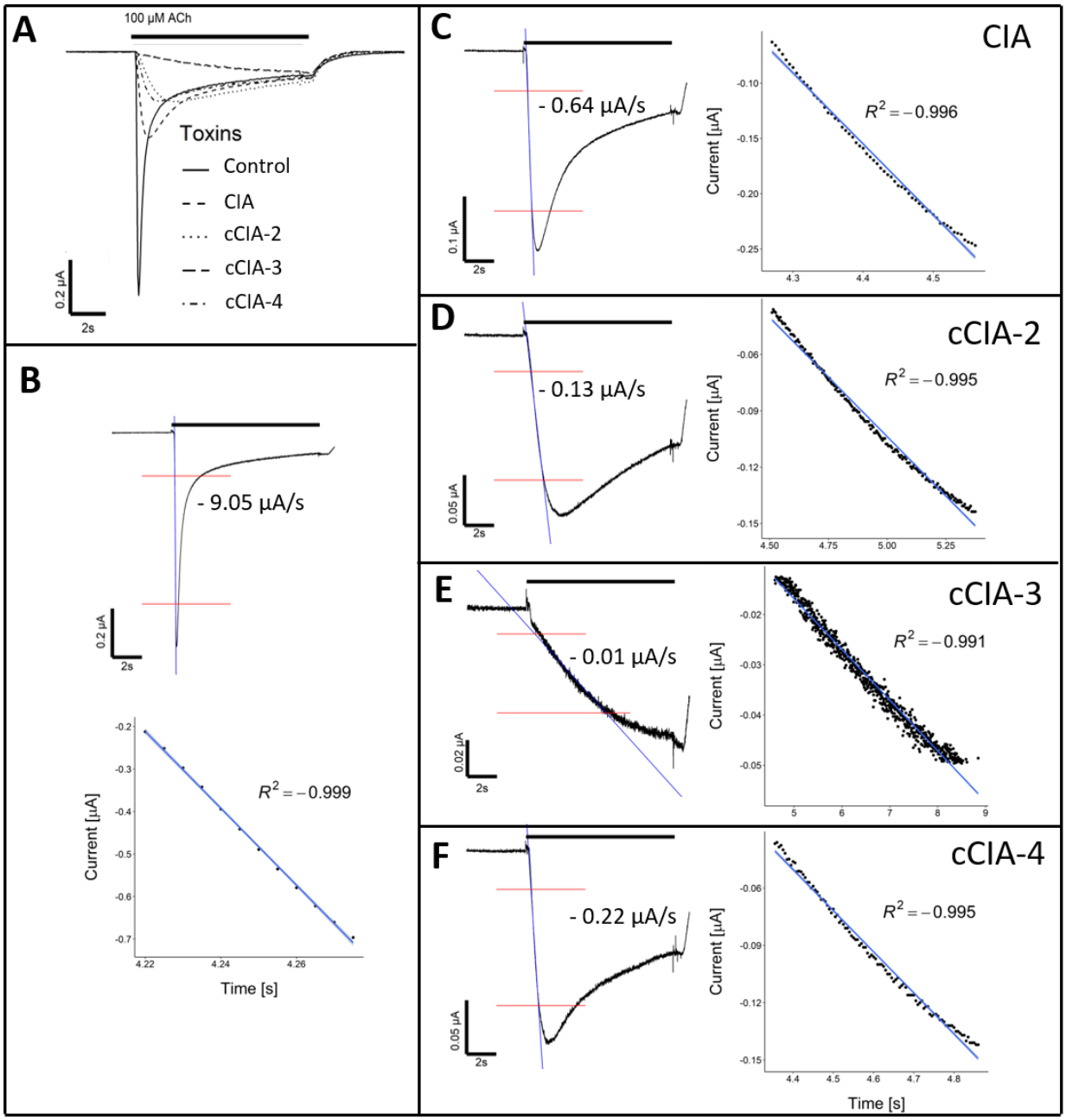


Figure S2: Real-time recovery of α3β2 currents after inhibition by CIA and its cyclic derivates. (A) Merged current responses during a 10s application of ACh (100 µM, black bar) evoked before (control) and after 3min incubation with the indicated peptides. Note the delayed current peak due to dissociation of the peptide during the agonist application. The steepness of the rising phase of the current was used as a surrogate for the off rate. Isolated peak traces are shown with red lines indicating 20% and 80% of the peak current and a linear regression (blue line) within these borders was performed for (B) control (ND96), (C) 1 µM native CIA, (D) 1 µM cCIA-2, (E) 100 nM cCIA-3 and (F) 1 µM cCIA-4. The slope of each linear regression is written in the graph and the linear regression in the chosen borders is shown in detail with regression coefficient (R^2).

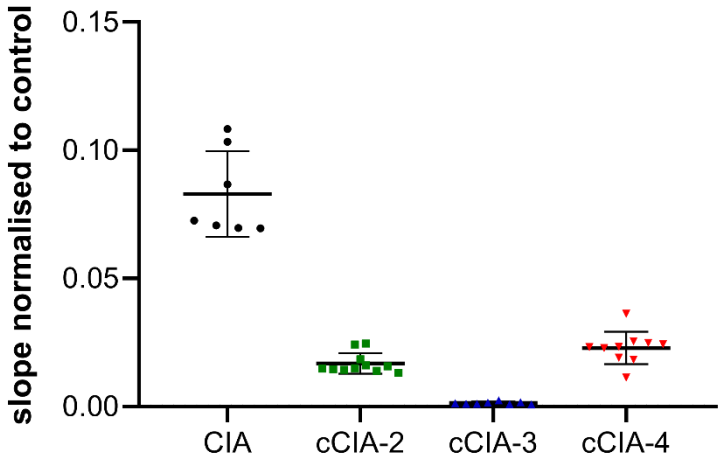


Figure S3: Comparison of the dissociation of CIA and its cyclic derivates from α3β2 using the slope of the rising phase of the ACh evoked current as surrogate. The slopes were determined as shown in Figure S2 and normalized to the slope of the control peak before incubation with the peptide (peak slope / control slope). The single values are shown with the mean as straight bar and S.D., $n \geq 7$. Note that higher values indicate a faster dissociation.

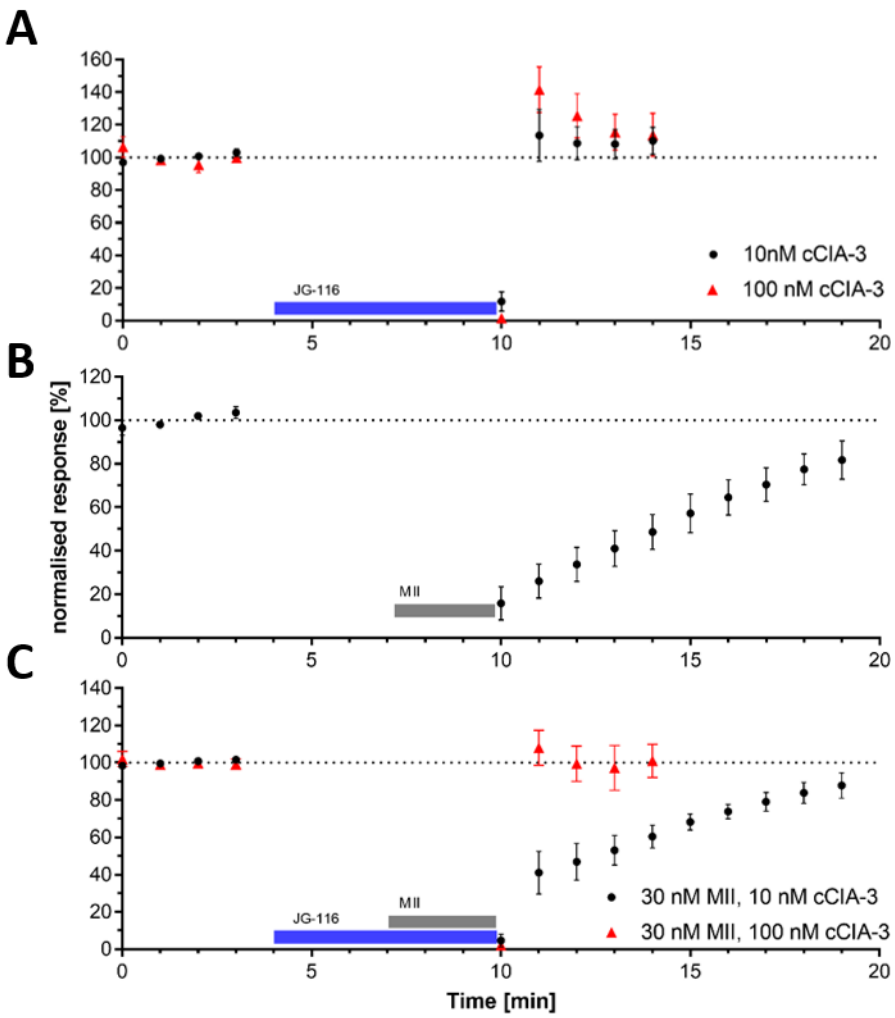


Figure S4: Functional competition between cCIA-3 and MII. Responses to 2s-ACh pulses were recorded at 1-min intervals and normalized to the mean of 4 current peaks recorded prior to peptide incubation in a static bath. Three different protocols were conducted: (A) 1 min after ACh wash-out, cCIA-3 (10 or 100nM) was added to the static bath for 6 min before ACh applications were continued, (B) 4 min after ACh wash-out, MII (30 nM) was added to the static bath for 3 min before ACh applications were continued (C) cCIA-3 (10 or 100 nM) was added after 1 min and MII (30nM) was added after 3 min. ACh applications were continued after 6 min. Data are represented as MEAN $\pm$ S.D. with $n \geq 5$.

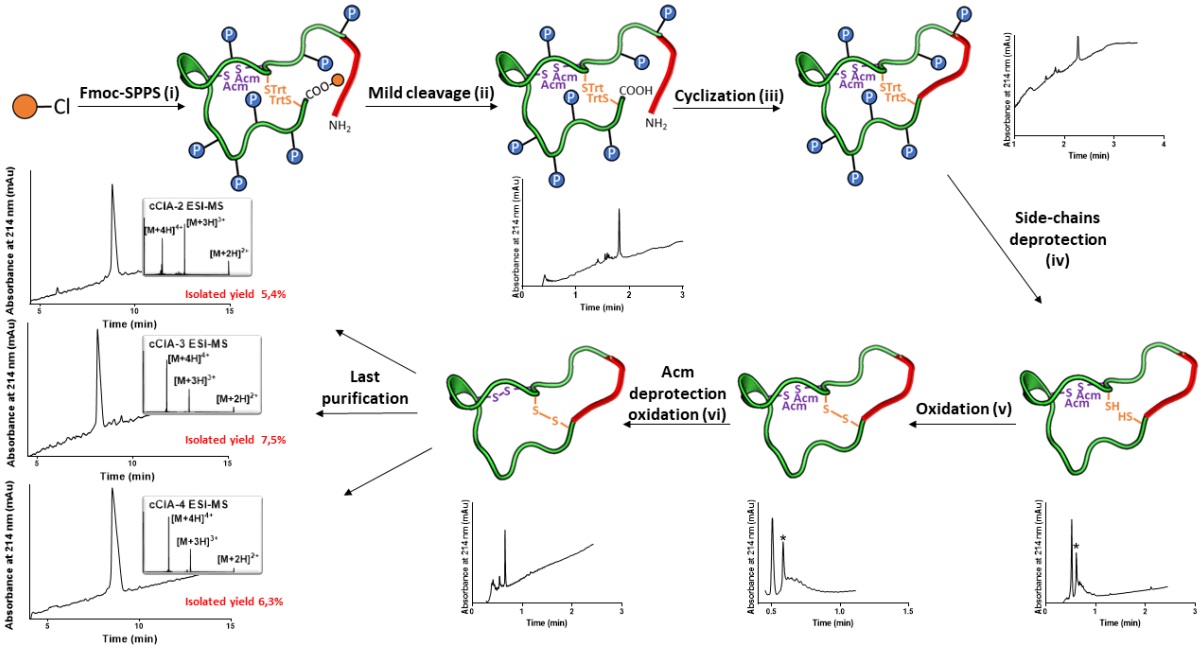


Figure S5: Synthesis strategy of the three cylic analogues of CIA. (i) Double coupling 10 min HATU (5eq)/DIPEA (10eq) (ii) 1% TFA in DCM (v/v) 10 x 5 min (iii) HATU (2,5eq)/DIPEA (5eq) 4 h (iv) TFA/TIS/H_2O (95/2.5/2.5 v/v/v) 2,5 hr (v) Tris-HCl 50 mM buffer pH 8, DTP (7eq), 10 min (vi) I_2 10 mM (20eq) in H_20/TFA/ACN (78/2/20 v/v/v), 10 min. The orange round circle represents a chloro-(2'-chloro) trityl resin and P represents common side chains protective groups used in Fmoc-SPPS (Solid Phase Peptide Synthesis). RP-HPLC-UV (ACN gradient from 0 to 30% in 30 min) coupled to ESI-MS analyses revealed the presence of dominant peaks of the expected masses.

Table S1: Structural statistics for the cCIA analogues.

Structural statistics for the cCIA ensembles			
Experimental restraints	cCIA-2	cCIA-3	cCIA-4
Interproton distance restraints	136	179	154
Intraresidue	36	53	49
Sequential	51	75	67
Medium range (i-j < 5)	23	33	19
Long range (i-j ≥5)	26	18	19
Dihedral-angle restraints	21	23	21
R.m.s. deviations from mean coordinate structure (Å)			
Backbone atoms (with linker)	0.91 ± 0.35	0.29 ± 0.17	1.05 ± 0.45
All heavy atoms (with linker)	1.78 ± 0.46	0.85 ± 0.24	1.59 ± 0.46
Backbone atoms (without linker)	0.72 ± 0.26	0.13 ± 0.06	0.27 ± 0.11
All heavy atoms (without linker)	1.73 ± 0.46	0.82 ± 0.22	1.17 ± 0.35
Ramachandran Statistics			
% in most favoured region	68.5%	61.10%	76.40%
% in additionally allowed region	31.5%	38.9%	23.6%

iii. Synthèse des toxines CIC et Δ-CIC

Dans la continuité de la découverte des toxines CIA et CIB, ainsi que la démonstration de l'utilité de l'analogue cyclique cCIA-3 en tant qu'outil pharmacologique, nous avons poursuivi l'étude pharmacologique des toxines de *Conus catus*. Une autre conotoxine nommée CIC identifiée à partir de la même étude protéo-transcriptomique du venin de *Conus catus* [341] a attiré notre attention par sa singulière extrémité N-terminale anormalement étendue pour une α-conotoxine. La toxine native CIC et un mutant privé de sa partie N-terminale Δ-CIC ont été synthétisés, et caractérisés pharmacologiquement et structurellement afin d'étudier l'influence de cette partie N-terminale.

Cette étude dans laquelle j'ai effectué la synthèse et la caractérisation par spectrométrie de masse est présentée ci-après sous la forme d'une publication non publiable en l'état. La discussion est basée sur des résultats de test d'activité biologique préliminaires. Les procédures expérimentales de RMN et d'électrophysiologie manquent également à la partie méthodes.

1. Introduction

A new α-conotoxin CIC was identified in the transcriptomic study of *Conus catus* and confirmed by proteomic as a component of the predation-evoked venom [341]. Therefore, CIC has been selected for further characterization considering its high sequence homology with PeIA and MII conotoxins, both potent antagonists of α3β2 nAChRs subtypes displaying IC_{50} values in the low nanomolar range [314,361]. However, in addition to displaying a typical 4/7 α-conotoxin cysteine pattern, CIC exhibits an unusual extended N-terminal tail outside of the bridging loops (Figure 1). So, far, GID was the only described 4/7 α-conotoxin showing an N-terminal extension and it was found to be a potent blocker of α3β2 and α7 (low nanomolar IC_{50}) nAChRs, but it also displays a significant and noteworthy activity towards α4β2 subtypes (IC_{50} 152 nM), known as an α-conotoxins-resistant receptor [20,276]. Deletion of the N-terminal sequence of GID (Δ-1-4) had a significant effect on the activity at the α4β2 subtype but no effect on α3β2 and α7 subtypes, suggesting these residues were important for both selectivity and affinity [276]. Interestingly, despite its influence on activity, the N-terminal tail of GID appeared to be disordered in solution.

Therefore, the aim of this study was to investigate the role of the unusual N-terminal part on the three-dimensional structure and on the pharmacological activity of this novel CIC conotoxin. To this end, native CIC conotoxin as well as a mutant variant with no N-terminal tail (Ala1-Thr6 removed) named Δ-CIC were synthesized and pharmacologically and structurally characterized.

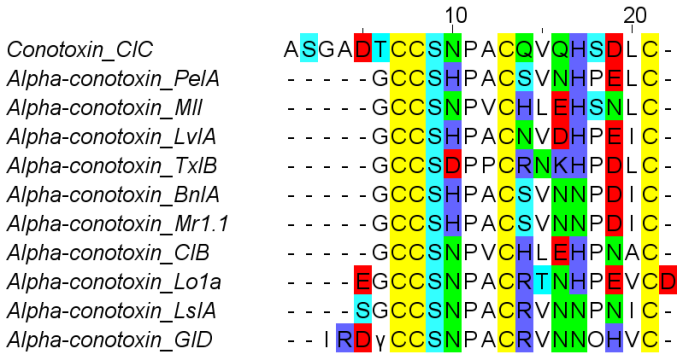


Figure 1: Alignment of CIC conotoxin with others closely related 4/7 α-conotoxins. γ symbol indicates a carboxyglutamate residue and O capital letter indicates an hydroxyproline residue.

2. Methods

2.1 Abbreviations

Acm, acetamidomethyl; ACN, acetonitrile; Boc, *tert*-butoxycarbonyl; DCM, Dichloromethane; DIPEA, diisopropylethylamine; DMF, *N,N'*-dimethylformamide; DTP, 2,2'-Dithiopyridine; ESI-MS, electrospray ionization mass spectrometry; Fmoc, fluorenylmethoxycarbonyl; HATU, 1[*Bis*(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate; LC/MS, liquid chromatography/mass spectrometry; MeOH, methanol; nAChR, nicotinic acetylcholine receptor; NMR, nuclear magnetic resonance; Pbf, pentamethyl-dihydrobenzofuran-5-sulfonyl; RP-HPLC, reversed phase high performance liquid chromatography; SPPS, solid phase peptide synthesis; t-Bu, *tert*-butyl;TFA, trifluoroacetic acid; TIS, triisopropylsilane; Tris, 2-Amino-2-(hydroxymethyl)propane-1,3-diol;Trt, trityl; UV, ultra-violet.

2.2 Chemical synthesis

DMF, DIPEA, ACN, TIS, TFA, piperidine and all others reagents were obtained from Sigma-Aldrich (Saint-Louis, MI, USA) or Merck (Darmstadt, Germany) and were used as supplied. Fmoc (L) amino acid derivatives and HATU were purchased from Iris Biotech (Marktredwitz, Germany). PS-2-Chlorotrityl chloride resin (100-200 mesh, 1.6 mmol/g) was purchased from Iris Biotech (Marktredwitz, Germany). The following side-chain protecting groups were used: Trt for Asn, II-IV Cys, His and Gln; *t*Bu for Ser, Thr and Asp; and Acm for I-III Cys. Peptides were manually synthesized using the Fmoc-based solid-phase peptide synthesis technique on a VWR (Radnor, PA, USA) microplate shaker. All Fmoc amino acids and HATU were dissolved in DMF to reach 0.5 M. The first amino acid was coupled onto the resin for 6 h in a 1/1 (v/v) mix of DMF and DCM, with a 2.5-fold excess of amino acid and 5-fold excess of DIPEA followed by addition of methanol and further mixing for 15 min to cap any remaining reactive functionalities on the resin. The resin was washed with DMF, DCM, MeOH, and DMF. Fmoc deprotection was carried out with piperidine in DMF (1/2 v/v) twice for 3 min. Subsequent amino acids

were coupled onto 0.1 mmol of prepared resin (determined loading value 0.73 mmol/g) twice for 10 min using an amino acid/HATU/DIPEA ratio of 5:5:10 relative to resin loading. DMF was used for resin washing between deprotection and coupling steps. After chain assembly was complete, the terminal Fmoc group was removed and the resin washed with DMF and DCM. Side-chain (except acm) deprotection and cleavage from the resin was carried out by adding 10 mL of TFA/TIS/H_2O (95/2.5/2.5 v/v/v) and stirring the mixture for 2.5 h at room temperature. Crude peptides were purified by preparative RP-HPLC and pure fractions were combined and freeze-dried. A twostep oxidation procedure was then carried out. The first disulfide bridge was formed between the free cysteine residues CysII-CysIV by dissolving the peptide at 0.2 mM in 50 mM Tris-HCl buffer adjusted to pH 8 and adding dropwise 7 equivalents of DTP at 10 mM in MeOH. When reaction was complete, the reaction mixture was acidified to pH 3 and loaded onto preparative RP-HPLC and pure fractions were combined. The second disulfide bride CysI-CysIII was formed by deprotection/oxidation of the Acm protecting group directly on the combined pure fractions of the mono bridged intermediates by treatment with 20 equivalents of 10 mM iodine in H_20/TFA/ACN (78/2/20 v/v/v). When the reaction was complete, the reaction mixture was quenched with 20 mM ascorbic acid until total discoloration of the solution, acidified and purified by preparative RP-HPLC. The combined pure fractions were freeze-dried and their purity were confirmed by LC/ESI-MS.

2.3 Mass spectrometry

Solvents used for LC/MS were of HPLC grade.

Intermediate peptides were characterized using a LC/MS system consisting of a Waters (Milford, OH, USA) Alliance 2695 HPLC, coupled to a Waters Micromass ZQ spectrometer (electrospray ionization in positive mode (ESI+) fitted with a quadrupole mass analyzer). All the analyses were carried out using a Chromolith (Fontenay sous Bois, France) HighResolution RP-18e (4.6 x 25 mm, 15 nm–1.15 µm particle size, flow rate 3.0 mL/min) column. A flow rate of 3 mL/min and a gradient of 0–100% B over 2.5 min for routine analyses and 0-30% B over 30 min for quality control of pure products were used. Eluent A was water/0.1% HCO_2H and eluent B consisted of acetonitrile/0.1% HCO_2H. UV detection was performed at 214 nm. Electrospray mass spectra were acquired at a solvent flow rate of 200 µL/min. Nitrogen was used for both the nebulizing and drying gas. The data were obtained in a scan mode ranging from 100 to 1000 m/z or 250 to 1500 m/z to in 0.7 s intervals.

Folded peptides were characterized using a Synapt G2-S high-definition MS system (Waters, Corp., Milford, MA, United States) equipped with an ESI source and an hybrid QTof mass analyzer configuration. Chromatographic separation was carried out at a flow rate of 0.4 ml/min on a Acquity H-Class ultrahigh performance liquid chromatography (UPLC) system (Waters, Corp., Milford, MA,

United States), equipped with a Kinetex C18 100Å column (100 mm x 2.1 mm, 2.6 mm particle size) from Phenomenex (France). The mobile phase consisted of water (solvent A) and ACN (solvent B) with both phases acidified by 0.1% (v/v) formic acid. Mass spectra were acquired in the positive ionization mode.

2.4 Preparative RP-HPLC

Preparative RP-HPLC was run on a Gilson PLC 2250 Purification system (Villiers le Bel, France) instrument using a preparative column (Waters DeltaPak C18 Radial-Pak Cartridge, 100 Å, 40 × 100 mm, 15 µm particle size, flow rate 50.0 mL/min). Buffer A was 0.1% TFA in water, and buffer B was 0.1% TFA in acetonitrile.

3. Results

3.1 Chemical synthesis

CIC and Δ-CIC were synthesized manually using Fmoc-SPPS after anchoring of the first C-terminal residue on a chloro-(2'-chloro) trityl resin. Considering its high sequence homology with others 4/7 α-conotoxins (Figure 1) which all display the canonical globular disulfide bond connectivity CysI-CysIII, CysII-CysIV, linear peptides were folded using a regioselective folding strategy. Briefly, the first disulfide bridge CysII-CysIV is formed by oxidation of the free cysteine residues in an aqueous basic buffer assisted by DTP (2,2'-Dithiopyridine) and the second disulfide bond CysI-CysIII is formed by deprotection/oxidation of the Acm protecting groups. The homogeneity of folded toxins was assessed by analytical RP-HPLC coupled to ESI-MS (Figure 2). ESI-MS(+) confirmed a monoisotopic mass of 1602.90 g/mol (calculated 1602.59 g/mol) for Δ-CIC and 2106.27 g/mol (calculated 2106.29 g/mol) for CIC. Overall, 7.7 mg (yield 3.1% from resin loading) and 9.9 mg (yield 3.2% from resin loading) of pure Δ-CIC and CIC were obtained, respectively.

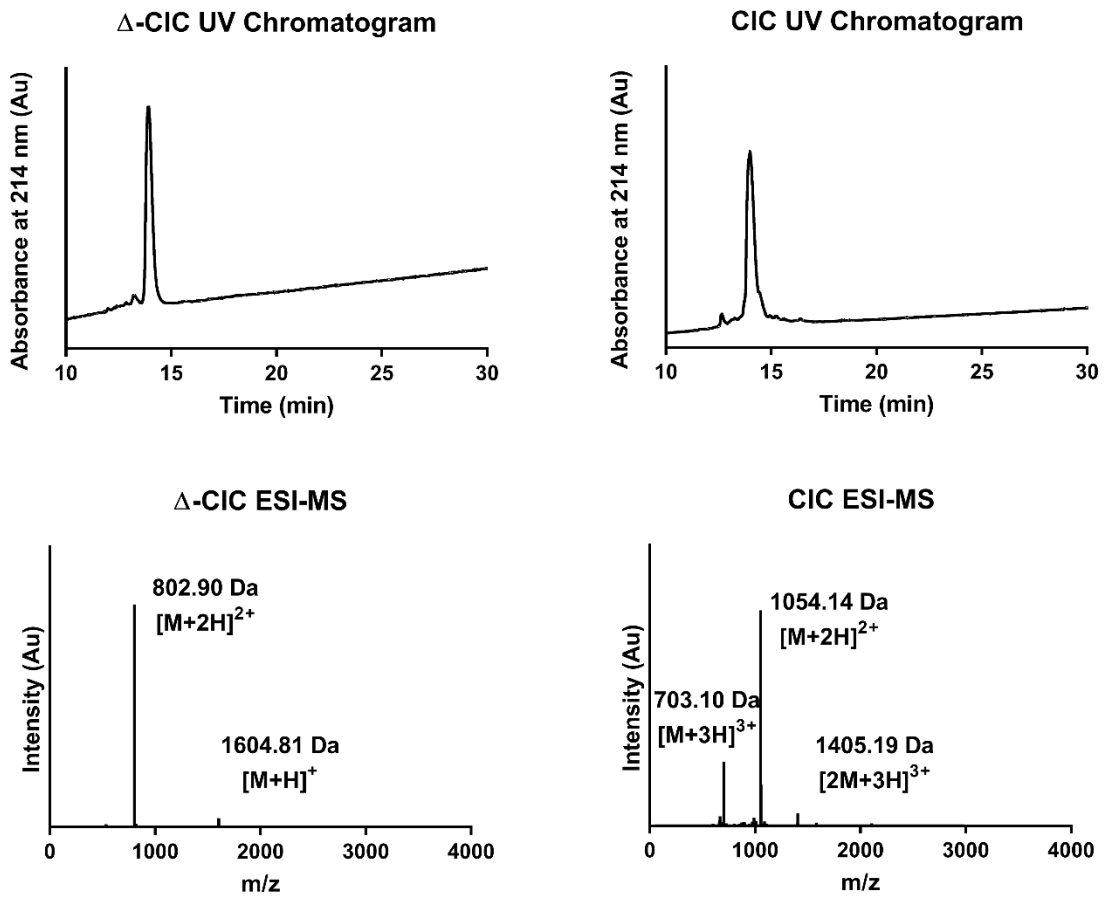


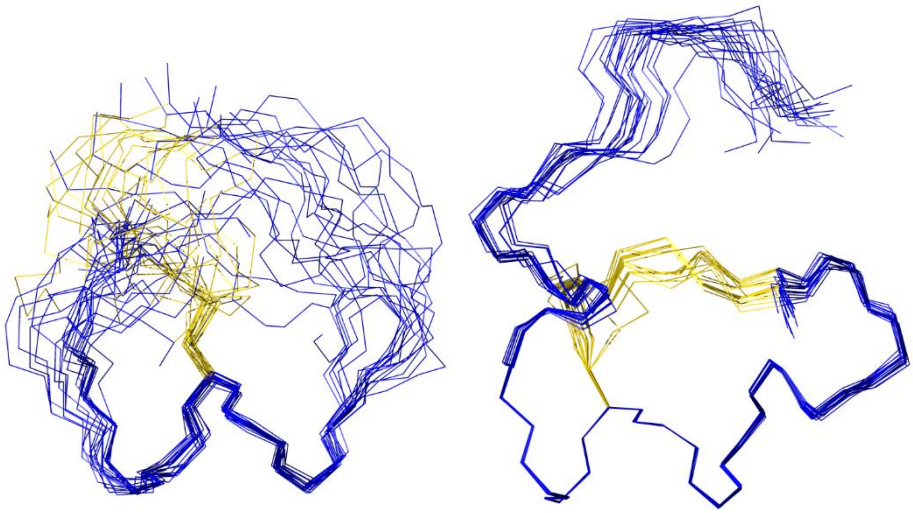
Figure 2: RP-HPLC/ESI-MS analyses of synthetic Δ-CIC and CIC. (A) Synthetic folded Δ-CIC UV chromatogram at 214 nm and ESI-MS analysis; (B) Synthetic folded Δ-CIC UV chromatogram at 214 nm and ESI-MS analysis. ACN gradient was from 0 to 30% over 30 min.

3.2 Electrophysiology

Since closely related PeIA, MII conotoxins are active on α3β2 nAChRs subtypes and N-terminally extended GID conotoxin are active on α3β2, α7 and α4β2, the biological activity of CIC and Δ-CIC was investigated using a two-electrode voltage clamp analysis on α3β2, α7 and α4β2 rat nAChRs subtypes. Unfortunately, according to the preliminary results, toxins only display a micromolar range IC_{50} on α3β2 and α4β2 nAChRs which is 1000-fold less potent than the expected range of potency of conotoxins ligands.

3.3 NMR spectroscopy

In order to determine if the extended N-terminal tail had any influence on the three-dimensional structures we also performed high-resolution NMR spectroscopy (Figure 3).



(A) (B)

Figure 3: Three-dimensional structures of Δ-CIC and CIC. The 20 lowest energy NMR structures, superimposed over the backbone atoms for (A) Δ-CIC and (B) CIC. The backbone is shown in line format and all the side-chains are hidden except the disulfide bonds, which are represented in yellow.

CIC structure is well defined, with low structures variation between the different states (Figure 3B). Interestingly the N-terminal tail seems to be pretty ordered. On the other hand, Δ-CIC exhibits much more structural variations suggesting structural stabilizing role of the extended N-terminal tail (Figure 3A).

4. Discussion

Given that, extended N-terminal tail of CIC could always explain the lack of activity compared to MII and PeIA conotoxins, the discussion will be based on the Δ-CIC sequence (Table 1).

Table 1: Sequence alignment and corresponding biological activities of Δ-CIC, PeIA and MII conotoxins. # indicates an amidated C-ter residue.

Name	Sequence	Activity	Refs
Δ-CIC	-CCSNPACQVQHSDLC	α3β2, α4β2 (micromolar)	
PeIA	GCCSHPACSVNHPELC#	α3β2=α6/α3β2β3>α9α10>α3β4 (nanomolar)	[235,314,399]
MII	GCCSNPVCHLEHSNLC#	α6/α3β2β3=α3β2=α7*α6β2*>α4β2 (nanomolar)	[233,246,361]

Hone *et al.* [235] by positional scanning identified Asn11 of PeIA as a critical residue interacting with α6 and α3 subunits, which is replaced with isosteric glutamine residue in Δ-CIC sequence. Everhart *et al.* performed ala-scan of MII conotoxin and showed that Asn5, Pro6 and His12 played critical role for conotoxin activity probably via Asn5 and His12 electrostatic interactions with β2E61 and βD169 as suggested by molecular modeling. Despite displaying all the sequence features required for the activity and high sequence homology with MII and PeIA, Δ-CIC conotoxin might not fit the receptor binding-pocket. Indeed, according to the structural determination, Δ-CIC displays high structural variations compared to native CIC, which might explain its lack of potency. Alternatively, it cannot be excluded that loss of activity is due lack of amidated C-terminal residue or conserved glycine at position 1. Interestingly, the N-terminal tail structure of CIC appears to be much conserved between the different states suggesting a structural stabilizing role although, α-GID N-terminal tail is very disordered in solution [276]. Overall, role of the extended N-terminal tail on biological activity could not be determined because of the lack of activity of the native toxin on tested receptors but we clearly see a role in stabilization of the three-dimensional structure. Since, CIC has been identified in predation-evoked venom of *Conus catus* it has to play a role in the evenomation process and displays a biological activity, additional bioassays will be conducted in the hope to find the appropriate target.

c. Synthèse de toxines de *Conus textile*

i. Synthèse des toxines TxVc et TxVD

Dans leur grande majorité, les α-conotoxines se présentent sous leur forme globulaire native, avec une connectivité des 4 cystéines 1-3, 2-4 (Figure 9). Contrairement aux toxines de *Conus catus* décrites dans les chapitres précedents, plusieurs α-conotoxines de *Conus textile* sont déjà largement utilisées comme outils pharmacologiques pour l'étude de certains sous types de nAChRs. Citons par exemple les α-conotoxines : TxIA qui bloque préférentiellement les sous-types α3β2 et α7 [166], TxIB qui est sélective du sous type α6/α3β2β3 [400] ou encore TxID qui bloque α3β4 [266]. De façon inattendue chez cette espèce, une autre famille de conotoxines ciblant les nAChRs a été mise en évidence, et celles-ci sont repliées différemment. En effet, cette famille, appelée « superfamille T » produit des toxines matures courtes contenant 4 cysteines repliées par paire dans la configuration « ruban », soit 1-4, 2-3. Nous avons donc poursuivi l'étude de cesconotoxines T aux propriétés pharmacologiques connues et identifiées chez *Conus textile* par la synthèse de leur isomère ruban. Cela représentait également l'opportunité de tester l'applicabilité de la stratégie de synthèse regioselective mise en place pour les toxines globulaires à deux ponts.

Dans un premier temps nous nous sommes intéressés à la conotoxine TxVC en raison de son activité et sélectivité unique décrite pour le sous type α4β2 des récepteurs nicotiniques, ainsi qu'à un nouveau variant TxVD identifié dans nos transcriptomes.

Cette étude est mise en forme comme un article, non publiable en l'état, dans lequel j'ai effectué la synthèse des toxines TcVC/D en version ruban et globulaire ainsi que la caractérisation par spectrométrie de masse. La discussion est basée sur des résultats de test d'activité biologique préliminaire par conséquent les procédures expérimentales d'électrophysiologie manquent à la partie méthodes.

1. Introduction

T-superfamily conotoxins display mostly two different cysteine patterns CC-CC- and CC-C-C, with the disulfide connectivity I-III, II-IV and I-VI, II-III respectively. To date over 300 T-superfamily conotoxins have been evidenced both at nucleic acid level and protein level [82,401]. Very few are pharmacologically important, they usually target sodium channels, presynaptic calcium channels and GPCRs [402,403]. For example, ε-TxIX reduces neurotransmitters release by decreasing the presynaptic influx of calcium [402,404] and Lt5d inhibit tetrodotoxin (TTX)-sensitive currents in adult rat dorsal root ganglion (DRG) neurons [405]. However, Wang *et al.* [278] isolated and sequenced [406] conotoxin TxVC which surprisingly inhibits α4β2 (IC_{50} 343 nM) and α3β2 (IC_{50} 1047 nM) nAChRs without significant

effects on DRG Na^+, K^+ or Ca^{2+} channels. Importantly this is the first reported conotoxin showing a positive selectivity profile for α4β2 nAChRs. To further investigate the mode of action and selectivity profile of T-superfamily related conotoxins, we report here on the synthesis and the pharmacological characterization of TxVD and a newly identified TxVD variant.

2. Methods

2.1 Abbreviations

Acm, acetamidomethyl; ACN, acetonitrile; Boc, *tert*-butoxycarbonyl; DCM, Dichloromethane; DIPEA, diisopropylethylamine; DMF, *N,N'*-dimethylformamide; DTP, 2,2'-Dithiopyridine; ESI-MS, electrospray ionization mass spectrometry; Fmoc, fluorenylmethoxycarbonyl; HATU, 1[*Bis*(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate; LC/MS, liquid chromatography/mass spectrometry; MeOH, methanol; nAChR, nicotinic acetylcholine receptor; NMR, nuclear magnetic resonance; Pbf, pentamethyl-dihydrobenzofuran-5-sulfonyl; RP-HPLC, reversed phase high performance liquid chromatography; SPPS, solid phase peptide synthesis; *t*Bu, *tert*-butyl;TFA, trifluoroacetic acid; TIS, triisopropylsilane; Tris, 2-Amino-2-(hydroxymethyl)propane-1,3-diol;Trt, trityl; UV, ultra-violet.

2.2 *Chemical synthesis*

DMF, DIPEA, ACN, TIS, TFA, piperidine and all others reagents were obtained from Sigma-Aldrich (Saint-Louis, MI, USA) or Merck (Darmstadt, Germany) and were used as supplied. Fmoc (L) amino acid derivatives and HATU were purchased from Iris Biotech (Marktredwitz, Germany). AmphiSpheres 40 RAM resin (0.37 mmol/g 75–150 µm) was purchased from Agilent Technologies (Les Ulis, France). The following side-chain protecting groups were used: Trt for Asn, II-IV Cys and His; *t*Bu for Ser and Asp; Boc for Lys; and Acm for I-III Cys. Peptides were manually synthesized using the Fmoc-based solid-phase peptide synthesis technique on a VWR (Radnor, PA, USA) microplate shaker. All Fmoc amino acids and HATU were dissolved in DMF to reach 0.5 M. Fmoc deprotection was carried out with piperidine in DMF (1/2 v/v) twice for 1 min. Subsequent amino acids were coupled onto 0.1 mmol of resin twice for 1 min using an amino acid/HATU/DIPEA ratio of 5:5:10 relative to resin loading. DMF was used for resin washing between deprotection and coupling steps. After chain assembly was complete, the terminal Fmoc group was removed and the resin washed with DMF and DCM. Side-chain (except acm) deprotection and cleavage from the resin was carried out by adding 10 mL of TFA/TIS/H_2O (95/2.5/2.5 v/v/v) and stirring the mixture for 2.5 h at room temperature. After the resin was removed by filtration and washed three times with dichloromethane. Dichloromethane and TFA are removed under vacuum then cold diethyl ether was added to precipitate the peptide. Crude peptides were purified by preparative RP-HPLC and pure fractions were combined and freeze-dried. A twostep

oxidation procedure was then carried out. The first disulfide bridge is formed between the free cysteine residues by dissolving the peptide at 0.2 mM in 50 mM Tris-HCl buffer adjusted to pH 8 and adding dropwise 7 equivalents of DTP at 10 mM in MeOH. When reaction was complete, the reaction mixture was acidified to pH 3 and loaded onto preparative RP-HPLC and pure fraction were combined. The second disulfide bridge was formed by deprotection/oxidation of the acm protecting group directly on the combined pure fractions of the mono bridged intermediates by treatment with 20 equivalents of 10 mM iodine in H_2O/TFA/ACN (78/2/20 v/v/v). When the reaction is complete, the reaction mixture is quenched with 20 mM ascorbic acid until total discoloration of the solution, acidified and purified by preparative RP-HPLC. The combined pure fractions were freeze-dried and their purity were confirmed by LC/ESI-MS.

2.3 Mass spectrometry

Solvents used for LC/MS were of HPLC grade. The LC/MS system consisted of a Waters (Milford, OH, USA) Alliance 2695 HPLC, coupled to a Waters Micromass ZQ spectrometer (electrospray ionization in positive mode (ESI+) fitted with a quadrupole mass analyzer). All the analyses were carried out using a Chromolith (Fontenay sous Bois, France) HighResolution RP-18e (4.6 x 25 mm, 15 nm–1.15 µm particle size, flow rate 3.0 mL/min) column. A flow rate of 3 mL/min and a gradient of 0–100% B over 2.5 min for routine analyses and 0-30% B over 30 min for quality control of pure products were used. Eluent A was water/0.1% HCO_2H and eluent B consisted of acetonitrile/0.1% HCO_2H. UV detection was performed at 214 nm. Electrospray mass spectra were acquired at a solvent flow rate of 200 µL/min. Nitrogen was used for both the nebulizing and drying gas. The data were obtained in a scan mode ranging from 100 to 1000 m/z or 250 to 1500 m/z to in 0.7 s intervals.

2.4 Preparative RP-HPLC

Preparative RP-HPLC was run on a Gilson PLC 2250 Purification system (Villiers le Bel, France) instrument using a preparative column (Waters DeltaPak C18 Radial-Pak Cartridge, 100 Å, 40 × 100 mm, 15 µm particle size, flow rate 50.0 mL/min). Buffer A was 0.1% TFA in water, and buffer B was 0.1% TFA in acetonitrile.

3. Results

Given that cone snail superfamily signal sequences are highly conserved [82], we search for TxVC sequence in our transcriptome of *Conus textile* (data not shown) based on T superfamily conotoxin signal sequence. We recovered a precursor sequence that contained a predicted mature conotoxin corresponding to the published TxVC, which is expected to have a C-terminal amidation [278] (Figure 1). Interestingly, we also recovered a second precursor sequence almost identical to TxVC, but with a

single N>S substitution at position 57 (or 9th position of the mature toxin). We named this variant TxVD.

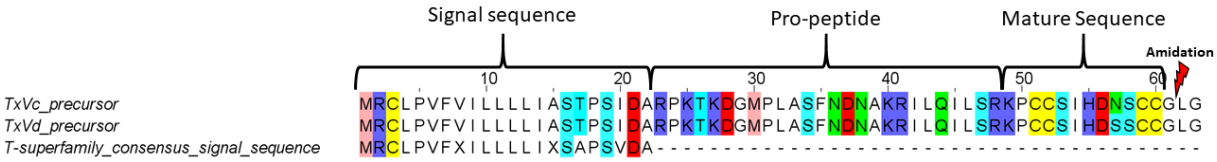


Figure 1: Alignment of TxVC/D precursors with T-superfamily consensus signal sequence. Higly conserved signal sequence and pro-peptide region greatly facilitated mature sequence and C-terminal amidation identification. Please note that TxVD is identical to TxVC except at position 57 where the Asn57 is substituted by Ser57 in TxVD sequence. X represents a non-conserved amino acid.

3.1 Chemical synthesis

TcVC and TxVD were synthesized manually using Fmoc-SPPS on rink amide resin. Considering the two most predominant disulfide connectivity from the T-superfamily *i.e.* globular CysI-CysIII, CysII-CysIV and globular CysI-CysIV, CysII-CysIII, we synthesized both isomer of each peptides. Linear peptides were folded using a regioselective folding strategy. Briefly, the first disulfide bridge is formed by oxidation of the free cysteine residues in an aqueous basic buffer assisted by DTP (2,2'-Dithiopyridine) and the second disulfide bond is formed by deprotection/oxidation of the Acm protecting groups. The homogeneity of folded toxins was assessed by analytical RP-HPLC coupled to ESI-MS (Figure 2). ESI-MS(+) confirmed an average mass of 1473.9 g/mol (calculated 1473.583 g/mol) for TxVC and 1447.2 g/mol (calculated 1447.709 g/mol) for TxVD. Overall, 16.1 mg and 14.05 mg (10.9% and 9.5% yield from resin loading respectively) of TxVC ribbon and globular were obtained respectively; 13.6 mg and 18.3 mg (9.4% and 12.7% yield from resin loading respectively) of TxVD ribbon and globular were obtained respectively.

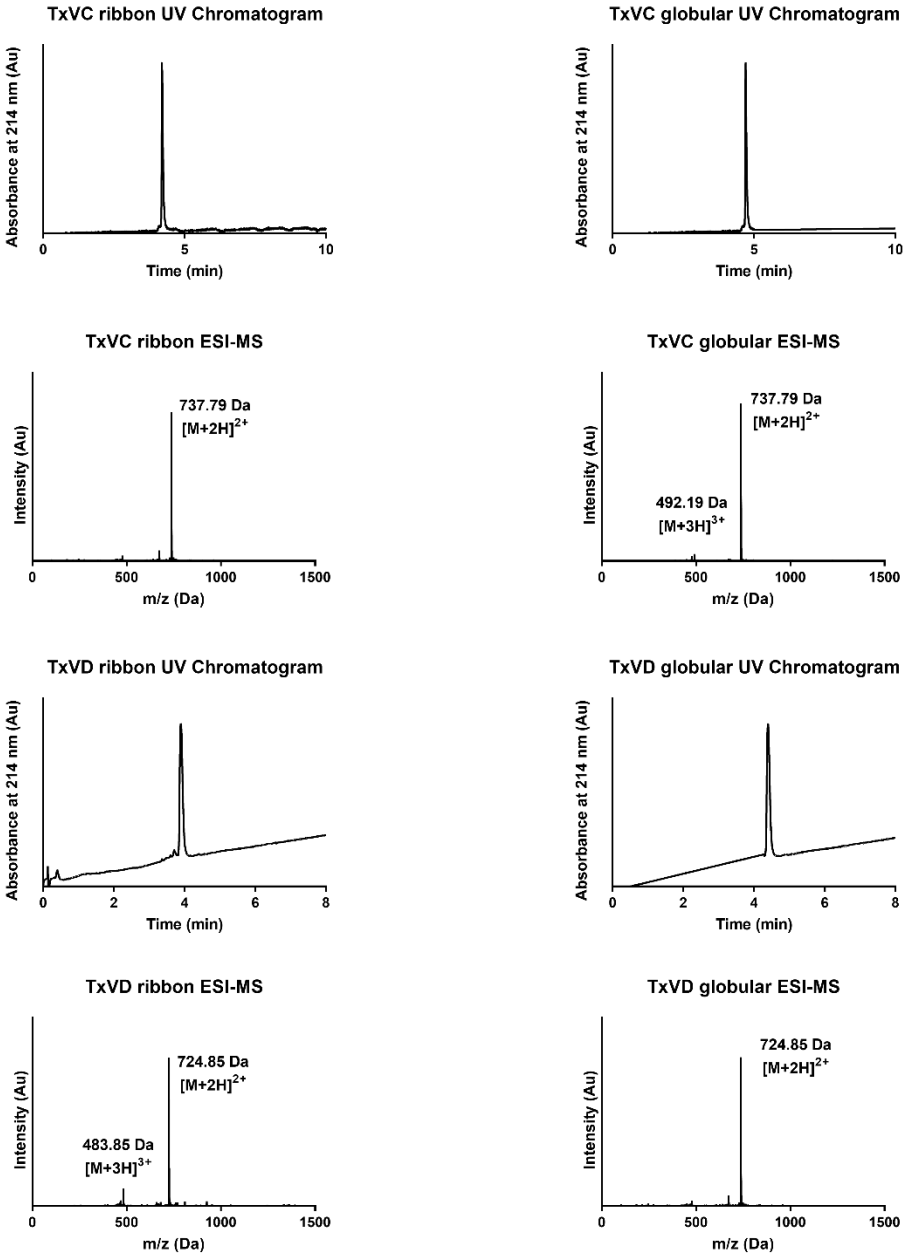


Figure 2: RP-HPLC/ESI-MS analyses of synthetic TxVC and TxVD. (A) Synthetic folded TxVC ribbon and (B) globular UV chromatogram at 214 nm and ESI-MS analysis; (C) Synthetic folded TxVD ribbon and (D) globular UV chromatogram at 214 nm and ESI-MS analysis. ACN gradient was from 0 to 30% over 10 min.

3.2 Electrophysiology

Since TxVD is almost identical to TxVC sequence at the exception of N57S substitution, their biological activities were investigated in parallel using a two-electrode voltage clamp analysis on α3β2, α7 and α4β2 rat nAChRs subtypes. Unfortunately, and contrary to what was originally described, none of the toxins showed activity up to 10 µM (Table 1). To further confirm this surprising lack of activity, a second

batch of TxVC synthesized by Australian colleagues (Pr. Alewood & Lewis, The University of Queensland) was tested using the same conditions, but again no activity was detected.

Table 1: Summary of the results obtained by electrophysiology. N.D = not detected.

Toxin	Concentration	α7 nAChRs	α4β2 nAChRs	α3β2 nAChRs
TxVC ribbon	10 µM	N.D	N.D	N.D
TxVD ribbon	10 µM	N.D	N.D	N.D
TxVC globular	10 µM	N.D	N.D	N.D
TxVD globular	10 µM	N.D	N.D	N.D

4. Discussion

The neuronal nAChRs α4β2 subtype is the most abundant in the human brain, playing a key role towards the effectiveness of synaptic connection by regulating the release of other neurotransmitters [209,269,270]. The α4β2 nAChR also plays an important role in nicotine addiction and cognitive performance [209,269–271], which makes it a valuable pharmacological target for improved cognition, smoking cessation, analgesia, and possibly for a wide range of neurological disorders such as Alzheimer's and Parkinson's disease, depression, and attention deficit disorders [272–274]. To date, the only described conotoxins targeting the α4β2 are α-GID [276], α-GIC [407] and α-MII [361], which all belong to the superfamily A and are structurally significantly different from TxVC/TxVD (superfamily T). Moreover, based on the original publication, TxVC was the first conotoxin showing a positive α4β2 selective profile with a α4β2 *vs.* α3β2 ratio of approximatively 3-fold, whereas GID, GIC and MII exhibit an inverted α3β2 *vs.* α4β2 ratio of approximatively 50-, 300- and 1000-fold, respectively. Such conotoxin would represent a unique and promising tool to investigate the α4β2 subtype and a good candidate for further SAR studies.

Unfortunately, our results suggest that most likely, the original study of Wang *et al*. had flaws and wrongly attributed nAChR activity to TxVC. Indeed, our attempts to reproduce their results failed. Although the synthesis of TxVC does not present any particular challenge and is rather straightforward, we obtained another batch of TxVC synthesized by a highly reputable and experienced group (Alewood PF, UQ, Australia) in order to exclude any potential issue on our side. Both peptides were found identical by LC-MS analysis (same molecular masses and retention times), and both were inactive at all nAChRs tested. Further biological tests on nAChRs were carried out by Australian colleagues using different methods (FLIPR, binding assays), without success. We hypothesized that Wang *et al*. may have investigated the ribbon isomer instead of the globular isomer, or possibly that traces of the ribbon

isomer had contaminated the original study and were producing the apparent activity. Thus, we have synthesized the ribbon forms of both TxVC and TxVD, but both were totally inactive. Different nAChR isoforms could also explain the discrepancy of the results, but since Wang *et al.* [278] also used rat nAChRs in their study and that we synthesized both ribbon and globular isomers, the observed lack of potency of TxVC/D could not be rationally explained. As a last point, it appears highly suspicious that all analogues (single alanine substitution) synthesized by Wang *et al*. show a complete loss of activity compared to the native TxVC. Indeed, the vast majority of such alanine-walk studies usually show some patterns of unaffected activity, reduced activity and rarely complete loss of activity for these analogues. Therefore, we suggest that the molecular target of TxVC is not the nAChRs and still remains to be identified. Thus, additional bioassays on other receptors (ion channels and other membrane receptors) will be carried out to elucidate the appropriate target.

ii. Synthèse de l'isomère ruban de l'α-conotoxine TxIA

De façon originale, la synthèse d'isomère ruban de toxine de *Conus textile* a été poursuivie dans le cadre de la caractérisation de produit obtenu par expression recombinante de l' α-conotoxine TxIA. En effet, celle-ci menait à l'obtention d'un produit moins actif que l'isomère natif initiallement caractérisé, et il était soupçonné que cela était dû à la formation d'un fold non natif dans la machinerie cellulaire. La synthèse chimique stéréo-dirigée de l'isomère ruban a servi de référence pour comparer les structures RMN et les activités biologiques afin de confirmer la production majoritaire de l'isomère ruban par expression recombinante. Cela était également l'occasion de confirmer par la structure RMN la validité de la méthodologie pour la synthèse d'isomères rubans non natifs.

Cette étude est présentée sous forme d'une publication parue [408] dans laquelle j'ai effectué la synthèse de l'isomère ruban de la conotoxine TxIA, et sa caractérisation par spectrométrie de masse.

Periplasmic Expression of 4/7 α-Conotoxin TxIA Analogs in E. coli Favors Ribbon Isomer Formation – Suggestion of a Binding Mode at the α7 nAChR

Yamina El Hamdaoui[1†‡], Xiaosa Wu[2‡], Richard J. Clark[3], Julien Giribaldi[4], Raveendra Anangi[2], David J. Craik[2], Glenn F. King[2], Sebastien Dutertre[4], Quentin Kaas[2*‡], Volker Herzig[2*‡] and Annette Nicke[1*‡]

[1] Walther Straub Institute of Pharmacology and Toxicology, Faculty of Medicine, Ludwig-Maximilians-Universität München, Munich, Germany, [2] Institute for Molecular Bioscience, The University of Queensland, Brisbane, QLD, Australia, [3] School of Biomedical Sciences, The University of Queensland, Brisbane, QLD, Australia, [4] CNRS, Institut des Biomolécules Max Mousseron, UMR 5247, Université de Montpellier, Montpellier, France

Abstract

Peptides derived from animal venoms provide important research tools for biochemical and pharmacological characterization of receptors, ion channels, and transporters. Some venom peptides have been developed into drugs (such as the synthetic ω-conotoxin MVIIA, ziconotide) and several are currently undergoing clinical trials for various clinical indications. Challenges in the development of peptides include their usually limited supply from natural sources, cost-intensive chemical synthesis, and potentially complicated regioselective disulfide-bond formation in the case of disulfide-rich peptides. In particular, if extended structure–function analysis is performed or incorporation of stable isotopes for NMR studies is required, the comparatively low yields and high costs of synthesized peptides might constitute a limiting factor. Here we investigated the expression of the 4/7 α-conotoxin TxIA, a potent blocker at α7β2 and α7 nicotinic acetylcholine receptors (nAChRs), and three analogs in the form of maltose binding protein fusion proteins in *Escherichia coli*. Upon purification via nickel affinity chromatography and release of the toxins by protease cleavage, HPLC analysis revealed one major peak with the correct mass for all peptides. The final yield was 1–2 mg of recombinant peptide per liter of bacterial culture. Two-electrode voltage clamp analysis on oocyte-expressed nAChR subtypes demonstrated the functionality of these peptides but also revealed a 30 to 100-fold potency decrease of expressed TxIA compared to chemically synthesized TxIA. NMR spectroscopy analysis of TxIA and two of its analogs confirmed that the decreased activity was due to an alternative disulfide linkage rather than the missing C-terminal amidation, a post-translational modification that is common in α-conotoxins. All peptides preferentially formed in the ribbon conformation rather than the native globular conformation. Interestingly, in the case of the α7 nAChR, but not the α3α2 subtype, the loss of potency could be rescued by an R5D substitution. In conclusion, we demonstrate efficient

expression of functional but alternatively folded. Ribbon TxIA variants in *E. coli* and provide the first structure–function analysis for a ribbon 4/7-α-conotoxin at α7 and α3α2 nAChRs. Computational analysis based on these data provide evidence for a ribbon α-conotoxin binding mode that might be exploited to design ligands with optimized selectivity.

1. Introduction

Nicotinic acetylcholine receptors (nAChRs) are members of the Cys-loop superfamily of pentameric ligand-gated ion channels. Whereas the muscle-type receptors consist of four different subunits (α1,β1, γ/ε,δ), the assembly and stoichiometry of the 11 cloned neuronal subunits (α2–α7, α9, α10, and α2–α4) into nAChR subtypes is considerably more diverse and not welldefined. These neuronal nAChRs represent important drug targets such as for the treatment of pain, Alzheimer's disease, and nicotine addiction [20,409–412]. Deciphering their specific compositions and location in nervous system tissues as well as their respective physiological roles requires highly selective ligands. Venoms from animals contain complex mixtures of small proteins and peptides that are often structurally constrained by disulfide bridges. Some of these peptides show high potency and specificity for certain ion channel subtypes. Cone snails are predatory marine animals that have evolved numerous peptide families that target various ligand- and voltage-gated ion channels [82]. The best-studied conotoxin family from a pharmacological perspective is the α-conotoxins, members of which are competitive antagonists of nAChRs [166]. Some of these peptides have analgesic activity in vivo and thus are important lead structures for drug development [82,313]. The majority of α-conotoxins are composed of 12–19 amino acid residues including four cysteine residues that form two disulfide bonds. The cysteines are arranged in a CC–C–C pattern that defines the conotoxin Cysteine Framework I [413]. This framework is characterized by vicinal Cys1 and Cys2 residues and two loops formed by Cys1–Cys3 and Cys2–Cys4 disulfide bridges (referred to as the "globular" conformation). Based on the number of amino acid residues within the two loops, the currently characterized α-conotoxins are further classified into 3/4, 4/4, 3/5, 4/6, and 4/7 α-conotoxin subfamilies. These subfamilies show some common specificity for certain nAChR subtypes, with for example, the 3/5 α-conotoxins targeting the muscle-type nAChR and most identified 4/7 α-conotoxins preferentially targeting α7 and/or α3β2* neuronal nAChRs (* indicates the potential presence of further subunits) [166]. Understanding the structure-activity relationships of conotoxins might aid in the development of optimized peptides with tailored selectivity. Usually, such studies employ chemical synthesis for the production of modified versions of the toxins. However, the production of multiple analogs or large quantities for automated application systems or preclinical treatment studies is costly, as is the production of large quantities of isotopically enriched samples for high resolution NMR spectroscopy studies or metabolic flux analysis

[414]. Chemical synthesis is also tedious if done manually and requires special equipment and experience that is not typically found in molecular biology laboratories. More generally, in the case of larger peptides (>40 aa), the yield from chemical synthesis is typically low. Finally, certain native peptides are inherently difficult to produce synthetically. Venom-peptide production in heterologous expression systems might provide an efficient and economical alternative to chemical synthesis for molecular biology laboratories [415]. It might also be suitable for large scale commercial toxin production. In the current study, we adapted an Escherichia coli periplasmic expression system [415] for the production of 4/7 α-conotoxin TxIA and three analogs. Unexpectedly, the functional and structural characterization of the expressed analogs indicated that they adopt a fold different from the native peptide (i.e., a 1–4, 2–3 "ribbon" rather than a 1–3, 2–4 "globular" disulfide conformation). These data demonstrate the need for careful structural analysis and confirm earlier findings that non-native folds could still be active. As an explanation for this activity and a basis for future structural studies, we provide the first NMR solution structure of a 4/7 α-conotoxin ribbon isomer and propose a binding mode for this peptide at the α7 nAChR. This information might be useful for the design of new lead structures based on the ribbon α-conotoxin scaffold.

2. Materials and methods

2.1 Preparation of plasmids

A nucleotide sequence encoding a tobacco etch virus (TEV) protease recognition site (ENLYFQG) followed by [R5N,I9H] TxIA and including KpnI and SacI restrictions sites at the 50 and 30 ends, respectively, was optimized for expression in *E. coli* and assembled from synthetic oligonucleotides (Life Technologies). This construct was cloned in-frame to an 5' His_6-maltose binding protein (MBP) sequence into the pLICC_D168 vector [416] using *KpnI* and *SacI* cloning sites. The conserved N-terminal glycine residue of the conotoxin sequence forms the last amino acid of the TEV protease recognition site and thus, upon cleavage, produces a native N-terminus in the conotoxin. Plasmids for production of TxIA as well as single mutated R5D and R5N analogs were obtained from this construct by site-directed mutagenesis. The sequences of all plasmids were confirmed by Sanger sequencing.

2.2 Recombinant expression of TxIA analogs in *E. coli*

α-Conotoxin expression in *E. coli* was performed using a previously described method [415,416]. BL21 strains of *E. coli* were transformed with the respective plasmids using a standard heat shock protocol. Five colonies were selected to inoculate 5 ml culture medium (Luria-Bertani plus ampicillin) and grown at 37°C until the optical density (OD_{600}) reached 1.0. To determine optimal conditions for toxin expression, 4 × 1 ml samples of each culture were then diluted and grown again in 5 ml cultures until

an OD of 0.6–0.8 was reached. Then heterologous protein expression was induced with IPTG. In doing so, culturing at 16°C and 37°C and induction with 0.5 and 1.0 mM IPTG were performed and protein expression in induced and non-induced cultures was compared after 4 or 12h by SDS-PAGE analysis of a cell pellet from 500 μl of culture medium. The clone with the highest protein expression was selected and used to start 1–2 l cultures (from 50–100 ml precultures), which were then grown at 37°C with shaking at ~220 rpm. After reaching an OD of 0.6–0.8, toxin expression was induced with 1 mM IPTG and cultures were grown overnight at 16° C. Cultures were then centrifuged (15 min at 10,500 g) and bacterial pellets frozen at –80°C.

2.3 SDS-PAGE analysis

500 μl of culture was centrifuged (5 min, 3,381 g (6,000 rpm) in a desktop centrifuge) and cell pellets were supplemented with 30 μl SDS running buffer and 30 μl 3× loading buffer. After heating (10 min at 95–100°C), 30 μl of the solubilized cells were separated on a 10% SDS-PAGE gel under reducing conditions (5% β-mercaptoethanol in the sample). To control the purification and cleavage process, equivalent volumes (30–200 μl) of lysate, flow-through, eluate, and cleaved sample were supplemented with SDS running buffer to obtain 200 μl. After addition of 200 μl 3× SDS running buffer, 50 μl of each sample were separated using SDS-PAGE.

2.4 Recombinant peptide purification and cleavage

Cell pellets were resuspended in 50 ml equilibration buffer (TN: 25 mM TrisHCl, 300 mM NaCl, pH 7.0) and lysed in a constant-pressure cell disrupter (27 kpsi, TS Series Cell Disrupter, ConstantSystems Ltd., Daventry, UK). After centrifugation (4°C, 30 min, 44,267 g), the supernatant was diluted with an equal volume of modified TN buffer (40 mM TrisHCl, 400 mM NaCl, pH 8.0), supplemented with DNAse (1 μg/ml lysate) and incubated for 30 min with 5 ml Ni-NTA Superflow resin (Qiagen Pty Ltd., Chadstone, VIC, Australia) in a gravity-fed column. The column was washed with 50 ml TN buffer containing 15 mM imidazole and then the His_6-MBP-TxIA fusion protein was eluted with 3 × 10 ml (each time 30 min) TN buffer containing 400 mM imidazole. The flow-through was concentrated to 5 ml (Amicon Ultra 30K filter) and diluted with 5 ml TN buffer. TxIA peptide was liberated from the fusion protein by cleavage overnight at room temperature using ~100 μg TEV protease. In order to maintain TEV protease activity, reduced/oxidized glutathione were added in a ratio of 1:5 w/w.

2.5 Recombinant peptide purification by HPLC

The protease-cleaved samples were acidified with trifluoracetic acid (TFA), centrifuged (10 min at 6,762 g) and filtered (Milipore ultrafree MC 0.2 μm) to remove protein precipitates. Then 50% acetonitrile (ACN) was added to a final concentration of 5% ACN and the peptides were separated on

a C18 semi-preparative reversed-phase (RP) HPLC column (Vydac protein/peptide C18 column Cat# 218TP1010) on a Shimadzu Prominence HPLC system (Shimadzu, Rydalmere, NSW, Australia). The following linear gradients of solvent B (90% ACN, 0.043 % TFA in water) in solvent A (0.5% TFA in water) were used at a flow rate of 3 ml/min: 5% B for 10 min, then 5–35% B for 30 min followed by 35–80% B over 5 min and 80% B for another 7 min. Absorbance was determined at 214 nm and 280 nm and collected fractions were lyophilized and stored at –20°C. The fractions containing the correct peptide masses as determined by matrix-assisted laser desorption/ionization mass spectrometry (MALDI MS) were then subjected to another HPLC fractionation using a hydrophilic interaction liquid chromatography (HILIC) column. Samples were dissolved in 95% solvent B and injected into a VisionHT HILIC column (5 µm particle size, 150 × 4.6 mm; Grace, Columbia, MD, USA) at a flow rate of 1 ml/min. The same solvents as during the RP-HPLC fractionation were used with the following linear gradients: 95% solvent B for the first 3 min, then 95%–75% B over 20 min. Absorbance was determined at 214 nm and 280 nm and the molecular masses of the peptides determined using MALDI MS.

2.6 Analysis Analysis by liquid chromatography/mass spectrometry

Solvents used for LC/MS were of HPLC-grade. Recombinant peptide masses were determined by MALDI time-of-flight (TOF) MS using a 4700 Proteomics Bioanalyzer model (Applied Biosystems, Carlsbad, CA). Peptides were dissolved in water and mixed 1:1 (v/v) with α-cyano-4-hydroxycinnamic acid matrix (7 mg/ml in 50% ACN, 5% formic acid) and mass spectra acquired in positive reflector mode. All reported masses are for the monoisotopic [$M+H^+$] ions.

For analysis of the synthetic TxIA ribbon analogue, the LC/MS system consisted of an Waters Alliance 2695 HPLC (Milford, OH, USA) coupled to a Micromass ZQ mass spectrometer (electrospray ionization mode, ESI+). All analyses were carried out using a Chromolith (Fontenay sous Bois, France) HighResolution RP-18e column (4.6 × 25 mm, 15 nm–1.15 µm particle size). A flow rate of 3 ml/min and a gradient of 0–100% solvent B over 2.5 min were used. Eluent A was water/0.1% formic acid while eluent B was ACN/0.1% formic acid. UV detection was performed at 214 nm. ESI mass spectra were acquired using a solvent flow rate of 200 µl/min. Nitrogen was used for both the nebulizing and drying gas. Data were obtained in a scan mode over the m/z range 100–1000 or 250–1500 in 0.7 s intervals. Fully folded synthetic TxIA ribbon was characterized using a Synapt G2-S high-definition MS system (Waters Corp., Milford, MA) equipped with an ESI source. Chromatographic separation was carried out at a flow rate of 0.4 ml/min on a Acquity H-Class ultrahigh performance liquid chromatography (UPLC) system (Waters Corp., Milford, MA), equipped with a Kinetex C18 100A column (100 × 2.1 mm, 2.6 µm particle size) from Phenomenex (France). The mobile phase consisted of water (solvent A) and ACN (solvent B) with both phases acidified by 0.1% (v/v) formic acid. Mass spectra were acquired in the positive ionization mode.

2.7 Two-electrode voltage clamp (TEVC) electrophysiology

cDNAs encoding rat nAChR subunits were provided by J. Patrick (Baylor College of Medicine, Houston, TX, USA) and subcloned into the oocyte expression vector pNKS2 [417]. cRNA was synthesized with the SP6 mMessage mMachine Kit (Ambion, Austin, TX, USA) and adjusted to a concentration of 0.5 μg/μl. nAChR subunit RNAs were mixed in the ratios 1:1 (α3/β2) or 5:1 (α4/β2). *Xenopus laevis* (Nasco International, Fort Atkinson, WI, USA) oocytes were injected with 50 nl aliquots of cRNA (Nanoject automatic oocyte injector, Drummond Scientific, Broomall, PA). Antagonist concentration-response curves were measured as described previously [246,279] in ND96 (96 mM NaCl, 2 mM KCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$, and 5 mM HEPES at pH 7.4). In brief, current responses to acetylcholine were measured at room temperature 1–6 days after cRNA injection and recorded at −70 mV using a Turbo Tec 05X Amplifier (NPI Electronic, Tamm, Germany) and Cell Works software. A standard concentration of 100 μM ACh or nicotine was used to activate the $\alpha\beta$ combinations and the α7 nAChR, respectively. A fast and reproducible solution exchange (<300 ms) was achieved with a 50-μl funnel-shaped oocyte chamber combined with a fast vertical solution flow fed through a custom-made manifold mounted immediately above the oocyte. Agonist pulses were applied for 2 s at 4-min intervals. Following 1 min of perfusion directly after the agonist application, peptides were applied in a static bath for 3 min. IC_{50} values were calculated from a nonlinear fit of the Hill equation to the data (Prism version 6.0; GraphPad Software, Inc., San Diego, CA). Data are presented as means ± S.E.M. from at least three oocytes.

2.8 Chemical synthesis of ribbon and native α-TxIA

ACN, TFA, dimethylformamide (DMF), *N*,*N*-diisopropylethylamine (DIPEA), triisopropylsilane (TIS), dichloromethane (DCM), piperidine and other reagents were obtained from Sigma-Aldrich (Saint-Louis, MI, USA) or Merck (Darmstadt, Germany) and used as supplied. Fmoc L-amino acid derivatives and 1[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate (HATU) were purchased from Iris Biotech (Marktredwitz, Germany). AmphiSpheres 20 HMP resin (0.6 mmol/g) was purchased from Agilent Technologies (Les Ulis, France). The side chain protecting groups for amino acids are *t*Bu for Asp and Ser; Acm for $Cys_{2,8}$; Trt for $Cys_{3,16}$ and Asn; Pbf for Arg. Ribbon α-TxIA was manually synthesized using the Fmoc-based solid-phase peptide synthesis technique on a VWR (Radnor, PA, USA) microplate shaker. All Fmoc amino acids and HATU were dissolved in DMF to reach 0.5 M. Chain elongation was performed step by step using 0.2 mmol of AmphiSpheres 20 HMP resin. Fmoc deprotection was performed with 20% piperidine in DMF two times, each for 1 min at room temperature, then the resin was washed three times with DMF. Each Fmoc-protected amino acid (5 eq.) was coupled twice in the presence of HATU (5 eq) and DIPEA (10 eq) in DMF at room temperature for 2 min. After completion of the coupling reaction, the resin was sequentially washed twice with DMF. Cleavage of peptide from the resin and removal of side-chain protecting groups were

carried out using 20 ml of a solution containing TFA/TIS/H_2O (95:2.5:2.5, v/v/v) for 3 h at room temperature. After the resin was removed by filtration and washed with DCM, the DCM and TFA were removed under vacuum and cold diethyl ether was added to precipitate the peptide. Crude peptide was purified by preparative RP-HPLC, and its purity determined using LC/ESI-MS. Trt groups were removed during the cleavage step while the Acm protective groups are resistant to cleavage conditions. A two-step oxidation procedure was then carried out. Briefly, the first disulfide bridge was formed from free cysteines with the use of 2,2'-dithiopyridine (DTP) [196,357] and then the second disulfide bridge was formed by concomitant deprotection and oxidation of the Acm groups with the use of iodine [356]. The fully folded peptide was purified by analytical RP-HPLC and purity confirmed by high-resolution mass spectrometry. During the purification process, the ribbon form of α-TxIA appears as a broad peak on analytical RP-HPLC (containing unseparable major and a minor peaks), which is possibly due to slow *cis-trans* interconversion of the proline residues.

Globular TxIA was synthesized as described (Dutertre et al., 2007).

2.9 Synthetic peptide purification

Preparative RP-HPLC was run on a Gilson PLC 2250 HPLC system (Villiers le Bel, France) using a preparative column (Waters DeltaPak C18 Radial-Pak Cartridge, 100 Å, 40 x 100 mm, 15 µm particle size, flow rate 50.0 mL/min). Buffer A was 0.1% TFA in water, and buffer B was 0.1% TFA in ACN. Fully folded synthetic TxIA ribbon was purified on a UltiMate 3000 UHPLC system (ThermoFischer scientific) using an Kinetex C18 100 A column (100×2.1 mm, 2.6 µm particle size) from Phenomenex (France). Buffer A was 0.1% HCOOH in water, and buffer B was 0.1% HCOOH in acetonitrile.

2.10 NMR spectroscopy analysis

2.10.1 Secondary shifts of recombinant peptides

Peptides were dissolved in 90% H_2O/10% D_2O (Cambridge Isotope Laboratories) at a concentration of 1 mg/ml and pH ~3.6. Spectra were recorded on a Bruker Avance 600 MHz spectrometer at 7°C and referenced to 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) at 0 ppm. Standard Bruker pulse programs were used for all two-dimensional spectra. Excitation sculpting with gradients was used to achieve water suppression for TOCSY and NOESY experiments [418]. NMR experiments included TOCSY [419] using a MLEV-17 spin lock sequence with a 80 ms mixing time, NOESY [420] with a 200 ms mixing time, ^{1}H- ^{13}C HSQC [421] and ^{1}H-^{15}N HSQC [421]. Spectra were recorded with 4096 data points in the F2 dimension and 512 increments in the F1 dimension for TOCSY, and NOESY experiments, 2048 x 240 for ^{1}H-^{13}C HSQC and 2048 x 128 for ^{1}H-^{15}N HSQC data points in the F2 dimension and increments in the F1 dimension, respectively. The t1 dimension was zero-filled to 1024 real data points, and the F1 and

F2 dimensions were multiplied by a sine-squared function prior to Fourier transformation. All spectra were processed using TopSpin 2.1 (Bruker) and assigned with CcpNmr Analysis [422] using the sequential assignment protocol [388]. Secondary shifts were calculated using the random coil values reported by Wishart et al. [423].

2.10.1 Structure of synthesized ribbon TxIA

A 2D NOESY spectrum was acquired with a mixing time of 200 ms at 7°C and interproton distance constraints were calculated from the relative intensities of NOE cross-peaks. Predictions of ϕ and ψ backbone angles were performed with TALOS-N [424]. Distance restraints for the disulfide bonds defined by the regioselective synthesis were used in the structure calculations. The ANNEAL function in CYANA was used to perform 10,000 steps of torsion angle dynamics to generate an initial ensemble of 100 structures from which the 20 structures with the lowest penalty function values were chosen for analysis. Several rounds of structure calculations were performed to resolve distance and angle constraint violations. Using protocols from the RECOORD database [425], an ensemble of 100 structures was subsequently calculated with CNS [426] using the force field distributed with Haddock 2.0 [427] and further refined in a water shell [428]. A set of 20 structures with the lowest energy and no NOE violations greater than 0.2 Å or dihedral-angle violations greater than 3° was selected for MolProbity analysis [429]. The structures were visualized and figures generated using MOLMOL [384] and figures were generated using MOLMOL and PyMOL.

2.11 Molecular modeling

Molecular models of the complexes between the ligand-binding domain of α7 nAChR and conotoxins were prepared by homology modeling and then refined using either molecular dynamics (MD) simulations or the ToxDock method [430]. Homology models were constructed using Modeller 9v18 [431] using three templates: (i) the NMR solution structure of ribbon TxIA (this study); (ii) the α4/α4 interface of the cryo-electron microscopy structure of the human α4β2 nAChR (PDB 6CNK) [432]; and (iii) the crystal structure of the complex between globular conotoxin [A10L]TxIA and the Aplysia californica acetylcholine binding protein (AChBP, PDB 2UZ6) [166]. A hundred models were generated for each system and the model with the lowest DOPE score [433] was selected for further refinement.

2.11.1 Refinement with MD simulations

Homology models of the interaction of ribbon deamidated TxIA, [R5D] ribbon TxIA, and [R5N] ribbon TxIA with rat α7 nAChR were obtained by homology modelling as described above. These initial models were then minimized and subjected to MD simulations in Gromacs 2018 using the Ambr99SB-ILDN forcefield [434]. The systems were embedded in a cubic box of 11. $\times$ 11 $\times$ 9 nm^3 and solvated with

~32,000 water molecules and several sodium ions to neutralize the systems. The system was then energy minimized using 10,000 steps of steepest gradient and submitted to an equilibration protocol during which position restraints on the receptor and toxin were progressively released; the constraints on the toxin atoms were released first over 6 ns, and the constraints on the receptor then over 2.4 ns. The systems were then simulated unrestrained for 50 ns. Long-range electrostatic interactions were computed using particle mesh Ewald with default parameters in Gromacs. All bonds involving hydrogens were constrained with the LINCS algorithm, and the time step of the simulation was set to 2 fs. The V-rescale thermostat [435] was used to maintain the temperature at 27°C and the pressure was maintained at one atmosphere using the Berendsen isotropic pressure coupling. The short range electrostatic and Van der Waals cutoffs were both set at 1 nm. The binding free energy of each system was then computed using the MMPB/SA method [436] as implemented in Amber18 by analyzing 10 frames extracted every ns from the last 10 ns of the unrestrained simulation.

2.11.2 ToxDock refinement

The ToxDock [430] refinement and energy computations were carried out using the online server Rosie [437]. Briefly, the Rosetta fast relax protocol [438] was used to generate 200 alternative models representing small structural fluctuation of the complex. The 50 models with lowest Rosetta total score were submitted for a FlexPepDock refinement [439], generating 500 alternative binding modes of the peptide bound to the receptor and computing the reweighted docking Rosetta score. The average of the 25 lowest reweighted score was taken as the relevant to the binding score of the complex, as previously described [430]. The Talaris2013 energy function was used for all Rosetta computations [440].

3. Results

3.1 Efficient production of recombinant TxIA analogues

Like most characterized 4/7 α-conotoxins, TxIA preferentially targets $\alpha7$ and $\alpha3\beta2$ nAChRs with nanomolar potency but has no detectable activity at the $\alpha4\beta2$ subtype [166]. Whereas some $\alpha3\beta2$-selective conotoxins such as MII [441] also have high potency at the closely related $\alpha6\beta2$ receptor, no conotoxin that selectively targets the $\alpha4\beta2$ nAChR has been yet identified, and only a few 4/7 α-conotoxins, such as MII [361], GID [276], and GIC [407] have been shown to block this receptor, albeit at high nanomolar or micromolar concentrations. Therefore, one aim of this study was to identify amino acid side chains that could prevent conotoxin binding to $\alpha4\beta2$ receptors by investigating if TxIA could be modified to obtain a conotoxin with at least weak activity at $\alpha4\beta2$. To this end, we compared the amino acid sequence of TxIA with those of MII, and GID and selected two positions, 5 and 9, in

which we substituted the amino acid side chains (Fig 1A). Position 5 was previously shown to be important for MII binding to $\alpha 3\beta 2$ nAChRs [248].

The DNA sequences corresponding to TxIA and analogues were optimized for expression in *E. coli* and fused N-terminally via a TEV protease cleavage site to MBP (Fig. 1B). Although the consensus TEV protease cleavage site is ENLYFQ/S, the protease also efficiently cleaves at ENLYFQ/G [442], which has the advantage that the N-terminal glycine residue that is present in most α-conotoxins can be used and remains after cleavage. Expression of the TxIA analogues in form of a His_6-MBP fusion protein was done to target the construct to the *E. coli* periplasmic space where the disulfide bond folding machinery of the bacterium (i.e., the Dsb proteins) is located and the oxidative environment is favorable for conotoxin folding and disulfide bond formation. Upon transformation of *E. coli* with these constructs and induction with IPTG, the respective fusion proteins represented the dominant proteins produced and they could be successfully purified and cleaved as indicated by the size shift of the MBP fusion construct (Fig. 1C, S1A). Further purification by RP-HPLC revealed one dominant peak for each of the TxIA analogues (Fig 2A, S1B). The corresponding fractions contained the calculated mass of the fully oxidized peptides (compare insets in Fig. 2C) and turned out to represent the sole functional fractions when tested by TEVC electrophysiological analysis on $\alpha 3\beta 2$ nAChRs expressed in *Xenopus laevis* oocytes (Fig. 2B, S1C). The obtained yields were 1–2 mg of conotoxin per liter of culture. An additional HILIC-HPLC purification step was employed to obtain pure recombinant peptides as determined using MALDI-MS (Fig. 2C, S2). These purified peptides were then used for preparation of concentration response curves (see below). The double mutant [R5N,I9H] was not further analyzed.

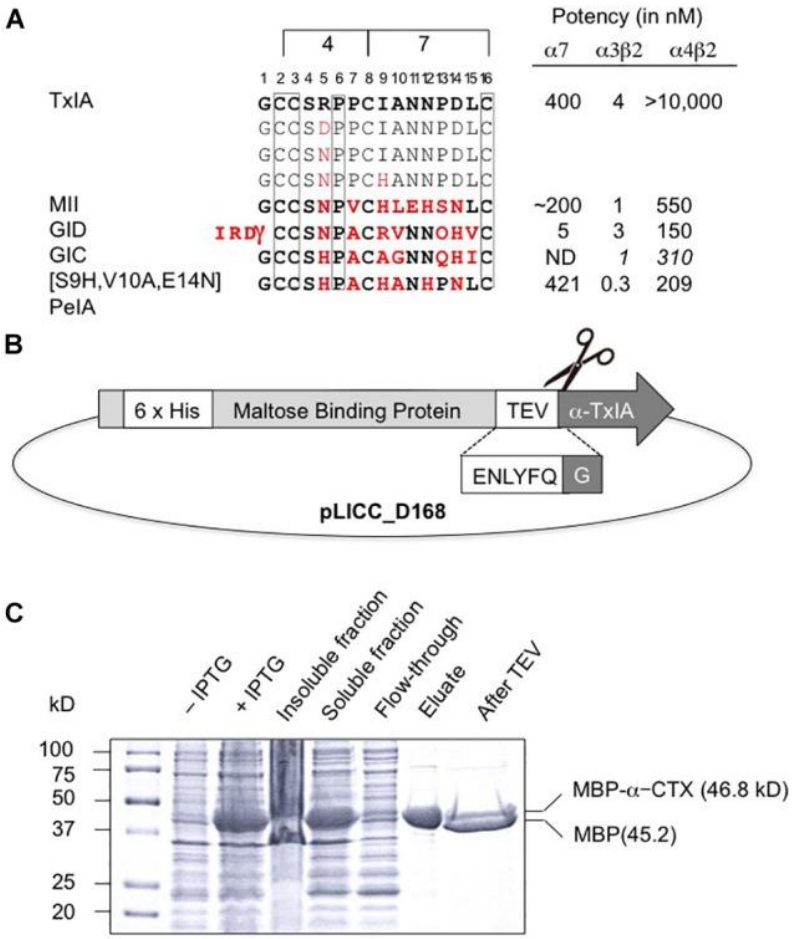


Figure 1: Recombinant production of TxIA. (A) Amino acid sequence alignment of native and mutant analogues of TxIA with α-conotoxins that show some affinity to α4β2 nAChRs, and their respective IC_{50} values (for references see (Dutertre et al., 2017)). Conserved cysteine residues are shown in boxes. The number of residues between the cysteines defines the 4/7 α-conotoxin subfamily. Residues in MII, GID, GIC, PeIA that differ from TxIA are shown in bold red. Mutated TxIA residues in position 5 and 9 are also highlighted in red. IC_{50} values refer to studies using *Xenopus laevis* oocytes. (B) Schematic representation showing the design of the pLic-MBP plasmid used for periplasmic expression of disulfide-rich peptides in *E. coli*. The construct encodes a His_6 tag for affinity purification, followed by a MBP fusion tag for solubility and a TEV protease recognition site between the MBP and the TxIA sequences. (C) Coomassie Blue stained SDS-PAGE gels showing samples of the fusion construct obtained during different steps of recombinant peptide production and purification. The difference in molecular mass between the concentrated eluate before (second last lane) and after (last lane) TEV protease cleavage indicates cleavage of TxIA from the His_6-MBP.

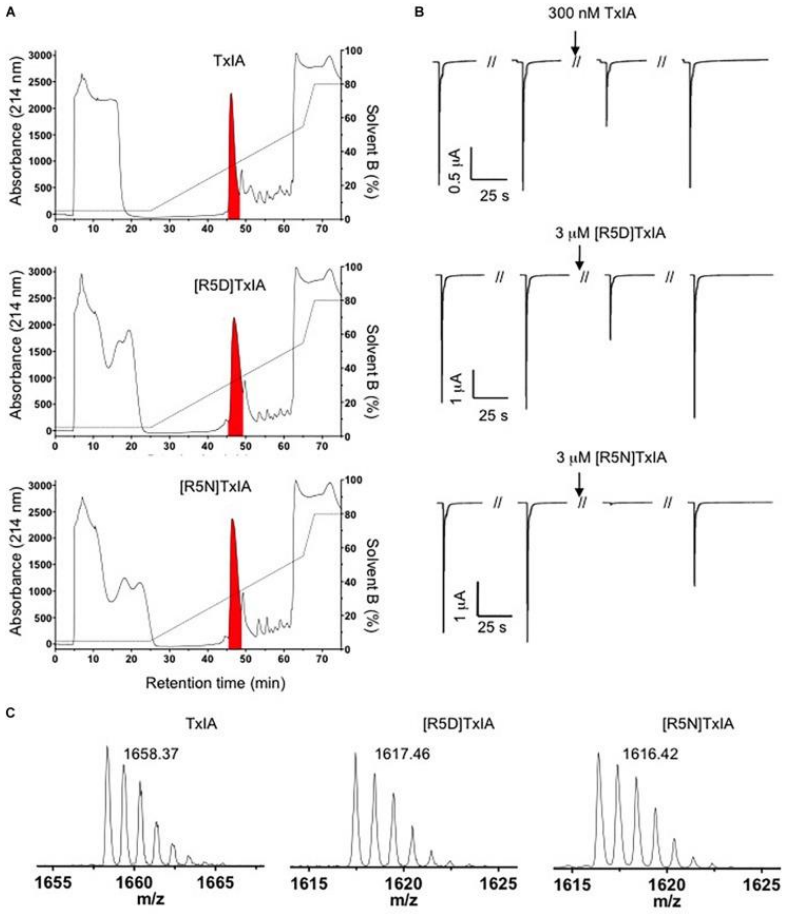


Figure 2: Purification of TxIA and mutant (R5D, R5N) analogues. (A) RP-HPLC chromatograms showing first step RP-HPLC fractionation with the peaks highlighted in red containing the correct molecular mass as determined by MALDI MS. (B) Antagonistic activity of the purified toxin peak (i.e. the red peak from fig. 2A) at $\alpha 3\beta 2$ nAChRs. Oocytes were clamped at -70 mV and 100 mM ACh were used to activate $\alpha 3\beta 2$ nAChRs. (**C**) MALDI MS traces after second step HILIC fractionation of the purified TxIA and analogues. (red coloured fractions in Fig. S2) indicate the monoisotopic molecular masses of the purified toxins. All reported masses are for $[M + H]^+$ ions.

3.2 Electrophysiological analysis of recombinant TxIA analogues reveals strongly reduced potency

We compared the potencies of the repurified expressed peptides TxIA, [R5D]TxIA, and [R5N]TxIA with synthetic TxIA at oocyte-expressed $\alpha 3\beta 2$, $\alpha 4\beta 2$, and $\alpha 7$ nAChRs. Unlike the *E. coli*-expressed α-conotoxins, but like native TxIA, the recombinant TxIA was C-terminally amidated. Not only were none of the analogues active at the $\alpha 4\beta 2$ nAChR but all recombinant peptides showed a marked drop in potency compared to synthetic TxIA (Fig. 3A). Preparation of full concentration-response curves revealed that in comparison to synthetic globular TxIA, which yielded IC_{50} values of ~5 nM and 2 μM at $\alpha 3\beta 2$ and $\alpha 7$ nAChRs, respectively, the recombinant TxIA was about 100- and 50-fold less active at

α3β2 and α7 nAChRs (Fig. 3B, Table 1). Interestingly, the potency decrease at the α7 receptor was almost completely reversed when the arginine residue in position 5 was substituted by an aspartic acid residue, but not when an asparagine residue was introduced in this position. Consequently, whatever caused the decrease in activity of the recombinant TxIA, lack of amidation, proline isomerization, and/or misfolding of the peptide, could be compensated by changes in the primary structure.

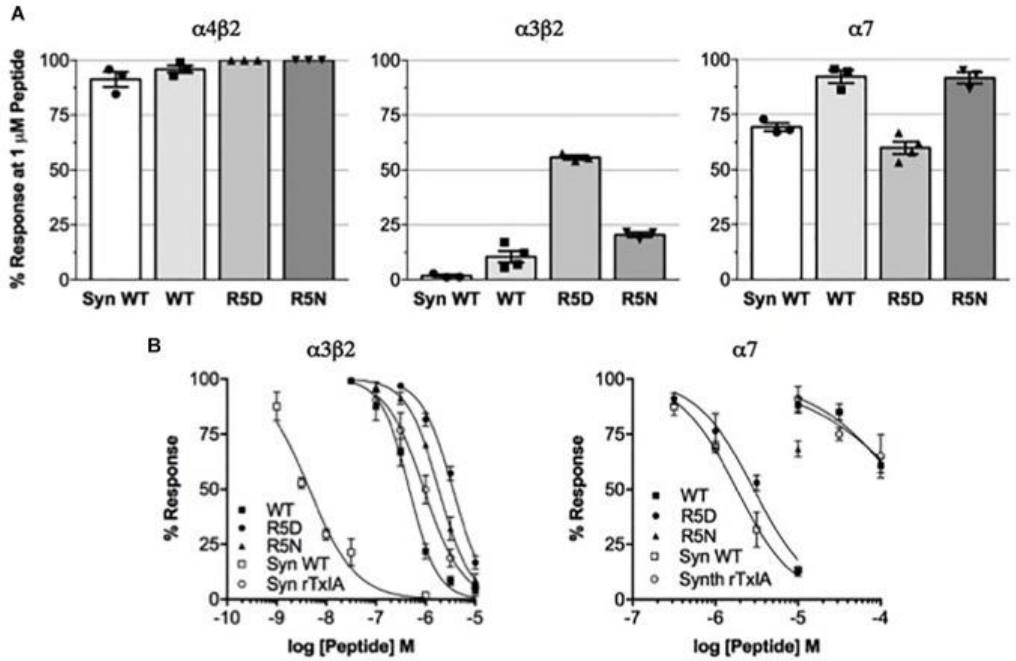


Figure 3: Potency comparison of synthesized amidated and expressed non-amidated TxIA on nAChR subtypes. (A) Reduction of control responses activated by 100 μM ACh or nicotine (in case of α7 nAChR) by 1 μM synthetic (syn) or recombinant TxIA (rec=wildtype; R5D and R5N amino acid substitutions) at the indicated nAChR subtypes expressed in *Xenopus* oocytes and measured using TEVC electrophysiology. (B) Concentration-response curves for synthetic and recombinant TxIA at α3β2 (left) and α7 (right) nAChRs. Error bars represent the S.E.M. from at least three recordings (except for high (>10 μM) concentrations of synthetic peptides, where not enough material was available). Oocytes were clamped at –70 mV. 100 μM ACh or 100 μM nicotine were used to activate α3β2 and α7 receptors, respectively.

Table 1: IC_{50} values and Hill coefficients (n_H) for synthetic TxIA and recombinant TxIA, [R5D]TxIA, and [R5N]TxIA at rat α3β2 and α7 nAChRs. Numbers in brackets are 95% confidence intervals. ND means not determined. Italic numbers are estimates.

	α3β2		α7	
	IC_{50} (nM)	n_H	IC_{50} (nM)	n_H
Synthetic TxIA	4.6 (3.6–5.8)	-0.93 (-1.15– -0.71)	1812 (1555–2111)	-1,23 -1,52– -0,99
Synthetic rTxIA	928 (810–1064)	-1.14 (-1.3– -0.98)	*≈ 200.000*	ND
Recombinant TxIA	470 (432–512)	-1.54 (-1.73– -1.35)	*≈ 200.000*	ND
Recombinant [R5D]TxIA	3543 (3166–3965)	-1.38 (-1.58– -1.17)	2915 (2406–3532)	-1.24 (-1.53– -0.96)
Recombinant [R5N]TxIA	1834 (1704–1974)	-1.36 (-1.49– -1.24)	ND	ND

3.3 Structural analysis of recombinant TxIA analogues confirms the preferential formation of ribbon isomers

Because all of the recombinant peptides showed at least some potency at α3β2 nAChRs and peptides in the other fractions turned out to be inactive or less active, we assumed that the dominant peaks corresponded to the native (i.e., globular) conformation of the peptides. However, because of the strong potency decrease that is clearly higher than the 2-fold or 5-fold potency reduction of non-amidated AnIB analogues at α3β2 and α7 nAChRs, respectively [443] we considered the possibility of non-native disulfide linkage in the E. coli-expressed TxIA analogues. To test this hypothesis, NMR structural analysis was performed. The three expressed TxIA analogues yielded well dispersed ^{1}H NMR resonances, implying that they adopt ordered structures in solution. These peptides were further analyzed using two-dimensional NMR so that the data could be compared to published data for Pn1.2, which has a related sequence and where NMR data for each of the disulfide isomers is available [444]. To assign the spectra, H_α-NH_{i+1} connectivities obtained from NOESY spectra were used in the sequential assignment of individual spin systems determined from the TOCSY spectra. For all of the TxIA analogues, sequential H_α-NH_{i+1} connectivities were observed for the entire peptide chain, except at the Pro residues that lack backbone amide protons. Analysis of the NOE data and the Cδ chemical shifts of Pro7 indicated that this proline is in the cis confirmation [445]. Analysis of secondary shift data can

supply information on secondary structural elements of peptides [446], and is useful for comparing structural frameworks in disulfide-rich peptides. Fig. 4B shows the Hα secondary shifts for TxIA and compares these values with previously published secondary shifts for the three possible disulfide isomers of Pn1.2. The Hα secondary shifts for the TxIA most closely resemble those of the ribbon isomer of Pn1.2, especially in the N-terminal region, which suggests that the expressed TxIA is in the ribbon form. Furthermore, a comparison of the secondary shifts for each expressed TxIA analogue (Fig 4A) and synthetic ribbon TxIA (see below and Fig. S3C) reveals that they are almost identical and all four peptides form the same disulfide isomer. Most of the secondary shifts for the TxIA analogues vary within the ±0.1 ppm range that is consistent with a random coil structure, however residues 9–11 have secondary shift values more negative than –0.1 ppm, suggesting this region has some helical character. Our structural and functional characterization of the TxIA analogues is in agreement with data from Wu et al. 2014 [447], where an 80-fold potency reduction (IC_{50} of 5.4 nM to 430 nM) of synthesized ribbon TxIA was observed at α3β2 nAChRs, which corresponds well to the 100-fold reduction found in our study. Importantly, this analysis further confirmed the purity of the expressed toxins as similar retention times for globular and ribbon TxIA have been described [447].

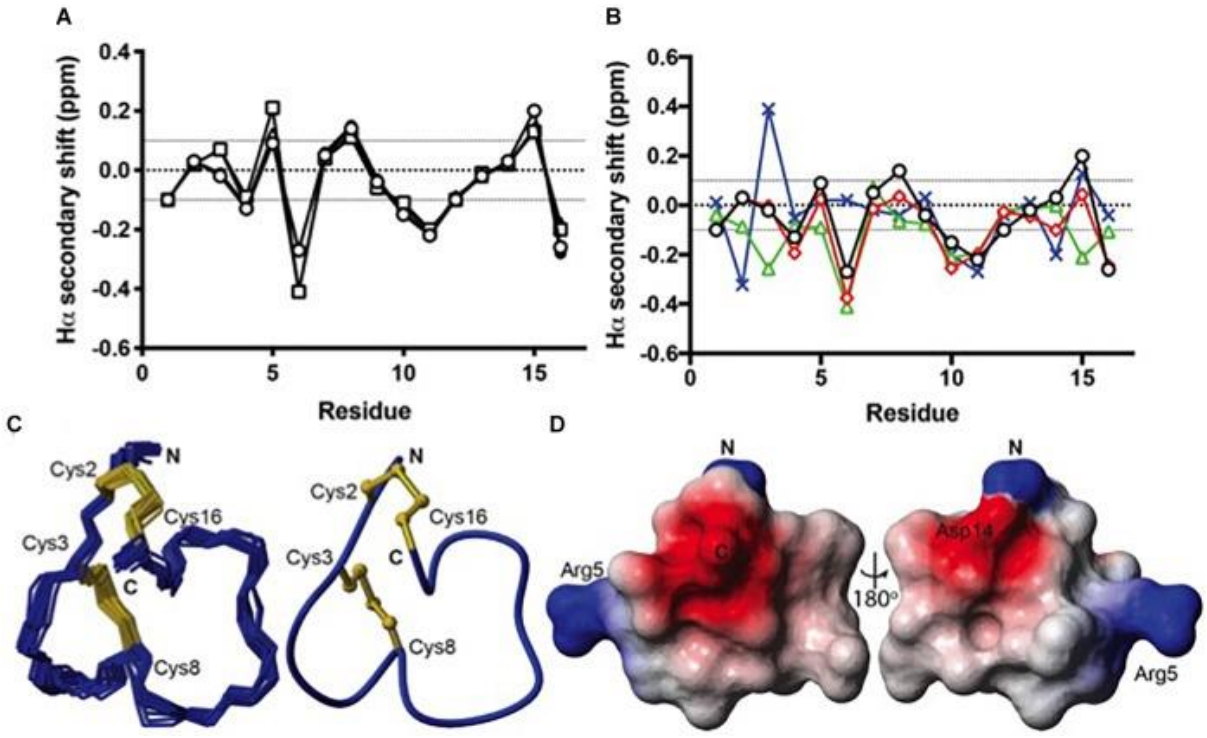


Figure 4: Structural analysis of the ribbon disulfide isomers of TxIA and analogues. (A) Hα secondary shift values for recombinant TxIA (open circles), [R5D]TxIA (squares), [R5N]TxIA (triangles), and synthetic ribbon TxIA (closed circles) at 7°C. Please note that due to the identity of the peptides, the secondary shifts overlap and most data points (filled circles) are hardly visible (B) A comparison of the Hα secondary shifts for ribbon TxIA (black) with the Hα

secondary shifts of ribbon (red), globular (green) and beads (blue) isomers of Pn1.2. The values for ribbon TxIA most closely match those of ribbon Pn1.2, especially in the N-terminal region. (C) Left panel: ensemble of the 20 lowest energy structures for ribbon TxIA. The peptide backbone is in blue and the disulfides are in gold. Right panel: The mean structure of ribbon TxIA calculated from the ensemble. The structure is shown in ribbon format with disulfide bonds in ball and stick. The N- and C-termini are indicated and the cysteines are labelled. (D) Electrostatic surface of ribbon TxIA highlighting the negatively charged regions (red) corresponding to the C-terminal acid and Asp14, and the positively charged regions (blue) corresponding to the N-terminus and the side chain of Arg5.

3.4 Solution structure and potential bindind mode of the ribbon isomer of TxIA

α-Conotoxins are generally considered to be weakly active or inactive if they are folded in non-native conformations such as "ribbon" or "bead" disulfide isomers. However, in case of the 4/6 conotoxin AuIB, the ribbon isomer turned out to be more potent at $\alpha3\beta4$ nAChRs than the globular form and seemed to bind to the orthosteric binding site [95]. Likewise, the ribbon form of the 4/3 α-conotoxin ImII was found to compete with binding of its natively folded isomer at $\alpha7$ nAChRs and also bound to an additional binding site at *Torpedo* nAChRs [448]. For AuIB, a binding mode has recently been defined in detail [449]. However, no ribbon 4/7 α-conotoxin structure has been determined to date [401]. Inspired by the fact that ribbon isomers of some conotoxins can competitively block ACh binding and that TxIA potency at the $\alpha7$ nAChR could be improved by the substitution R5D, we set out to determine the structure of the ribbon conformer of TxIA and its binding mode at $\alpha7$ nAChR. To this end, a non-amidated ribbon isomer of TxIA was synthesized (Fig. S3) by a two-step oxidation method and its NMR structure calculated using CYANA 3.0 [358] and refined in CNS [426] using TALOS-N [389] derived dihedral angle restraints, and hydrogen bond restraints derived from D_2O exchange experiments.

The structural statistics for the ensemble of the 20 lowest energy structures for each isomer are shown in Table 2, and the structural ensemble of the 20 representative structures are shown in Fig. 4C (left). The mean structure of the ribbon isomer of TxIA (Fig. 4C, right), although well defined (backbone RMSD = 0.58 $\pm$ 0.3 Å), possesses no defined secondary structure elements. However, several of the structures in the ensemble have a 3_{10}-helical segment spanning residues Cys8 to Asn11, which is generally consistent with the Hα secondary shift data. Furthermore, loop 1 of ribbon TxIA (Cys2– Cys8) is well defined (backbone RMSD = 0.21 ± 0.07 Å) and resembles loop 1 of globular TxIA, but loop 2 (Cys8–Cys16) is less well defined (backbone RMSD = 0.48 ± 0.29 Å) and adopts a conformation that is distinct from loop 2 of globular TxIA. Analysis of the electrostatic surface features in MOLMOL reveals negatively charged regions on opposite faces, corresponding to the C-terminal acid and the sidechain

of Asp14, and a positively charged region due to the sidechain of Arg5 and N-terminus (Fig. 4D). This structure was subsequently used to probe the binding mode in a MD-refined model and a model based on the recently defined ToxDoc application.

Table 2: Structural statistics for the final ensemble of 20 structures for the ribbon isomer of TxIA with the highest overall MolProbity score.

Energies (kcal/mol)	
Overall	–475.3 ± 3.1
Bonds	6.51 ± 0.67
Angles	20.11 ± 2.48
Improper	9.21 ± 2.26
Van der Waals	–36.82 ± 3.04
NOE	0.02 ± 0.01
cDih	0.98 ± 0.67
Dihedral	65.78 ± 0.89
Electrostatic	–541 ± 6.3
MolProbity Statistics	
Clashes (>0.4 Å / 1000 atoms)	14.3 ± 4.66
Poor rotamers	0.05 ± 0.22
Ramachandran Outliers (%)	0.0 ± 0.0
Ramachandran Favoured (%)	85.5 ± 6.19
MolProbity score	2.29 ± 0.31
Residues with bad bonds	0.00 ± 0.00
Residues with bad angles	0.00 ± 0.00
Atomic RMSD (Å)	
Mean global backbone	0.58 ± 0.26 (residues 1–16)
Mean global heavy	1.19 ± 0.37 (residues 1–16)
Distance Restraints	
Intraresidue (|i-j| = 0)	35
Sequential (|i-j| = 1)	44
Medium range (1 < |i-j| < 5)	16
Long range (|i-j| > 5)	8
Hydrogen bonds	3
Total	106
Dihedral angle restraints	
ϕ	11
ψ	8
χ^1	9
Total	28
Violations from experimental restraints	
Total NOE violations exceeding 0.3 Å	0
Total Dihedral violations exceeding 3.0^0	0

3.5 Molecular modelling

3.5.1 Binding modes in the ToxDock- and MD-refined models

The binding modes of ribbon TxIA at the α7 nAChR generated by MD simulations and by ToxDock are globally similar in terms of the orientation of the toxin in the binding site, with some molecular interactions shared between the two models but also a range of interactions that are different (Fig. 5A, B). Pro6 in Loop 1 of wild-type or variant ribbon TxIA occupies a similar position in the models as P6 of the globular TxIA in the TxIA/AChBP experimental structure. P6 is an important determinant of the interaction of α-conotoxins as it is embedded in the aromatic box, which is at the bottom of the orthosteric binding site and is recognized by acetylcholine. Another important conserved feature of the interaction of globular α-conotoxin with AChBP is the stacking between the vicinal disulfide bridge of the C-loop with the first disulfide bond of the α-conotoxin. The MD-refined model displays this interaction but not the ToxDock-refined model in which the side chain of D14 stacks between the two disulfide bonds. This is interesting because the conformation of toxin in the two models is globally preserved (the backbone RMSD of the toxins in the two models is only 1.5 Å). The binding mode in the ToxDock model is slightly tilted compared to the MD model, and this leads to a change in reorientation of the N-terminus (from G1 to S4) and of Loop 2 in general. The conformation of the wild-type and variants in the context of the binding site was highly stable, with a backbone RMSD < 1.5 Å and on average 1 Å from the initial homology model. The binding modes of the toxins were also similarly stable, with Loop 1 being highly rigid and Loop 2 undergoing larger, albeit small, conformational fluctuation. The binding mode of the ToxDock model was not sampled during the MD refinement, and they should therefore be considered as two hypothetical binding modes. The location and interaction of position 5 in the models of the wild-type TxIA are similar, and the interactions of this position in TxIA and variants will now be described using the MD-refined models (Fig. 5C–E). Figure S4 in Supplementary Material describes the interaction as suggested in the ToxDock-refined models.

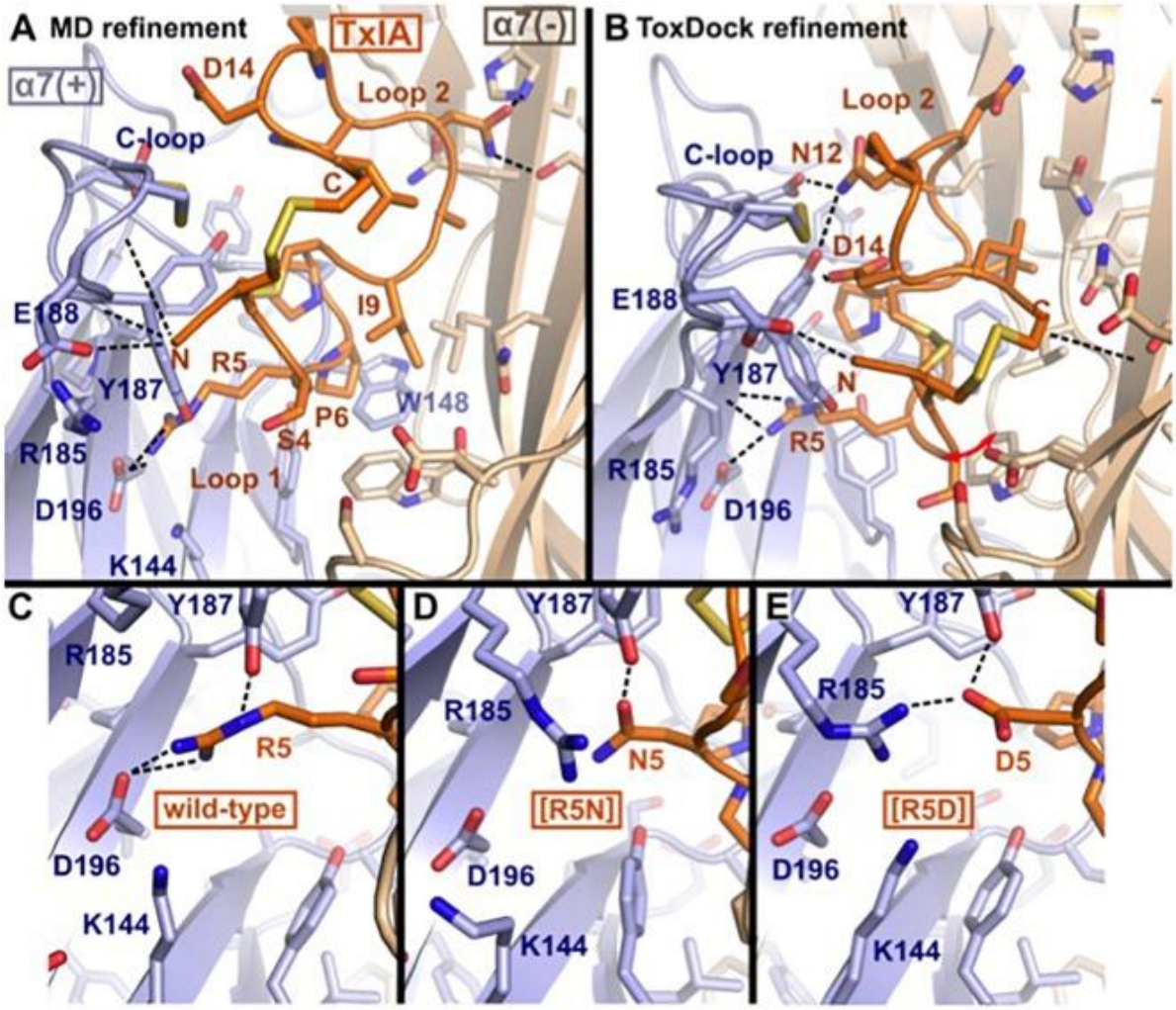


Figure 5: Molecular models of the binding modes of ribbon TxIA at the α7 nAChR: (A) MD-refined molecular model, (B) ToxDock refined molecular model, and interaction of position 5 in the MD-refined model of (C) ribbon TxIA, (D) ribbon [R5N]TxIA or (E) ribbon [R5D]TxIA. The last frame of the 50 ns MD simulation is shown in panel A. The model with the lowest re-weighted Rosetta score is shown in panel B. The backbone of the receptor and peptides are shown using cartoon representation (blue and wheat color for the principal (+) and complementary (–) subunits, respectively), the side chains of the toxins are shown as sticks (orange) and the side chain of receptor residues located at less than 5 Å from the toxin are shown in stick representation. Hydrogen bonds are indicated by dashed black lines.

3.5.2 Molecular interactions of position 5

TxIA R5 is located in Loop 1 and establishes a salt-bridge with α7 D196 in the wild-type model, similar to the interaction of globular TxIA R5 with AChBP D197 in the TxIA/AChBP crystal structure. The model also suggests the proximity of two principal subunit basic residues, α7 R185 and K144, which could create an unfavorable electrostatic environment for Arg 5. In contrast, the model of ribbon [R5D]TxIA

shows that D5 establishes a salt bridge with α7 R185. In the ToxDock model of [R5D]TxIA, D5 also forms a salt-bridge but with α7 K144. Position 5 therefore seems to be better suited for a negatively charged residue than a basic residue. Asn and Asp are isosteric, but Asn is neutral whereas Asp is negatively charged. In contrast to [R5D]TxIA, the mutant displaying an Asn at position 5, *i.e.* [R5N]TxIA, has similar activity to the parent peptide, strongly suggesting that the charge of the side chain at position 5 drives the increase in activity at the α7 nAChR (Fig. 5D, E). This observation also suggests that the charge interactions of TxIA R5 with the complementary subunit balanced out and do not contribute to affinity. Interestingly, this analysis suggests that an R5D substitution of the globular TxIA would increase affinity for the α7 nAChR.

3.5.3 Mutational energy predictions

The interaction energies predicted by ToxDock were similar for the three complexes, with the interaction energy of wild-type ribbon TxIA in Rosetta units (RU) being –592 for TxIA, –594 for [R5D]TxIA and –595 for [R5N] TxIA. The mutational energies predicted by MMPB/SA suggest that the [R5N] mutant should have increased affinity by –4 kcal/mol compared to the parent peptide, while the [R5D] substitution was predicted to decrease affinity by 13 kcal/mol. None of the energy predictions therefore correlated with the experimental IC_{50} values, possibly because the main determinants of the affinity change involve evaluating the interaction of salt bridges in a partially solvated environment in the vicinity of several oppositely charged residues, which is challenging. Current force fields still have difficulty in reproducing the strength of solvated salt bridges [450]. In addition, the estimation of binding energy is highly dependent on the accuracy of the three-dimensional structures of the complex. The incomplete consensus between the ToxDock- and MD-refined models suggest that these binding modes of ribbon α-conotoxins are not accurate enough for predicting the impact of mutations using score or energy computations. The determination of an experimental structure of a ribbon conotoxin bound to either AChBP or a nAChR should dramatically increase our ability to predict the interaction of ribbon conotoxins and the impact of mutations.

4. Discussion

Here we describe for the first time the production of a recombinant α-conotoxin (4/7 α-conotoxin TxIA and three analogues) by expression in the periplasm of *E. coli*. We characterized the structure of three of these recombinant toxins and determined that, contrary to expectation, they adopt a ribbon conformation, which is not the dominant disulfide connectivity obtained during random oxidation of synthetically produced α-conotoxins. Ribbon TxIA experienced a significant decrease in activity at the $\alpha3\beta2$ and $\alpha7$ nAChRs, and we discovered that the mutant [R5D]TxIA could rescue the activity at the

α7 nAChR at a similar level to that of globular TxIA, possibly because the location of position 5 in the α7 binding site is more electropositive.

4.1 Expression of α-conotoxins in bacteria

So far, three α-conotoxins, the 4/7 α-conotoxins Vc1.1, LvIA, and TxIB, have been recombinantly expressed [245,451,452]. Vc1.1 was expressed in the non-pathogenic *Salmonella enterica* serovar Typhimurium strain LT2. It was fused via a TEV protease cleavage site to a flagellar secretion substrate FlgM, enabling export of recombinant nonflagellar peptides through the flagellum and into the surrounding medium [451]. No functional or structural data of *Salmonella*-expressed Vc1.1 were provided in that study.

LvIA was expressed in tandem repeats of various lengths in *E. coli* and purified from inclusion bodies which are generated during conventional recombinant protein expression in *E. coli* [245]. To allow chemical cleavage with cyanogen bromide, the individual peptide sequences were linked by methionine residues, which resulted in the N-terminal addition of a methionine residue to the cleaved peptides. Using this procedure and folding by air oxidation, a 18-fold decrease in activity from 9 to 160 nM was observed for the recombinant peptide compared to synthetic LvIA [228,245]. This decrease in activity was suggested to be due to the additional N-terminal methionine and/or the missing C-terminal amidation. The latter would be in agreement with the 2- and 5-fold potency reduction of non-amidated AnIB analogues at α3β2 and α7 subtypes respectively [443]. However, the disulfide connectivity of the *E. coli*-expressed LvIA was not determined and our study of recombinant TxIA suggests that alternative disulfide connectivity could also explain the activity change.

In a more recent study from the same group [452], monomeric TxIB was expressed in *E. coli* as a ketosteroid isomerase (KSI)–TxIB(M)–His_6–fusion protein. KSI was used to help stabilize the peptide in inclusion bodies and the insoluble KSI could be easily separated upon cleavage from the peptide. A C-terminal methionine residue (M) had to be added to allow subsequent release of the peptide from KSI and the His_6-tag by cleavage with cyanogen bromide. Interestingly, the recombinant TxIB retained selectivity for α6β2 receptors and showed only a moderate 2-fold decrease in potency. Based on these properties, a globular fold was inferred because ribbon and bead isomers of TxIB were reported to be inactive.

Encouraged by reports on bacterial expression of other toxins and in an attempt to bypass the problem of aggregation in inclusion bodies, we exploited a system that allows expression of venom peptides in the form of periplasmic MBP fusion proteins. The idea of using periplasmic expression is to hijack the disulfide-bonding machinery in *E. coli* for producing natively folded heterologous peptides [415]. Here, intramolecular disulfide bonds should form as the polypeptide chain exits the reductive environment

of the bacterial cytoplasm and enters the oxidizing periplasmic environment. With a success rate of 75% based on the expression of 31 venom peptides (ranging from 17 to 76 residues in length with 2–6 disulfide bonds) from spiders, scorpions, sea anemone, and cone snails, this method mostly produces correctly folded peptides [415]. However, the 25% of peptides that failed to express using this method indicate that the folding machinery in *E. coli* is not always a good surrogate for the mechanisms found in the venomous organisms from which the respective peptides were sourced. In the case of venom peptides from cone snails, the only other peptide that was tried (and failed) in this expression system was MVIA [415]. In regard to TxIA, we could also not produce any native (globular) isomer, as all of the recombinant TxIA formed the ribbon isomer. In summary, for the only two *Conus* venom peptides that have been investigated so far, periplasmic expression in *E. coli* failed to produce the native disulfide-bond isomer. However, additional data with other *Conus* peptides is required to make a conclusion about the general suitability of this system for producing natively-folded cone snail venom peptides.

4.2 Does the primary structure affect the folding?

One surprising finding was that all expressed variants eluted in one dominant peak. In contrast to the present results and in agreement with data on other α-conotoxins, chemically synthesized and randomly oxidized α-conotoxin TxIA folds in all three possible conformations, globular, ribbon, and beads [447]. This suggests the presence of factors that aid the formation of ribbon isomers in the periplasm of E. coli in contrast to the venom gland of the cone snail, in which the globular isomer is formed. In that case, misfolding could be a general problem for *E. coli* expression of α-conotoxins. Alternatively, TxIA could have features in its amino acid sequence that favor formation of the ribbon fold, such as non-covalent interactions that stabilize the respective cysteine positioning. For example, a study by Kang et al. [453] found that C-terminal amidation can cause a preferential (but not absolute) formation of the globular form of the 4/3-α-conotoxin ImI, supposedly by different hydrogen bonding interactions of the C-terminus. In this context, the potential effect of a Pro-Pro motif present in the first loop of TxIA instead of the more common Pro-Ala motif might also be important due to the structural constraints induced by proline residues. However, TxIB, which appeared to fold preferentially in the globular fold, [452] has a first loop with amino acid sequence identical to [R5D]TxIA, which favored the ribbon conformation in our study. Thus, factors in loop 2 of 4/7 α-conotoxins would be more likely to affect the disulfide formation.

The inability to create posttranslational modifications of α-conotoxins in *E. coli* (besides disulfide bonds) is potentially a problem as some of these modifications are important for activity in some conotoxins. Nevertheless, certain modifications such as tyrosine sulfation, glutamate carboxylation or C-terminal amidation had relatively minor effects in previous studies and might be dispensable for

some toxins [276,443]. In vitro modifications such as C-terminal amidation, might also be feasible on purified recombinantly expressed toxins [454,455].

4.3 Structural comparison of 4/7 α-conotoxin ribbon and globular isomers

The solution structures of four ribbon α-conotoxins have been determined and are publicly available; these include the 3/5 α-conotoxin GI (PDB 1XGB) [97], the 4/4 α-conotoxin BuIA (PDB 2NS3) [456], the 4/6 α-conotoxin AuIB (PDB 1MXP) [261], and the 4/6 α-conotoxin Pu1.2 [457]. The structures of ribbon conotoxins are typically less well-defined than that of the corresponding globular isomers, with Loop 1 being relatively rigid whereas Loop 2 adopts multiple conformations, presumably indicating a greater flexibility (Fig. 6). The four ribbon conotoxins with four residues in Loop 1, i.e. BuIA, rAuIB, Pu1.2, and TxIA, have a very similar conformation of Loop 1. Their Loop 2 conformations are different, reflecting the larger structural variability of Loop 2 in globular conotoxins [82].

The ribbon isomers of α-conotoxins have naturally occurring structural analogues, the cysteine Framework X, with which they share two disulfide bridges in a ribbon isomer connectivity, a similar size, and four cysteine residues in similar pattern along the amino acid sequence (i.e. CC-C-C) [458]. Loop 2 of cysteine framework X conotoxins contains only two amino acids, with the second residue of this loop being a hydroxyproline. The Loop 1 of framework X toxins has four residues, similar to the cysteine Framework I α-conotoxins, but it does not contain the Pro residue that is conserved in the α-conotoxins, and which occupies the orthosteric binding site of nAChRs as suggested by the crystal structures of complexes between α-conotoxins and AChBP [252]. Framework X conotoxins are not active at nAChRs but rather target the neuronal noradrenaline transporter, and hence they are classified in the χ pharmacological family [163]. Interestingly, the ribbon isomer of an α-conotoxin was even identified in a cone snail venom, indicating that for some α-conotoxins the ribbon isomer is also native [459].

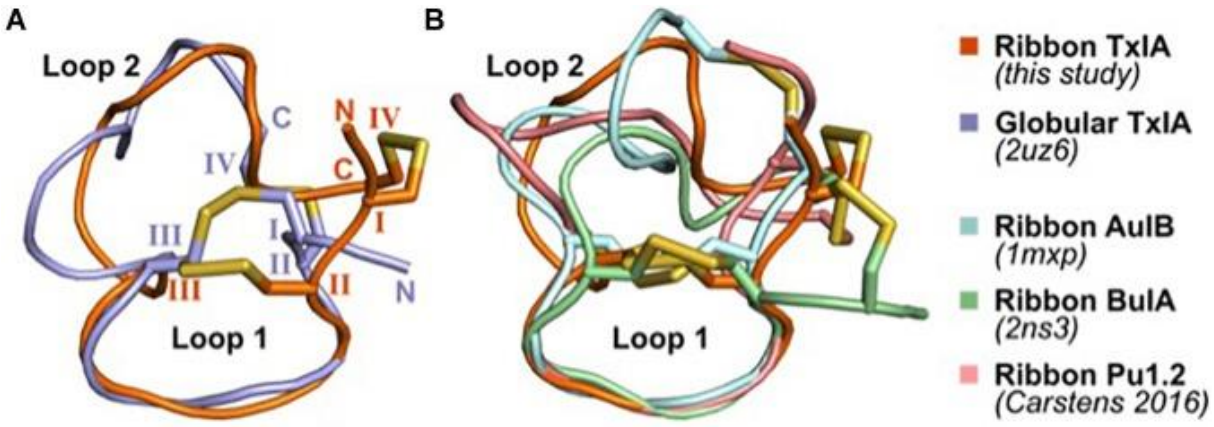


Figure 6: Comparison of the NMR solution structure of ribbon TxIA with (A) globular TxIA and (B) other ribbon contoxin structures. The first model from the NMR ensemble was used in this figure, with the caveat that the 20 models of ribbon AuIB display a high conformational variability of Loop 2. The peptides are shown in cartoon representations and the disulfide bonds are indicated as yellow sticks. The N- and C-termini are indicated in panel A, and the hemi-cystine residues are numbered using Roman numerals. The PDB identifier of the experimentally determined three-dimensional structures is indicated where available.

4.4 Comparison of proposed binding mode of ribbon and globular α-conotoxins

A binding mode was recently proposed for the ribbon α-conotoxin AuIB at the α3β4 nAChR [449], and this binding mode is similar to the MD-refined binding mode of ribbon TxIA, as illustrated in Fig. 7. It was demonstrated that the sole binding site of ribbon AuIB is at the α3(+)α3(–) interface based on its subunit stoichiometry-dependent activity at $\alpha3\beta4$ nAChR [95]. Ribbon AuIB was recently suggested to adopt a similar binding mode to that of the globular α-conotoxin on the basis of an Ala scan which revealed that most of the residues that are important for the activity of the globular isomer are also important for the ribbon isomer [449]. Crystal structures of α-conotoxin/AChBP complexes showed that the first loops of the globular α-conotoxins overlap well, whereas the second loops adopt different conformations and interactions [166]. Like the structures of globular α-conotoxins belonging to different loop-length subgroups bound to the AChBPs, the conformations of the Loop 1 of ribbon 4/7 TxIA and 4/6 AuIB overlay well in the binding sites but their second loop adopts different conformations. We have previously been able to explain the impact of more than 30 mutants of the complex between globular α-conotoxin ImI and the rat α7 nAChR, suggesting that our method can generate a reasonably accurate model of the receptor [460,461]. The differences of binding modes observed after the two refinement methods seem therefore to originate from modelling the flexibility and interaction of the peptide. Interestingly, ribbon [P7A]AuIB has a structure similar to globular AuIB, but despite this change of conformation this mutation is innocuous [449], suggesting that ribbon AuIB

would adopt a helical conformation by conformational selection. Similarly, ribbon TxIA also becomes more helical in both the MD- and ToxDock-refined models, suggesting that a helical conformation is optimal for interaction. The orientations of the loop 2 of ribbon TxIA and AuIB are different in the binding site, although position 9 (I9 of ribbon TxIA and Y9 of ribbon AuIB) was suggested to be buried at the interface with the complementary subunit for both peptides. Position 9 has been identified as a key determinant that modulates the activity of ribbon AuIB and it could also be involved in modulation of the activity of ribbon TxIA.

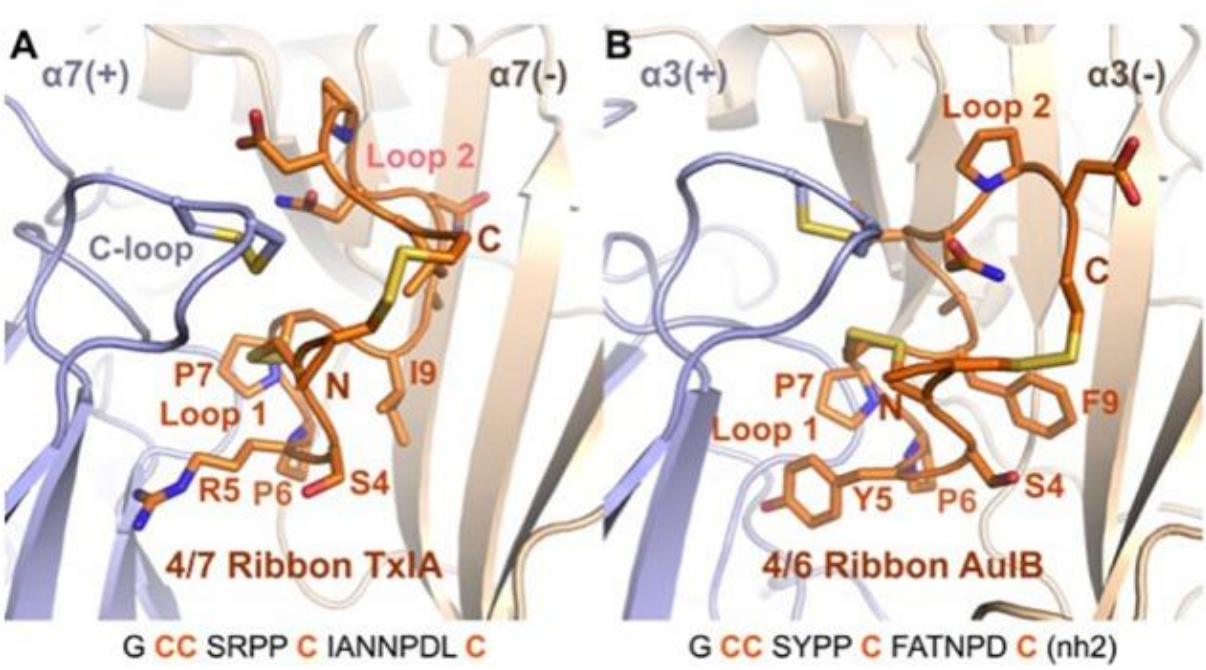


Figure 7: Comparison of the binding mode of ribbon TxIA (A) and ribbon AuIB (B) at the interface between two α7 or α3 subunits, respectively. The toxins are in orange, with their backbone in cartoon representation and their side chains in stick representation. The binding site principal (+) and complementary (–) subunits are shown in blue and wheat color, respectively. The vicinal disulfide bond of the C-loop is shown in stick representation. The two binding sites have been oriented similarly for convenient comparison. The sequence of the two toxins is shown at the bottom of the figure.

Because of their exceptional selectivity for certain ion channel or receptor subtypes, many conopeptides are used as tools in neuroscience and pharmacological research. Some of them have even been developed for clinical applications such as α-conotoxin MVIIA (ziconitide), which is an FDA-approved analgesic for the treatment of intractable pain [200,462]. In the case of α-conotoxins, the 4/7 α-conotoxins Vc1.1 and MII, the 4/3 α-conotoxin RgIA, and the 4/6 α-conotoxin AuIB, were found to have potent analgesic properties. Although there still exists some controversy regarding their

physiological target (for a recent review see [202] these conotoxins are important lead structures. New methods for large-scale production of these conotoxins and development of a deeper understanding of their structure-activity relationships will accelerate their development into optimized tools and, hopefully, novel drugs.

5. Supporting information

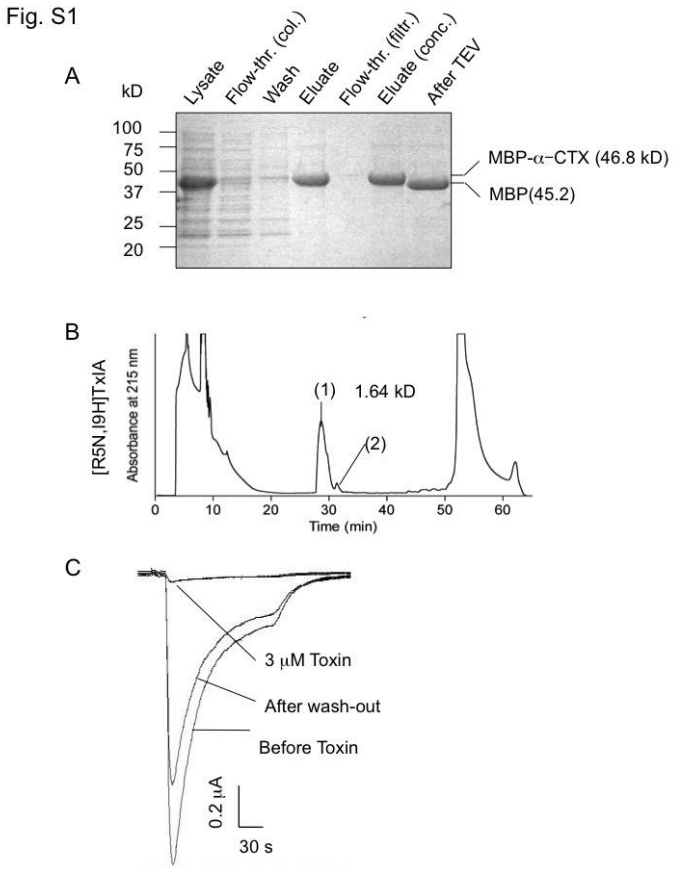


Figure S1: Expression, purification, and analysis of antagonistic potency at $\alpha 3\beta 2$ nAChR of the [R5N,I9H]TxIA double mutant.

Fig. S2

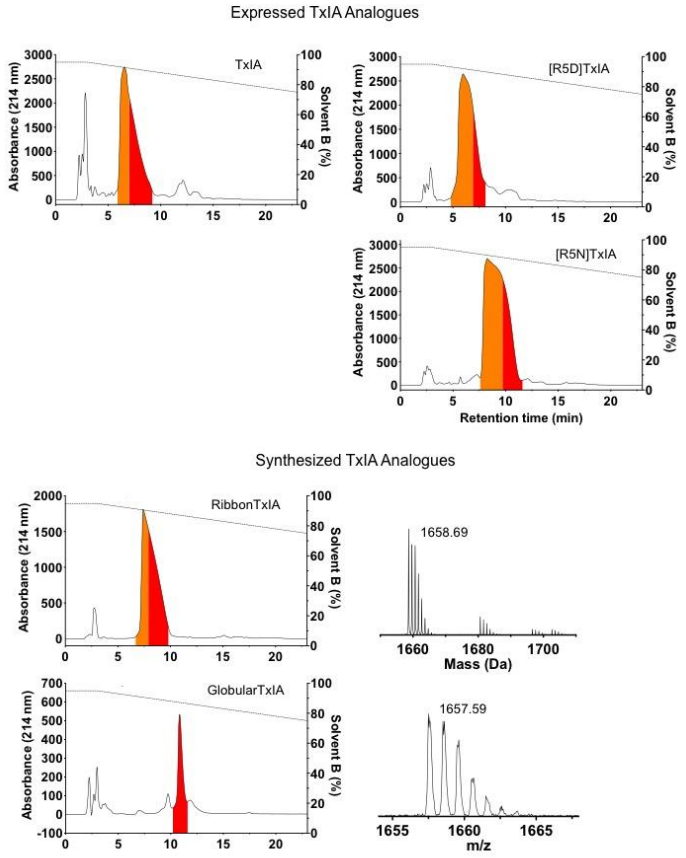


Figure S2: Second step HILIC fractionation of purified expressed TxIA analogues and synthetic globular and ribbon TxIA isomers. Fractions containing the correct molecular mass are highlighted in colour. The red coloured fractions were generally used for all further experiments, orange coloured fractions contained additional salt ions. The dashed line indicates the gradient of solvent B (90% ACN/0.043% TFA). MALDI MS traces indicate the monoisotopic molecular masses of the synthesized toxins. All reported masses are for [M + H]+ ions. Note that the synthetic globular TxIA was C-terminally amidated, explaining the difference in molecular mass to the synthetic ribbon and recombinant TxIA (Fig 2C).

Fig. S3

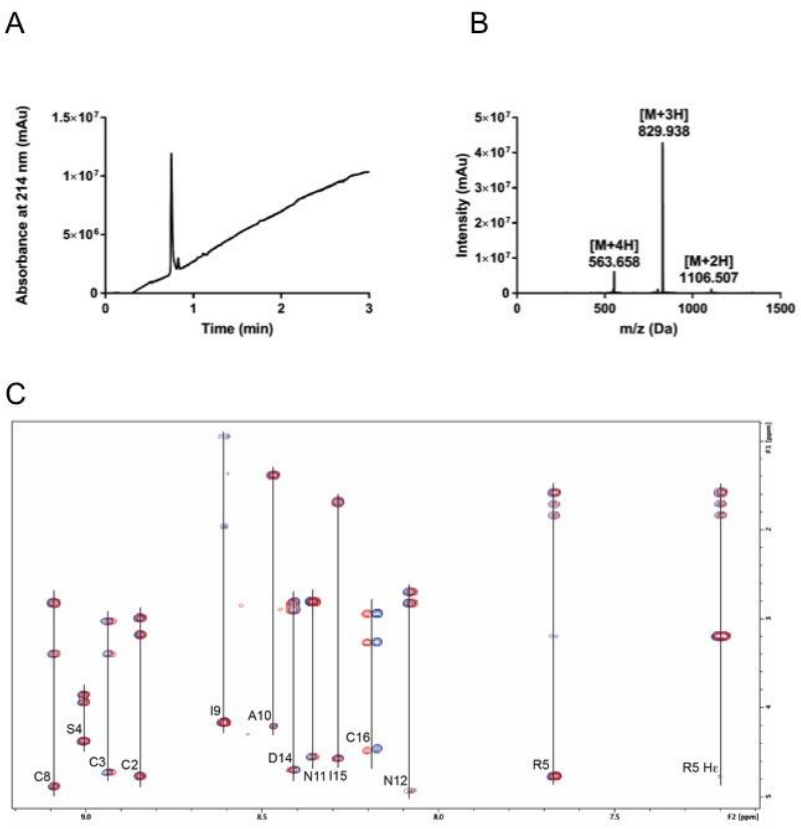


Figure S3: Characterization of chemically synthesized ribbon TxIA. (A) UV absorbance chromatogram at 214 nm of synthetic TxIA ribbon. The minor peak is probably due to conformational exchange since the merging of the two peaks is ovserved at 55°C. ACN gradient from 0 to 100% in 2.5min. (B) ESI-MS of synthethic TxIA ribbon. (C) Overlay of the amide region of the TOCSY spectra for recombinantly expressed TxIA (red) and synthetic ribbon TxIA (blue). Small differences in chemical shift for the HN protons are observed for Cys3 and Cys16 between samples presumably due to a slight variation in pH between samples.

Fig. S4

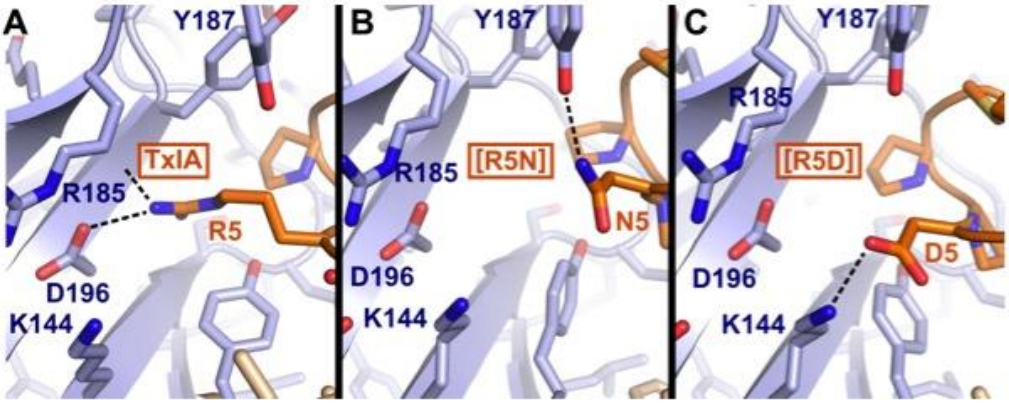


Figure S4: Comparison of the interactions between the position 5 of ribbon TxIA and variants at the α7 nAChR obtained from the ToxDock-refined molecular model.

d. Synthèse d'un peptide endogène humain ciblant le récepteur à la ghréline

Ce chapitre s'inscrit dans le cadre d'un projet collaboratif avec plusieurs équipes de l'IBMM (F9 et F13) hors synthèse de toxine, mais toujours dans le cadre de la synthèse de produits naturels comportant des ponts disulfures. La synthèse du peptide humain antimicrobien exprimé par le foie (LEAP-2) comportant 40 acides aminés et deux ponts disulfures m'a ainsi permis de tester l'applicabilité de la stratégie mise en place pour la synthèse de peptides longs.

Cette étude est présentée ci-après sous la forme d'une publication parue [463] dans laquelle j'ai effectué la synthèse du peptide total et la synthèse de la partie structurée.

N-Terminal Liver-Expressed Antimicrobial Peptide 2 (LEAP2) Region Exhibits Inverse Agonist Activity toward the Ghrelin Receptor

Céline M'Kadmi,† Agustina Cabral,‡ Franco Barrile,‡ Julien Giribaldi,† Sonia Cantel,† Marjorie Damian,† Sophie Mary,† Séverine Denoyelle,† Sébastien Dutertre,† Sylvie Péraldi-Roux,† Jérémie Neasta,† Catherine Oiry,† Jean-Louis Banères,† Jacky Marie,† Mario Perello,‡ and Jean-Alain Fehrentz*,†

†Faculté de Pharmacie, Institut des Biomolécules Max Mousseron (IBMM), UMR 5247, CNRS, Université de Montpellier, Ecole Nationale Supérieure de Chimie de Montpellier, 15 avenue Charles Flahaut, BP 14491, 34093 Montpellier cedex 5, France.
‡Laboratory of Neurophysiology of the Multidisciplinary Institute of Cell, La Plata, 1900 Buenos Aires, Argentina

Abstract

The ghrelin receptor or growth hormone secretagogue receptor (GHSR) is a G-protein-coupled receptor that controls growth hormone and insulin secretion, food intake, and reward-seeking behaviors. Liver-expressed antimicrobial peptide 2 (LEAP2) was recently described as an endogenous antagonist of GHSR. Here, we present a study aimed at delineating the structural determinants required for LEAP2 activity toward GHSR. We demonstrate that the entire sequence of LEAP2 is not necessary for its actions. Indeed, the N-terminal part alone confers receptor binding and activity to LEAP2. We found that both LEAP2 and its N-terminal part behave as inverse agonists of GHSR and as competitive antagonists of ghrelin-induced inositol phosphate production and calcium mobilization. Accordingly, the N-terminal region of LEAP2 is able to inhibit ghrelin-induced food intake in mice. These data demonstrate an unexpected pharmacological activity for LEAP2 that is likely to have an important role in the control of ghrelin response under normal and pathological conditions.

1. Introduction

Ghrelin [464] is a stomach-derived 28-amino acid peptide that plays a major role in the regulation of body energy homeostasis. It controls, among others, growth hormone secretion, blood glucose homeostasis, gastrointestinal tract motility, and reward-related behaviors [465]. Notably, ghrelin is the only known orexigenic peptide hormone and its administration to humans or rodents potently increases food intake [465]. The biological effects of ghrelin are due to its interaction with a prototypical class A G-protein-coupled receptor (GPCR), named the ghrelin receptor or growth hormone secretagogue receptor (GHSR) [466]. GHSR is a peculiar receptor as it displays a strikingly high constitutive activity [467,468]. Observations in animal models as well as in a family with a naturally occurring mutation of GHSR that selectively abolishes constitutive activity without altering ghrelin-

evoked activity suggest that constitutive GHSR activity plays a role in vivo, independent of ghrelin action [469–471].

Liver-expressed antimicrobial peptide 2 (LEAP2) [472] is a bicyclic, cationic peptide predominantly expressed in the liver and small intestine. This peptide is initially produced as a 77- residue precursor and subsequently processed, through different steps, into a mature 40-residue peptide. NMR-based structural analysis indicates that the mature peptide is composed of a disordered, hydrophobic N-terminal region and a compact central part with two disulfide bridges connected in a I–III, II–IV pattern. This peptide was originally reported to display antimicrobial activities with inhibitory concentrations in the hundreds of µM range [473]. However, such effective antimicrobial concentrations are much higher than the physiological levels of LEAP2 [473]. It has also been suggested that LEAP2 modulates fibroblast growth factor signals [474]. Very recently, LEAP2 was reported to also have noncompetitive allosteric antagonistic activity toward GHSR [475]. This antagonist activity was directly related to the inhibition of ghrelin effects in vivo. Notably, at physiological levels, LEAP2 is able to bind to GHSR and impair the action of ghrelin. Thus, it is now recognized that endogenous LEAP2 plays a major role as an antagonist of ghrelin action.

To illuminate the mechanisms responsible for the biological actions of LEAP2 on ghrelin signaling, we carried out a detailed analysis of the pharmacological properties of this peptide and of its different structural domains, namely, the unstructured N-terminal part and the compact central core. For this, we evaluated different peptide fragments in a panel of model systems ranging from purified GHSR in lipid nanodiscs to GHSR-expressing HEK293T cells, ex vivo and in in vivo rodent models. By doing so, we found that LEAP2 is a competitive antagonist of ghrelin at GHSR and an inverse agonist toward GHSR constitutive activity. Moreover, we demonstrated that the activities of LEAP2 are localized to its N-terminal region.

2. Methods

2.1 Abbreviations

GHSR, growth hormone secretagogue receptor; DIPEA, *N,N* diisopropylethylamine; DMEM, Dulbecco's modified Eagle's medium; HEPES, 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid; HTRF, homogeneous time-resolved fluorescence

2.2 Materials and methods

Ghrelin (1–28) was purchased from PolyPeptide Laboratories. [D-Arg^{1}-D-Phe^{5},D-Trp^{7,9},Leu^{11}]substance P (SPA) was from Bachem and K-(D-1Nal)-FwLL-NH_2 was synthesized at Institut des Biomolecules Max

Mousseron as described previously. IP-One HTRF kit, benzyl guanine-$Tb3^{+}$-cryptate and Insulin Ultra Sensitive kit were provided by CisBio. Cell line: HEK293T stable cell line expressing a SNAP-Tag-GHSR (SNAP-GHSR) was a generous gift of Eric Trinquet (Cisbio Bioassays, Codolet, France). BODIPY® FL GTPγS was from Invitrogen.

Chemical reagents used for buffer preparation, collagenase V, bovine serum albumin (BSA), phosphate buffered saline (PBS) with $MgCl_2$ and $CaCl_2$, were obtained from Sigma-Aldrich (St. Louis, MO, USA).

2.3 Peptide synthesis and characterization

Peptides were synthesized by solid-phase peptide synthesis using Fmoc chemistry, a 1 [bis(dimethylamino)-methylene]-1H-1,2,3-triazolo[4,5-b]-pyridinium 3-oxide hexafluorophosphate /diisopropylethylamine system for coupling, and piperidine/DMF for deprotection. A double coupling with 5 equivalents of reagents was used for each step, and the final deprotection was performed with a trifluoroacetic acid/triisopropylsilane/H2O (95:2.5:2.5) mixture for 2−3 h. After purification by preparative RP-HPLC, the identity of all peptides was evaluated by LC−MS, MALDI-MS, and MS−MS analyses and UV purity was assessed to be >95% (see in the Supporting Information).

2.4 Cell Culture

HEK293T cells were maintained in DMEM Glutamax (Invitrogen) supplemented with 1 mg/ ml Geneticin (Gibco) and containing 50 µg/ml penicillin, 50 µg/ml streptomycin, , 2mM HEPES, 1% non-essential amino acids and 10% heat-inactivated fetal calf serum.

2.5 Ligand Binding Assay

K_i values were determined from binding competition experiments performed on intact HEK293T cell line expressing the SNAP-GHSR using a Homogenous Time Resolved Fluorescence (HTRF) assay previously described [476]. HTRF signal was collected in a PHERAstar microplate reader (BMG LABTECH). K_i values were obtained from binding curves using GraphPad Prism 5.0 software (GraphPad Software, Inc., San Diego).

2.6 Inositol Phosphate Assay

Inositol phosphate accumulation assay was carried out on adherent HEK293T cell line expressing the SNAP-GHSR in a 96-well plate format at a density of 50,000 cells/well. IP1 production was measured using the IP-One HTRF kit (Cisbio Bioassays Ref. 621PAPEC) as described previously [476]. Briefly, cells were stimulated for 30 min at 37 °C with the ligand to be tested in 70 µl of IP1 stimulation buffer. An anti-IP1 antibody labeled with Lumi4-Tb (15 µl) and an IP1-d2 derivative (15 µl) were added to the

cells. The medium was incubated for 1 h at room temperature. Signals at 665 and 620 nm were detected using a PHERAstar (BMG LABTECH) fluorescence reader. EC_{50} values from dose response curves were obtained using GraphPad Prism 5.0 Software.

2.7 Intracellular Ca^{2+} measurement

HEKTSA cells stably expressing GHSRwere seeded at 100 µl/well of fresh media with 10^5 cells into 96 wells black-wall/clear-bottom plate (Corning 3603) precoated with poly-D-lysine. 24 h later, cells were washed and incubated with 100 µL/well of assay buffer containing Hanks' balanced salt solution, 0.5% BSA, 20 mM HEPES, 1 mM $MgSO_4$, 3.3 mM Na_2CO_3, 1.3 mM $CaCl_2$, 2.5 mM probenecid, pH 7.4 supplemented with 1 μM fluo 4-AM (Interchim). Cells were incubated at 37°C, 5% CO_2 with fluo 4-AM for 60 min followed by a 15 min incubation with LEAP2 peptides or control buffer. Then, cells were washed with 2 × 100 µl of assay buffer containing LEAP2 peptides and 50 µl/well of assay buffer was added. The plate containing the cells was placed into the FlexStation (benchtop scanning fluorometer, Molecular Devices, Sunnyvale, CA, U.S.A.) and the reading started by automatic addition of 50 µl/well of assay buffer containing ligands. Fluorescence output (Exc 495 nm/ Em 520 nm) was measured for 1 min. Data were expressed as relative fluorescence units (RFU) corresponding to the signal obtained by subtracting the baseline to the max calcium signal. EC_{50} values were obtained from dose-response curves using GraphPad Prism (version 5.0).

2.8 Receptor preparation and labeling

For GTPγS binding assays, the monomeric wild type GHSR in lipid nanodiscs was prepared as described [477]. For FRET measurements, the ghrelin receptor with a TAG amber codon at the position encoding F^{71} and a single reactive cysteine at position 255, was produced in lipid nanodiscs and subsequently labeled with Click-IT Alexa Fluor 488 DIBO Alkyne (LifeTechnologies) and Alexa Fluor 350 maleimide, as described [478].

2.8.1 G protein preparation

$G\alpha_{13}$ was purchased from Kerafast. His_6-tagged $G\alpha_q$ was expressed in Sf9 cells and purified using Ni-nitrilioacetic acid (Ni-NTA) affinity chromatography [479]. $G\beta_1$ was expressed with the His_6-Tagged Gγ in Sf9 cells and the Gβγ dimer purified using Ni-NTA chromatography combined to ion exchange chromatography [479].

2.8.2 GTPγS binding

GTPγS binding experiments were carried out using the fluorescent Bodipy-FL GTPγS analog [480]. Association of Bodipy-FL GTPγS to the G protein was monitored using a fluorescence spectrophotometer (Cary Eclipse, Varian) with the excitation wavelength set at 500 nm and the emission wavelength at 511 nm. Reaction conditions were 100 nM $G\alpha_q$ or $G\alpha_{13}$, 500 nM $G\beta_1\gamma_2$, 100 nM Bodipy-FL GTPγS, and 20 nM receptor in lipid discs. The receptor, the ligands (10 µM final concentration) and the G protein were first incubated for 30 minutes at room temperature and fluorescence was measured after addition of Bodipy-FL GTPγS and 10 minutes incubation at 15°C.

2.9 FRET measurements

For FRET measurements, the receptor, the ligand and $G\alpha_q\beta_1\gamma_2$ protein were incubated for 30 minutes at room temperature. GHSR and ligand concentrations were in the 0.5 µM and 10 µM range, respectively, with a 1:5 receptor-to-G protein molar ratio. Fluorescence emission spectra were recorded between 360 and 700 nm on a Cary Eclipse spectrofluorimeter (Varian) with an excitation at 346 nm. Buffer contributions were systematically subtracted. The proximity ratio was calculated from the emission spectra as described [481].

2.10 Animals

In vivo testing of LEAP2-related peptides was performed using 3-4 month old C57BL/6J male mice, which were single-housed in a 12-h light/dark cycle with regular chow and water available ad libitum. Studies were carried out in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Research Council, USA and received approval from the IACUC committee of the IMBICE (ID:10-0112).

2.11 Insulin secretion experiments in isolated rat pancreatic islets

All animal care and experiments adhered to the Directive 2010/63/EU of the European Parliament and of the Council. Male Wistar rats (Charles River, France) were housed in a temperature and humidity-controlled room under a 12 h light/dark cycle and had food and water ad libitum. Rats were weighing 280-320 g at the time of the experiments. Rats were anesthetized with isoflurane and killed by decapitation. The common bile duct was cannulated and the pancreas was filled by injection of collagenase V solution (10 mL, 1 mg/mL), excised and digested at 37°C for 10 min. After three successive sedimentations using PBS plus 5 mM glucose, islets of Langerhans were hand-picked under a microscope. Then, islets were stabilized for 1 h (5% CO_2, 37°C) in buffer A (120 mM NaCl, 4.7 mM KCl, 1.2 mM KH_2PO_4, 1.2 mM $MgSO_4$, 2.5 mM $CaCl_2$, 24 mM $NaHCO_3$, 0.1% (w/v) BSA, pH 7.4) containing

2.8 mM glucose. Batches of 5 islets were incubated for 1 h at 37°C, in 1 mL of buffer A with 0.1% (w/v) BSA containing 2.8 mM or 8.3 mM glucose in the absence or presence of the tested ligands. At the end of the incubation period, supernatants were collected and stored at -20°C until insulin assay. Experiences were performed in quintuplicate. Insulin was quantified using the CisBio Insulin Ultra Sensitive kit according to the manufacturer's instructions.

2.12 Food intake in mice

The ability of LEAP2-related peptides to affect ghrelin-induced food intake was tested using peripheral of the GHSR ligands. Mice were subcutaneously injected with saline alone or containing 600 pmoles/g of BW of LEAP2, LEAP2 (1-12-NH_2) or LEAP2 (15-40-OH) and ten minutes later subcutaneously injected with saline alone or containing 60 pmoles/g of BW of ghrelin (n=5-9 per group). The dose of ghrelin and LEAP2-related peptides was chosen based on previous studies showing that this ghrelin dose reliably increases food intake and that a 1:10 ghrelin:LEAP2-related peptides ratio would be enough to unmask the effect of LEAP2 [466,481]. In all cases, food intake was assessed at 10:00 a.m. in *ad libitum* fed mice, which were subjected to the different treatments and exposed to a single pre-weighed chow pellet in the floor of the home cages. Mice remained undisturbed for 2-hour, when chow pellets were collected and weighed. Weighing was performed in calibrated scales that had a precision of 1 mg. Two-hour food intake was calculated subtracting the remaining weight of the pellet to the initial weight, and expressed in mg.

3. Results and discussion

In all synthesized peptides (Figure 1), the N-terminal methionine residue was replaced by its norleucine isostere to avoid oxidation of the methionine residue. In a first set of experiments, due to the presence of a glycyl residue at position 14, we decided to synthesize three peptides: (1) full-length LEAP2, (2) an N-terminal LEAP2 fragment (1−14-OH), hereafter named LEAP2 (1−14-OH), and (3) a C-terminal LEAP2 structured fragment (15−40-OH) with the two disulfide bonds, hereafter named LEAP2 (15−40-OH). In a second set of experiments, we downsized the sequence of the N-terminal part of LEAP2. Peptides were assembled on a solid support using Fmoc chemistry and starting from 2-chlorotrityl resin or an Agilent Amphisphere 40 RAM resin for acid peptides or amide peptides, respectively. All peptides were purified by reversed-phase high-performance liquid chromatography (RP-HPLC) and characterized by liquid chromatography− mass spectrometry (LC−MS) and matrix-assisted laser desorption ionization-mass spectrometry (MALDI-MS)/MS with a purity >95% (see the Supporting Information).

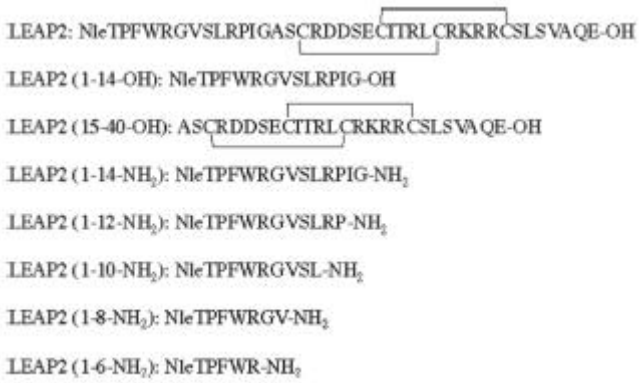


Figure 1: Sequences of the synthesized peptides.

3.1 Binding and activity of LEAP2-related peptides

3.1.1 Binding

We first analyzed whether the different LEAP2 variants directly bind to the ghrelin receptor (GHSR) expressed in HEK293T cells. Ki values were thus determined from binding competition experiments using the homogenous time-resolved fluorescence (HTRF) assay previously described [482]. As a reference, we included the peptide K-(D-1- Nal)-FwLL-NH_2, which is described as an inverse agonist [483]. As shown in Figure 2 and Table 1, LEAP2 displaced labeled ghrelin from its binding site with a Ki value in the nanomolar range (1.26 ± 0.05 nM). Interestingly, LEAP2 (1−14-OH) also displayed high affinity for GHSR (Ki 3.66 ± 0.64 nM), whereas LEAP2 (15−40-OH) was unable to compete with ghrelin up to a 10^{-5} M concentration.

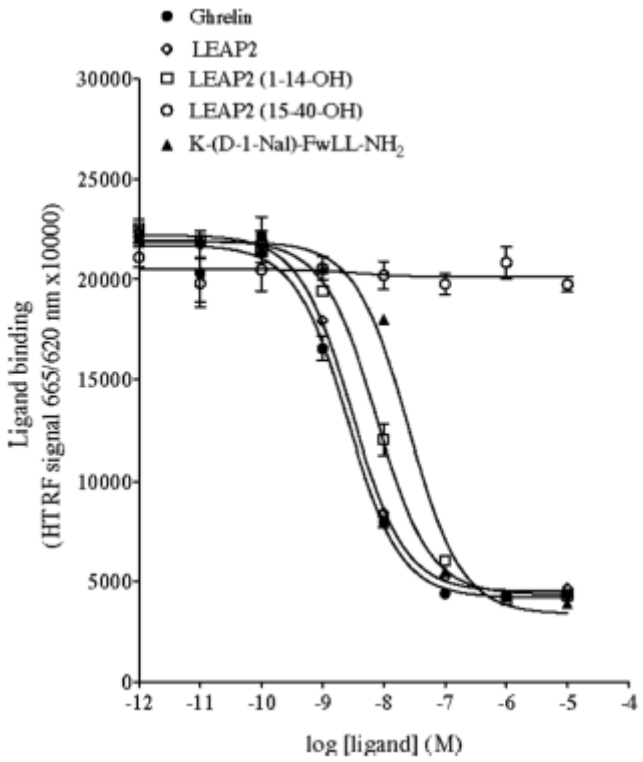


Figure 2: Competition binding curves obtained with LEAP2-related peptides, ghrelin, and the K-(D-1-Nal)-FwFwLL-NH_2 inverse agonist. Data are representative of two to three independent experiments, each performed in triplicate. Error bars correspond to standard deviation from the mean of triplicates. Ki values are collected in Table 1.

3.1.2 Intrinsic activity

As LEAP2 and some of its structural domains appeared to compete with ghrelin for binding to GHSR, we subsequently investigated their possible impact on ghrelin-independent GHSR signaling. To this end, we first tested the ability of LEAP2 variants to affect the basal level of inositol phosphate 1 (IP1) production in HEK293T cells expressing GHSR (Figure 3 and Table 1) [468,484]. As seen in Figure 3, the high basal level of IP1 production in GHSR-expressing cells, representing 60−70% of the maximal IP1 production promoted by ghrelin, resulted from the constitutive activity of GHSR as previously shown [467,484]. LEAP2 decreased the basal level of IP1 production by 50% (Emax: 52.3 ± 6.6%) promoted by GHSR with an EC_{50} in the 20 nM range (22.8 ± 7.8 nM). This indicates that LEAP2 displays an inverse agonist activity toward basal IP1 production promoted by GHSR. Interestingly, the maximal inverse agonist effect of LEAP2 was very similar to that of the reference inverse agonist compound K- (D-1-Nal)-FwLL-NH_2 [483] (Emax: 48.7 ± 1.2%). Similarly, LEAP2 (1−14-OH) reduced GHSR-promoted basal IP1 production with an EC_{50} of 76.4 ± 6.3 nM and a maximal effect of 50.8 ± 6.7%. In contrast, LEAP2 (15−40-OH) did not modify the basal level of IP1 production, indicating that this fragment had no activity of its own, in agreement with the absence of binding of this peptide to GHSR. These data indicate that LEAP2 behaves as an inverse agonist toward GHSR signaling at least on the GHSR-

promoted IP1 pathway and that its linear N-terminal part is both necessary and sufficient for its binding and its activity.

To our knowledge, current data are the first indication that LEAP2 acts as an inverse agonist of GHSR. Strikingly, the study by Ge and colleagues [475] found that LEAP2 did not affect GHSR-mediated β-arrestin recruitment and concluded that LEAP2 had no inverse agonist activity. Unfortunately, the referred study did not report the use of known inverse agonists of GHSR as positive controls to confirm that the sensitivity of the β-arrestin recruitment assay was sufficient to unmask the inverse agonist activity of LEAP2 [475]. Such controls are important because β-arrestin recruitment at GHSR is low under basal conditions [468,484]. In contrast, our conclusions are based on the data derived from the IP1 assay, which is recognized as a very sensitive assay for measuring the constitutive activity of GPCRs and studying the inverse agonist activity of compounds [473,485].

Table 1: Binding and activity of LEAP2-related peptides. Values reported in the table were obtained from experiments performed in HEK293T cells stably expressing GHSR. Ki values were determined from HTRF competition binding assays. EC_{50} and Emax values were obtained from dose–response curves of inositol phosphate (IP1) production in HEK293T cells expressing GHSR. Values are mean ± S.E. of two and three independent experiments for binding and activity, respectively, each experiment being performed in triplicate.

ligand	binding K_i (nM)	inositol phosphate 1 EC_{50} (nM)	inositol phosphate 1 E_{max} (% basal inhibition)
LEAP2 (1–14-OH)	3.66 ± 0.64	76.4 ± 6.3	50.8 ± 6.7
LEAP2 (15–40-OH)	not detected	not detected	0
LEAP2	1.26 ± 0.05	22.8 ± 7.8	52.3 ± 6.6
ghrelin	1.28 ± 0.36	1.51 ± 0.85	
K-(D-1-Nal)-FwLL-NH_2	3.7 ± 0.05	5.5 ± 2.6	48.7 ± 1.2

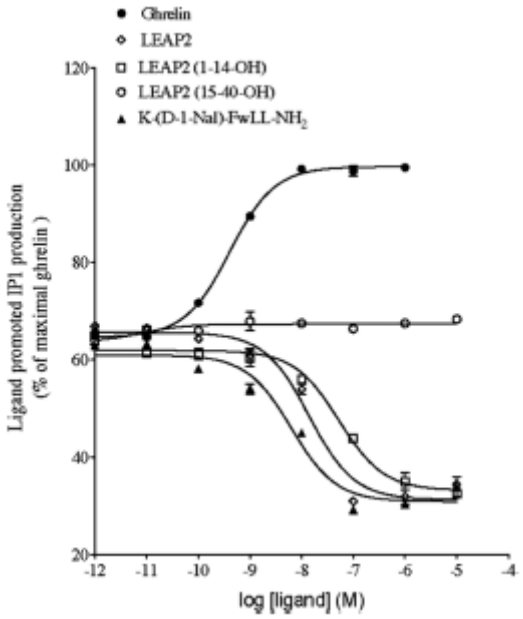


Figure 3: Dose−response curves of ligands at promoting IP1 production. Data are expressed as the percentage of maximal effect of ghrelin on HEK293T cells expressing GHSR. In this experiment, basal IP1 production of GHSR-expressing cells represents 64.5 ± 1.5% of the maximal effect triggered by ghrelin. K-(D-1-Nal)-FwLLNH$_2$ described in the literature as an inverse agonist is given as reference. Data are representative of two experiments, each performed in triplicate.

3.2 Antagonist effect of LEAP2 peptides on Ghrelin-induced activation of GHSR

3.2.1 IP1 Production

We first tested the antagonist effect of LEAP2-related peptides on ghrelin-induced IP1 production. To evaluate whether the interaction between LEAP2 and ghrelin is competitive or noncompetitive, we constructed dose−response curves of ghrelin in the absence or presence of increasing concentrations of LEAP2 or LEAP2 (1−14-OH). As shown in Figure 4, increasing the concentration of LEAP2 increased the EC_{50} value of ghrelin from 0.36 ± 0.01 to 16.6 ± 5.5 nM in the absence or the presence of 1 µM LEAP2, respectively (Figure 4A). Similarly, increasing the concentration of LEAP2 (1−14- OH) up to 1 µM increased the EC_{50} value of ghrelin from 0.40 ± 0.05 to 2.5 ± 0.3 nM (Figure 4B). pA2 values of 8.83 ± 0.22 and 7.46 ± 0.21 were obtained for LEAP2 and LEAP2 (1−14- OH), respectively. In contrast, neither LEAP2 nor LEAP2 (1−14-OH) changed the maximal effect elicited by ghrelin (Figure 4). These data strongly suggest that LEAP2 and its N-terminal (1−14) domain are competitive antagonists for ghrelin with regard to IP1 production.

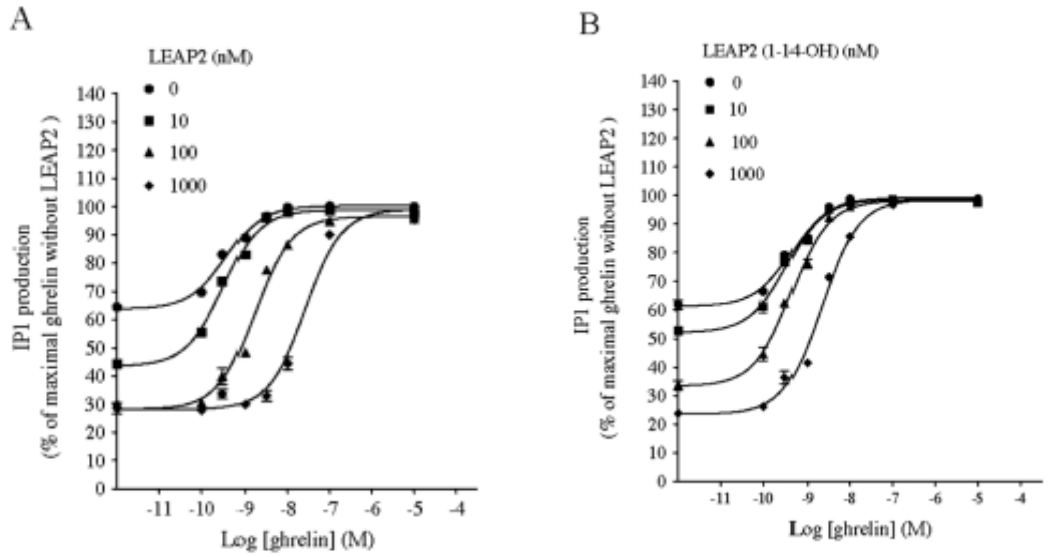


Figure 4: Antagonist effect of LEAP2 peptides on ghrelin-induced IP1 production in HEK293T cells expressing GHSR. Dose−response curves of ghrelin-induced IP1 production in the presence of increasing concentrations of LEAP2 (A) and LEAP2 (1−14-OH) (B). Data are expressed as the percentage of maximal ghrelin response in HEK293T cells stably expressing GHSR, where zero represents basal IP1 production of mock-transfected HEK293T cells. Data are representative of two experiments, each performed in triplicate.

3.2.2 Calcium release

We then tested whether the antagonist competitive character of LEAP2-related peptides existed also on calcium release following activation of GHSR by ghrelin. As for the IP1 assay, we constructed dose−response curves of ghrelin-induced calcium release in the presence of increasing concentrations of LEAP2-related peptides in the same GHSRexpressing cell system. As shown in Figure 5A, the EC_{50} value of ghrelin increased in parallel with the increase of LEAP2 concentration from 0.41 ± 0.13 to 27.7 ± 13 nM. Similarly, increasing concentrations of LEAP2 (1−14-OH) up to 1 µM increased the EC_{50} value of ghrelin from 0.34 ± 0.12 to 18 ± 4 nM (Figure 5B). pA2 values of 7.99 ± 0.15 and 7.05 ± 0.05 were calculated for LEAP2 and LEAP2 (1−14-OH), respectively. Moreover, as it was for IP1 production, LEAP2 and LEAP2 (1−14-OH) did not affect the maximal effect elicited by ghrelin (Figure 5). It has to be mentioned that, in contrast to IP1 production, no reduction of the basal signal was observed for calcium levels. This is not inconsistent with the fact that LEAP2-related peptides are inverse agonists. Indeed, it is well established that the inverse agonist activity of compounds cannot be detected with this calcium assay because of both the inability of this assay to produce an elevated basal signal and the nonequilibrium experimental conditions of the assay [486].

Overall, these data indicate that both LEAP2 and LEAP2 (1−14-OH) are competitive antagonists of ghrelin at GHSR, as increasing concentrations of LEAP2 and its N-terminal domain resulted in an increase of EC_{50} values of ghrelin with no change in its maximal effect for both IP1 production and calcium mobilization. The reason for the difference between our data and those previously reported on the competitive behavior of LEAP [475] is not clear so far, and the exact nature of this behavior is still an open question at the present stage of the analysis.

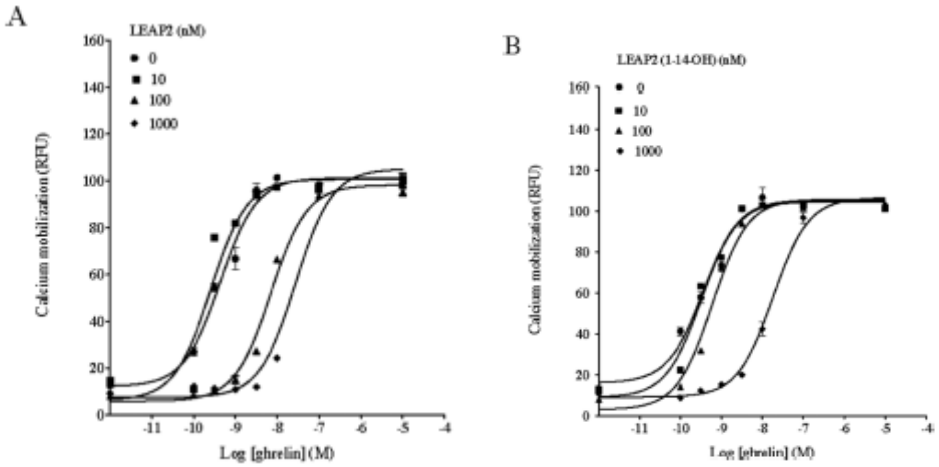


Figure 5: Antagonist effect of LEAP2-related peptides on ghrelininduced calcium release in HEK293T cells expressing GHSR. Dose−response curves of ghrelin-induced GHSR activation at promoting intracellular calcium release in the presence of increasing concentrations of LEAP2 (A) and LEAP2 (1−14-OH) (B). Data are expressed as the percentage of maximal ghrelin response in HEK293T cells expressing GHSR, where zero represents the basal IP1 production of mock-transfected HEK293T cells. Data are representative of two experiments, each performed in triplicate.

3.3 Binding and biological activity of N-terminal truncated fragments

Our data above indicate that LEAP2 (1−14-OH) contains the structural determinants for recognition and activation of GHSR. To assess the minimal active sequence within LEAP2, we prepared different Nterminal truncated LEAP2-related variants. These were synthesized as their terminal amide analogues to mask the carboxylic charge that is not present in LEAP2. We then evaluated their binding properties and their ability to modulate the GHSR activity on IP1 production (Table 2). LEAP2 (1− 14-OH) and LEAP2 (1−14-NH_2) exhibited almost the same affinity, and their inverse agonist potency and efficacy were similar. All LEAP2 N-terminal fragments displayed Ki values in the same range. Only LEAP2 (1−6-NH_2) displayed a 20-fold increased Ki value compared to that of LEAP2 (1−14-NH_2). In terms of IP1 production, truncation of LEAP2 did not affect the ability of the peptides to act as inverse agonists

of GHSR; however, their potency decreased with the reduction in their size. Indeed, a significant increase of EC_{50} was observed for LEAP2 (1−10-NH_2) and LEAP2 (1−8-NH_2), whereas their efficacy remained unchanged, as compared to those for both LEAP2 (1−14-OH) and LEAP2 (1−14-NH_2) as well as to those for the reference inverse agonists substance P analogue (SPA) [467] and K-(D-1-Nal)-FwLL-NH_2. Downsizing of only the N-terminal fragment of LEAP2 to six residues dramatically decreased both its potency and efficacy (EC_{50}: 6344 ± 2155 nM, Emax: 26 ± 7%), whereas its binding affinity was much less affected. These data suggest that the binding determinants of LEAP2 are contained within the (1−8) N-terminal part of LEAP2, whereas a longer N-terminal sequence encompassing at least the 1−12 region is required for maximal potency and efficacy.

Table 2: Binding affinities and Inverse agonist activities of truncated N-terminal peptides of LEAP2. Ki values were determined by the HTRF competition binding assay on HEK293T cells expressing GHSR. For activity, EC_{50} and Emax values were obtained from dose−response curves of inositol phosphate (IP1) production in HEK293T cells expressing GHSR. Values are mean ± S.E. of two and three independent experiments for binding and activity, respectively, each performed in triplicate.

ligand	binding K_i (nM)	inositol phosphate 1 EC_{50} (nM)	inositol phosphate 1 E_{max} (% basal inhibition)
LEAP2	1.26 ± 0.05	22.8 ± 7.8	52.3 ± 6.6
LEAP2 (1−14-OH)	3.66 ± 0.64	76.4 ± 6.3	50.8 ± 6.7
LEAP2 (1−14-NH_2)	3.05 ± 0.15	48.0 ± 5.4	50.3 ± 1.5
LEAP2 (1−12-NH_2)	2.7 ± 0.1	46 ± 18	49.3 ± 1.5
LEAP2 (1−10-NH_2)	3.4 ± 0.01	134 ± 48	50.3 ± 1.5
LEAP2 (1−8-NH_2)	6.45 ± 0.25	368 ± 63	45.3 ± 1.5
LEAP2 (1−6-NH_2)	55 ± 4	6633 ± 2212	26 ± 7.2
K-(D-1-Nal)-FwLL-NH_2	3.7 ± 0.05	5.5 ± 2.6	48.7 ± 1.2
rPKPfQwFwLL-NH_2 (SPA)	76.7 ± 12.6	98.7 ± 20.0	51 ± 13
ghrelin	1.28 ± 0.36	1.51 ± 0.85	

3.4 LEAP2 activity on the purified ghrelin receptor

We then used the purified ghrelin receptor assembled into lipid nanodiscs to evaluate whether the impact of LEAP2 on GHSR signaling observed in HEK293T cells was associated with a direct effect of this peptide on the receptor. To this end, the ability of LEAP2 and its N-terminal fragments to affect GHSR activation was assessed by the ability of the purified receptor to activate purified G proteins in a GTPγS binding assay. The purified apo GHSR in lipid nanodiscs displayed a significant constitutive activity at both G_q and G_{13} (Figure 6). Ghrelin further enhanced these G-protein-activation properties, whereas the K-(D-1-Nal)-FwLL-NH_2 inverse agonist decreased GTPγS binding for Gq and G_{13} [487]. Finally, the neutral antagonist JMV 3011 [484] did not change basal GTPγS binding for both G proteins. As shown in Figure 6, the full-length LEAP2 decreased the constitutive activity of GHSR for both G_q and G_{13} in the absence and in the presence of ghrelin. The degree of decrease in the basal activity with

LEAP2 was similar to the one detected with K-(D-1-Nal)-FwLL-NH_2. This indicates that LEAP2 by itself exerts an inverse agonist activity on GHSR-dependent basal activation of G_q and G_{13}. Importantly, this inverse agonist activity was preserved with all of the N-terminal truncated fragments of LEAP2, although it was slightly diminished with LEAP2 (1–6-NH_2) (Figure 6). Thus, these data suggest that the inverse agonist activity of LEAP2 observed in HEK293T cells is directly related to its interaction with the ghrelin receptor independent of the cellular context.

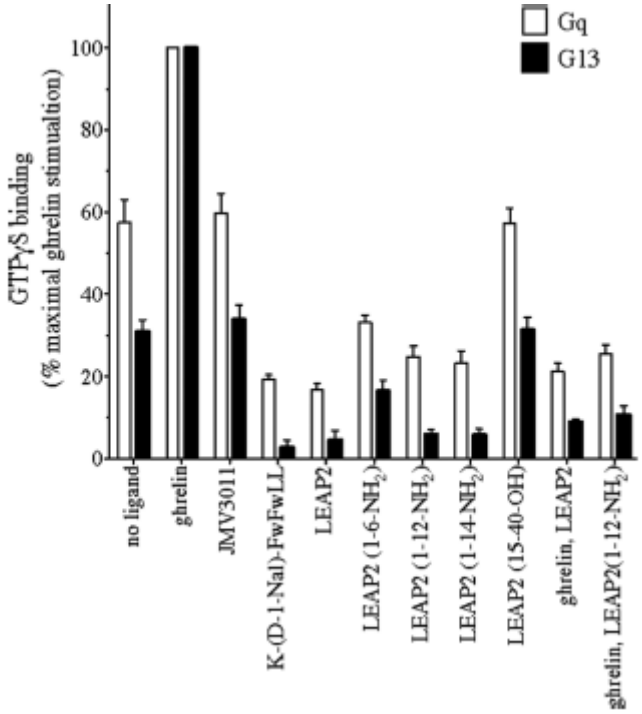


Figure 6: Effect of LEAP2 peptides on GHSR-catalyzed G protein activation. BODIPY-FL GTPγS binding to G_q and G_{13} proteins induced by GHSR in lipid disks in the absence of ligand; in the presence of ghrelin, JMV 3011, and K-(D-1-Nal)-FwLL-NH_2 (all ligands at 10−6 M); and in the presence of LEAP2, its N-terminal amide peptides, and C-terminal LEAP2 (15−40-OH) (10^{-5} M). Data are presented as the percentage of maximum BODIPY-FL fluorescence change measured in the presence of ghrelin and represent the mean ± S.E. from three independent experiments.

3.5 Impact of LEAP2 on GHSR conformation

We previously reported that the pharmacological properties of GHSR ligands are intricately related to their ability to stabilize specific receptor conformations [488]. To assess whether the inverse agonist activity of LEAP2 is indeed associated with its ability to stabilize an inactive conformation similar to that observed in the presence of K-(D-1-Nal)-FwLL-NH_2, we analyzed the conformational features of GHSR using a fluorescence resonance energy transfer (FRET)-based approach (Figure 7) [489]. To this

end, the purified monomeric receptor in nanodiscs was labeled with a fluorescence donor (AF350) and an acceptor (AF488) in the cytoplasmic ends of TM1 and TM6. As shown in Figure 7, addition of ghrelin or K- (D-1-Nal)-FwLL-NH_2 at saturating concentrations induced a significant change in the FRET-monitored proximity ratio, indicative of a change in the distance between the two fluorescent probes. These variations were not due to changes in the mobility or orientation of the fluorophores, as the anisotropy of fluorophores attached to the receptor was not altered by treatment with ligands. Specifically, the full agonist ghrelin triggered a decrease in the proximity ratio, whereas the neutral antagonist JMV 3011 did not change this ratio and the inverse agonists triggered an increase in this ratio. As shown in Figure 7, all LEAP2-related peptides also triggered an increase in the proximity ratio very similar to that observed with K-(D- 1-Nal)-FwLL-NH_2, suggesting that these compounds stabilize similar GHSR conformations. As in the case of the GTPγS binding assays, the proximity ratio was essentially unaffected by the deletion of the central domain of LEAP2, indicating that the interaction with the N-terminal region of LEAP2 is sufficient to stabilize the conformation of GHSR associated with its reduced basal activity. Altogether, these data indicate that the N-terminal domain of LEAP2 exerts an intrinsic inverse agonist activity on GHSR-catalyzed G protein activation through the stabilization of a specific, inactive receptor conformation.

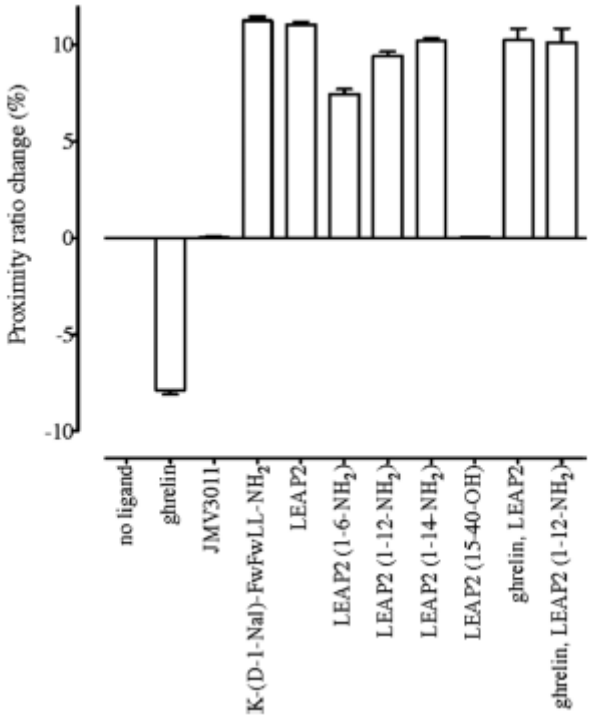


Figure 7: Effect of LEAP2 on GHSR conformation. Proximity ratio changes induced by ghrelin, JMV 3011, K-(D-1-Nal)-FwLL-NH_2, LEAP2, its N-terminal amide peptides, and C-terminal LEAP2 (15− 40-OH). These changes were calculated from the FRET signal between the fluorophores in TM1 and TM6 of the purified ghrelin receptor in lipid disks (see the Materials and Methods section). The data are representative of two experiments performed in duplicate, and the error bar represents the S.E.

3.6 Effect of LEAP2 (1-12 NH_2) on "ex vivo" endogenously expressed GHSR

Next, we aimed at determining whether the N-terminal domain of LEAP2 was also active on endogenously expressed GHSR. To do so, we tested rat pancreatic islets that express GHSR [487,490] Indeed, ghrelin is known to inhibit glucose-induced insulin secretion in pancreatic islets [491]. Therefore, insulin release was quantified after incubating isolated rat pancreatic islets with a stimulating (8.3 mM) glucose concentration in the presence or absence of ghrelin and LEAP2 (1−12-NH_2). LEAP2 (1−12-NH_2) was chosen because its inverse agonist potency and efficacy were equivalent to those of LEAP2 (1−14-NH_2) (Figure S1 in the Supporting Information). As expected, ghrelin attenuated glucose-stimulated insulin secretion (Figure S2 in the Supporting Information). Importantly, co-treatment of pancreatic islets with ghrelin and LEAP2 (1−12-NH_2) abolished the insulinostatic action of ghrelin. These data indicated that LEAP2 (1−12-NH_2) binds to endogenous GHSR and blocks its activation by ghrelin.

3.7 In vivo evaluation of LEAP2-related peptides

Finally, we tested the ability of LEAP2, LEAP2 (1−12-NH_2), and LEAP2 (15−40-OH) to affect the orexigenic effects of ghrelin in mice. We performed this experiment using similar experimental conditions as those used by Ge and colleagues [475] to compare the outcomes of the studies. Thus, we tested a ghrelin to LEAP2-related peptides ratio of 1:10 and assessed cumulative food intake 2 h after treatment, as the entire orexigenic response to the hormone occurs within this time window [492]. In particular, mice were initially subcutaneously (s.c.) injected with the vehicle alone or containing the different LEAP2 variants (0.6 nmol/g BW) and injected with ghrelin (0.06 nmol/g BW, Figure 8) 10 min later. Food intake was assessed 2 h after treatment. As compared to that in vehicle treatment, ghrelin increased food intake, whereas LEAP2, LEAP2 (1−12-NH_2), and LEAP2 (15−40-OH) alone did not affect food intake. LEAP2 fully inhibited ghrelin-induced food intake as previously described. More interestingly, LEAP2 (1− 12-NH_2) also inhibited ghrelin-induced food intake in a similar fashion as seen for LEAP2, whereas LEAP2 (15−40-OH) did not affect ghrelin-induced food intake, as expected based on our in vitro data. These in vivo data confirmed that the N-terminal part of LEAP2 preserves the biological effects of the whole peptide on food intake.

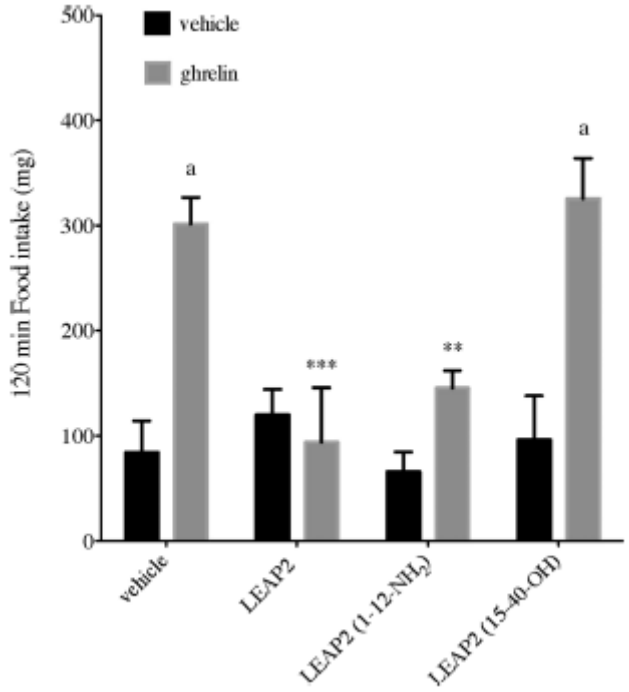


Figure 8: Food intake monitored at 2 h after acute subcutaneous administration of LEAP2-related peptides in the presence or absence of ghrelin. ***p < 0.001; **p < 0.01 versus vehicle + ghrelin group. [a]*p* < 0.05 versus the same initial treatment (either vehicle, LEAP2, LEAP2 (1−12-NH_2), or LEAP2 (15−40-OH)) + vehicle. n = 5−9 per group.

4. Conclusions

We delineated here the active part of LEAP2 and showed that the activity of this peptide toward the ghrelin receptor is restricted to its N-terminal region. Moreover, our data provide evidence that both LEAP2 and its N-terminal part act as inverse agonists of GHSR by stabilizing an inactive conformation of the receptor. LEAP2 and its 1−14 N-terminal region also appeared to be competitive antagonists of ghrelin at promoting IP1 production and calcium mobilization in HEK293T cells expressing GHSR. Accordingly, LEAP2 (1− 12-NH_2) blocks ghrelin-induced insulin level reduction in rat pancreatic islets, indicating that LEAP2 antagonizes ghrelin action *ex vivo* at endogenous GHSR. Moreover, s.c. administration of LEAP2 and LEAP2 (1−12-NH_2) in mice was associated with a decrease of ghrelin-induced food intake. One can hypothesize from our in vitro pharmacological data that LEAP2 and/or its N-terminal part may behave as endogenous inverse agonists of GHSR. Indeed, GHSR displays very high constitutive activity and it may be possible that LEAP2 regulates physiologically the tonic activity of GHSR, a hypothesis that certainly deserves to be tested in future studies. Interestingly, besides our LEAP2 inverse agonist of the GHSR, another endogenous inverse agonist, the agouti-related protein, has been reported for the melanocortin receptor [493]. In this line, LEAP2-related peptides developed in the present study should be useful tools to further investigate the physiological role of these potential endogenous inverse agonists of GHSR and to design new potential antiobesity drugs. Altogether, our observations shed light on the control of ghrelin-mediated regulation of food intake and certainly pave the way for future studies.

5. Supporting information

5.1 Preparative RP-HPLC

Preparative RP-HPLC was run on a Gilson PLC 2250 Purification system (Villiers le Bel, France) instrument using a preparative column (Waters DeltaPak C18 Radial-Pak Cartridge, 100 Å, 40-100 mm, 15 um particle size, flow rate 50.0 mL/min). Buffer A was 0.1% TFA in water, and buffer B was 0.1% TFA in acetonitrile.

5.2 LC/MS analyses

The LC/MS system consisted of a HPLC-micromassZQ (Waters) or UPLC Acquity H-Class (Waters) coupled to a Synapt G2-S (Waters) equipped with an ESI source. Analyses were carried out using a Phenomenex Kinetex column (C18, 100Å, 100 x 2.1 mm, 2.6 µm). A flow rate of 0.5 mL/min and a gradient of 0-100% B in 5 min were used: eluent A, water/0.1% HCO_2H; eluent B, ACN/0.1% HCO_2H.

Positive-ion electrospray (ESI^+) mass spectra were acquired from 100 to 1500 m/z with a scan time of 0.2 s. Nitrogen was used for both the nebulizing and drying gas. Retention time (RT) is given in minutes.

5.3 MALDI MS and MS/MS Analyses

MALDI mass spectrometry experiments: Analyses were operated on the Rapiflex instrument from Bruker Daltonics©. Each sample was analysed from CHCA or SA matrix deposits, in positive ion mode. A pulsed Nd:YAG laser at a wavelength of 355 nm was operated with a laser focus of 29 %. Data were acquired with the Flex Control software (version 4.1, Bruker Daltonics©). Spectra were integrated with the Flex Analysis software (version 4.0, BrukerDaltonics©), the centroid algorithm was used to assign peaks.

An acceleration voltage of 20.0 kV (IS1), 2.45 kV (PIE) and a lense voltage of 11.6 kV was applied. The reflectron mode was used for the ToF analyser (voltages of 20.8 kV and 8.6 kV). The delayed extraction time was 170 ns. MS data were processed with the Flex Analysis software (version 4.0, Bruker Daltonics©). External calibration was performed with commercial peptide mixture (Calibration peptide standard II, Calibration protein I, Bruker Daltonics, Wissembourg, France).

Fragmentation experiments were performed under laser induced dissociation (LID) conditions with the LIFT cell voltage parameters set at 14.0 kV (LIFT 1) and 18.7 kV (LIFT 2) for a final acceleration of 29.5 kV (reflector voltage) and a pressure in the LIFT cell around 5×10^{-7} mbar. The precursor ion selector was set manually to the first monoisotopic peak of the molecular ion pattern for all analyses. MS/MS data were processed with the Flex Analysis software (version 4.0, Bruker Daltonics©). Mass lists were generated according to the following parameters: SNAP as peak detection algorithm, S/N threshold 3.

LEAP2 (1-6-NH_2), H-Nle-Thr-Pro-Phe-Trp-Arg-NH_2

Mol. Formula: $C_{41}H_{59}N_{11}O_7$

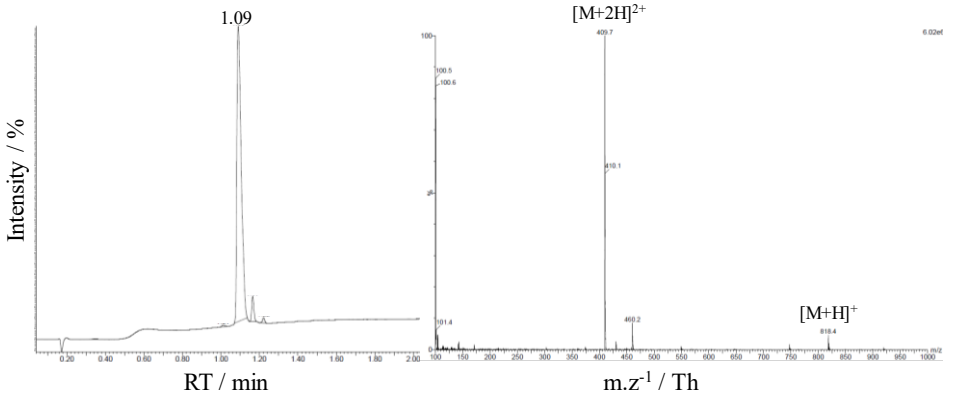
MW= 817.9 g.mol^{-1}

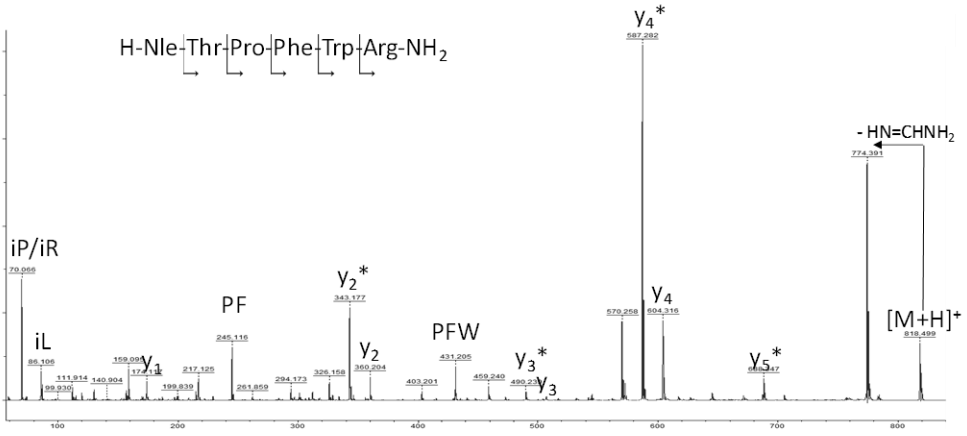
UV Purity: 95%

LC/MS data

RT: 1.09 min

ESI(+): Mcalculated 817.4 Da; observed m/z 409.7 Th $[M+2H]^{2+}$, m/z 818.4 Th $[M+H]^+$





LEAP2 (1-8-NH_2), H-Nle-Thr-Pro-Phe-Trp-Arg-Gly-Val-NH_2

Mol. Formula: $C_{48}H_{71}N_{13}O_9$

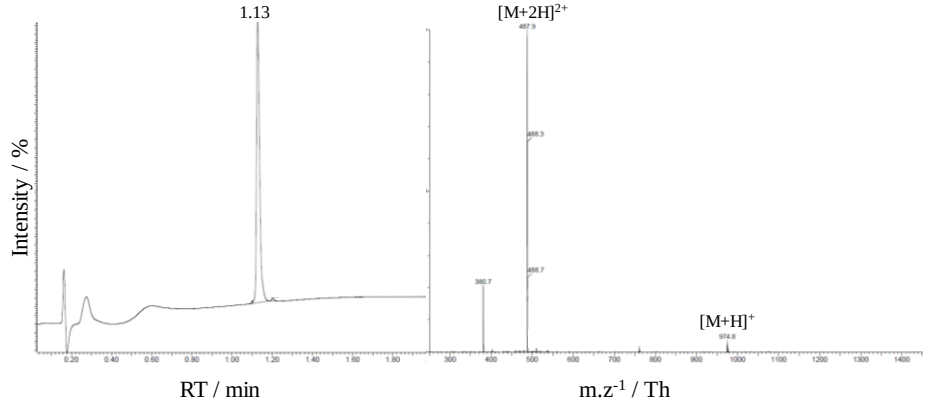
MW= 974.2 g.mol^{-1}

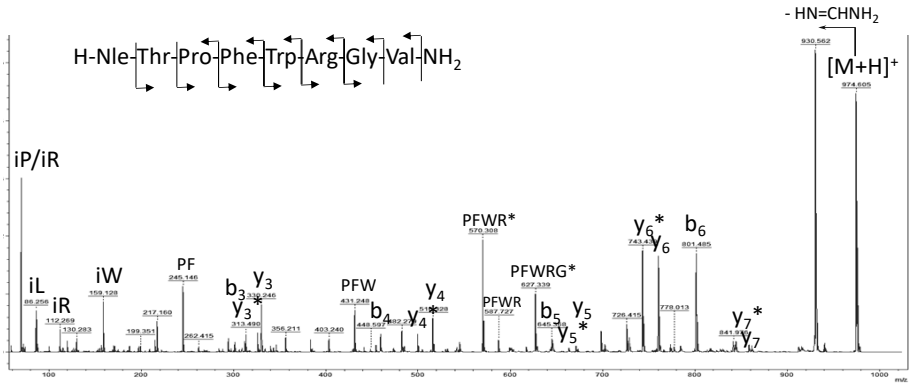
UV Purity: 99%

LC/MS data

RT: 1.13 min

ESI(+): Mcalculated 973.5 Da; observed m/z 487.9 Th $[M+2H]^{2+}$, m/z 974.6 Th $[M+H]^+$

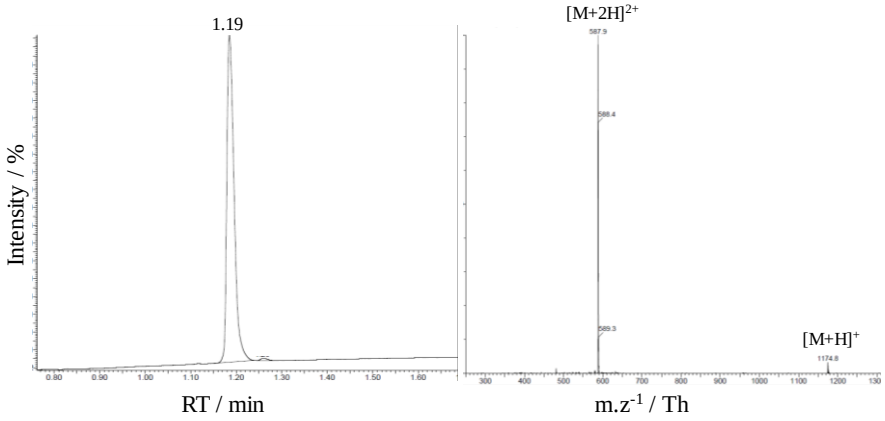


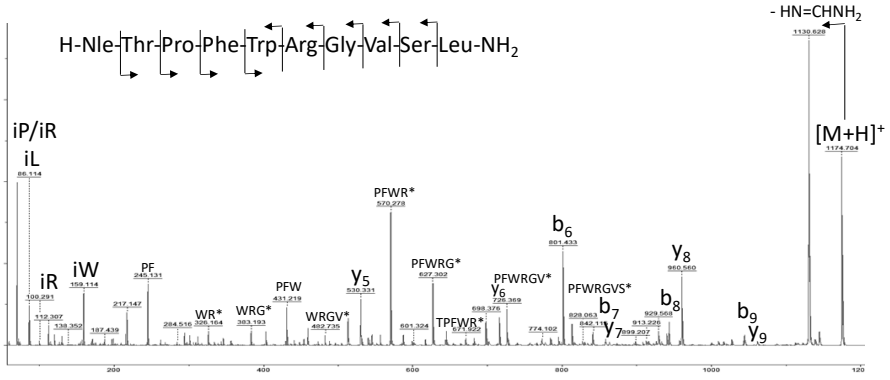


LEAP2 (1-10-NH_2), H-Nle-Thr-Pro-Phe-Trp-Arg-Gly-Val-Ser-Leu-NH_2

Mol. Formula: $C_{57}H_{87}N_{15}O_{12}$
MW= 1174.4 g.mol^{-1}
UV Purity: 99%

LC/MS data
RT: 1.19 min
ESI(+): Mcalculated 1173.7 Da; observed m/z 587.9 Th $[M+2H]^{2+}$, m/z 1174.8 Th $[M+H]^+$





LEAP2 (1-12-NH_2), H-Nle-Thr-Pro-Phe-Trp-Arg-Gly-Val-Ser-Leu-Arg-Pro-NH_2

Mol. Formula: $C_{68}H_{106}N_{20}O_{14}$

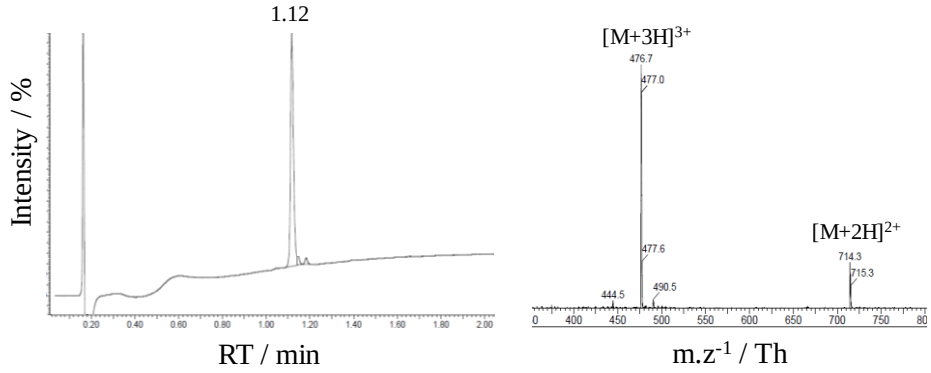
MW= 1427.7 g.mol^{-1}

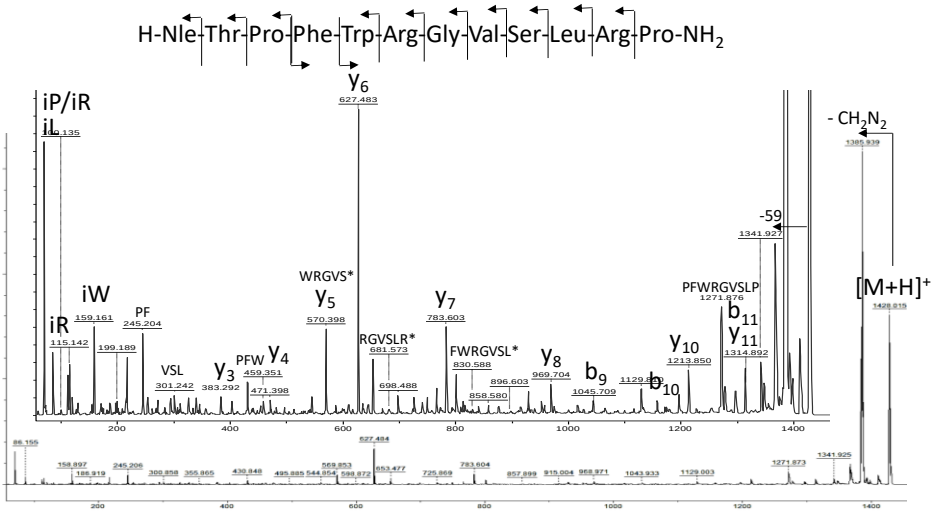
UV Purity: 95%

LC/MS data

RT: 1.12 min

ESI(+): Mcalc 1426.8 Da; obs m/z 476.7 Th $[M+3H]^{3+}$, m/z 714.3 Th $[M+2H]^{2+}$





LEAP2 (1-14-NH_2), H-Nle-Thr-Pro-Phe-Trp-Arg-Gly-Val-Ser-Leu-Arg-Pro-Ile-Gly-NH_2

Mol. Formula: $C_{76}H_{120}N_{22}O_{16}$

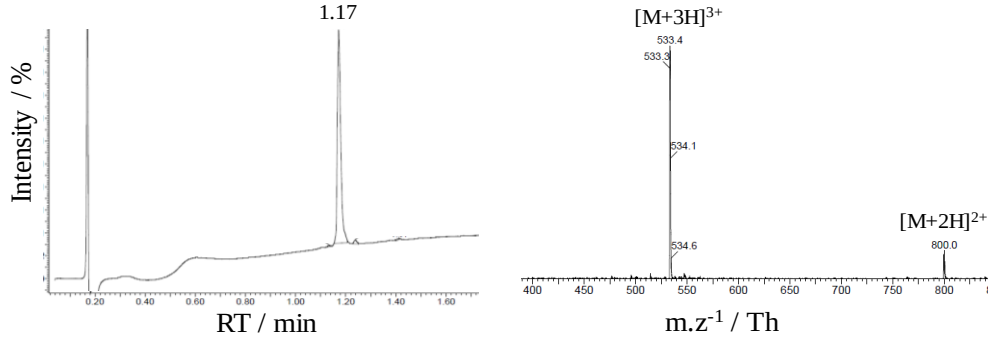
MW= 1597.9 $g.mol^{-1}$

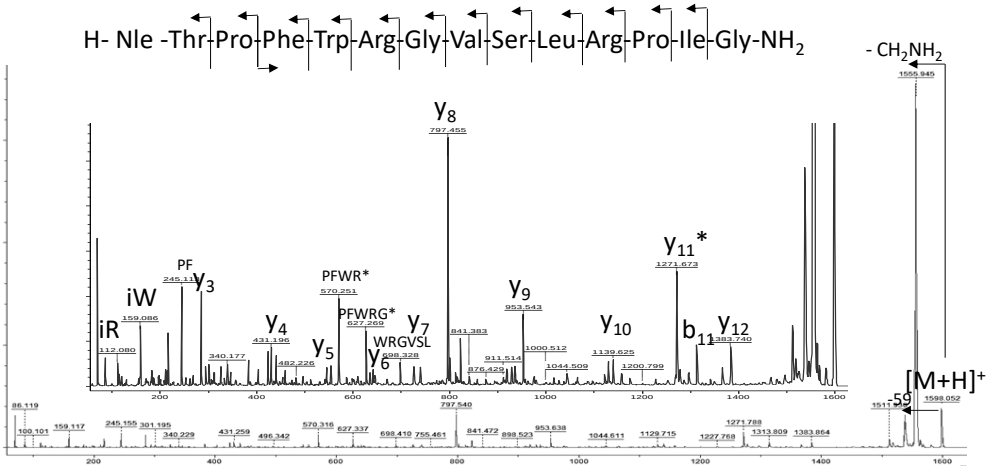
UV Purity: 98%

LC/MS data

RT: 1.17 min

ESI(+): Mcalc 1596.9 Da; obs m/z 800.0 Th $[M+3H]^{3+}$, m/z 533.4 Th $[M+2H]^{2+}$





LEAP2 (1-14-OH), H-Nle-Thr-Pro-Phe-Trp-Arg-Gly-Val-Ser-Leu-Arg-Pro-Ile-Gly-OH

Mol. Formula: $C_{76}H_{120}N_{21}O_{17}$

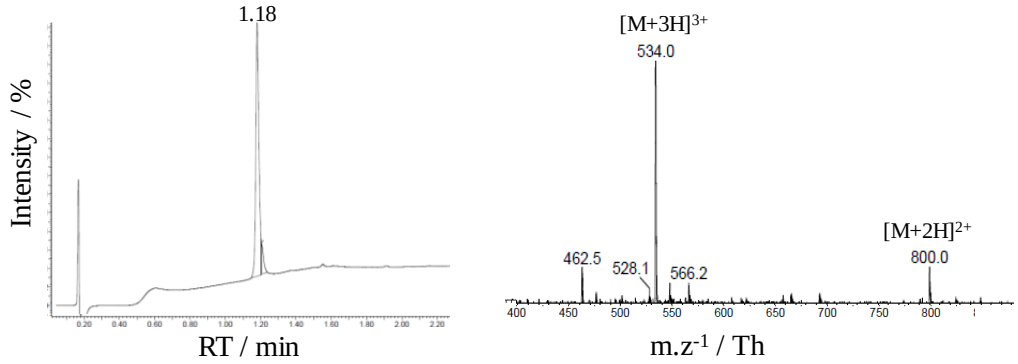
MW= 1598.9 g.mol^{-1}

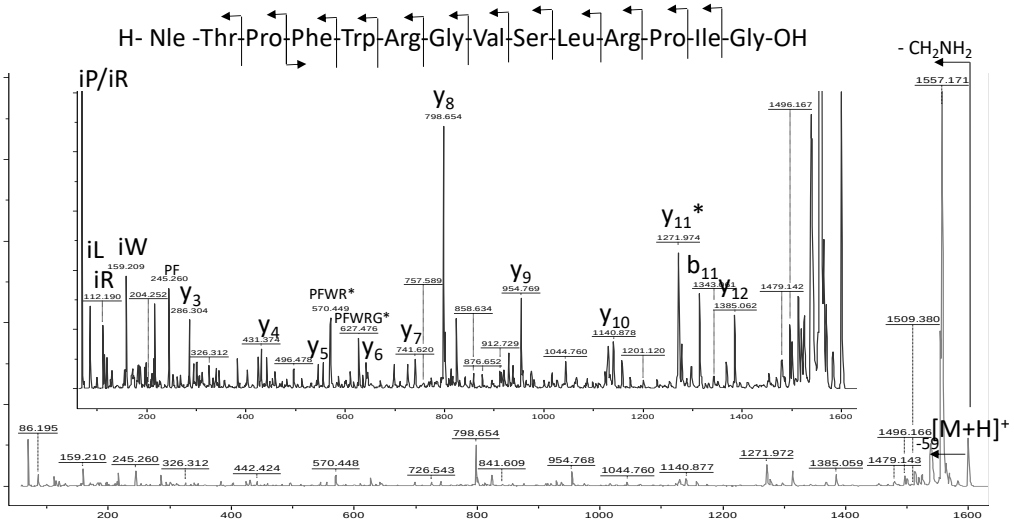
UV Purity: 93%

LC/MS data

RT: 1.18 min

ESI(+): Mcalc 1597.9 Da; obs m/z 800.0 Th $[M+3H]^{3+}$, m/z 533.4 Th $[M+2H]^{2+}$

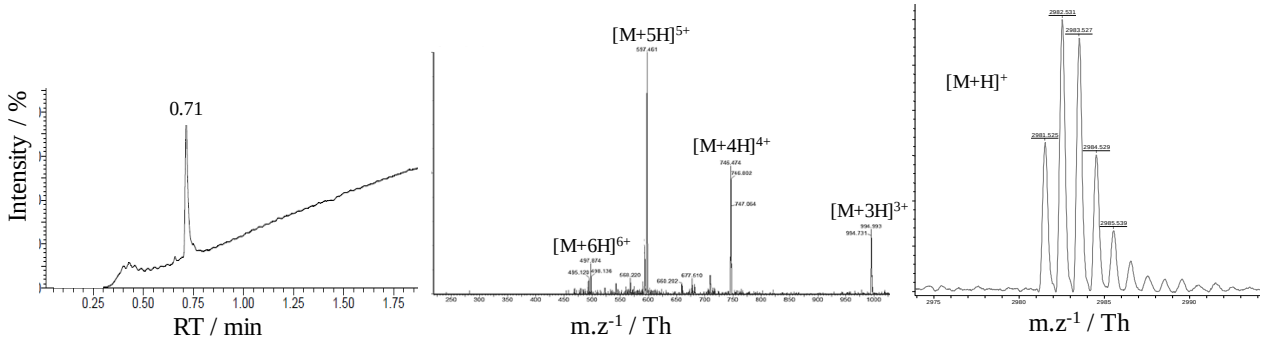


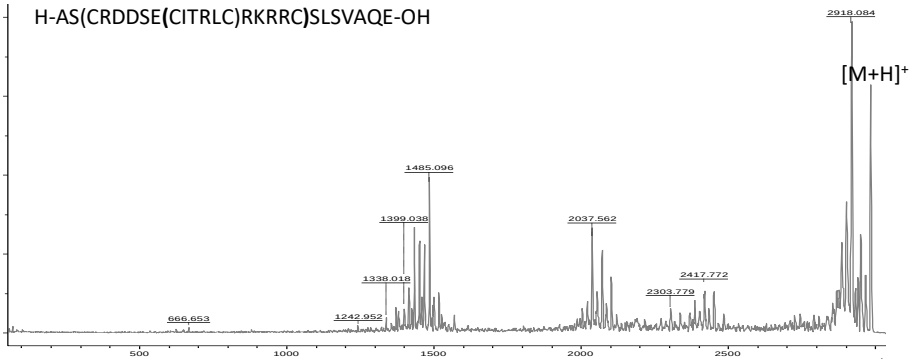


LEAP2 (15-40-OH), H-AS(CRDDSE(CITRLC)RKRRC)SLSVAQE-OH)

Mol. Formula: $C_{116}H_{202}N_{43}O_{41}S_4$
MW= 2982.5 g.mol^{-1}
UV Purity: 95%

LC/MS data
RT: 0.71 min
MALDI(+) (CHCA, reflectron mode): Mcalc 2980.4 Da; obs m/z 2981.5 Th $[M+H]^+$
ESI(+):obs m/z 497.9 Th $[M+6H]^{6+}$, 597.4 Th $[M+5H]^{5+}$, 746.4 Th $[M+4H]^{4+}$, 994.9 Th $[M+3H]^{3+}$



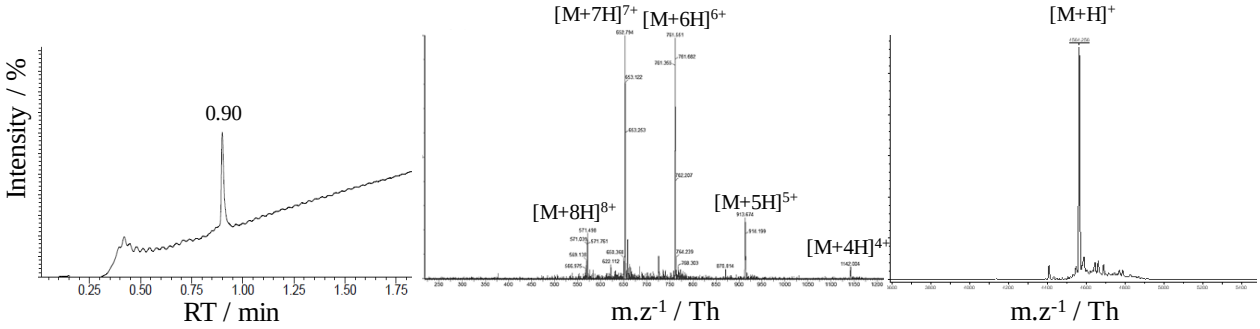


LEAP2(1-40-OH),

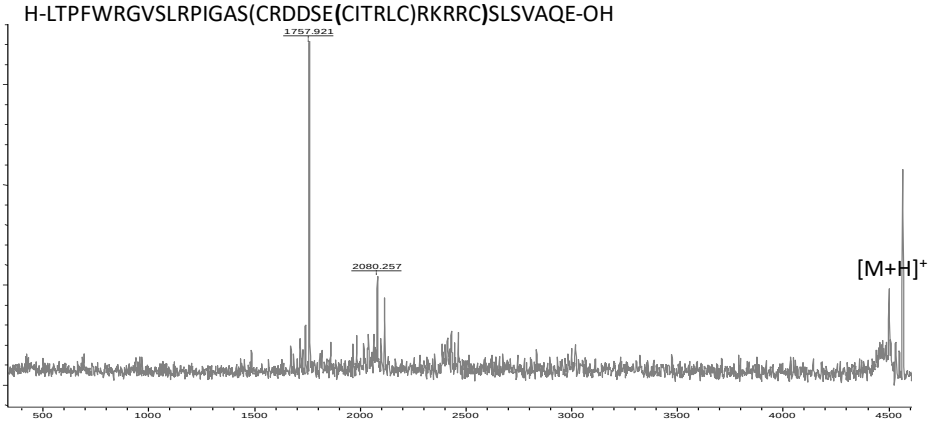
H-NleTPFWRGVSLRPIGAS*(*CRDDSE(CITRLC*)*RKRRC)SLSVAQE-OH)

Mol. Formula: $C_{192}H_{319}N_{64}O_{57}S_4$
MW= 4563.4 g.mol^{-1}
UV Purity: 95%

LC/MS data
RT: 0.90 min
MALDI(+) (SA linear mode): obs m/z 4564.2 Th $[M+H]^+$
ESI(+): obs m/z 571.5 Th $[M+8H]^{8+}$, 662.7 Th $[M+7H]^{7+}$, 761.5 Th $[M+6H]^{6+}$, 913.6 Th $[M+5H]^{5+}$, 1142.0 Th $[M+4H]^{4+}$



H-LTPFWRGVSLRPIGAS(CRDDSE**(**CITRLC)RKRRC**)**SLSVAQE-OH



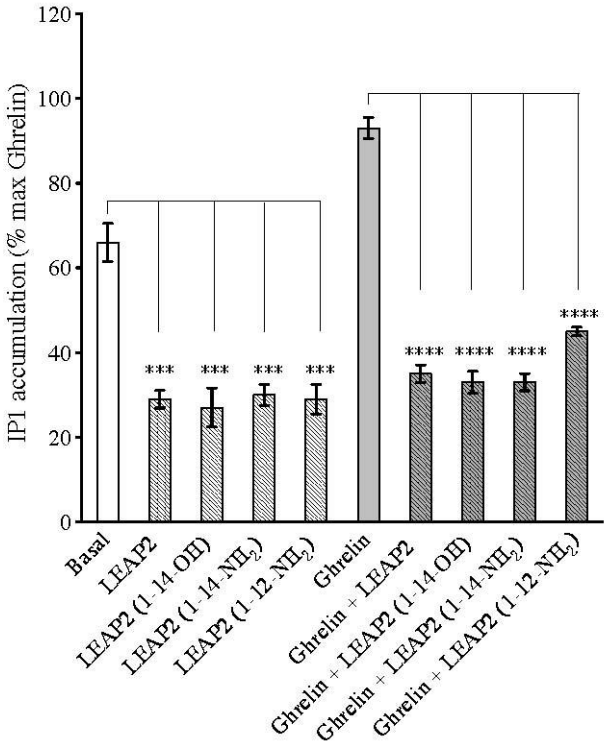


Figure S1. Activity of LEAP2 peptides on ghrelin-independent and ghrelin-dependent stimulation of IP1 production in HEK293T cells expressing GHSR. IP1 production was measured by IPone-HTRF in HEK293T cells expressing GHS-R1a stimulated with LEAP2 peptides 10^{-5}M, Ghrelin 10^{-8}M or Ghrelin 10^{-8}M + LEAP2 peptides 10^{-5}M. The data are mean ± SD from one experiment performed in triplicate. Statistical significance between basal and LEAP2-stimulated cells or between ghrelin and ghrelin + LEAP2 peptides was assessed using one-way ANOVA followed by Dunnett's post hoc test (****, $p < 0.0001$, ***, $p < 0.001$).

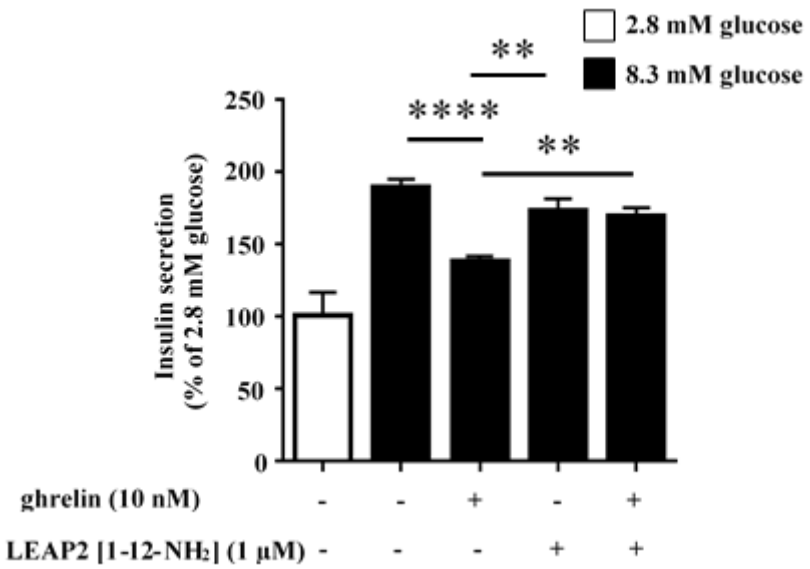


Figure S2: Effect of LEAP2 (1-12-NH_2) on ghrelin-induced inhibition of insulin secretion in rat pancreatic islets. Rat pancreatic islets were incubated under basal (2.8 mM glucose) or 8.3 mM glucose-stimulated conditions in the presence or absence of ghrelin (10 nM) and LEAP2 (1-12-NH_2) (1 µM). After 1h of incubation, the extracellular medium was collected and insulin concentration was quantified. The histogram depicts the mean ± SEM of insulin secretion expressed as the percentage of response obtained under basal condition. Two-way ANOVA, performed with 8.3 mM glucose data points, showed no effect of LEAP2 (1-12-NH_2) but showed an effect of ghrelin [$P = 0.001$] as well as an interaction between both factors [$P < 0.001$] . **** $p<0.0001$, ** $p<0.01$, Holm-Sidak multiple comparison test, n = 5.

Conclusions et perspectives

La cyclisation de la séquence CIA a conduit à l'obtention de peptides avec de fortes affinités pour le récepteur à l'acétylcholine nicotinique neuronal α3β2 et musculaire, permettant ainsi l'étude de diminution du TOF. Il serait intéressant de poursuivre ce travail avec des études de docking afin de déterminer quels changements dans les interactions avec le récepteur sont responsables du gain d'affinité, de plus de 50 fois sur le sous-type neuronal α3β2. Cela pourrait également permettre de donner des pistes pour la conception de ligands cycliques sélectifs de l'un ou de l'autre des récepteurs étudiés. L'influence de l'extrémité N-ter sur la séquence CIC apparait évidente pour la stabilisation de la structure d'après les données RMN, contre-intuitivement c'est le mutant ne possédant pas l'extrémité N-terminale qui apparait le plus désordonné en solution. De manière générale, le pro-peptide précèdant la toxine mature est suspecté de jouer en rôle dans l'étape de repliement du peptide. Ainsi, étant donné que cette extrémité N-terminale semble impliquée dans la stabilisation de la structure, il serait intéressant de déterminer si elle joue aussi un rôle de catalyseur dans le repliement du peptide. Les résultats des tests d'électrophysiologie permettront de déterminer si cette extrémité a également une influence sur l'activité. La mise en évidence du rôle essentiel de la partie N-terminale du peptide LEAP 2 pour l'activité sur le récepteur à la ghréline a permis deux autres études [494,495] dans lesquelles la partie N-terminale, moins contraignante à synthétiser, est utilisée comme un analogue du peptide LEAP-2.

Il est important de noter que malgré une formation des ponts disulfures quasiment totale en se basant sur les chromatogrammes LC-ESI/MS, les rendements totaux sont rarement supérieurs à 10%, cela est probablement dû à des pertes par adsorption sur la colonne lors des purifications. En effet il est préférable pour des quantités de produits de l'ordre de la dizaine de milligramme d'utiliser des colonnes de types semi-préparative avec des diamètres internes (3-5 mm) plus petit que la colonne préparative majoritairement utilisée dans notre laboratoire (40 mm). Récemment, Brik *et al.* [496] ont décrit la déprotection/oxydation concomitante des groupements Acm en présence de $PdCl_2$ et diethyldithiocarbamate, il serait intéressant de comparer si cette méthode donne des meilleurs rendements que la déprotection à l'iode.

Une grande partie des toxines comportant 3 ponts disulfures, la suite de ce travail a été l'élaboration de stratégies de synthèse pour la production de toxines à 3 ponts disulfures.

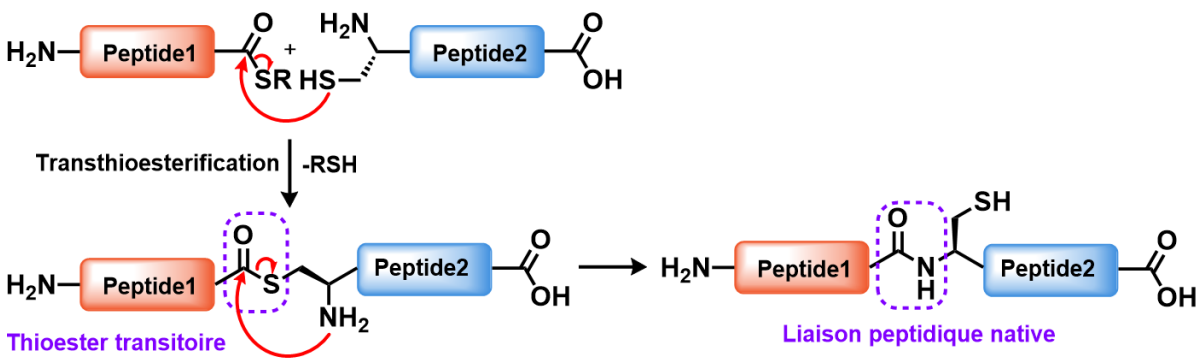
4. Chapitre 3 : Synthèse de toxines à 3 ponts disulfures

La stratégie de synthèse des peptides à deux ponts disulfures ayant été validée sur des peptides de différentes tailles, sur différents folds et même sur des peptides modifiés cycliques, la suite logique était de l'adapter à la synthèse de toxines comportant trois ponts disulfures. Celles-ci sont typiquement composées de 30 à 50 résidus et ciblent prioritairement les canaux ioniques. Leurs synthèses ont été réalisées par ligation chimique native et le folding a été réalisé par voie oxydante non dirigée quand les conditions le permettaient et par voie dirigée dans le cas contraire.

a. Introduction

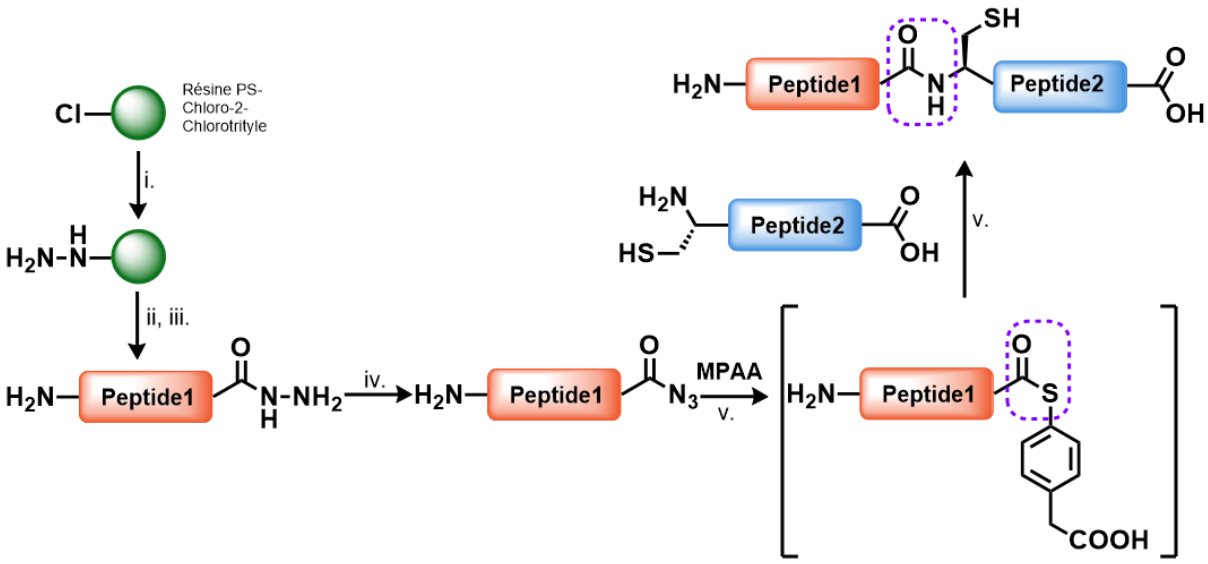
Le procédé de synthèse peptidique en phase solide et le couplage de fragments peptidiques protégés [497,498] ont considérablement allongé la taille des polypeptides accessibles par synthèse chimique. Cependant, la nécessité de surmonter les difficultés de manipulation des peptides protégés, qui sont notoirement difficiles à solubiliser dans des solvants organiques, ainsi que la recherche d'une complexité toujours croissante des protéines pouvant faire l'objet d'une synthèse chimique ont conduit à l'émergence de nouvelles stratégies de synthèse basées sur une chimie effectuée dans l'eau avec des segments peptidiques non protégés. Parmis ces techniques dites de ligation, la ligation chimique native (NCL) est une technique de choix pour la synthèse de produits naturels difficilement faisable en utilisant une stratégie linéaire classique de synthèse peptidique. Introduite par Kent et ses collaborateurs en 1994 [498], cette réaction implique une réaction de transthioesterification entre thioester peptidique C-terminal et un peptide cystéinyle N-terminal. Le thioester transitoire généré se réarrange spontanément par transfert intramoléculaire de S, N-acyle pour générer une liaison peptidique au site de ligation (Schéma 6). Il existe de nombreuses méthodes de synthèse de peptide thioester, ici seules les stratégies utilisées dans les travaux de ce manuscrit seront détaillées.

Schéma 6 : Principe de la ligation chimique native (NCL).



Dans leurs récents travaux, Fang *et al.* [499] ont constaté que les peptides hydrazides constituaient de bons précurseurs stables de thioester pour la NCL (Schéma 7). Grâce à une simple opération d'activation au nitrite de sodium et de thiolyse généralement via l'acide 4-mercaptophénylacétique (MPAA) [500], introduit sous forme de thiol externe à une concentration d'environ 100 mM, un peptide hydrazide peut être rapidement et proprement converti en un thioester peptidique sans épimérisation et ensuite être engagé dans une autre réaction de ligation sans purification. De plus, les peptides hydrazides ne réagissent pas dans la réaction de NCL à moins d'être activés par le nitrite de sodium. Ils peuvent donc être utilisés comme thioesters « masqués » dans la synthèse convergente de protéines. Dans ces études, la stratégie de synthèse implique généralement deux étapes clés : premièrement, la Fmoc-SPPS du peptide hydrazide en utilisant la résine 2-Cl-(Trt)-$NHNH_2$; deuxièmement, la NCL des peptides hydrazides par l'activation au nitrite de sodium et la thiolyse *in situ*.

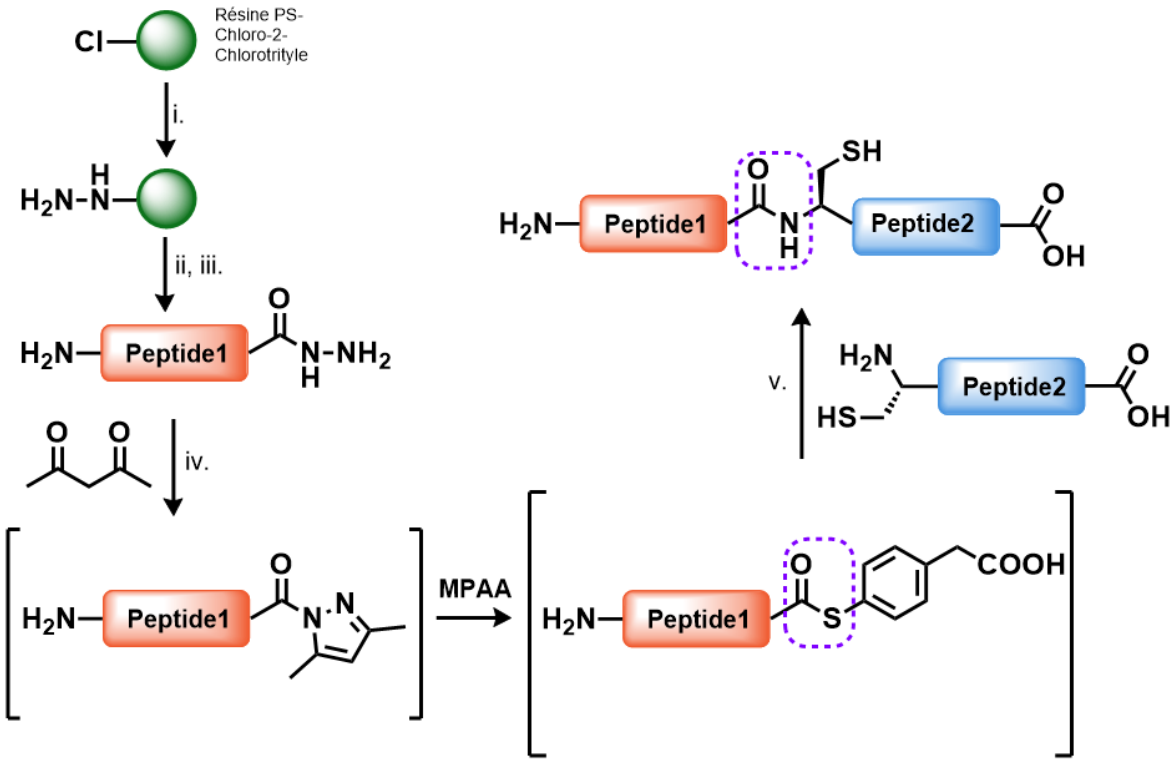
Schéma 7 : Méhode de NCL avec les peptides hydrazides décrite par Fang et al. [499]. (i) 5% NH_2-NH_2/DMF. (ii) Fmoc-SPPS. (iii) Clivage. (iv) $NaNO_2$ (10 éq), pH 3, -15°C (v) Peptide2 (1 éq), MPAA 100 éq, pH 6.8-7.



Récemment une autre méthode mettant en jeu les peptides hydrazides et revisitant la synthèse de Knorr des pyrazoles a été développée (Schéma 8). Cette méthode d'activation élégante est plus simple à mettre en place que la méthode d'activation des peptides hydrazides au nitrite de sodium explicitée précédemment. Flood *et al.* [501] ont montré que les peptides hydrazides peuvent être convertis efficacement en peptides thioesters par la formation intermédiaire de N-acyle pyrazoles qui sont des précurseurs activés de peptides thioester. Ces derniers sont produits *in situ* par une réaction de condensation mettant en jeu la fonction hydrazide et l'acétylacétone (acac) à pH acide, puis déplacés par l'additif thiol pour générer le thioester. Le criblage de divers thiols a donné les meilleurs résultats

pour la MPAA. Ce mode d'activation du peptide hydrazide diffère des autres approches existantes et fournit un processus chimiosélectif et doux pour générer des thioesters en l'absence de réactifs acylants ou oxydants.

Schéma 8 : NCL utisant les N-acyles pyrazoles décrite par Flood *et al.* [501]. (i) 5% NH_2-NH_2/DMF. (ii) Fmoc-SPPS. (iii) Clivage. (iv) 10% acac/H_2O (2.5 éq), MPAA (25 éq), pH 3 (v) Peptide2 (1 éq), pH 7-7.4.



b. Synthèse de la toxine U-Asilidin1-Mar1a

Dans le but de tester l'applicabilité de la méthode aux peptides hydrazides activés par le nitrite de sodium pour la synthèse de toxine à 3 ponts, nous nous sommes tout d'abord intéressé à la synthèse de la U-Asilidin1-Mar1a, un nouveau peptide isolé à partir du transcriptome des glandes salivaires d'un insecte prédateur. En effet celle-ci constituait un bon modèle car elle présente un motif de type « inhibitor cystine knot » (ICK) rencontré également dans de nombreuses autres toxines à 3 ponts disulfures.

Cette étude est présentée ci après sous forme d'une publication parue [502] dans laquelle j'ai effectué la synthèse par ligation chimique native (NCL) de la toxine U-Asilidin1-Mar1a et sa caractérisation par spectrométrie de masse.

A Dipteran's Novel Sucker Punch: Evolution of Arthropod Atypical Venom with a Neurotoxic Component in Robber Flies (Asilidae, Diptera)

Stephan Holger Drukewitz [1,*], Nico Fuhrmann [2], Eivind A. B. Undheim [3], Alexander Blanke [4,5], Julien Giribaldi [6], Rosanna Mary [6], Guillaume Laconde [6], Sébastien Dutertre [6] and Björn Marcus von Reumont [1,7,*]

[1] *Institute for Biology, University of Leipzig, Talstr. 33, 04103 Leipzig, Germany*
[2] *Max Planck Institute for Evolutionary Biology, August-Thienemann-Str. 2, 24306 Plön, Germany; fuhrmann@evolbio.mpg.de*
[3] *Centre for Advanced Imaging, The University of Queensland, St. Lucia, Brisbane, QLD 4072, Australia; e.undheim@uq.edu.au*
[4] *Institute for Zoology, Biocenter, University of Cologne, Zuelpicher Str. 47b, 50674 Cologne, Germany; a.blanke@uni-koeln.de*
[5] *Medical and Biological Engineering Research Group, School of Engineering and Computer Science, University of Hull, Hull HU6 7RX, UK*
[6] *Institute for Biomolecules Max Mousseron, UMR 5247, University of Montpellier—CNRS, Place Eugène Bataillon, 34095 Montpellier Cedex 5, France; julien.giribaldi@umontpellier.fr (J.G.); rosanna.mary@umontpellier.fr (R.M.); guillaume.laconde@umontpellier.fr (G.L.); sebastien.dutertre@umontpellier.fr (S.D.)*
[7] *Department of Life Sciences, Natural History Museum, Cromwell Rd, London SW7 5BD, UK*
* *Correspondence: bmvr@arcor.de (B.M.v.R.); stephan.drukewitz@uni-leipzig.de (S.H.D.); Tel.: +49-034-1973-6742 (B.M.v.R.); +49-034-1973-6726 (S.H.D.)*

Abstract: Predatory robber flies (Diptera, Asilidae) have been suspected to be venomous due to their ability to overpower well-defended prey. However, details of their venom composition and toxin arsenal remained unknown. Here, we provide a detailed characterization of the venom system of robber flies through the application of comparative transcriptomics, proteomics and functional morphology. Our results reveal asilid venoms to be dominated by peptides and non-enzymatic proteins, and that the majority of components in the crude venom is represented by just ten toxin families, which we have named Asilidin1–10. Contrary to what might be expected for a liquid-feeding predator, the venoms of robber flies appear to be rich in novel peptides, rather than enzymes with a putative pre-digestive role. The novelty of these peptides suggests that the robber fly venom system evolved independently from hematophagous dipterans and other pancrustaceans. Indeed, six Asilidins match no other venom proteins, while three represent known examples of peptide scaffolds convergently recruited to a toxic function. Of these, members of Asilidin1 closely resemble cysteine inhibitor knot peptides (ICK), of which neurotoxic variants occur in cone snails, assassin bugs, scorpions and spiders. Synthesis of one of these putative ICKs, U-Asilidin1-Mar1a, followed by toxicity assays against an ecologically relevant prey model revealed that one of these likely plays a role as a neurotoxin

involved in the immobilization of prey. Our results are fundamental to address these insights further and to understand processes that drive venom evolution in dipterans as well as other arthropods.

Keywords: Asilidae; neurotoxins; cysteine inhibitor knot peptide; arthropod venom evolution; functional morphology; synchrotron micro computed tomography; Asilidin

Key Contribution: This study provides the first comprehensive description of the venom system of two robber flies (Asilidae). We reveal a complex venom apparatus and an unusual, enzyme depleted venom with unique proteins, including also a new, neurotoxic ICK peptide.

1. Introduction

Venoms are key adaptations that have evolved on numerous occasions in animal lineages to serve a range of ecological roles including defense, predation, communication and competition [503–505]. Venoms constitute complex cocktails of proteins, peptides, salts and different organic molecules, collectively referred to as toxins [503,506]. These toxins are expressed in venom glands and delivered via venom ducts to structures that finally inject the composed venom from the venom delivery system via a wound into the prey. Proteins and peptides usually comprise the main venom components, and these have evolved from ancestral molecules with basic physiological ”every-day” functions into highly potent and chemically stable toxins [504,507]. However, the processes that drive this functional transition are understood only fragmentary [508–510]. Moreover, the majority of all known toxins from animal venoms can be classified into a limited number of structural classes [511]. This extreme level of convergence of venoms implies that toxins can provide insights into fundamental processes of protein functional evolution and biochemical adaptations.

Understanding convergent evolution requires a broad taxonomic sampling to accurately identify adaptive traits through comparative evolutionary studies [512]. However, only a few venomous lineages have so far been studied in-depth. Although a number of animal lineages previously considered potentially venomous have been described in recent years [505,513–516], the taxonomic sampling remains fragmented, particularly among venomous insect lineages. One of these venomous insect groups are robber or assassin flies (Asilidae), which are a species-rich family within one of the largest and most diverse insect groups, the flies (Diptera) [517,518]. Assassin flies have a worldwide distribution (except Antarctica) and comprise more than 500 genera including over 7000 species [519].

In contrast to most other dipterans, asilids are also predatory in their adult life stage [519], and have been known since the 19th century [520,521] as major predators of other arthropods including flies, beetles, grasshoppers, dragonflies, hymenopterans, and even spiders [522–525]. Their ecological and

economic impact as predators is illustrated by their ability to significantly affect grasshopper populations, while other species that feed on wasps and bees are known to take out entire beehives [526,527]. Asilids are adapted to a predatory lifestyle on the wing with their slender but robust body in combination with a heavily sclerotized proboscis and a needle-like hypopharynx, large eyes and good flight capabilities [528,529]. The predation on larger or defensive prey, which is almost instantly paralyzed, led to early suspicions that robber flies utilize some kind of venom to overpower venomous or larger prey [520,521,530].

Early experiments by Whitfield showed that grasshoppers attacked by robber flies indeed died dramatically faster compared to grasshoppers stabbed in a similar way with needles [530]. In the same study, Whitfield also described two separate gland systems consisting of a pair of smaller labial glands and a pair of thoracic glands for the asilid *Machimus atricapillus*. Kahan (1964) tested for the proteolytic activity of extracts from the thoracic and labial glands and compared these to the activity of the stomach content of *Promachus griseiventris* and the stomach only of *Philonicus dorsiger* by injection experiments. These experiments concluded that neurotoxic activity was present in gland extracts due their paralyzing effects on the legs of locusts and mice. However, although these early experiments demonstrated toxic effects of asilid venoms, the general composition of the venom and the mechanisms by which it is secreted has remained unknown [520,531].

In the present study, we apply a combined transcriptomic and proteomic approach to provide detailed insight into the composition of the venom of two common European asilid species (*Eutolmus rufibarbis* and *Machimus arthriticus*) (Figure S1 (Supplementary File 1)). We also use high-resolution synchrotron micro computed tomography (SRμCT) to provide a characterization of the morphology of the venom system of asilids. Our results reveal that asilid venoms are not the protease-dominated venoms expected from liquid-feeding lineages, but that they are instead peptide-rich cocktails originating in the thoracic glands and expelled through an elaborate high-speed venom delivery system. Furthermore, we show that at least one of these peptides indeed plays a neurotoxic role. This peptide toxin likely assumes a fold common to spider and cone snail neurotoxins, illustrating the value of virtually unknown venomous lineages in identifying molecular adaptations through evolution-based structure-function relationships. Our results thus provide the foundation for understanding venom evolution in not only flies but also other venomous groups of insects through comparative studies. The high novelty of the putative toxins in robber fly venoms also highlights their potential as sources of new therapeutic and agrochemical approaches [532,533].

2. Materials and Methods

2.1. Specimen Collection and Determination

In total 40 individuals of common robber fly species *Eutolmus rufibarbis* (25 specimens) and *Machimus arthriticus* (15 specimens) were collected in 2014 near Altenrath, Germany, (Figure S1 (Supplementary File 1)). Species were morphologically determined using the identification key by Fritz Geller-Grimm [534]. For each species one voucher sample was stored in 96% Ethanol. Additionally, barcodes of the mitochondrial cytochrome c oxidase subunit I (COI) were sequenced for all specimens except for two in Bouin solution sampled individuals (Supplementary Figure S9). DNA was extracted following standard procedure with the NucleoSpin® Tissue Kit (MACHERY-NAGEL, Düren, Germany). PCR was performed with the primers LCO1490F and HCO2198R [57]. PCR products were cleaned with the NucleoSpin® Gel and PCR Clean-up Kit (MACHERY-NAGEL, Düren, Germany). Sequencing was performed at GATC Biotech AG (Konstanz, Germany). Barcode sequences are accessible in GeneBank (NCBI) with accession numbers KY485001–KY485038.

2.2. Specimen Dissection and Sample Preservation

For transcriptome sequencing thoracic glands of 6 *Machimus arthriticus* and 10 *Eutolmus rufibarbis* specimens were immediately dissected on ice in RNAse free TBE buffer and water. From three of those individuals, for each species body tissue samples (muscle tissue) were preserved to analyze complementary paralog or ancestral variants of venom proteins from body tissue. All samples were stored in RNAlater at −4 °C. To preserve crude venom for proteomics analyses 8 glands from *Machimus arthriticus* and 12 glands of *Eutolmus rufibarbis* were dissected on ice and squeezed out in sterile proteinase inhibitor buffer. The proteinase inhibitor buffer was prepared following the manufacturer protocol (Complete Ultra, ROCHE, Mannheim, Germany). The venom extracts were then lyophilized and stored at −20 °C until proteomic analysis. Remaining parts of all dissected specimens were preserved as vouchers in 96% ethanol. Two specimens of the more frequent *Eutolmus rufibarbis* were stored in Bouin solution and critical point dried for later synchrotron-based micro-computer tomography (SRμCT) reconstruction.

2.3. RNA Extraction, Transcriptome Sequencing and Assembly

RNA-Extraction (Trizol method), construction of cDNA libraries (Illumina TruSeq kit, San Diego, CA, USA and sequencing was performed at the BGI Sequencing facility (Beijing, China), using the Illumina® HiSeq 2000 platform with 100 bp paired end. Gland tissue samples were sequenced on one-third, body tissue samples on one-fourth Illumina lane. All data of venom gland and body tissue transcriptomes are accessible in GeneBank under the BioProject PRJNA361480, SRA accession numbers: SRR5185499, SRR5185498, SRR5185497, SRR5185496 and TSA entries: GFGA00000000, GFFZ00000000, GFZR00000000, GFZQ00000000.

Raw reads were pre-processed and quality checked before assembly. First, all raw reads were visually inspected to check for overall quality and for overexpressed sequences using FastQC v0.11.2 [535]. Afterwards, Trimmomatic v0.33 [536] was applied to exclude low quality reads below a phred value of 32 (sliding window size 4, HEADCROP 10, minimum length 50 bp). A modified and extended template file with known adapter and vector sequences was used to screen and to remove adapter and vector contaminations. Reads were assembled using Trinity v2.0.2 [537] applying standard settings, except for a transcript minimum length of 101.

2.4. Assessing Coding Regions and Expression Levels for Transcripts

Coding regions within the assembled transcripts were identified with TransDecoder v3.0.1 [538]. To include shorter putative toxins such as neurotoxins, the minimal open reading frame length was set to 40 amino acids. BlastP v.2.4 search against the UniProt database [539] and HMMscan v.3.1 [540] against the Pfam database [541] were performed to provide additional information for the identification of potential protein coding regions. In the case that several open reading frames per transcript are equally likely, TransDecoder retains all open reading frames with the same likelihood for this transcript.

To assess the expression level or abundance of the identified coding regions and transcripts, two slightly different approaches were chosen, first the read mapper Segemehl v.0.2.0 [542], and second the RNASeq quantification tool Kallisto v.0.43.1 [543]. Both approaches result in a normalized expression value for all transcripts. Segemehl was used to map the reads back to the coding region predicted with TransDecoder. In contrast to other read mappers, segemehl reports multiple hits for a read if several alignments are equally likely and fulfill the set alignment parameters. This strategy in combination with the restriction to map reads only for identified coding regions gives the most accurate approximation of expression levels for putative toxins. Segemehl was applied with 95 percent accuracy and the default setting for all other options.

After mapping, the number of mapped reads was normalized to account for the length of the coding region and for the number of overall mapped reads. All coding regions with a TPM smaller than 1 were excluded for subsequent analyses, similar to [544,545]. As an alternative approach to calculate the abundance of transcripts, the RNA-seq quantification program Kallisto was used (Table 1). As an input, the processed paired end libraries were provided. Kallisto calculates the normalized expression value TPM, which takes the library size and the effective length of a transcript into account. Is it important to note, that Kallisto uses an approximation of effective length of every transcript for the calculation of the TPM, while in the Segemehl approach the complete length of the coding region is used for length normalization. To minimize the false positive aligned reads, the bootstrap value was set to 100.

2.5. Identification of Venom Protein Classes and Putative Toxins via Transcriptomics

To search unspecific for protein families of putative toxins and noticeable high expressed proteins in the toracic glandsystem, BlastP searches (*e*-value 10-4) against the Toxprot Database [546] were carried out with the protein sequences of the coding regions. Additionally, hmmer 3 [540] was used to perform hmm searches with an own, customized and hand curated alignment database of over 40 known venom proteins (Supplementary Files 2–5). The alignments to train hmm models were compiled with sequences from the UniProt database, covering non-venomous and venomous species, following [505,515]. As a cut-off value a bit score of 20 was chosen, potential homologs with a lower bit score were excluded. Finally, the TPM values were linked to all identified coding regions that match venom protein classes or putative toxins for body tissue and putative venom gland tissue (Supplementary Files 2–5). Identified sequences were extracted by a customized python script and manually optimized and curated in Geneious R9. Finally, sequences were annotated with the InterProScan v.1.1.0 Plug-in Genieous R9 [547]. For all transcript sequences the presence of a signal peptide was tested using SignalP 4.1 [548]. This search strategy was performed for both venom gland and body tissue transcriptomes. Identical but also venom gland unique transcripts in body and venom gland tissue were identified using the CD-HIT-2D tool (sequence identity 0.99) of the CD-HIT software package v.4.6 [549] and additionally by visual inspection of alignments in Geneious R9 (Figure 9 and Supplementary File 7).

2.6. Sample and Data Processing for Proteome Analyses

Lyophilized venom in protease inhibitor was dissolved in water to a concentration of 1 mg/mL by repeated pipetting, before 20 µg protein was precipitated by addition of 5:1 volume ratio of −20 °C acetone, incubation for 1 h at −20 °C, and centrifugation at 20,000 rcf for 20 min at 0 °C. The acetone was then removed, the pellet washed (ice cold absolute ethanol), centrifuged, and the ethanol removed (pipetting and air drying). Protein was then dissolved to a concentration of 5 mg/mL in 4 M urea, 15% acetonitrile (ACN), 100 mM ammonium bicarbonate, before cystines were reduced by incubation with 5 mM dithiothreitol at 60 °C for 5 min and then alkylated with 10 mM iodoacetamide at 37 °C for 60 min. The reduced and alkylated sample was then digested by incubating with 20 ng/µL trypsin overnight at 37 °C in 2 M urea 10% ACN 100 mM ammonium bicarbonate. The digested sample was desalted using a C18 ZipTip (Thermo Fisher, Waltham, MA, USA) and dried prior to LC-MS/MS analysis using a vacuum centrifuge.

The digested protein was dissolved in 0.5% formic acid (FA) and 2 µg analyzed on an AB Sciex 5600 TripleTOF equipped with a Turbo-V source heated to 550 °C. Venom was fractionated on a Shimadzu (Kyoto, Japan) Nexera UHPLC with an Agilent Zorbax stable-bond C18 column (2.1 × 100 mm, 1.8 µm particle size, 300 Å pore size), using a flow rate of 180 µL/min and a gradient of 1–40% solvent B (90%

ACN 0.1% FA) in 0.1% FA over 60 min. MS1 survey scans were acquired at 300–1800 *m*/*z* over 250 ms, and the 20 most intense ions with a charge of +2 to +5 and an intensity of at least 120 counts/s were selected for MS2. The unit mass precursor ion inclusion window mas ± 0.7 Da, and isotopes within ±2 Da were excluded from MS2, which scans were acquired at 80–1400 *m*/*z* over 100 ms and optimized for high resolution.

For protein identification, MS/MS spectra were searched against the translated combined venom gland and body transcriptomes using ProteinPilot v5.0 (AB Sciex, Framingham, MA, USA). Searches were run as thorough identification searches, specifying urea denaturation, tryptic digestion and cysteine alkylation by iodoacetamide. Amino acid substitutions and biological modifications were allowed in order to identify potential post translational modifications and to account for chemical modifications due to experimental artefacts. Decoy-based false discovery rates (FDR) was estimated by ProteinPilot v5.0, and only protein identifications with a corresponding local FDR of <0.5% were considered significant. Spectra were also examined manually to further eliminate any false positives.

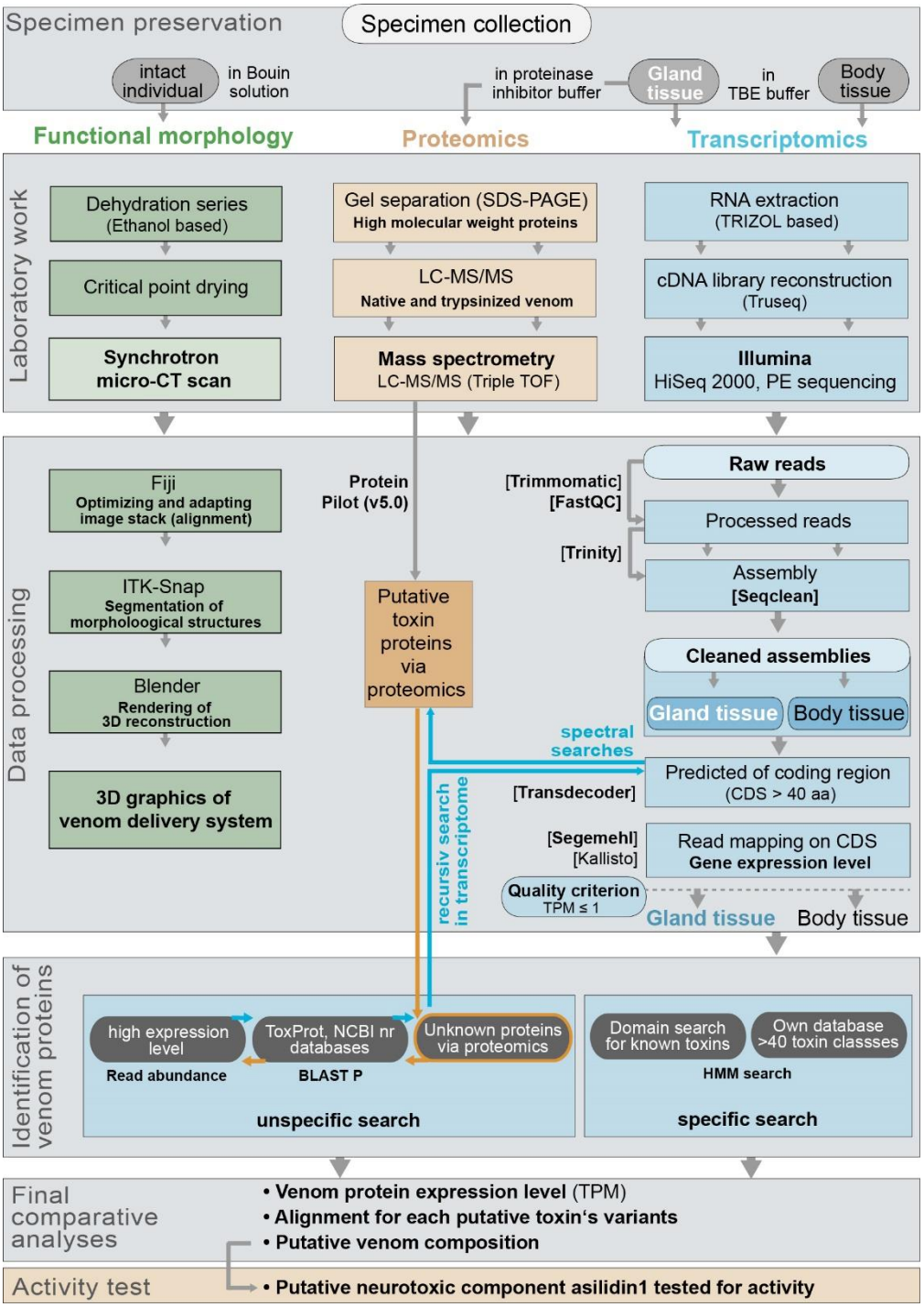


Figure 9. Workflow that was developed and applied in this study.

2.7. Chemical Synthesis of U-Asilidin$_1$- Mar1a

All solvents used were of HPLC grade. DMF, DIPEA, ACN, TIS, TFA, piperidine and all other reagents were obtained from Sigma-Aldrich or Merck and were used without further purification. Fmoc amino acid derivatives, HATU and 2-Chorotrityl chloride resin (100–200 Mesh, 1.44 mmol Cl/g resin) were purchased from Iris Biotech (Marktredwitz, Germany). AmphiSpheres 20 HMP resin (0.6 mmol/g 75–150 µm) was purchased from Agilent Technologies (Les Ulis, France). Peptides were manually

synthesized using Fmoc-based solid-phase peptide synthesis (SPPS). All Fmoc amino acids and HATU were dissolved in N,N-dimethylformamide (DMF) to reach 0.5 M. Fmoc deprotection was performed with 20% piperidine in DMF twice (1 min at room temperature), then the resin was washed three times with DMF. Each Fmoc-protected amino acid (5 eq) was coupled twice in the presence of HATU (5 eq) and DIPEA (10 eq) in DMF at room temperature for two min. After completion of the coupling reaction, the resin was sequentially washed twice with DMF. Cleavage of peptide from the resin and removal of side-chain protecting groups were carried out using 10 mL of a solution containing TFA/triisopropylsilane/H_2O (95:2.5:2.5, *v*/*v*/*v*) for 2 h 30 min at room temperature. Then, the resin was removed by filtration and washed three times with DCM. DCM and TFA were removed under vacuum and cold diethyl ether was added to precipitate the peptide, followed by two steps of centrifugation and removal of the supernatant. Crude peptide was purified by preparative RP-HPLC, and their purity were confirmed by LC/ESI-MS. Monoisotopic mass of folded Mar1a was confirmed by MALDI-HRMS analysis.

2-Cl-(Trt)-$NHNH_2$ was prepared from 2-Chorotrityl chloride resin (100–200 Mesh, 1.44 mmol Cl/g resin) [550]. Segment 1 of the Mar1a peptide was synthetized on 250 mg of 2-Cl-(Trt)-$NHNH_2$ resin, whereas segment 2 of the Mar1a peptide was synthetized on 250 mg of HMPA-Pro-Fmoc resin. After cleavage from the resin and treatments, peptides were purified by preparative RP-HPLC. Elution fractions of the targeted peaks were pooled and lyophilized. The homogeneity of purified peptides was assessed by analytical RP-HPLC and ESI-MS confirmed the correct mass with an observed *m*/*z* of 819.5 Da for $[M + 2H]^{2+}$ (calculated *m*/*z* 819.33 Da) for segment 1 and an observed *m*/*z* of 745.8 Da for $[M + 2H]^{2+}$ (calculated *m*/*z* 745.77 Da), see Figure 1 activity. Overall, 24.4 mg of a white solid were obtained for the first segment (yield of 7.46%) and 12 mg for the second (yield = 5.37%) (Figure S10 (Supplementary File 1)).

Native Chemical Ligation between segment 1 and segment 2 was carried out with 9.83 mg (6 µmol) of segment 1 and 8.97 mg (6 µmol) of segment 2 [550]. After formation of a single major product, the peptide was purified by preparative RP-HPLC. Eluted fractions of the targeted mass were pooled and lyophilized. The homogeneity of purified peptide was assessed by analytical RP-HPLC and ESI-MS confirmed the correct mass with an observed *m*/*z* of 1033.6 Da for $[M + 3H]^{3+}$ (calculated *m*/*z* 1033.12 Da). 4 mg (yield = 21.53%) of a white solid were obtained (Figure S11 (Supplementary File 1)).

Folding of linear Mar1a peptide (3 mg, 0.97 µmol, 1 eq) was carried out by stirring in the presence of GSH (5.96 mg, 19.39 µmol, 20 eq) and GSSG (1.19 mg, 1.94 µmol, 2 eq) into a solution of 0.1 M Tris-HCl (pH 8.0, 19.41 mL) at room temperature for 16 h. After formation of a single major product, the peptide was purified by preparative RP-HPLC. Elution fraction of the targeted peak was pooled and lyophilized. The homogeneity of purified peptide was assessed by analytical RP-HPLC. ESI-MS: observed *m*/*z* 1031 for $[M + 3H]^{3+}$ (calcd *m*/*z* 1031.12). 1.58 mg (yield = 52.67%) of a white solid were

obtained. MALDI-HRMS: observed *m*/*z* 3089.12 (monoisotopic mass) for $[M + H]^+$ (calcd *m*/*z* 3089.14) (Figure S12 (Supplementary File 1)).

2.8. HPLC Purification and Activity Tests of U-Asilidin$_1$- Mar1a 1a with Bioassay

Preparative RP-HPLC was run on a Gilson PLC 2250 Purification system (Villiers le bel, France) instrument using a preparative column (Waters DeltaPak C18 Radial-Pak Cartridge, 100 Å, 40 × 100 mm, 15 µm particle size, flow rate 50.0 mL/min). Buffer A was 0.1% TFA in water, and buffer B was 0.1% TFA in acetonitrile. Analytical analyses were carried out using a Chromolith (Fontenay sous Bois, France) Flash 25 × 4.6 mm C18 reversed-phase column. A flow rate of 3 mL/min and a gradient of 0–100% B over 2.5 min were used. Eluent A: water/0.1% HCO_2H; eluent B: acetonitrile/0.1% FA. UV detection was performed at 214 nm.

Domestic honeybees (*Apis mellifera*) were caught at the hive entrance and placed at 4 °C prior to injection. 30 µM Mar1a, 100 µM ω-atracotoxin (positive control) or MilliQ water (negative control) were injected into the median ocellus. To perform the injection, the lens of the median ocellus was perforated with a pulled borosilicate-glass capillary. Using a second capillary, 1 µL of toxin or control solution was aspirated and 200 nL were injected per bee (5 individual per experiment). Following the injection, bees were placed in a petri dish containing a small cup filled with water + honey solution. Behavioral observations were recorded for 60 min and typical toxin effects included slow movements, disorientation and paralysis. The results were analyzed with Prism (GraphPad, San Diego, CA, USA) using a one-way analysis of variance (ANOVA) (Figure 6, Supplementary File 6).

2.9. Morphological Work and Data Processing

The anatomy of specimens was investigated using synchrotron micro-Computer Tomography (SRµCT) [546]. Prior to scanning, samples were critical point dried (Samdri-PVT-3D) and mounted on beamline-specific specimen holders. Generally, X-ray imaging has a high penetrating power and allows visualization of large specimens without the need for sectioning. SRµCT offers a true 3D spatial resolution of up to 1 µm with moderate resolving power of tissues and tissue interfaces. Specimens were scanned at the Swiss Light Source electron synchrotron accelerator (SLS, Villigen, Switzerland; Stampanoni et al., 2010). The SLS X-ray sources were optimized for high-density and spatial resolution (1–10 um) imaging with monochromatic X-rays. The raw data are available upon request from the corresponding author.

Segmentation and rendering of single structures was carried out using ITK-snap v3.6.0 [551] and Blender v2.78 [552]. Both software packages are distributed under the General Public License. The terminology for asilid musculature and other described structures is based on Owsley (1946).

3. Results

3.1. Assembled Transcripts and Numbers of Assigned Coding Regions

In order to identify both putative toxins and their ancestral “house-keeping” variants, transcriptomes of body and venom gland tissue were sequenced on Illumina HiSeq 2000 platforms for both species, with gland tissue on 1/3 and body tissue on 1/4 Illumina channel. The resulting libraries showed almost equal magnitudes of raw and processed read numbers between the two different tissue types (Table 1). The contig numbers of cleaned assemblies vary slightly between species but show a larger number of contig sequences for the thoracic gland of *Machimus arthriticus* compared to *Eutolmus rufibarbis*.

The expression levels of transcripts based on coding regions identified with Transdecoder were assessed with two different methods, the read mapper Segemehl [542] and the RNA quantification tool Kallisto [543]. To prevent over-interpretation of our data, only coding regions that show a TPM (transcripts per million kilobase) value larger or equal 1 were included in our subsequent analyses. Both methods found overall similar results, which shows the robustness of our analyses considering that both methods use different approaches (Supplementary File 2–5). However, for our data interpretation we rely on Segemehl as an established and precise read mapper [553] (Material and Methods).

Table 1. Library size (numbers raw and processed reads), assembled contig sequences and numbers of identified coding regions for the analyzed species *Eutolmus rufibarbis* and *Machimus arthriticus.* Library size (processed reads) refers to the read numbers that survive the processing (trimming, adapter removing and filtering of quality scores) before they were assembled with Trinity [537]. The final contig numbers are given (contigs of cleaned assembly) for each transcriptome assembly after screening and trimming the contigs versus remaining adaptor and contaminating sequences (VecScreen), see also material and methods (Figure 9).

Species	Tissue	Library Size (Raw Reads)	Library Size (Processed Reads)	Contigs of Cleaned Assembly	Extracted Coding Regions	Coding Regions (TPM ≥ 1, Kallisto)	Coding Regions (TPM ≥ 1, Segemehl)
Eutolmus rufibarbis	Thoracic gland	108,632,880	87,187,856	56,640	33,218	16,049	15,523
	Body tissue	64,751,420	50,968,759	70,281	42,919	32,920	28,827
Machimus arthriticus	Thoracic gland	106,668,752	83,421,201	69,849	41,816	17,798	15,346
	Body tissue	64,651,716	44,208,096	67,504	42,784	34,478	30,629

3.2. Transcriptomic and Proteomic Profiles Differ in Thoracic Glands

Our proteomic results revealed a major discrepancy between predicted venom components based on homology to other known toxins and relative expression levels of transcripts (transcriptomes) compared to the observed secreted proteins (proteomes) in the thoracic glands. For both species, the total numbers clearly differ between identified highly expressed or venom gland unique transcripts (16 for *E. rufibarbis* and 17 for *M. arthriticus*) and the number of finally secreted proteins in the crude venom cocktail. Surprisingly, the crude venoms of the two species also differed with regards to the proteomically detected putative toxin families. Only eight groups of secreted proteins were detected in the crude venom of *E. rufibarbis*, compared to 13 for *M. arthriticus* (Figures 1–3, Figure S2 (Supplementary File 1)).

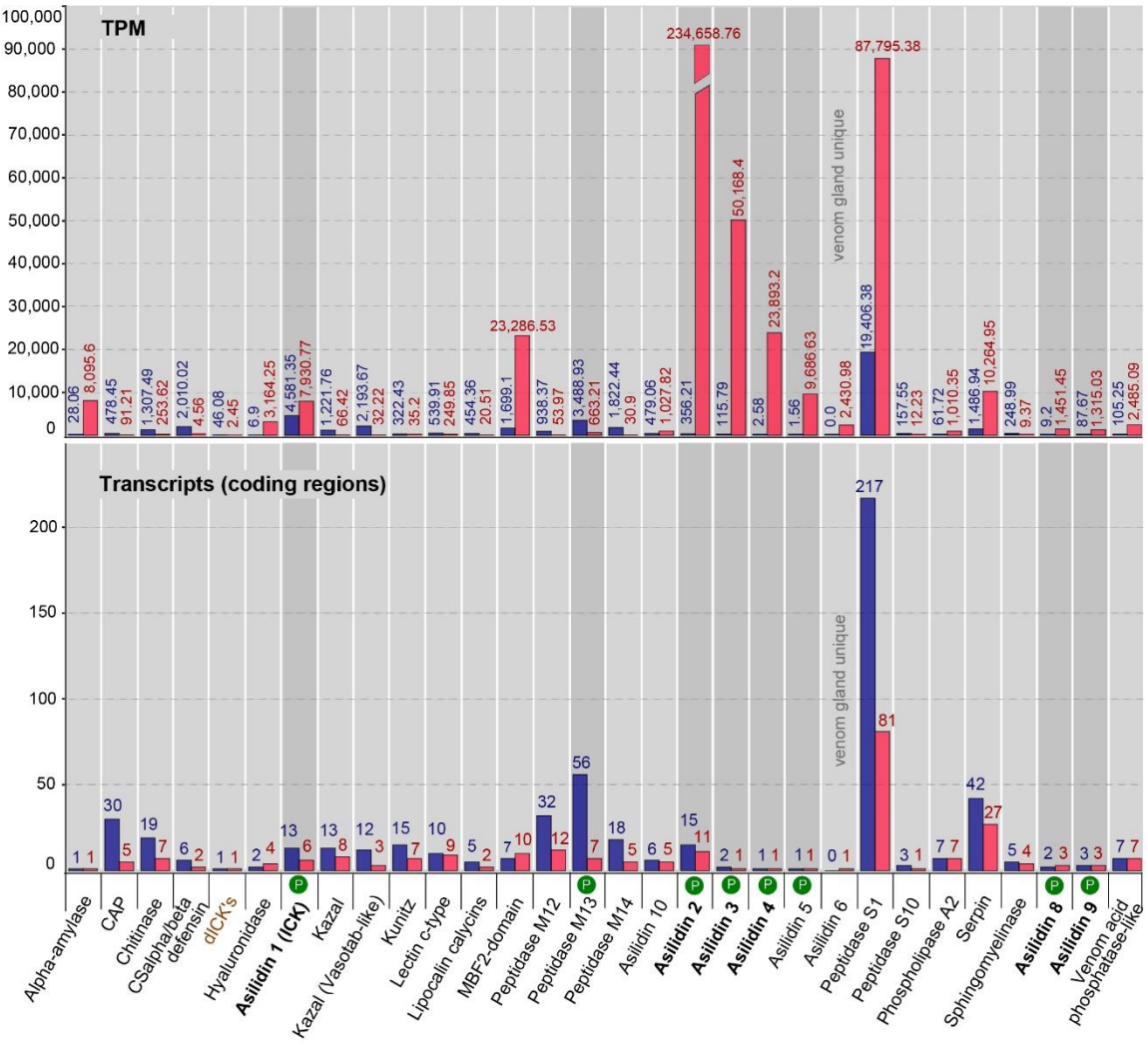


Figure 1. Comparative bar charts of 30 putative toxin protein classes from body and thoracic gland tissue from *Eutolmus rufibarbis*. Results from transcriptomics and proteomics (P) are summarized. The lower chart shows on the Y-axis the number of coding regions per protein class that passed the expression threshold. The relative expression level of each protein class in transcripts per million (TPM) in the putative venom gland (red) and in the body tissue (blue), is shown on the Y-axis in the upper part of the chart. The presence of the protein family in the proteome of the gland is marked with a white P in the green circle. Proteins that are present in transcriptomes and proteomes of both species are printed in bold, the two species unique protein classes are colored in brown.

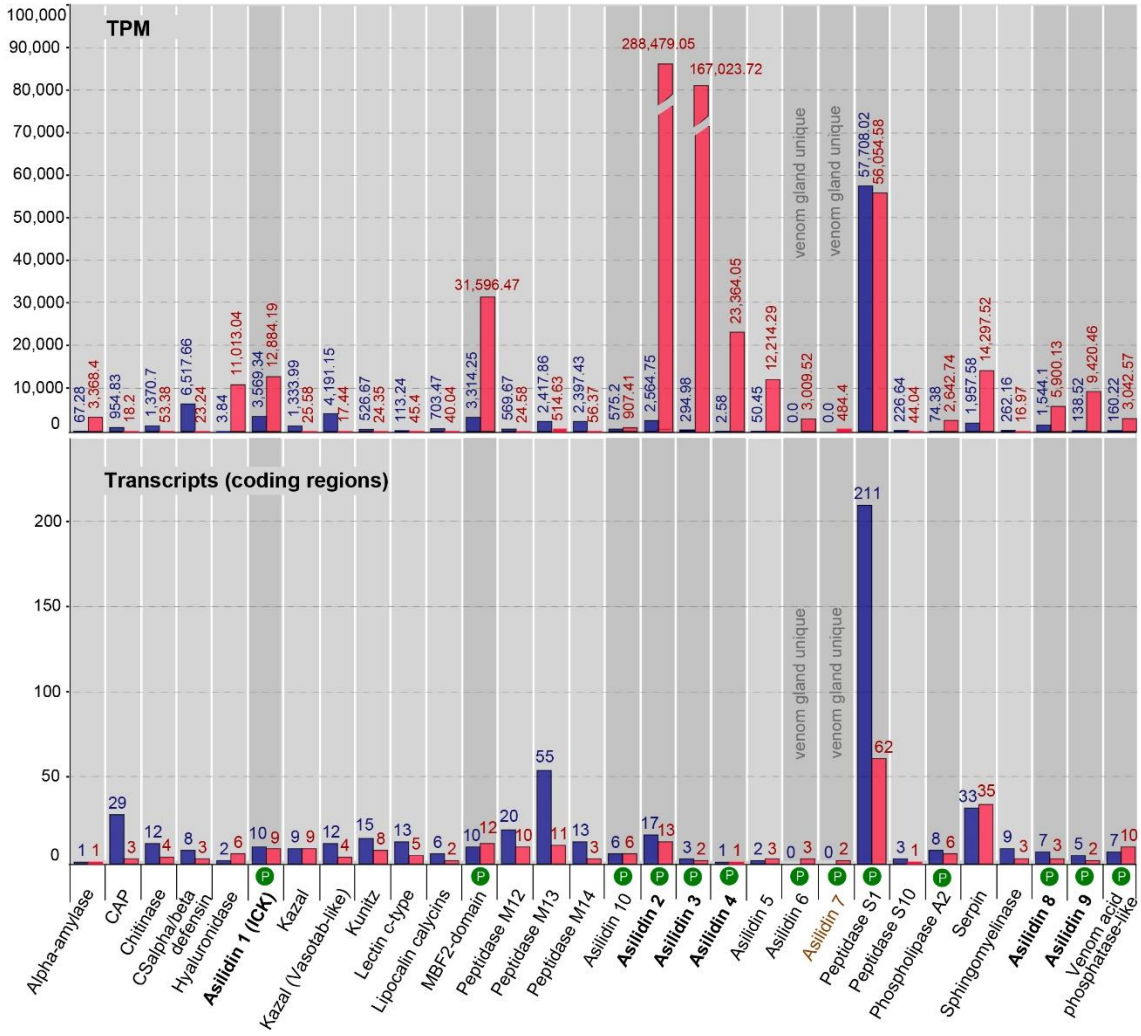


Figure 2. Comparative bar charts of 30 putative toxin protein classes from body and thoracic gland tissue from *Machimus arthriticus*. Results from transcriptomics and proteomics (P) are summarized. The lower chart shows on the Y-axis the number of coding regions per protein class that passed the expression threshold. The relative expression level of each protein class in transcripts per million (TPM) in the putative venom gland (red) and in the body tissue (blue), is shown on the Y-axis in the upper part of the chart. The presence of the protein family in the proteome of the gland is marked with a white P in the red circle. Proteins that are present in transcriptomes and proteomes of both species are printed in bold, the two species unique protein classes are colored in brown.

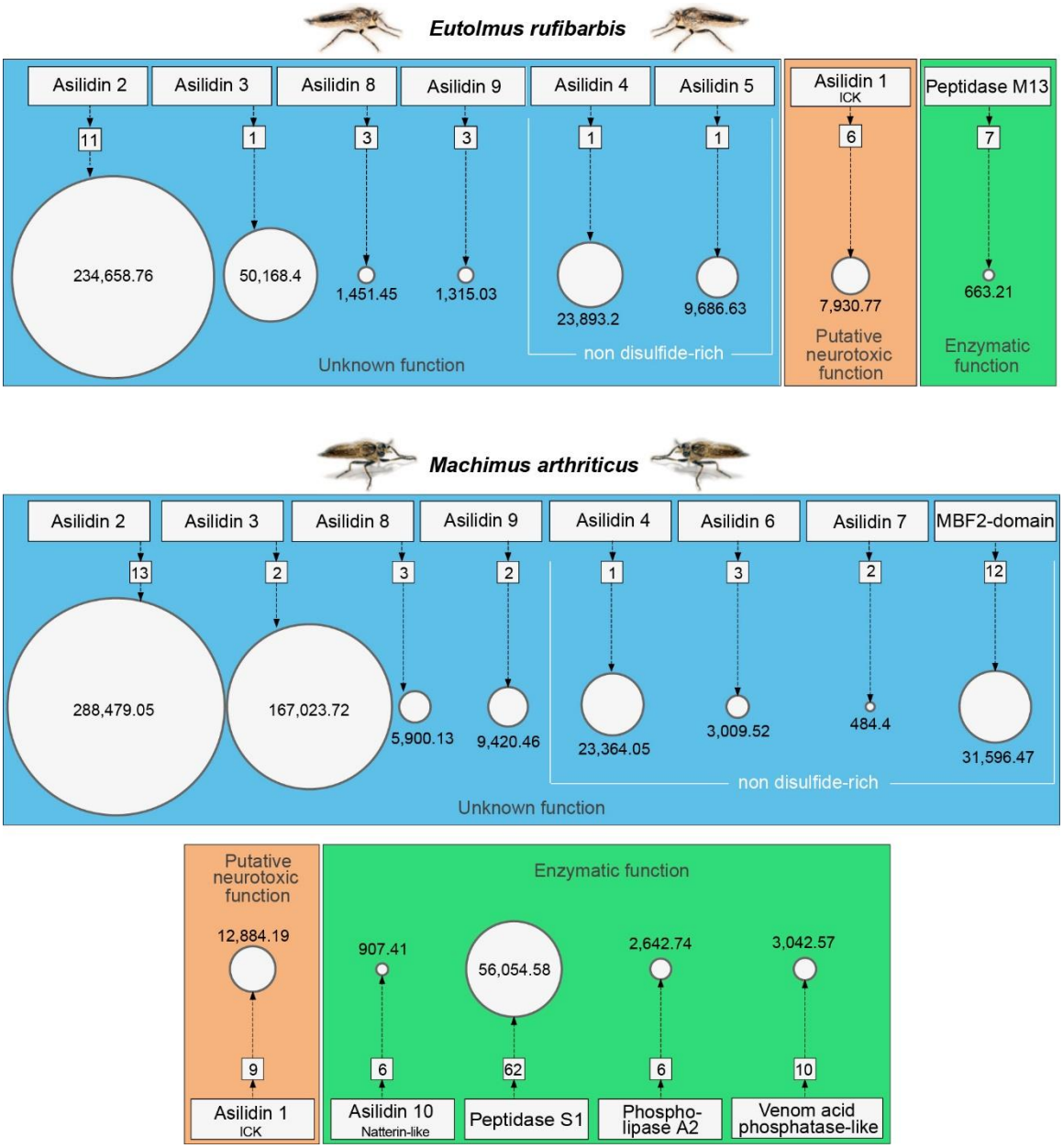


Figure 3. Bubble charts depicting the protein families secreted in the thoracic glands of *Eutolmus rufibarbis and Machimus arthriticus* and their expression levels. The number of assigned coding regions for each protein class is plotted in squares while related TPM values are plotted in circles. The size of the circle relates to the sum of all TPM values of all coding regions assigned to this protein class.

3.3. Similar Transcript Diversity but Different Expression Levels in Thoracic Gland and Body Tissue Transcriptomes

In total, we identified 31 putative venom protein classes that were expressed in the transcriptomes of the thoracic glands. 10 of these putative toxin families are novel, henceforth named Asilidins 1–10 (Table 2), according to the rational nomenclature of centipede toxins [514,554]. Of these, 29 of the 31

proteins were expressed by both robber fly species. Both species expressed only one protein class that was uniquely expressed in the gland transcriptome but that was not identified in the other species: Double cysteine inhibitor knots (dICKs) were identified only in *E. rufibarbis* (but not in the venom proteome), while Asilidin7 was exclusive to *M. arthriticus*. However, both unique proteins are expressed at low levels, the dICKs with one transcript with 2.45 (TPM), the U_7-Asilidin by two transcripts with 484.4 (TPM) (Figures 1 and 2).

The relationship between the numbers of identified coding regions and their expression levels is very similar for both species and the thoracic gland and body tissue transcriptomes, respectively. The numbers of coding regions are generally higher for the body tissue transcriptomes. Exceptions for which the numbers of thoracic gland transcripts were higher are the hyaluronidase and the MBF2-domain-like proteins for both species and the venom acid phosphatase, Asilidin5, Asilidin6, Asilidin7 and serpin in *M. arthriticus*. For almost all identified putative venom proteins similar variants were also complementary recovered in the body tissue. Only one protein class without matching variants in the body tissue could be found in the glands of *E. rufibarbis* (Asilidin6, one transcript, TPM 484.4) while two protein classes were unique to the glands of *M. arthriticus* (Asilidin6 and Asilidin7).

Table 2. Novel proteins and peptides identified in the proteomic analyses of the thoracic glands o *E. rufibarbis* and *M. arthriticus*. Proteomic hits specify the number of transcripts matching the fragments from mass spectrometry for each peptide and protein. P-Distance-based networks are provided in Figures S3–S8 (Supplementary File 1) for proteins with more than two sequences.

Protein Family	Structural Fold	Scaffold	Number of Residues	Molecular Weight (Average)	Proteome Hits *E. rufibarbis*	Proteome Hits *M. arthriticus*
Asilidin1	ICK	$Cx_{3-6}Cx_{5-9}CCx_{3-10}Cx_{4-6}C$	51–65	6.2 kDa	1	3
Asilidin2	unknown	$Cx_{35}Cx_{8-13}Cx_{36-39}C$	267–339	33.7 kDa	5	8
Asilidin3	unknown	$Cx_{7-8}Cx_{22}Cx_{14}C$	88–104	11.36 kDa	1	2
Asilidin4	unknown	no cysteine scaffold	86	18.47 kDa	1	1
Asilidin5	unknown	no cysteine scaffold	273	30.2 kDa	1	0
Asilidin6	unknown	no cysteine scaffold, two P rich domains	115	13.12 kDa	0	2
Asilidin7	unknown	no cysteine scaffold	139	15.43 kDa	0	1
Asilidin8	SVWC-domain	$Cx_{16}Cx_{4}Cx_{10}Cx_{7}Cx_{14}CCx_{5}C$	101	11.29 kDa	1	1
Asilidin9	SVWC-domain	$Cx_{22}Cx_{4}Cx_{10}Cx_{8}Cx_{12}CCx_{4}C$	118–119	13.5 kDa	1	1
Asilidin10	Natterin-like	no cysteine scaffold, P rich domain	146–226	20.38 kDa	0	2

The expression levels (TPM values) showed profound differences compared to the numbers of identified coding regions of putative toxins (Figures 1 and 2). In contrast to the body tissue, 16 up-regulated protein classes for *E. rufibarbis*, and 17 higher expressed coding regions in *M. arthriticus* were identified for the thoracic gland (including Peptidase S1 which was slightly lower expressed compared to the body tissue from *M. arthriticus*). Major differences between gland transcripts of both species is the sequence of the three most highly expressed proteins (Figures 1 and 2). Asilidin2 was the most highly expressed gland protein in both species with 234,658.76 (TPM) in *E. rufibarbis* and 288,479.05 (TPM) in *M. arthriticus*. Asilidin3 was the second most up-regulated protein in *M. arthriticus*, while it ranked third in *E. rufibarbis*. Peptidase S1, however, was the second highly up-

regulated protein in *E. rufibarbis* while it was not present in *M. arthriticus*, (Figure 3). Except for peptidase S1 in *E. rufibarbis* all other proteins were also present in the crude venom of both species and supported by our proteomic data.

3.4. Anatomy of the Venom Delivery System

The thoracic gland system (magenta colored, Figure 4) is composed of a pair of elongated sac-like glands located in the dorsal parts of the first and second thoracic segments. Each gland opens into a separate duct (Figure 4C), which fuse ventrally of the food channel just before entering the head capsule. The single duct continues ventral of the food channel into a salivary pump composed of a non-return valve-like mechanism and two associated muscles (M5+6, Figure 4B) responsible for opening the valve. This salivary pump is attached to the base of the hypopharynx and consists of two plates, one for M6 attachment and another one to which muscle M5 attaches. The single duct continues further after the salivary pump and opens into the hypopharynx near the apex of the proboscis. The hypopharynx is indirectly moved by the strong paired expansor muscle M4 which is attached to lateral apophyses of the basipharynx on both sides. The pharyngeal pump is operated by a strong ring or sphincter muscle engulfing the pharynx and a dorsal expansor which attaches to the roof of the pharynx. The volume of the cibarium is controlled by the muscle group M7 (Figure 4).

Our results support descriptions by Whitfield (1925) that the labial glands (Figure 4) open into the lumen between the theca and the labium near the apex of the proboscis. They are therefore structurally separated from the hypopharynx at the tip of the proboscis where the duct of the thoracic glands opens into the lumen of the hypopharynx. Since the hypopharynx is the only structure entering the prey (Whitfield 1925) and an elaborate pumping apparatus is missing, we conclude that these labial glands are not part of the venom production system. Additionally, the structure of the labial glands and their location are variable within asilids [528]. It was suggested that they secret a lubricant that facilitates the mechanical sting process and their opening between the theca and the labium supports this idea.

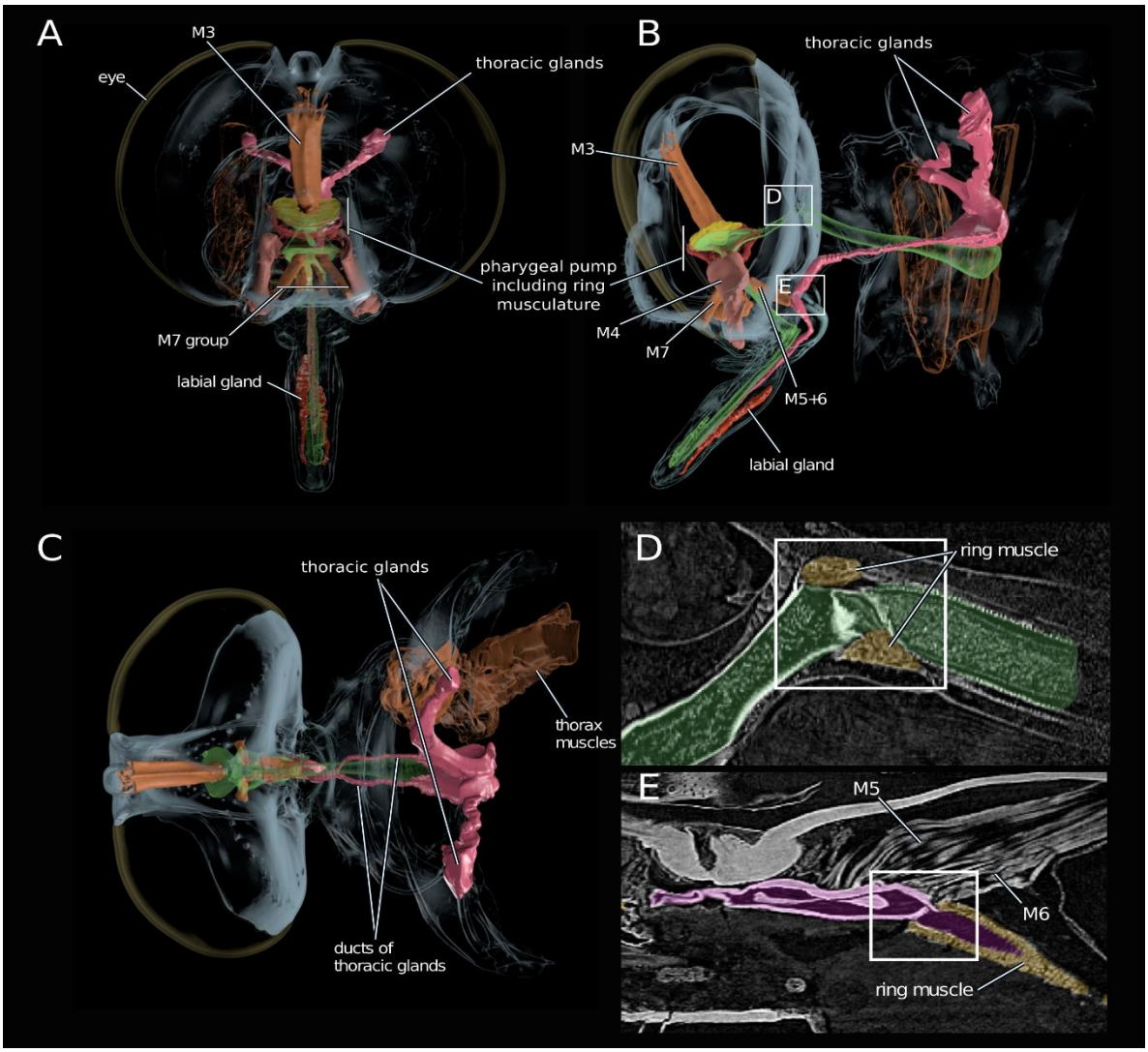


Figure 4. 3D reconstruction of the gland and head morphology in *E. rufibarbis.* Reconstruction based on synchrotron micro-computed tomography data. (**A**) frontal view (**B**) lateral view (**C**) dorsal view. (**D**) and (**E**) Images of the original scan data showing the digestive tract with the ring musculature to close the pharyngeal tract (**D**) and the musculature attaching at the salivary pump system (**E**). Note the "closed" position of the valve at the center of the white frame.

3.5. Both Species Exhibit Similar Venom Proteomes

The venoms of both *E. rufibarbis* and *M. arthriticus* are dominated by a few highly expressed, novel proteins with unknown function. In *M. arthriticus* these are Asilidin2–4 and Asilidin6–10, while Asilidin2 and Asilidin3 show a similar pattern in *E. rufibarbis.* Asilidin2–5, Asilidin8 and Asilidin9 were also detected in the venom of *E. rufibarbis*, however showed lower relative expression than in *M. arthriticus* (Figures 1 and 2). For both species two forms of cysteine-rich single von Willebrand factory

type C (SVWFC) domain-containing peptides are expressed within the group without known function (Asilidin8 and Asilidin9).

A lower expressed integrant comprises proteins with enzymatic function. For *M. arthriticus* this fraction was composed of peptidase S1, phospholipase A2, venom acid phosphatase-like proteins and a low expressed, putative natterin-like protease known from a toadfish [555–557]. Despite its high expression level (third rank) peptidase S1 is not detected in the proteome of *E. rufibarbis*. The only detected enzymatic component in its gland system was a low expressed metalloprotease M13.

Putative neurotoxic peptides with a cysteine inhibitor knot (ICK)-like structure (Asilidin1) [558] constitute another integral venom part. These peptides were identified in both gland proteomes of *M. arthriticus* and *E. rufibarbis*, and exhibit a cysteine scaffold typical of ICK peptides, which constitutes the bulk of the molecular diversity and ion channel modulating in spider venoms (Figure 5).

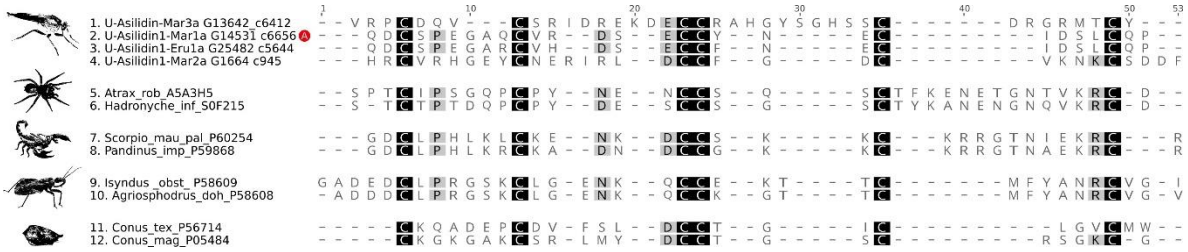


Figure 5. Sequence alignment of the mature Asilidin1 sequences identified in the proteome and transcriptome of the glands from *M. arthriticus* and *E. rufibarbis*. Mature peptide sequences of known and activity-tested ICK's were included from venom of the cone snails *Conus textile* and *Conus magus*, the assassin bugs *Isyndus obscurus* and *Agriosphodrus dohrni*, the funnel web spiders *Hadronyche infensa* and *Atrax robustus* and the scorpions *Scorpio palmatus* and *Pandinus imperator*. The cysteine residue backbone shared by all sequences is highlighted in yellow. The red circle with a white A marks the activity tested, synthesized ICK sequence.

3.6. Neurotoxic Activity of U-Asilidin$_1$- Mar1a

In order to test the functional convergence between venom ICKs present in asilid and other venoms, the peptide corresponding to the mature peptide encoded by the highest expressed Asilidin1 transcript, U-Asilidin$_1$-Mar1a (henceforth Mar1a) (Figure 6 and Figure S3 (Supplementary File 1)) was synthesized and injected into *Apis mellifera*. This revealed a clear neurotoxic effect on honeybees, including slow movements, disorientation and paralysis, similar to that of the positive control group treated with the potent insectidal spider toxin ω-atracotoxin from *Hadronyche versuta* [559,560] (Supplementary File 6).

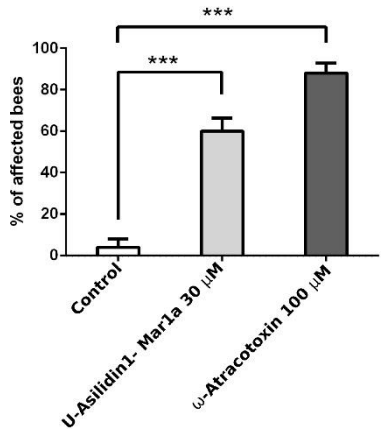


Figure 6. Neurotoxicity of Mar1a. Injection of 30 µM of Mar1a produced significant neurotoxic effects in bees (60% of injected bees were affected, see Supplementary File 6), compared to the negative control (MilliQ water, Supplementary File 6), including slow movements, disorientation and paralysis. The overall effect observed with Mar1a was comparable to our positive control, the potent insectidal spider toxin ω-atracotoxin (100 µM, >80% of bees were affected, Supplementary File 6). This bar graph represents the average of 5 experiments (5 bees per experiment), error bars represent S.E.M. and *** indicates the results were significantly different from control ($p < 0.001$).

4. Discussion

This study shows that robber flies (Asilidae) feature a complex venom delivery system and secrete venom from their thoracic venom glands. Like most venoms, asilid venoms consist of non-disulfide-rich peptides, disulfide-rich peptides novel proteins of unknown function, and enzymes. Contrary to our expectations for liquid-feeding predators, however, asilid venoms appear to be largely devoid of enzymatic proteins and instead consist of a large fraction of novel peptides and proteins. Moreover, we show that one of these peptides, Mar1a, produces neurotoxicity effects in the ecologically relevant prey species, the honey bee.

4.1. Highly Expressed Novel Proteins

Novel peptides and proteins with unknown function are by far most abundantly recovered by transcriptomics and proteomics in both species (Figure 3). This group comprises over 97% (*E. rufibarbis*) and 87% (*M. arthriticus*) of the TPM values assigned to all secreted putative toxins. Novel putative toxins are not unusual in proteomic/transcriptomic studies of poorly characterized venoms, such as in remipede crustaceans, where five unknown proteins accounted for ~15% of all FPKM values [561,562]. However, asilids seem to contain an unusually high percentage of unknown venom proteins. This is unexpected given that some dipterans such as *Drosophila* and *Aedes* are well studied model organisms of which several genomes have been sequenced and annotated. Without potential

homologues outside the *E. rufibarbis* and *M. arthriticus* no speculation about the origin or putative function of these proteins seems feasible (Table 2 and Figure S4–S8 (Supplementary File 1)).

4.2. Asilid Venoms Contain Putative ICK Neurotoxins

While previous studies have observed neurotoxic activity of crude gland extracts [520,531,563], we here show that asilid venoms do indeed contain neurotoxic peptide toxins. This protein family, Asilidin1, is characterized by a cysteine pattern that is strongly suggestive of a cystine inhibitor knot (ICK) structural motif, which is defined by an antiparallel β sheet stabilized by a cystine knot. The cystine knot consists of a ring formed by two disulfide bonds and their intervening sections of the peptide backbone that is pierced by a disulfide bond that generally stabilizes the C-terminal region of the peptide. The ICK is probably the most widely recruited peptide fold in animal venoms, and ICK-like toxins are known from the venom of cone snails, scorpions, spiders and assassin bugs [564]. The neighbor network of Asilidin1 splits in to three distinct clades (Figure S3 (Supplementary File 1)). One clade includes U-Asilidin$_1$-Mar2a, which is a unique sequence exclusively present in the thoracic glands with no matches in the body tissue. The other two clades also comprise identical body tissue sequences, although the expression values between gland sequence and body tissue sequence differ substantially. The most highly expressed peptides U-Asilidin$_1$-Mar1a and U-Asilidin$_1$-Eru1a are around 3000 times higher expressed in the thoracic glands compared to their respective body tissues. Activity test of a manually synthesized Mar1a protein showed motor effects in honey bees, which suggests Asilidin1 represents a family of neurotoxic ICK toxins (Figure 5 and Supplementary File 6).

4.3. Missing "Usual Suspect" Enzymes in Putative Asilid Venom?

Enzymes are known as important venom components in a vast number of venomous predators such as snakes, cephalopods, centipedes, assassin bugs, stinging insects and remipede crustaceans. Several classes of enzymatic proteins are often abundant in venoms, such as chitinases or serine proteinases, particularly in liquid-feeding arthropods where they likely play an important role in pre-digestion of prey [514–516,562]. Interestingly, the venom of *M. arthriticus* and *E. rufibarbis* differs from other liquid-feeding venomous insect predators by being largely devoid of enzymes. Although the transcriptome data of both species show high expression levels of M13 and S1 proteases as well as PLA2 and acid phosphatase (see Figures 1–3), our proteomic analyses indicate that these enzymes constitute only minor components in the crude venom. This is in contrast to previous studies, which assumed that asilids secrete venom composed of proteolytic and neurotoxic activity [520,531,563]. Dipterans are generally known to digest orally by refluxing digestive enzymes from their gut to achieve extra oral digestion, a strategy applied in different variants also in other predatory arthropods [565,566]. Earlier studies showed that the stomach of the tested asilids shows higher proteolytic

activity but also a larger range of pH-values in which enzymatic activity is observed [520]. Our proteomic results are in accordance with these findings.

4.4. Scenario for Envenomation of Prey by Asilids

Based on our proteomic and activity test result, and in combination with morphological data derived from SRμCT, we are now able to draw a more precise and partly new picture of the envenomation process in asilids. Envenomation begins with the transport of the secreted venom including the neurotoxic Mar1a into the lumen of the hypopharynx (Figure 7). This is on the one hand achieved by the salivary pump muscles but might be supported through an increase of the cibarial and pharyngeal lumen. Since the musculature directly inserting at the different parts of the gland system is minimal and only serves to control the non-return valve of the salivary pump, the larger volume changes of the cibarial pump are needed to effectively transport a larger fraction of venom into the prey with the first sting. Once the venom is loaded into the lumen of the hypopharynx, it can be injected in a second step against the hemolymph pressure of the prey through contraction of the ring musculature of the pharyngeal and cibarial pump. At the same time, the ring musculature of the salivary duct and the alimentary canal is contracted in order to prevent unwanted backflow. Literature data on the prey capture process indicates that prey is paralyzed within seconds and this could not be achieved with a passive inflow of saliva into the prey [520,530]. The prey's inner compartments are then liquefied in a third step presumably by pumping stomach liquids through the proboscis into the lumen of the prey. The fourth and last step is the uptake of food. The liquefied body tissue is sucked up by contraction of M3 and M7 muscles which again increases the lumen of the cibarial and pharyngeal pump and thereby creates a negative pressure with respect to the lumen of the prey (Figures 4 and 7).

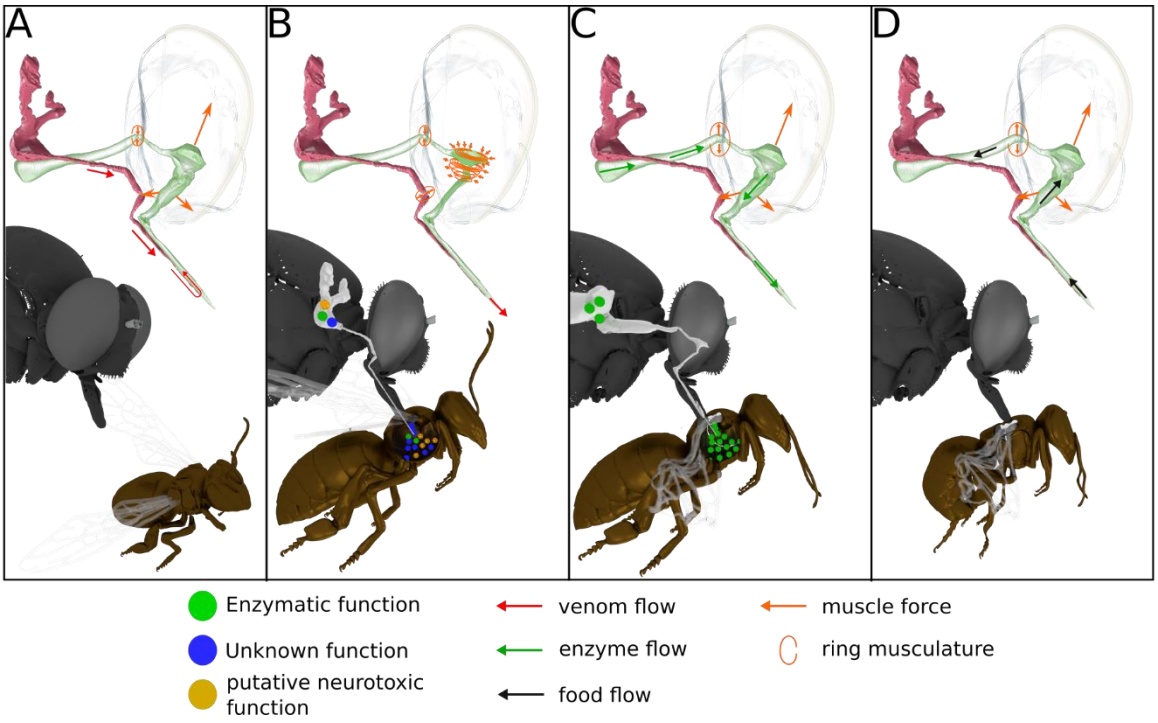


Figure 7. Envenomation and feeding process in asilids. Summary of the hypothesized feeding process for robber flies. (**A**) Closed valve of the digestive tract, open valve of the thoracic glands, increase of the pharyngeal pump volume leads to transport of the venom from the thoracic glands to the hypopharynx; (**B**) Closed valve of digestive tract, closed valve of the thoracic glands, decrease of the pharyngeal pump volume leads to powerful injection of the venom via the hypopharynx; (**C**) Open valve of the digestive tract, increase of the pharyngeal pump volume, transport of enzymes from the digestive tract to the hypopharynx and injection in the paralyzed prey; (**D**) Open valve of the digestive tract, increase of the pharyngeal pump volume, feeding on the liquefied prey contents.

4.5. Implications in the Context of Fly and Insect Venom Evolution

Many dipteran groups have convergently evolved a hematophagous life style and are thus per definition also venomous. Due to their role as disease vectors and their impact on humans or live-stock, mosquitoes and tabanid species have received increased attention [567–573] (Figure 8). Surprisingly, however, even for these comparably well-studied groups, only a few species have been covered by recent –omics approaches or detailed morphological studies using modern technology (Figure 8) [506]. This situation hinders detailed comparative analyses of the evolution of venoms and the morphological dynamics of venom delivery systems within flies. Our study comprehensively describes the venom system of asilids, which, unlike most other dipterans, are a truly predatory lineage. Venom is not restricted to blood feeding females, as is the case in hematophagous dipterans, but occurs in both genders. Blood feeding evolved early in dipterans, long before the rapid radiation

in the more basal groups of Brachycera to which asilids belong (Figure 8). The results from the present study support an independent evolution of asilid venom in which a separate suite of proteins and peptides compared to those of hematophagous lineages were functionally recruited as toxins. Reflecting this independent evolution of a predatory venom, the asilid the venom delivery system appears highly adapted to high-speed predation and envenomation that is facilitated by a derived muscle system.

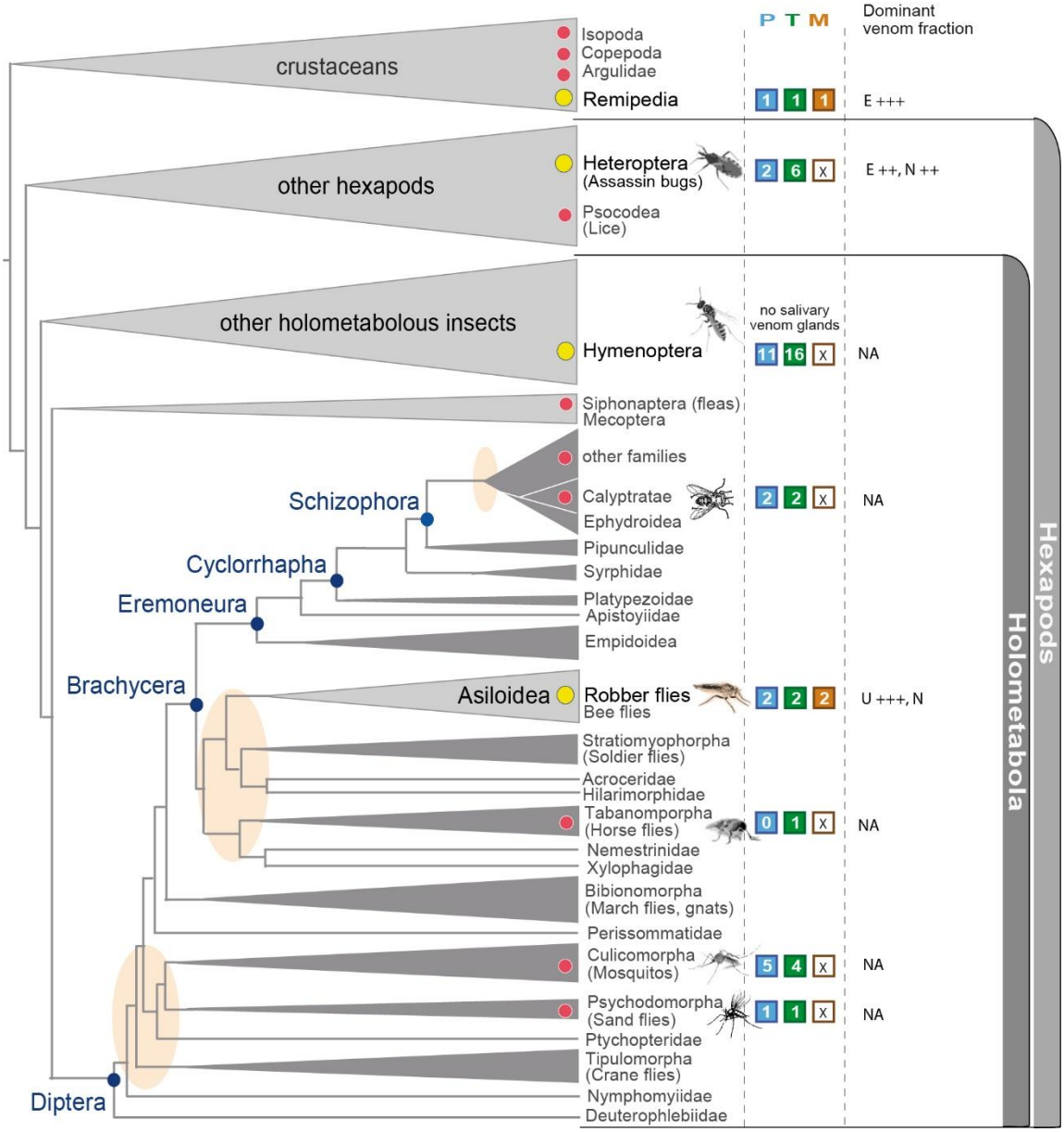


Figure 8. Current picture of venomics studies and phylogenetic implications for dipteran and insect venom evolution. The number of species that are covered by studies applying modern transcriptomic, proteomic and morphologic analyses are shown in the colored squares. Blue: Proteomics; Green: Transcriptomics; Brown: Reconstructions of the venom delivery system. Dominant venom fractions are roughly summarized with E: Enzymes; N: Neurotoxins; U:

Unknown proteins. The plus sign indicates overexpressed fractions in the venom. Yellow circles indicate venomous predatory lineages, red circles hematophagous groups. Only studies based on Illumina transcriptome data and proteomic profiles of complete venom gland were included. See also a general overview on available salivary/venom gland transcriptomes of arthropods until 2014 in von Reumont et al., 2014 (Toxins)[506]. The tree is based on Wiegmann et al., 2011 and Misof et al., 2014 [518,574]. The branches do not reflect precise distances. Oval circles in beige indicate disputed branching events, blue dots represent major clades in Diptera.

Diptera is one of the three groups within insects besides hymenopterans and heteropterans that is known to feature venomous species, however, their venoms remain largely uncharacterized. Yet, in the quest to better understand venom evolution in insects in a larger framework, dipterans play an eminent role as the youngest insect group, having evolved in the ending Jurassic period around 150 million years ago [574]. Moreover, the venoms studied here indicate asilid venoms differ substantially from other venomous liquid feeding insect lineages. Future studies on a taxonomically wider sample may therefore provide insight into the key processes that govern toxin evolution [575].

5. Conclusions

This study provides a detailed description of the venom system of robber flies (Asilidae). Their venom is produced in the thoracic glands, and injected through a complex venom delivery system that agrees well with their agile, predatory lifestyle. Surprisingly, their venoms appear to be largely devoid of enzymatic proteins and instead consist of a large fraction of novel peptides and proteins. This high degree of molecular novelty suggests dipteran-specific groups of proteins and peptides were recruited as toxins into asilid venom. Moreover, we show that one of these peptides, Mar1a, produces neurotoxicity effects in honey bees, suggesting the other, novel peptide families may also harbor neurotoxic activities. Due to the attractiveness of peptides as easily synthesized compounds with a high degree of pharmacological potency and selectivity, asilid venoms therefore appear to be a good source of molecular tools and potential lead molecules for development into therapeutic and agrochemical products [532]. Our results thus provide the foundation for future studies to not just help understand mechanisms of toxin recruitment and evolution within and outside dipteran arthropods, but also mine this source of novel, bioactive tools and potential lead molecules.

6. Supporting information

All the supporting information are not included in the manuscript but are available online at www.mdpi.com/2072-6651/10/1/29/s1, Supplementary File 1: Summary and description of the identified enzymatic components, with a short discussion of the putative functions in the robber fly

venom. Figure S1: Pictures of the analyzed species and the habitat, where specimens were captured, Figure S2: SDS-PAGE gel of crude venom of *Machimus arthriticus* and *Eutolmus rufibarbis*, Figure S3: Neighbor Networks for the Asilidin1, Figure S4: Neighbor Networks for the Asilidin2, Figure S5: Neighbor Networks for the Asilidin3, Figure S6: Neighbor Networks for the Asilidin8, Figure S7: Neighbor Networks for the Asilidin9, Figure S8: Neighbor Networks for the Asilidin10, Figure S9: CO1 barcode alignment, Figure S10: Overall strategy for synthesis of U-Asilidin1-Mar1a, Figure S11: UV chromatogram of linear U-Asilidin1-Mar1a, Figure S12: UV chromatogram of folded U-Asilidin1-Mar1a, Supplementary File 2: Summarizing table of the identified classes of putative toxins for the thoracic gland system transcriptome of Eutolmus rufibarbis. The TPM values for each identified coding region are shown, the presence of a signal peptide and presence of proteomic data is marked, Supplementary File 3: Summarizing table of the identified classes of putative toxins for the body tissue transcriptome of *Eutolmus rufibarbis*. The TPM values for each identified coding region are shown, the presence of a signal peptide is marked, Supplementary File 4: Summarizing table of the identified classes of putative toxins for the thoracic gland system transcriptome of *Machimus arthriticus*. The TPM values for each identified coding region are shown, the presence of a signal peptide and presence of proteomic data is marked, Supplementary File 5: Summarizing table of the identified classes of putative toxins for the body tissue transcriptome of Machimus arthriticus. The TPM values for every identified coding region are shown, the presence of signal peptides is marked, Supplementary File 6: Video of observed effects of U-Asilidin1-Mar1a, ω-atracotoxin and negative control on honey bees based on the bioassay of the synthesized U-Asilidin1-Mar1a, Supplementary File 7: Protein alignments of all identified classes of putative toxins. For each class the coding regions identified in the thoracic gland system and the body tissue of both robber flies are aligned together with toxins known from other species. For some classes the alignments were cut to align only the domains, in these cases the name of the pruned alignment is marked with “cut”.

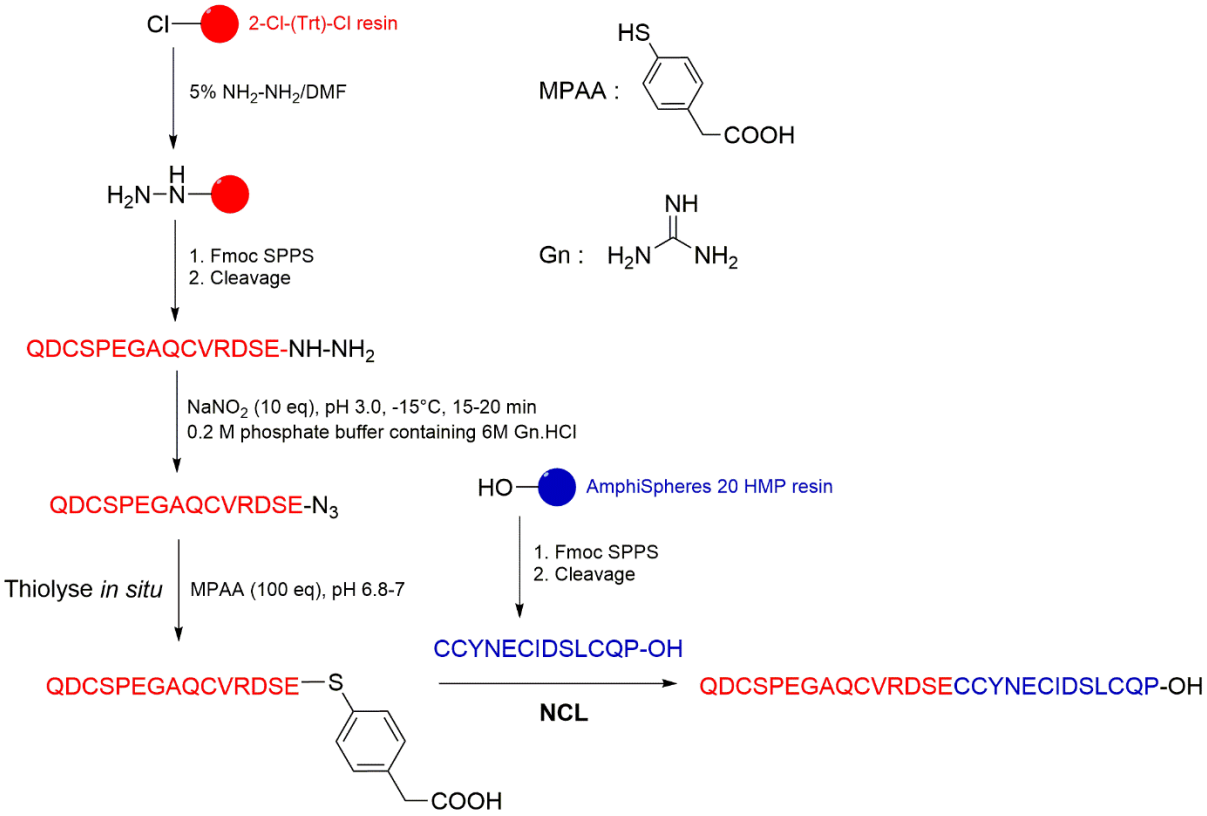


Figure S10: Overall strategy for asilidin1 peptide synthesis.

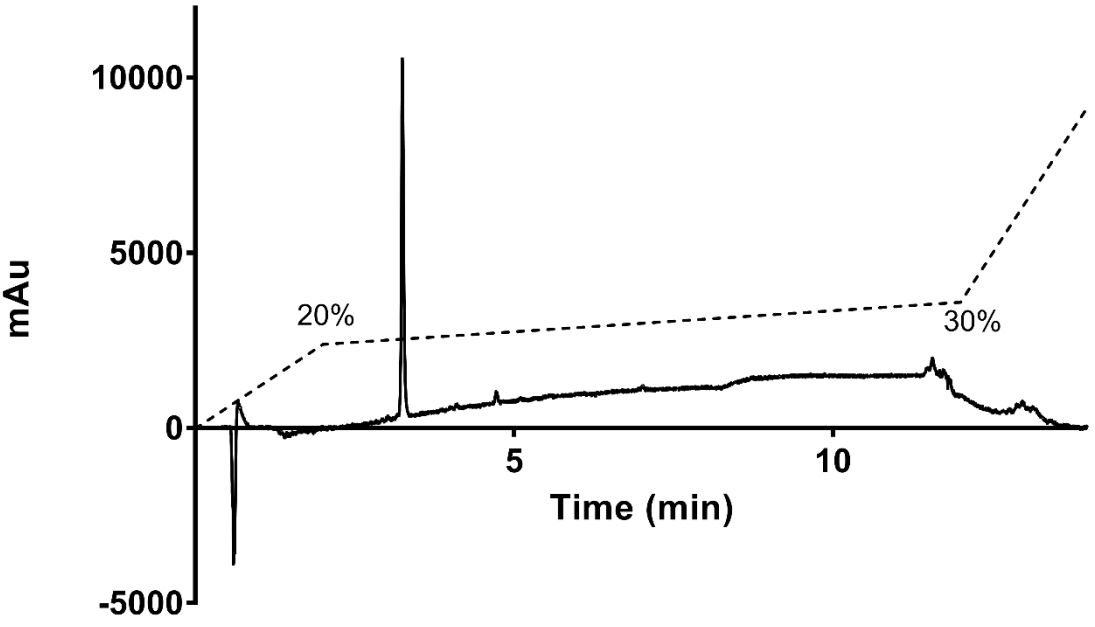


Figure S11: UV chromatogram of linear toxin (dashed line indicates acetonitriles gradient).

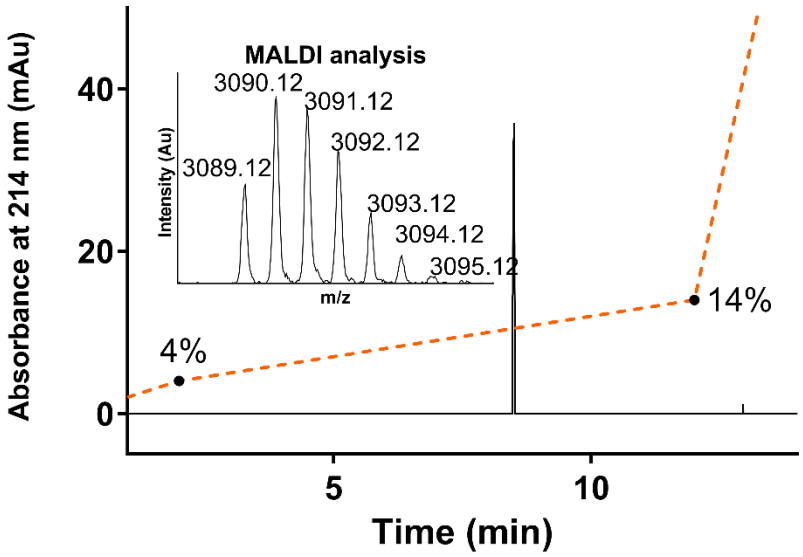


Figure S12: UV chromatogram of folded toxin and MALDI analysis (dashed line indicates acetonitrile gradient).

c. Synthèse de la toxine PmuTx1

Encouragés par le folding supposé de type ICK réussi de la toxine U-asilidin1-Mar1a, nous nous sommes intéressés à la synthèse d'une nouvelle toxine qui bloque les canaux Cav3.2, appelée PmuTx1 et isolée du venin de la mygale *Pterinochilus murinus*. Cela était également l'occasion de vérifier si les conditions de folding pouvaient également s'appliquer à cette séquence.

Cette étude dans laquelle j'ai effectué la synthèse et la caractérisation par spectrométrie de masse est présentée ci-après sous forme de publication, non publiable en l'état.

1. Introduction

Essential physiological functions, including muscle contraction, hormone secretion, neurotransmitter release and nociceptive transmission are controlled by calcium (Ca^{2+}) currents in mammalian excitable cells. Voltage-gated calcium channels (Ca_v) are composed of various subunits assuming different functions such as pore-forming/voltage-sensing α_1 subunit and several regulatory $\alpha_2\delta$, β, γ subunit isoforms [576]. To date, 10 pore forming (α_{1A}-α_{1I} and α_{1S}) and several splice variants have been identified and characterized. $Ca_v3.2$ are T-type calcium channel subtypes which are localized in various subpopulations of primary afferent neurons [577,578], altogether suggesting a role of these channels in pain processing. As exemplified by analgesic effect induced by systemic or intrathecal delivery of T-type calcium channel blockers such as ethosuximide and mibefradil in rodents [579,580].

Most venomous animals have evolved through millions years of evolution to efficiently neutralize their prey with finely tuned venoms. Accordingly, their toxins target a large range of ion channels expressed in the neuronal and neuromuscular systems of prey and predators. As a result, many of the most selective ion channel blockers originate from venoms [42]. Although spider venoms have provided many Ca_v-inhibitor toxins selective for various isoforms [581], spider toxins active at $Ca_v3.2$ have not yet been reported. Indeed, the majority of these peptides displays activity at Ca_v1, $Ca_v2.1$, $Ca_v2.2$ and $Ca_v2.3$ [582–596]. The other Ca_v channels subtypes, especially $Ca_v2.1$ and $Ca_v2.3$ also mediate significant neurotransmitter release at neuronal synapses [597,598] therefore, the non-selective block of these Ca_v isoforms induce severe side effects due to inhibition of neurotransmitter release in non-nociceptive neurons. Since there is yet no selective blocker of $Ca_v3.2$ T-type calcium channels, its role in the nociceptive pathway is still poorly understood. Thus, selective blocker of $Ca_v3.2$ would be highly valuable to delineate their role in nociceptive pathway and would represent potential therapeutic lead for neuropathic pain treatment.

Here we report on the discovery of a new toxin targeting $Ca_v3.2$, PmuTx1 from *Pterinochilus murinus* tarantula and subsequent attempts of regioselective synthesis strategy.

2. Methods

2.1 Abbreviations

Acm, acetamidomethyl; ACN, acetonitrile; Boc, *tert*-butoxycarbonyl; DCM, Dichloromethane; DIPEA, diisopropylethylamine; DMF, *N,N'*-dimethylformamide; DMSO, Dimethyl sulfoxide; DTP, 2,2'-Dithiopyridine; ESI-MS, electrospray ionization mass spectrometry; Fmoc, fluorenylmethoxycarbonyl; HATU, 1[*Bis*(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate; IPA, isopropanol; LC/MS, liquid chromatography/mass spectrometry; MeOH, methanol; nAChR, nicotinic acetylcholine receptor; NMR, nuclear magnetic resonance; Pbf, pentamethyl-dihydrobenzofuran-5-sulfonyl; RP-HPLC, reversed phase high performance liquid chromatography; SPPS, solid phase peptide synthesis; *t*Bu, *tert*-butyl;TFA, trifluoroacetic acid; TIS, triisopropylsilane; Tris, 2-Amino-2-(hydroxymethyl)propane-1,3-diol;Trt, trityl; UV, ultra-violet.

2.2 Chemical synthesis

DMF, DIPEA, ACN, TIS, TFA, piperidine and all others reagents were obtained from Sigma-Aldrich (Saint-Louis, MI, USA) or Merck (Darmstadt, Germany) and were used as supplied. Fmoc (L) amino acid derivatives and HATU were purchased from Iris Biotech (Marktredwitz, Germany). AmphiSpheres® (0.6 mmol/g) was purchased from Iris Biotech (Marktredwitz, Germany). The following side-chain protecting groups were used: Trt for Asn, Cys, and His; Acm for Cys; *t*Bu for Thr, Tyr, Cys, Asp and Glu; Pbf for Arg; Boc for Lys and Trp; and S*t*Bu for Cys. Peptides were manually synthesized by using the Fmoc-based solid-phase peptide synthesis technique on a VWR (Radnor, PA, USA) microplate shaker. All Fmoc amino acids and HATU were dissolved in DMF to reach 0.5 M. the first residue was anchored on 20 HMP resin using the method described by *Grandas et al*. [99]. The resin was washed with DMF, DCM, MeOH, and DMF. Fmoc deprotection was carried out with piperidine in DMF (1/2 v/v) twice for 3 min. Subsequent amino acids were coupled onto 0.1 mmol of prepared resin (determined loading value 0.44 mmol/g) twice for 25 min using an amino acid/HATU/DIPEA ratio of 5:5:10 relative to resin loading. DMF was used for resin washing between deprotection and coupling steps. After chain assembly was complete, the terminal Fmoc group was removed and the resin washed with DMF and DCM. Side-chains (except acm) deprotection and cleavage from the resin was carried out by adding 10 mL of TFA/TIS/H_2O (95/2.5/2.5 v/v/v) and stirring the mixture for 2.5 h at room temperature except for the regioselective 2 where the cleavage was carried out using TFA/DMB/H_2O (95/2.5/2.5 w/w/w) [599]. Crude peptides were purified by preparative RP-HPLC and pure fractions were combined and

freeze-dried their purity were confirmed by LC/ESI-MS (isolated yields ≈ 2%).

2.2.1 Regioselective strategy 1

Linear peptide was dissolved in isopropanol/water (1/2 v/v) adjusted to pH 8.5, to reach 0.05 mM concentration and stirred at room temperature until completion of the reaction.

2.2.2 Regioselective strategy 2

Linear peptide was dissolved in ACN/DMSO/water (60/25/15 v/v/v) to reach 3.4 mM concentration and 30 eq of PBu_3 are added. The reaction mixture was stirred at room temperature for two hours leading to 86% removal of S*t*Bu protecting group based on UV chromatogram. The reaction mixture was acidified to pH 3 and loaded onto preparative RP-HPLC and pure fraction were combined and freeze-dried (isolated yield = 70%). The first disulfide bridge is formed between the free cysteine residues by dissolving the peptide at 0.2 mM in 50 mM Tris-HCl buffer adjusted to pH 8 and adding dropwise 7 equivalents of DTP at 10 mM in MeOH. The reaction mixture was stirred at room temperature until completion of the reaction.

2.2.3 Regioselective strategy 3

Linear peptide was dissolved in CH_3COOH/water (1/1 v/v) to reach 0.2 mM concentration and 1.1 eq of I_2 at 2 mM in MeOH was added dropwise. The reaction mixture was stirred at room temperature for 20 min leading to total formation of the first disulfide bridge based on UV chromatogram. The reaction mixture was loaded onto preparative RP-HPLC and pure fraction were combined and freeze-dried (isolated yield = 80%). The mono-bridged peptide is dissolved in CH_3COOH/ACN/ water (2/1/1 v/v/v) to reach 0.2 mM concentration and 3 eq (1.5 eq/Acm) of 0.1 M HCl are added. The second disulfide bridge was formed by treatment with 10 eq (5 eq/Acm) of 50 mM I_2 in MeOH. After 30 min the reaction was complete (based on UV chromatogram) and quenched with 20 mM ascorbic acid until total discoloration of the solution, then purified by preparative RP-HPLC. The combined pure fractions were freeze-dried (isolated yield = 50%) and dissolved in TFA/DMSO (95/5 v/v) to reach 0.1 mg/mL concentration. The reaction mixture was stirred at room temperature for 1 hour.

2.3 Mass spectrometry

Solvents used for LC/MS were of HPLC grade.

Intermediate peptides were characterized using a LC/MS system consisting of a Waters (Milford, OH, USA) Alliance 2695 HPLC, coupled to a Waters Micromass ZQ (electrospray ionization in positive mode (ESI+) fitted with a quadrupole mass analyzer). All the analyses were carried out using a Chromolith

(Fontenay sous Bois, France) HighResolution RP-18e (4.6 x 25 mm, 15 nm–1.15 µm particle size, flow rate 3.0 mL/min) column. A flow rate of 3 mL/min and a gradient of 0–100% B over 2.5 min for routine analyses and 0-30% B over 30 min for quality control of pure products were used. Eluent A was water/0.1% HCO_2H and eluent B consisted of acetonitrile/0.1% HCO_2H. UV detection was performed at 214 nm. Electrospray mass spectra were acquired at a solvent flow rate of 200 µL/min. Nitrogen was used for both the nebulizing and drying gas. The data were obtained in a scan mode ranging from 100 to 1000 m/z or 250 to 1500 m/z to in 0.7 s intervals.

2.4 Preparative RP-HPLC

Preparative RP-HPLC was run on a Gilson PLC 2250 Purification system (Villiers le Bel, France) instrument using a preparative column (Waters DeltaPak C18 Radial-Pak Cartridge, 100 Å, 40 × 100 mm, 15 µm particle size, flow rate 50.0 mL/min). Buffer A was 0.1% TFA in water, and buffer B was 0.1% TFA in acetonitrile.

3. Results and discussion

PmuTx1 have been identified by screening of *Pterinochilus murinus* tarantula venom over $Ca_v3.2$ channel subtypes and subsequently sequenced by Edman degradation revealing a 36 residues-long sequence comprising six cysteines. PmuTx1 displays high structural homology with several U4-theraphotoxins identified by proteo-transcriptomic analysis of *P. Formosa* [600], which exhibits a putative inhibitor cysteine knot motif (ICK). Moreover, two others toxins, namely ω-grammotoxin SIA and pn3a, were structurally characterized as ICK and are also very closely related to PmuTx1 sequence (Figure 1) [601,602]. However, none of these toxins have been pharmacologically characterized.

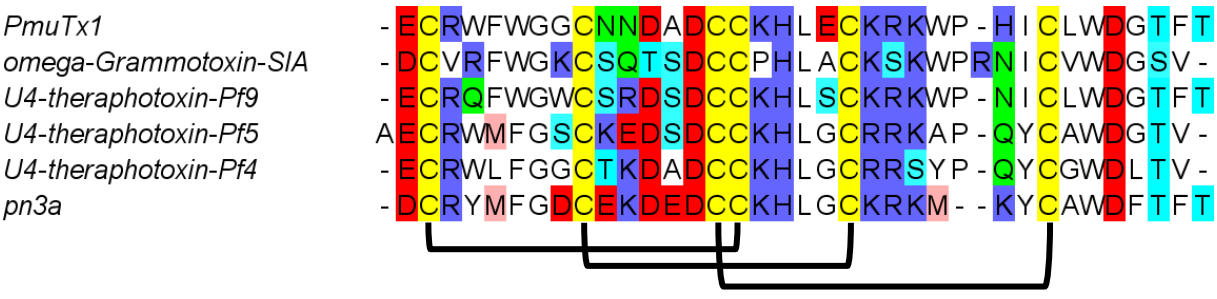


Figure 1: Alignment of newly identified PmuTx1 and other related toxins. Disulfide bonds are indicated in black lines. ω-grammotoxin SIA and pn3a structures have been elucidated by NMR spectroscopy and confirmed to display an ICK fold. U4-theraphotoxins have been identified by proteo-transcritptomic analysis of *P. Formosa* venom and classified as ICK toxins based on sequence homology.

3.1 Chemical synthesis

Linear peptides have been synthesized by Fmoc-SPPS chemistry and subsequent attempts of oxidative folding (see methods) led to the formation of several isomers not individually isolable. Therefore, we opted for a regioselective folding strategy assuming an ICK fold as represented in Figure 1.

Regioselective folding strategy of ICK toxins has been successfully applied to MfVIA [603] and gurmarin [604] synthesis in Boc-chemistry. Both groups used the same disulfide bond formation order: first, the CysIII-CysVI then the CysII-CysV and finally the CysI-CysIV. Although the rationale for this is not detailed, it might influence the folding yield.

3.1.1 Regioselective strategy 1

The first regioselective strategy used a combination of *t*Bu and Acm protected cysteines however, the formation of the first disulfide bridge between the free cysteine residues in aqueous basic conditions led to the formation of an oxidation byproduct (Figure 2A). This side reaction is increasing over time and starts before the formation of the first disulfide bridge is complete. Considering that, tryptophan residues are known to be oxidation-sensitive; the +16 and +32 peaks observed on the mass analysis is most likely due to hydroxy-tryptophan formation [605–607].

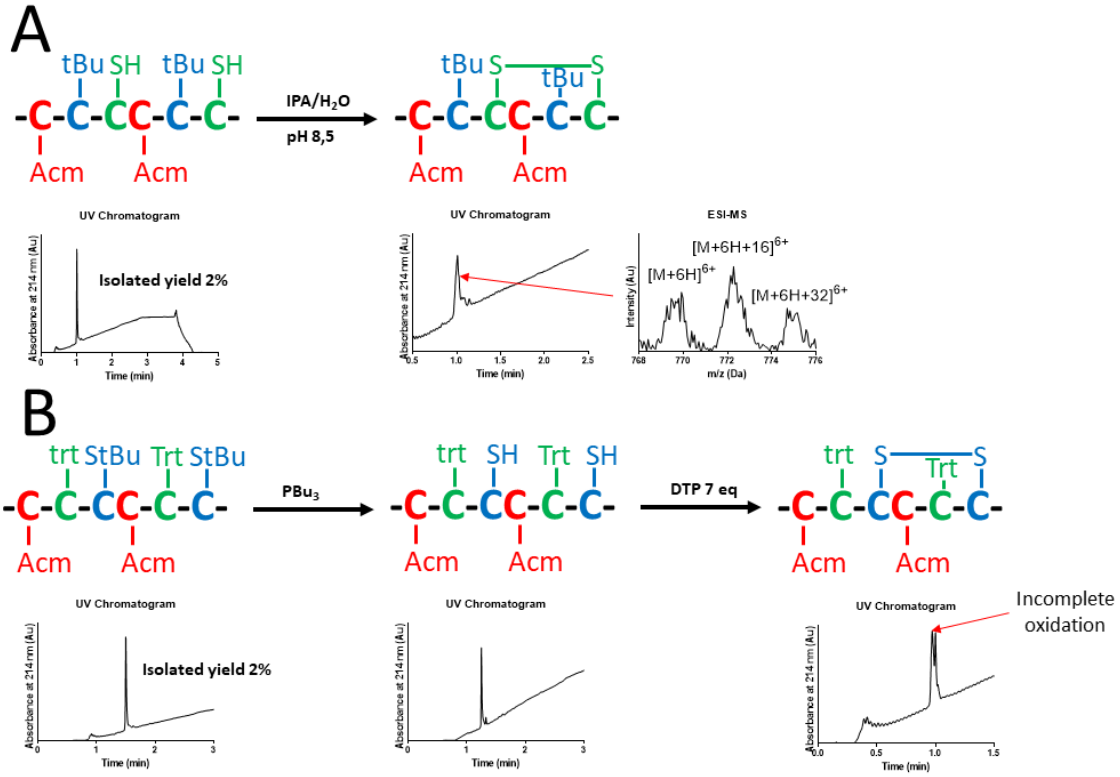


Figure 2: Schematic representation of regioselective folding strategies 1 and 2.

3.1.2 Regioselective strategy 2

With the objective of developping a regioselective folding strategy that would be theoretically suitable for the synthesis of four disulfide bridges-containing peptides, we used the combination of Trt/S*t*Bu and Acm protecting groups described by *Mochizuki et al.* for the μ-conotoxin SIIIA synthesis [599]. Briefly, this method is based on the use of 1,3-dimethoxybenzene (DMB), which is a less effective Trt cation scavenger than the Cys sulfhydryl group due to steric hindrance allowing the cysteine residues to re-capture the Trt groups after cleavage.

Unlike the strategy 1, we used DTP (2,2'-Dithiopyridine) to induce the formation of the first disulfide bridge [196,357]. This method has been widely employed by our group and others and always showed good results for the regioselective synthesis of peptides containing two disulfides bridges. We hoped that the formation of the bridge would be fast enough to prevent the oxidation of the tryptophan residues. However, the oxidation reaction was incomplete (50% based on the UV chromatogram) after 10 min: adding more DTP led to the partial deprotection of Acm protecting group [608] and allowing the reaction to continue, leading eventually to the oxidation of tryptophan residues in the same way that strategy 1 does. Although in the previously described syntheses DTP induced total bridge formation in 10 min, the incomplete formation of the first bridge after 10 min suggest that it might be limited to the predisposition of the sequence to favor the formation of the bridge.

3.1.3 Regioselective strategy 3

We hypothesized that oxidation of tryptophan residue was due to long exposure to alkaline conditions. Consequently, we attempted the formation of the first disulfide bridge in acidic conditions (Figure 3). To prevent an additional purification step, we switched back to the *t*Bu/Acm system (regioselective strategy 1).

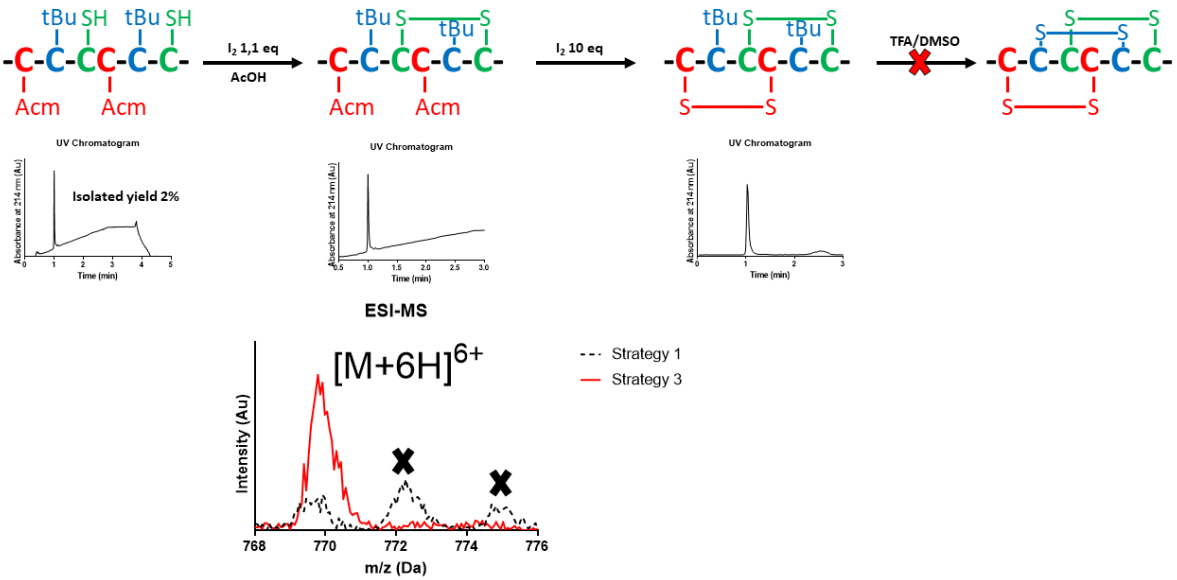


Figure 3: Regioselective strategy 3.

The formation of the first disulfide bridge in acidic condition does not yield to oxidation products (Figure 3) suggesting that it was due to alkaline condition of strategy 1 and 2. Iodine deprotective oxidation of the Acm protecting groups proceeds with good efficiency based on the UV chromatogram. Although the deprotection/oxidation step using TFA/DMSO was used with success for the synthesis of a α-SI dimer [117], in our case it led to the complete degradation of the product with several unidentified peaks suggesting that it might be sequence-dependent.

An alternative strategy would be to do the random oxidation of the first two disulfide bridges that should lead to three (or less) isomers easily separable by RP-HPLC and to perform the third disulfide bridge formation by Acm deprotection/oxidation in the presence of iodine.

d. Synthèse de la toxine ShK et de ses analogues

Comme nous pouvons le constater sur l'exemple de la synthèse de la PmuTx1, les rendements sur la synthèse de peptide peuvent être très faibles. Dans l'optique de tester la ligation chimique native comme technique alternative pour la synthèse du peptide linéaire, nous nous sommes intéressés à la synthèse de la toxine ShK et à des analogues identifiés par transcriptomique des glandes à venin du gastéropode *Cumia reticulata*.

Cette étude dans laquelle j'ai effectué la synthèse et la caractérisation par spectrométrie de masse est présenté ci-après sous forme de publication, non publiable en l'état.

1. Introduction

Membrane potential and calcium signaling in diverse cell types, including those involved in both innate and adaptive immune response, are regulated by potassium channels. A potent potassium channel blocker named ShK toxin and comprising 35 amino acid residues and six cysteines was originally isolated from the venom of the Caribbean sea anemone *Stichodactyla helianthus* [609] and chemically synthesized for further pharmacological characterization [610]. Thanks to its synthesis, K_v1.3 have been identified as the primary target of ShK toxin with a remarkable high potency of 133 pM and its three dimensional structure has been elucidated revealing the three disulfide bonds with connectivity CysI-CysVI, CysII-CysIV and CysIII-CysV [611,612]. K_v1.3 are functionally dominant channels in terminally differentiated effector memory T-cells, which are involved in multiple autoimmune pathologies. Therefore, K_v1.3 inhibition is a promising therapeutic strategy for the treatment of T cell-mediated autoimmune diseases such as multiple sclerosis, type 1 diabetes mellitus and rheumatoid arthritis. However, ShK toxin also displays subnanomolar affinity for K_v1.1 and K_v1.6 subtypes, which play a major role in brain and cardiac tissues. To increase its therapeutic window, approximatively 380 ShK analogs have been synthesized over the last decade leading to ShK-186 and ShK-192 analogs [613–615] as promising therapeutic candidates for autoimmune diseases, with ShK-186 in phase I clinical trials for multiple sclerosis treatment and other autoimmune diseases [616–619]. Thus, the identification of new ShK-like peptides with better selectivity profile would represent new molecular scaffold for the future development of potential therapeutic peptides.

Here we report on the synthesis of the ShK toxin using a recently described native chemical ligation (NCL) method never applied to the synthesis of ShK toxins and that could be applied to the synthesis of other ShK toxin analogs with better selectivity profile.

2. Methods

2.1 Abbreviations

Acac, acetylacetone; Acm, acetamidomethyl; ACN, acetonitrile; Boc, *tert*-butoxycarbonyl; DCM, Dichloromethane; DIPEA, diisopropylethylamine; DMF, *N,N'*-dimethylformamide; DTP, 2,2'-Dithiopyridine; ESI-MS, electrospray ionization mass spectrometry; Fmoc, fluorenylmethoxycarbonyl; GndCl, guanidinium chloride HATU, 1[*Bis*(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate; LC/MS, liquid chromatography/mass spectrometry; MeOH, methanol; MPAA, mercaptophenylacetic acid; nAChR, nicotinic acetylcholine receptor; NMR, nuclear magnetic resonance; Pbf, pentamethyl-dihydrobenzofuran-5-sulfonyl; RP-HPLC, reversed phase high performance liquid chromatography; SPPS, solid phase peptide synthesis; *t*Bu, *tert*-butyl;TFA, trifluoroacetic acid; TIS, triisopropylsilane; Tris, 2-Amino-2-(hydroxymethyl)propane-1,3-diol;Trt, trityl; UV, ultra-violet.

2.2 Chemical synthesis

2.2.1 Synthesis of linear peptides

DMF, DIPEA, ACN, TIS, TFA, piperidine and all others reagents were obtained from Sigma-Aldrich (Saint-Louis, MI, USA) or Merck (Darmstadt, Germany) and were used as supplied. Fmoc (L) amino acid derivatives and HATU were purchased from Iris Biotech (Marktredwitz, Germany). PS-2-Chlorotrityl chloride resin (100-200 mesh, 1.6 mmol/g) was purchased from Iris Biotech (Marktredwitz, Germany). The following side-chain protecting groups were used: Trt for Cys, Gln and His; *t*Bu for Thr, Ser, and Asp; Boc for Lys and Trp; and Pbf for Arg. Peptides were manually synthesized using the Fmoc-based solid-phase peptide synthesis technique on a VWR (Radnor, PA, USA) microplate shaker. All Fmoc amino acids and HATU were dissolved in DMF to reach 0.5 M. For acid peptides the first amino acid was coupled onto the resin for 6 h in a 1/1 (v/v) mix of DMF and DCM, with a 2.5-fold excess of amino acid and 5-fold excess of DIPEA followed by addition of methanol and further mixing for 15 min to cap any remaining reactive functionalities on the resin. For hydrazide peptides the PS-2-Cl-(Trt)-$NHNH_2$ resin was prepared by adding 3 times (1 mL/mmol resin) of 10% hydrazine solution in DMF. The resin was washed with DMF, DCM, MeOH, and DMF. Fmoc deprotection was carried out with piperidine in DMF (1/2 v/v) twice for 3 min. Subsequent amino acids were coupled onto 0.1 mmol of prepared resin (determined loading value 0.73 mmol/g) twice for 10 min using an amino acid/HATU/DIPEA ratio of 5:5:10 relative to resin loading. DMF was used for resin washing between deprotection and coupling steps. After chain assembly was complete, the terminal Fmoc group was removed and the resin washed with DMF and DCM. Side-chains deprotection and cleavage from the resin was carried out by

adding 10 mL of TFA/TIS/H_2O (95/2.5/2.5 v/v/v) and stirring the mixture for 2.5 h at room temperature. Crude peptides were purified by preparative RP-HPLC and pure fractions were combined and freeze-dried.

2.2.2 Thioester preparation

Hydrazide peptide were dissolved in 6M GndCl containing 200 mM MPAA to reach 8 mM concentration. Thioester was formed by treatment with 2.5 eq of 10% acac solution in water. The reaction was stirred at room temperature for 2-4 h until completion of the reaction.

2.2.3 Native chemical ligation

An equimolar amount of the Cys-fragment peptide (the C-terminal fragment of the ligated product) was dissolved in 6 M Gnd HCl with 200 mM Na_2HPO_4 and 50 mM TCEP (pH 8.5) to a volume equal to that which the thioesterification reaction occured. These two solutions were mixed together upon which the resultant pH was around 5-7 and the MPAA emulsion dissolves. The solution should be clear or slightly yellow. The pH of the combined solution was then adjusted to pH 7-7.4 by the addition of 1 M NaOH. The ligation reaction was left to stir for 18 hours. The reaction mixture was acidified to pH 3 and purified by RP-HPLC, pure fractions were combined and freeze-dried.

2.2.4 Oxidative folding

The linear ShK is dissolved in 50 mM ammonium acetate buffer adjusted to pH 8 to reach 0.1 mM concentration and left at room temperature without agitation for 42 hours. The reaction mixture was acidified to pH 3 and purified by RP-HPLC, pure fractions were combined and freeze-dried.

2.3 Mass spectrometry

Solvents used for LC/MS were of HPLC grade.

Intermediate peptides were characterized using a LC/MS system consisting of a Waters (Milford, OH, USA) Alliance 2695 HPLC, coupled to a Waters Micromass ZQ spectrometer (electrospray ionization in positive mode (ESI+) fitted with a quadrupole mass analyzer). All the analyses were carried out using a Chromolith (Fontenay sous Bois, France) HighResolution RP-18e (4.6 x 25 mm, 15 nm–1.15 µm particle size, flow rate 3.0 mL/min) column. A flow rate of 3 mL/min and a gradient of 0–100% B over 2.5 min for routine analyses and 0-30% B over 30 min for quality control of pure products were used. Eluent A was water/0.1% HCO_2H and eluent B consisted of acetonitrile/0.1% HCO_2H. UV detection was performed at 214 nm. Electrospray mass spectra were acquired at a solvent flow rate of 200 µL/min. Nitrogen was used for both the nebulizing and drying gas. The data were obtained in a scan mode ranging from 100 to 1000 m/z or 250 to 1500 m/z to in 0.7 s intervals.

Folded peptide was characterized using a Synapt G2-S high-definition MS system (Waters, Corp., Milford, MA, United States) equipped with an ESI source and an hybrid QTof mass analyzer configuration. Chromatographic separation was carried out at a flow rate of 0.4 ml/min on a Acquity H-Class ultrahigh performance liquid chromatography (UPLC) system (Waters, Corp., Milford, MA, United States), equipped with a Kinetex C18 100Å column (100 mm x 2.1 mm, 2.6 mm particle size) from Phenomenex (France). The mobile phase consisted of water (solvent A) and ACN (solvent B) with both phases acidified by 0.1% (v/v) formic acid. Mass spectra were acquired in the positive ionization mode.

2.4 Preparative RP-HPLC

Preparative RP-HPLC of linear peptides was run on a Gilson PLC 2250 Purification system (Villiers le Bel, France) instrument using a preparative column (Waters DeltaPak C18 Radial-Pak Cartridge, 100 Å, 40 × 100 mm, 15 µm particle size, flow rate 50.0 mL/min). Buffer A was 0.1% TFA in water, and buffer B was 0.1% TFA in acetonitrile.
Folded ShK was purified on an UltiMate 3000 UHPLC system (ThermoFischer scientific) using a Kinetex C18 100 Å column (100×2.1 mm, 2.6 µm particle size, flow rate 0.4 mL/min) from Phenomenex (France). Buffer A was 0.1% HCOOH in water, and buffer B was 0.1% HCOOH in acetonitrile.

3. Results

Recently, seven distinct CAP-ShK proteins specifically expressed in the salivary gland tissue of *Cumia reticulata* have been identified by transcriptomic analysis with an *in silico* assessment of their correctness by Gerdol *et al.* [620]. They comprise a single CAP (Cysteine-rich secretory proteins, Antigen 5, and Pathogenesis-related 1) and several additional domains categorized into ShK toxin-like, EGF-like and cysteine rich domains. Interestingly the identified CreCAP ShK3 protein contain three ShK-like domains with structural homology to ShK native toxin (Figure 1).

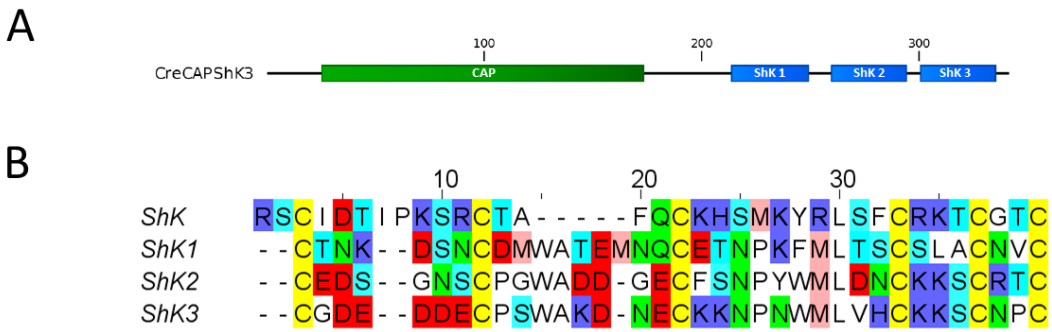


Figure 1: (A) Schematic representation of CreCAP ShK3 protein identified by *Gerdol et al*. [620] (B) Alignment of the three ShK-like domains from the CreCAP ShK3 protein with the native ShK toxin. The cysteine pattern is conserved between ShK and the ShK-like domains.

3.1 Chemical synthesis

Although, ShK toxin has been successfully synthesized by Pennington *et al.* [610] using stepwise Fmoc-SPPS chemistry our successive attempts were unsuccessful (difficulty in isolating the desired peptide from the purification of the cleavage solution, due to large amount of deletion peptides and a small quantity of desired product). Therefore, we envisaged its chemical synthesis using the native chemical ligation method described by Flood *et al.* [501] (Figure 2A).

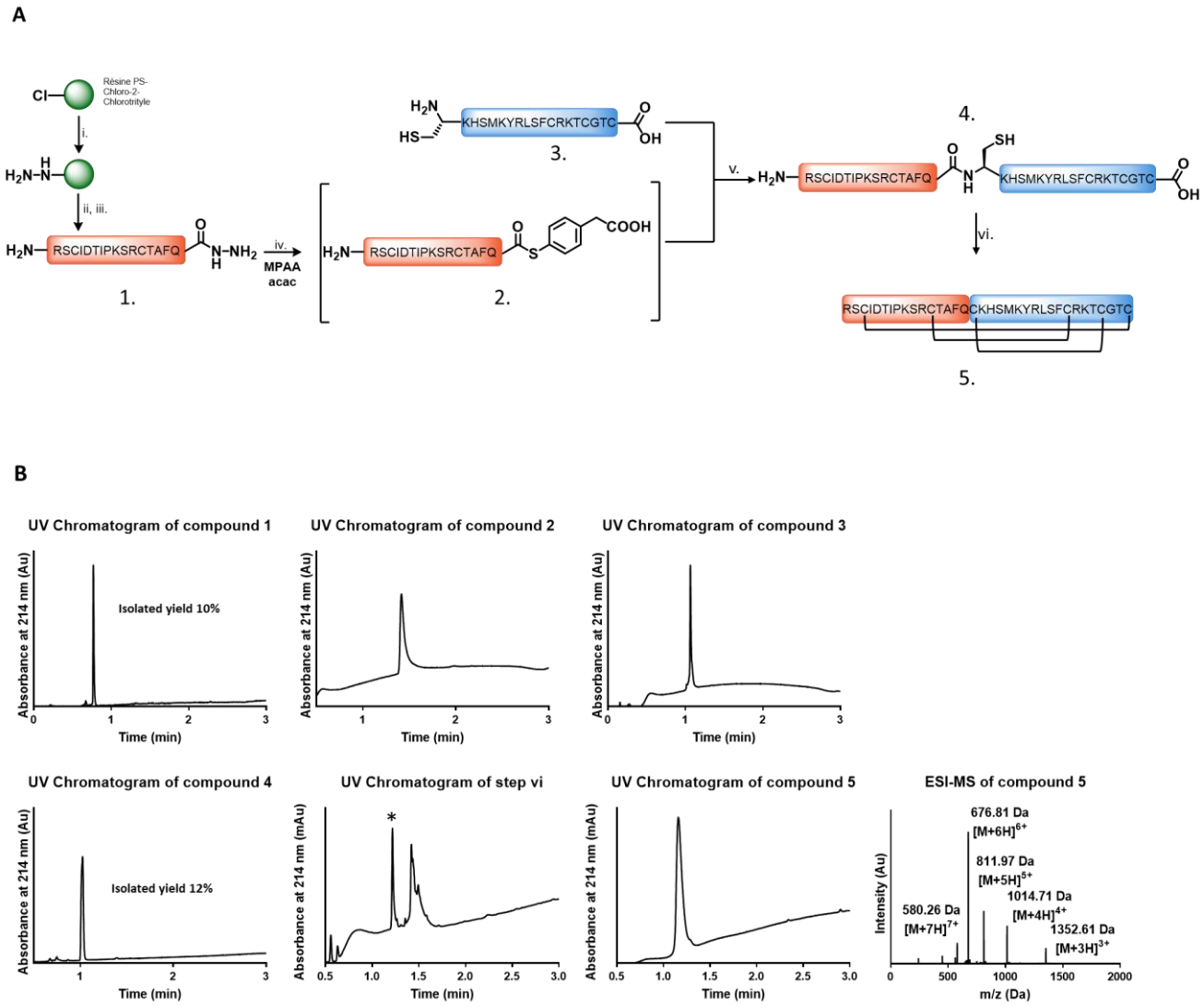


Figure 2: (A) Synthesis strategy of ShK peptide. (i) 5% NH_2-NH_2/DMF. (ii) Fmoc-SPPS. (iii) Cleavage. (iv) 10% acac/H_2O 2.5 eq, MPAA 25 eq, pH 3 (v) Peptide 1 eq, pH 7-7.4. (vi) NH_4OAc 50 mM, pH 8. Overall yield is calculated from isolated linear ShK toxin. MPAA means mercaptophenylacetic acid and acac means acetyl acetone. (B) UV chromatograms of intermediate compounds and UV/ESI-MS of the final product. Asterisk indicates the supposed correctly folded isomer based on Dang *et al*. UV chromatogram profile [621].

Thioester formation is fully complete in 2 hours based on UV chromatogram showing the high efficiency of hydrazide peptide activation using acetyl acetone. Crude ligation mixture is very clean after 18 hours showing limited amounts of remaining reactants. However, the isolated yield is only 50% probably due to loss during the purification process. Folding yield was less than folding yield calculated by Dang *et al*. however, it is based on UV chromatogram and not the amount of isolated folded peptide. Overall, this strategy is giving crude reaction mixtures at least as well as the previously

described ShK synthesis by native chemical ligation [621,622] and is less dangerous and easier to set up.

The synthesis of the ShK analogue peptides is still in progress but we should be able to use the same strategy for the synthesis of the linear peptides, on the other hand, the folding is unpredictable considering that these are new sequences. Considering the significant losses during purification processes, we might better use C8 or C4 preperative columns.

Conclusions et perspectives

La mise en place d'une méthodologie de synthèse des peptides à 3 ponts disulfures s'est révélée beaucoup plus délicate. En effet l'ordre de formation des ponts s'est avéré important, ce qui requiert au préalable de tester les différents ordres possibles, limitant ainsi fortement l'intérêt de l'utilisation d'une stratégie régiosélective en première intention. Néanmoins, aucune règle générale ne peut être dégagée pour le repliement oxydatif, certaines conditions peuvent s'avérer efficace pour une séquence et non pour une autre, comme nous avons pu le constater pour le repliement de la séquence U-asilidin-Mar1a efficace et rapide comparé à celui de la PmuTx1 où aucun isomère majoritaire n'a pu être isolé. Une alternative potentielle est de combiner repliement oxydatif sous controle thermodynamique et repliement régiosélectif, par exemple pour la toxine PmuTx1, seules les cystéines II et V pourraient être protégées par un groupement Acm et les ponts I-IV et III-VI pourraient être formés par repliement oxydatif sous controle thermodynamique. Cette méthode devrait conduire dans un premier temps à l'obtention de 3 isomères, dans le pire des cas séparables par chromatographie, et au peptide final par formation du dernier pont.

5. Chapitre 4 : Analyse protéo-transcriptomique de venin

Dans l'optique de finaliser ma formation sur la découverte de nouvelle séquences bioactives dans les venins, une formation dans le laboratoire de spectrométrie de masse à l'université de Liège encadré par le Pr. Loic Quinton et Dr. Fernanda Gobbi Amorim m'a permis d'appréhender les aspects pratiques de l'analyse protéomique et l'utilisation du logiciel PEAKS Studio. Dans le cadre de cette formation, l'étude protéomique du venin de *Vipera aspis aspis* (vipère aspic, serpent le plus redouté de France) a été effectuée à l'aide du logiciel PEAKS studio, ce qui a permis de confimer un grand nombre de séquences transcriptomiques. Le fonctionnement du logiciel PEAKS Studio étant essentiel à une interprétation éclairée des résultats, son fonctionnement est détaillé ci-après.

a. Introduction

L'identification des peptides utilisant la spectrométrie de masse en tandem [623] est une tâche essentielle de la protéomique. L'exactitude et la sensibilité de cette tâche ont une incidence directe sur les performances de l'identification des protéines à partir des résultats obtenus avec les peptides, ainsi que sur d'autres analyses en aval. De nombreux logiciels ont été développés pour l'identification des peptides et ces outils peuvent être divisés en deux catégories: (i) le séquençage *de novo* et (ii) la recherche dans les bases de données. Le séquençage *de novo* déduit la séquence peptidique directement du spectre MS/MS, alors qu'une recherche dans les bases de données va déterminer le meilleur peptide qui va correspondre aux pics du spectre MS/MS. La recherche dans les bases de données est généralement considérée comme une approche plus simple car les bases de données de séquences de protéines offrent un espace limité au logiciel pour la recherche. Par conséquent, lorsqu'une base de données de séquences de protéines est disponible, une recherche dans cette base est la méthode la plus courante d'identification de peptides. Cependant, les outils existants rencontrent encore des problèmes de faibles taux d'identification (sensibilité faible) et de taux de fausse découverte (FDR) élevés (faible précision) [624,625]. Contrairement à la méthode classique de recherche dans les bases de données, le logiciel PEAKS DB s'appuie fortement sur les résultats de séquençage *de novo* pour améliorer la filtration et la fonction de « scoring ». Cette combinaison entraîne une sensibilité et une précision nettement améliorées par rapport aux logiciels de recherche dans les base de données existants. Outre les deux objectifs (précision et sensibilité), la génération à haut débit de données protéomiques nécessite la validation automatique des résultats de recherche dans les bases de données. Actuellement, cette validation est généralement réalisée par la méthode du « target decoy » [626,627]. Cette méthode introduit les protéines leurres à rechercher par le même moteur de recherche et utilise les résultats du moteur sur les protéines leurres pour estimer le nombre de faux positifs. Le but de PEAKS DB est d'identifier des peptides dans une base de données à partir

des données MS/MS. En tant que tel, PEAKS DB appartient à la catégorie de recherche dans les bases de données. Cependant, PEAKS DB utilise le séquençage *de novo* comme sous-programme et exploite les résultats de ce séquençage pour améliorer à la fois la rapidité et la précision de la recherche dans la base de données. Les principales étapes algorithmiques du logiciel PEAKS DB sont les suivantes : 1) Séquençage *de novo* de chaque spectre MS/MS en utilisant l'algorithme PEAKS [623]. 2) Les séquences *de novo* sont utilisées pour trouver des correspondances approximatives avec les protéines de la base de données générant une liste restreinte de protéines. 3) Tous les peptides dérivés de la liste restreinte de protéines sont comparés avec les spectres MS/MS. 4) un score est attribué à chaque peptide identifié, en fonction de sa correspondance avec les candidats *de novo.* 5) Une approche « target decoy » modifiée est utilisée pour déterminer le nombre de peptide identifié correspondant à la valeur de FDR requise. 6) Les peptides identifiés avec une grande confiance par les étapes ci-dessus sont utilisés pour déduire les protéines présentes dans l'échantillon.

Pour bien comprendre le fonctionnement du logiciel chacune des 6 étapes précédemment décrites va être détaillées.

- Etape 1 : séquençage *de novo*

  L'algorithme PEAKS est utilisé pour effectuer le séquençage *de novo* pour chaque spectre MS/MS. Les mêmes paramètres (tolérance d'erreur de masse et PTM) spécifiés par l'utilisateur pour la recherche dans la base de données sont également utilisés pour le séquençage *de novo*. Pour chaque spectre, seul les 5 premiers peptides issus du séquençage *de novo* décrit par PEAKS sont utilisés. L'algorithme PEAKS calcule également un score de confiance en % pour chaque acide aminé dans la séquence *de novo* (Figure 14). PEAKS convertie cela en « sequence tag » en remplaçant les acides aminés de faible confiance par leurs valeurs de masse. Plus spécifiquement, chaque segment de résidus d'acides aminés adjacents avec une confiance <30% est remplacé par un « segment de masse » qui est égal à la masse totale des résidus (Figure 15).

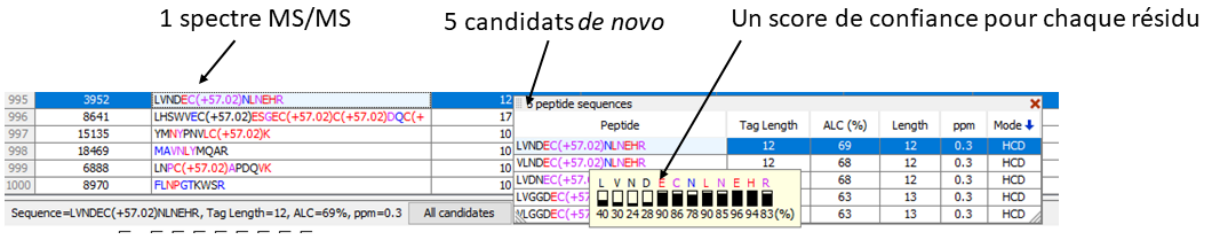


Figure 14 : Etape du séquençage *de novo*. Un spectre MS/MS génère 5 candidats *de novo* notés par rapport à la somme de leur score de confiance pour chaque résidu.

Figure 15 : Une séquence de novo calculée avec PEAKS a un indice de confiance local pour chaque acide aminé, représenté par les hauteurs des barres verticales. En utilisant un seuil de 30%, les acides aminés consécutifs au-dessous du seuil de confiance sont remplacés par leur masse totale de résidus [74].

- Etape 2 : génération de la liste restreinte de protéines

Dans cette étape, l'algorithme utilise les « *de novo* sequence tag » pour sélectionner une liste restreinte de protéines dans la base de données. Les prochaines étapes du processus ne seront effectuées que sur cette courte liste restreinte afin de réduire le temps de calcul total. La qualité de la correspondance entre une « *de novo* sequence tag » et un peptide de la base de données est mesurée par le nombre d'acides aminés communs (score CAA). La Figure 16 illustre le calcul du score CAA. Les protéines sont classées selon le score CAA le plus élevé obtenu par les peptides de chaque protéine.

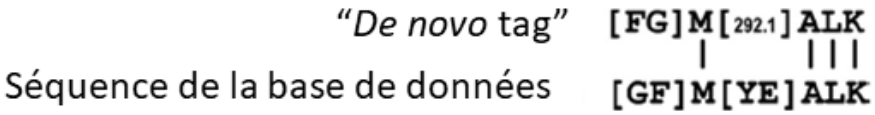


Figure 16 : Une « *de novo* sequence tag » est comparée à un peptide de la base de données. L'alignement garantit que la masse de chaque bloc aligné (entre crochets) est égale pour les deux séquences. Le score CAA est le nombre d'acides aminés communs dans l'alignement, qui est de 4 dans cet exemple [74].

- Etape 3 : génération de la liste restreinte de peptides

Toutes les séquences peptidiques digérées *in silico* (en prenant en compte l'enzyme spécifiée) à partir de la liste restreinte de protéines sont comparées aux spectres MS/MS pour trouver les « peptide spectrum matching » (PSM). Chaque séquence peptidique peut produire de multiples peptides modifiés en prenant en compte toutes les combinaisons possibles des PTM variables spécifiées par l'utilisateur. Pour chaque séquence peptidique (modifiée ou non), la masse est calculée et les spectres MS/MS avec la masse de précurseur correspondante sont comparés à la séquence. Un « quick scorer » est utilisé pour calculer le score du PSM. Le « quick scorer » est dérivé de la même fonction de notation de séquençage que celle utilisée dans le séquençage *de novo* PEAKS [623].

- Etape 4 : Attribution de scores aux peptides

Les 9 paramètres suivants sont pris en compte dans le calcul du score : 1) le nombre d'acide aminés correspondant au « *de novo* sequence tag » (score CAA). 2) Le « protein feature » : chaque protéine obtient un score en ajoutant ces trois plus haut score CAA, et le « protein feature » d'un peptide est le score maximal des protéines contenant ce peptide. 3) la longueur du peptide. 4) La longueur moyenne de séquence par clivage raté dans le peptide. 5) La longueur moyenne de séquence par PTM dans le peptide. 6) L'erreur de masse sur l'ion précurseur. 7) L'état de charge. 8) La longueur maximale d'ions fragments identifiés de manière consécutive. 9) Le nombre d'extrémités enfreignant les règles de la digestion enzymatique. Ces 9 caractéristiques sont combinées avec le « ion match score » normalisé de manière pondérée (« weighted sum score »). Le poids de chaque paramètre a été déterminé pour chaque instrument par un processus itératif sur un grand nombre de données LC-MS/MS de façon à maximiser le nombre de PSM respectant le FDR fixé. Le « weighted sum score » est converti en une valeur *p* pour une interprétation plus facile. Pour un score donné la valeur *p* correspondante est définie comme la probabilité qu'une identification fausse obtienne un score supérieur ou égal. La valeur *p* permet d'estimer la probabilité d'un faux positif (ce qui est différent du FDR). Le score d'un peptide identifié par PEAKS est affiché comme -10lg(*p).*

- Validation des résultats

Une approche de « target decoy » modifiée, appelée « decoy fusion », est utilisée pour estimer le FDR à un seuil de score donné. L'approche « target decoy » plus conventionnelle nécessite la génération d'une séquence protéique leurre pour chaque séquence protéique cible dans la base de données [628]. Les bases de données cible et leurre sont ensuite recherchées séparément ou ensemble par le logiciel, et le FDR est calculé par le rapport entre le nombre de correspondances du leurre et de la cible. Cependant, dans PEAKS DB, les séquences cibles et leurres ne sont pas traitées comme des entrées distinctes dans la base de données. En effet, ils sont concaténés ensemble pour chaque protéine. Ainsi, la base de données nouvellement générée contient le même nombre d'entrées de protéines, mais la longueur de chaque protéine est doublée. Le logiciel recherche cette base de données nouvellement générée. Après la recherche, les identifications cibles et leurres sont séparées en vérifiant si elles proviennent de la première ou de la seconde moitié de chaque séquence concaténée. Pour chaque seuil de score spécifié par l'utilisateur, le FDR est calculé comme le rapport entre le nombre de résultats sur la partie leurre et le nombre de résultats sur la partie cible, au-dessus du seuil de score.

- Identification et regroupement des protéines

Les protéines sont regroupées selon leurs peptides en commun. Le score final de chaque protéine est calculé à partir de ses peptides identifiés comme il suit : premièrement, les peptides redondants sont

éliminés ; si le même peptide est identifié plusieurs fois à partir de spectres différents, seul le peptide dont le score est le plus élevé est retenu. Deux peptides sont considérés comme identiques s'ils sont identiques ou ne diffèrent que par l'emplacement d'une PTM, mais considérés comme différents si la séquence d'acides aminés ou les PTM sont différentes. Deuxièmement, tous les scores non redondants des peptides sont triés comme suit : s1 ≥ s2 ≥…≥ sk. Enfin, le score de la protéine est égal à s1 + (1/2)s2 +…+ (1/k)sk.

Pour conclure le logiciel PEAKS Studio est un outil bioinformatique permettant d'identifier les séquences présentes dans une base de données par interprétation automatisée des spectres MS/MS, tout en corrigeant ses séquences à l'aide de son interprétation *de novo* des spectres MS/MS ce qui permet l'identification des PTMs, mutations (délétions, substitutions, insertion etc.) qui sont fréquentes dans les venins d'animaux. De plus, la possibilité de pouvoir utiliser sa propre base de données *i.e.* les transcriptomes pour l'étude venomique, font de PEAKS un outil de choix pour l'analyse protéo-transcriptomique de venins.

b. Protéo-transcriptomique du venin de Vipera aspis aspis

Les objectifs de l'analyse du venin de *Vipera aspis aspis* étaient triple : dans un premier temps appréhender l'utilisation du logiciel PEAKS Studio pour l'intégration de donnée transcriptomique dans l'analyse protéomique, la découverte de nouvelles séquences et enfin caractériser le venin de cette espèce potentiellement mortelle en France.

Cette étude est présentée ci-après sous forme d'une publication en preparation dans laquelle j'ai analysé les données protéo-transcriptomique, cependant les procédures de la récolte du venin et de la partie transcriptomique manquent encore à la partie méthodes.

Venomics of the asp viper *Vipera aspis aspis* from France

Julien Giribaldi[a], Taline Kazandjian[b], Fernanda G. Amorim[c], Christine Enjalbal[a], Loic Quinton[c], Nicholas R. Casewell[b] and Sebastien Dutertre[a]*

[a] *Institut des Biomolécules Max Mousseron, Département des Acides aminés, Peptides et Protéines, Unité Mixte de Recherche 5247, Université Montpellier - Centre Nationale de la Recherche Scientifique, Place Eugène Bataillon, 34095 Montpellier CEDEX 5, France.*
[b] *Alistair Reid Venom Research Unit, Parasitology Department, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, L3 5QA UK*
[c] *Laboratory of Mass Spectrometry, Department of Chemistry, University of Liège, Liège, Belgium.*
**Corresponding author: Sebastien Dutertre. Tel: +33 4 67 14 38 09*
Email address: sebastien.dutertre@umontpellier.fr

Abstract

The asp viper *Vipera aspis aspis* is a common venomous snake found in France and is responsible for several hundred bites each year. Despite its medical importance and potential neurotoxic venom, the complete toxin repertoire produced by the venom gland is unknown. In this study, we used an integrated venomics approach to decipher the complex composition of its venom. Transcriptomic analysis revealed 80 venom-annotated sequences grouped into 16 families. Among the most represented venom proteins were snake venom metalloproteases (23%), phospholipases A2 (15%), serine proteases (13%), snake venom metalloprotease inhibitors (13%) and C-type lectins (12%). Other components identified in the transcriptome include Kunitz-type proteins, disintegrins, vascular endothelial growth factors, nerve growth factors, L-amino acid oxidases, cysteine rich secretary proteins, 5'-nucleotidases, renin-like aspartic proteases, phospholipases B, hyaluronidases and phosphodiesterases. Proteomic analysis of the crude venom validated 57 of these transcriptomic sequences (>70%), including at least one sequence for each of the 16 families, but also identified 7 additional sequences not initially annotated as venom proteins in our curated transcriptome. These include a serine protease, a disintegrin, a glutaminyl-peptide cyclotransferase, a proactivator polypeptide-like and 3 aminopeptidases. Thus, in total, 87 sequences were retrieved from *Vipera aspis aspis* transcriptome and proteome, constituting a valuable resource that may help in the design of more efficient anti-venom and for the mining of useful pharmacological compounds.

Keywords: Proteomics; transcriptomics; venom; toxin; snake; Vipera

1. Introduction

Several venomous snakes occur in France, including a rear-fanged snake (Lamprophiidae, *Malpolon monspessulanus*) and four species of vipers (Viperidae). Among the latter, *Vipera berus berus* (*Vbb*) and *Vipera aspis aspis* (*Vasp*) are the most common and are responsible for the majority of reported

envenomations. In particular, the asp viper *Vipera aspis aspis* is of medical importance in France, with 100-1000 recorded bites each year, resulting in one death per year on average [629]. Symptoms following envenomation include local signs such as pain, œdema, phlyctenosis combined in severe cases with extensive swelling and/or systemic signs such as gastrointestinal disorders and hypotension [630]. A French grading system for viper envenomation has been established according to several clinical report descriptions [631]. Most cases stabilize at grade 1 (minor), with only 15 to 20% progressing to grade 2 (moderate), whereas grade 3 (severe) is defined as grade 2 symptoms that continue to progress further [631]. Consistent with these symptoms, the venom of these vipers is known to contain some procoagulant compounds, kinin-releasing factors, hypotensive factors, proteases and hyaluronidases [631–635].

From the 1990s, the Marseille Poison Center repeatedly witnessed unusual cases of envenomation after *Vasp* bites that included neurological signs such as ptosis, drowsiness, diplopia and ophtalmoplegia [636]. Given these neurotoxic symptoms, De Haro *et al.* made the hypothesis that the toxins causing neurological damages in *Vasp* bites might be similar to the PLA_2-I found in *Vipera aspis zinnikeri* and/or ammodytoxin [636]. The phospholipase PLA_2-I is also highly homologous to the vipoxin of the *Vipera ammodytes meridionalis*, and these toxins are indeed known to block the neuromuscular junction at the presynaptic or postsynaptic level [637]. Accordingly, the same group later identified two neurotoxins in the venom of *Vasp,* namely ammodytoxin B (monomeric PLA_2 originally found in the venom of *Vipera ammodytes ammodytes, Vasp*) and a novel non-covalent PLA_2 heterodimer named vaspin, which is analogous to vipoxin and PLA_2-I [638]. Ferquel *et al*. screened for venom PLA_2 at the genome and transcriptome levels and identified three different subpopulations of *Vasp*, based on their neurotoxic venom PLA_2 content [639]. The first group is defined by the presence of genes encoding AmI1 (Ammodytoxin B Isoform 1) and AmI2, the second group by gene encoding AmI1, AmI2, vasA (Vaspin A), vasB and the third group by gene encoding AmI1 and vaspin. Overall, it was concluded that the genome of *Vipera aspis aspis* comprises a repertoire of PLA_2 neurotoxins ready to be expressed under stimuli that remain to be identified.

Importantly, the venom variability within the different *Vasp* populations has significant implications for the design of antivenom. Indeed, intraspecific variations in venom composition is well known to cause major issues in the preparation of effective antivenoms [640]. Viperfav has been approved for marketing and is the only antivenom currently available in France. It contains F(ab')2 fragments produced by horses against *Vipera berus*, *Vipera aspis* and *Vipera ammodytes* venoms [631,641]. However, the efficiency of Viperfav regarding the neutralization of the neurotoxic effects was found variable [642]. Furthermore, the apparent intra-species variation might be also be due to hybridization

phenomenon, which could be an important mechanism of the evolution of the composition of snake venoms [639].

In this study, we report the investigation of the *Vasp* venom composition through a venomics approach that combines transcriptomic and proteomic analyses [643]. The powerful technology of next generation sequencing allows for the rapid identification of transcripts produced in the venom gland, facilitating the matching of MS/MS spectra to protein/toxin sequences [644]. This work will help provide the medical and research communities with a better understanding of the asp viper venom toxicity and evolution, and could lead to the design of more effective antivenom [46,58]. Moreover

2. Methods

2.1 Bioinformatic analysis

PEAKS Studio 8.5 (Bioinformatics solutions, Waterloo, ON, Canada) a *De novo* assisted database software [74,623] was chosen to analyze MS/MS data from *Vasp* venom. Indeed, by using the PEAKS studio software we were able to match MS/MS spectra obtained from proteomic analysis of *Vasp* venom to our own database resulting from the venom gland transcriptome of the same specimen. Briefly PEAKS studio start sequencing from MS/MS spectra with no database at all obtaining the most theoretically plausible peptides for each MS/MS spectra. The confidence score of peptides obtained by this process is given by an ALC (Average Local Confidence) score. These *de novo* sequences are then corrected by the database to provide additional information about PTM's, mutations, homologous peptides and novel peptides. Raw data were loaded in PEAKS 8.5 software. Carbamidomethylation was set as fixed modification, while oxidation (M) was set as variable modifications, with maximum missed cleavages at 3 for trypsin digestion. Parent mass and fragment mass error tolerance were set at 5 ppm and 0.015 Da, respectively. A -10lgP (where P represents the probability of a false positive result) score is attributed to peptides matching with the database search. Most of proteomics studies using PEAKS Studio software are made by setting FDR at 1% [645,646]. The PEAKS protein -10lgP score is calculated as the weighted sum of the -10lgP scores of the protein's supporting peptides. To limit the impact of low score identified peptides and after manual reviewing of the lowest protein identified score with 1% FDR, we decided that proteins identified with a -10lgP ≥ 120 were identified by enough reliable peptides MS/MS spectra. In order to identified more relevant sequences Spider algorithm [647] from PEAKS Studio software was used to find mutations or to correct the sequences. This algorithm correct the sequences stored in transcriptomic database with *de novo* sequences based on MS/MS spectra, which allowed us to identify PTM's and mutations. For Spider search, -10lgP peptide score was set according to 1% FDR and -10lgP proteins scores was set ≥ 120. A minimum number of 2 unique peptides was set to identify a protein. Minimum ion intensity for mutation and PTM's was set to 5%, and ALC

score ≥ 90 for *de novo* sequences leading to low precursor mass error in order to identify reliable PTM's and potential mutations.

2.2 Bottom-up mass spectrometry

2,5 µL of crude *Vipera aspis aspis* venom at 20,5 mg/mL in milli-Q water were diluted 20 times in 100 mM ammonium bicarbonate buffer then reduced and alkylated by 5 µL of DTT (Dithiothreitol) 500 mM in milli-Q water (1 h at 56°C) and 10 µL of IAA (Iodoacetamide) 500 mM in milli-Q water (1 h at room temperature, in darkness), respectively. Then, the venom sample was submitted to trypsin digestion by adding 5 µL of trypsin Pierce™ mass spectrometry grade (Thermo Scientific, Waltham, MA, USA) at 0.1 mg/mL in 100 mM ammonium bicarbonate buffer (4 h at 37°C). Reaction was stopped by adding 6 µL of TFA 10% in milli-Q water to the reaction mixture. 0,5 µg of the digested material was analysed in the Acquity UPLC® M-Class (Waters, Milford, MA, USA) coupled to the Q-Exactive™ Plus Hybrid Quadrupole-Orbitrap™ Mass Spectrometer (Thermo Scientific, Bremen, Germany). The chromatographic system is equipped with a monolithic PepSwift Capillary column 100 µm x 25 cm (Thermo Scientific, Waltham, MA, USA). Peptides were eluted using a gradient of 3–50% of solution B in 100 min and 50-80% of solution B in 30 min (A: water/0.1% formic acid; B: acetonitrile), at a flow rate of 0.6 mL/min, and data were acquired in the positive-ion mode. Concerning mass spectrometry, all the analyses were performed in data dependent analysis (DDA) mode that automatically triggers the MS/MS experiments. The automatic gain control (AGC) target values were 3.10^6 for MS spectra and 2.10^5 for MS/MS spectra. The maximum injection times were set at 200 ms for the MS step and 1000 ms for MS/MS events. The seven most intense peaks of each MS scan were fragmented by high-energy dissociation (HCD) and their corresponding MS/MS spectra were acquired in the Orbitrap analyzer.

Relative abundance of proteins families were calculated according to method described by Zybailov *et al.* [648]. Briefly, the method is based on the principle that most abundant proteins will give more tryptic peptides and consequently more MS/MS spectras. Therefore, spectral count of each proteins families can be related to its relative abundance. However, it is also needed to take in account that big proteins will give more proteotryptic peptides, which is why Zybailov et al. normalized the spectral count, by the protein masses. This method has been used to estimate protein abundance in many label-free shotgun proteomic [649–651]. However, results have to be taken carefully because it assumes that linearity of response is the same for every protein and it does not take into account the theoretical number of observable peptide per proteins.

2.3 Liquid chromatography coupled mass spectrometry

RP-UPLC was operated on an Acquity H-Class ultrahigh performance liquid chromatography (UPLC) system (Waters, Corp., Milford,MA, United States) fitted with a UV detector (PDA detector) under the control of Waters MassLynx software (version 4.1). Separation of the crude *Vasp* venom was achieved using a Kinitex C18 100 Å column (2.1 x 150 mm, 3 µm) fitted with a pre-column. Elution was carried out using a gradient of 0-80% B (0.1% formic acid in ACN) in 80 min. Samples eluting from the UPLC were introduced into the mass spectrometer at a flow rate of 50 µl/min. Acquisitions were carried out over the range 50 Da to 1800 Da *m/z* every second on a Synapt-G2-S high-definition MS system (Waters, Corp., Milford, MA, United States). To obtain the molecular masses of the venom components eluting between 0 and 70 min, each peak from the total ion current (TIC) chromatogram was analysed with Waters Mass Lynx software (version 4.1).

3. Results and discussion

3.1 Toxins identification from transcriptomic analysis

Analysis of the assembled transcriptome revealed that 1172 contigs were annotated as unidentified sequences (63.42%), 596 contigs were non-toxin elements (32.25%) and 80 contigs were toxins (4.33%) (Figure 1). However, in terms of expression levels, toxin-type contigs represented 53.83% of the whole venom gland transcriptome mRNA, non-toxin contigs (including housekeeping proteins) 24.51% and unidentified contigs 21.65% (figure 1).

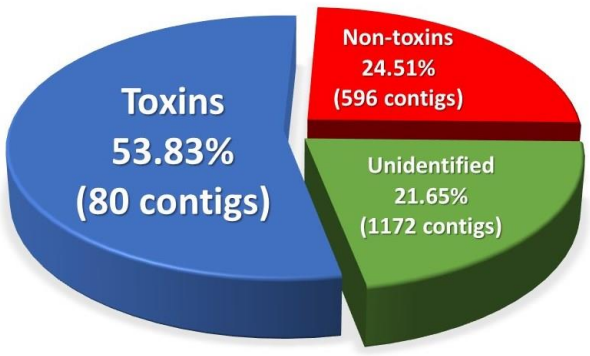


Figure 1: Representation of *Vasp* venom gland transcriptome total transcripts distribution and expression levels. Despite a relatively low number of contigs, the toxins abundance represents more than 50% of the total venom gland expressed mRNA.

3.1.1 Toxins perturbing hemostasis

Snake venom metalloproteases (SVMPs) are largely responsible for the impaired hemostasis in envenomed victims [652]. SVMPs are classified into 3 different subclasses; P-III which is constituted of

a protease domain, a disintegrin and a cysteine rich domain, whereas P-II lost the cysteine rich domain and P-I lost the disintegrin and the cysteine rich domains [653]. The loss of non-protease domains led to a decrease of the hemorrhagic potency of these toxins, and consequently, SVMP-PIII are more hemorrhagic than P-I and P-II forms [654,655]. These metalloproteases degrade the basal membrane components and the extracellular matrix such as type collagen IV, fibronectin, nidogen as well as plasma proteins involved in blood coagulation, including fibrinogen, prothrombin and FX leading to hemorrhage [652]. From the 17 different SVMP expressed in the transcriptome of *Vasp* representing 23.34% of toxin expression levels (Figure 2), 11 belongs to the P-III class, 2 to the P-II class and 4 to unknown class, which is consistent with the hemorrhagic effect of *vipera aspis* bites [631]. Indeed Sajevic *et al.* showed that most of the hemorrhagic activity of *Vipera ammodytes ammodytes* (*Vaa*) venom is due to a P-IIIc metalloprotease [656].

Snake venom serine proteases (SVSPs), by influencing blood coagulation, fibrinolysis, platelet aggregation and vasoconstriction, also cause an imbalance of the hemostatic system in mammalian cells [657,658]. SVSPs are expressed at lower levels than SVMPs in *Vasp* transcriptome, consistent with the demonstration that fibrinogenolytic activity of the *Vbb* and *Vasp* venoms is mainly caused by SVMPs [657]. From the 14 different serine proteases expressed in the transcriptome of *Vasp* representing 13.31% of toxin expression levels (Figure 2), none has been described as fibrinogenolytic, in agreement with *Vbb* and *Vasp* venom activities.

Snake venom C-type lectin like proteins (CTLs), also known as SNACLEC, are built around a lectin scaffold, although venom CTLs diverged significantly through evolution and lost their ability to bind carbohydrates [655,659]. SNACLEC are non-enzymatic proteins that can affect cell adhesion, thrombosis, hemostasis, endocytosis, or pathogen neutralization [660–662]. From the 19 different snaclec expressed in the transcriptome of *Vasp* representing 12.28% of toxin expression levels (Figure 2), the most expressed ones are related to hemostasis impairing toxins influencing platelet aggregation. Although no agonist of platelet aggregation has been reported in *Vasp* and *Vbb* venoms, the SNACLECs of *Vasp* might play a role in inducing the reported thrombocytopenia [657].

Disintegrins are mono or dimeric proteins present in the venoms of various vipers that selectively block the function of integrin receptors [657,663]. They can be divided into five different groups according to their length and their number of disulfide bonds [663,664]. Integrins are involved in cell-cell and cell-extracellular matrix interactions, serving as the final common pathway leading to aggregation. In snake venoms, these small cysteine-rich disintegrins strongly inhibit platelet aggregation and thereby prevent blood coagulation in bite victims [61,665]. From the 2 different disintegrins expressed in the transcriptome of *Vasp* representing 4.97% of toxin expression levels (Figure 2), both are closely related

to disintegrin labein isolated from *Macrovipera lebetina* and belongs to class II disintegrin comprising about 70 amino-acids with 6 disulfide bridges.

Several other toxins were found at lower levels of expression in the *Vasp* transcriptome, including L-amino acid oxidases (LAAOs) described as enzyme that inhibits ADP- and collagen-induced platelet aggregation [666], CRISPs, which are 20 to 30 kDa proteins containing sixteen highly conserved cysteine residues hypothesized to participate in envenoming by disrupting homeostasis [667,668] and a phosphodiesterase similar to venom phosphodiesterase 1 from *Crotalus adamanteus*, which inhibits platelet aggregation.

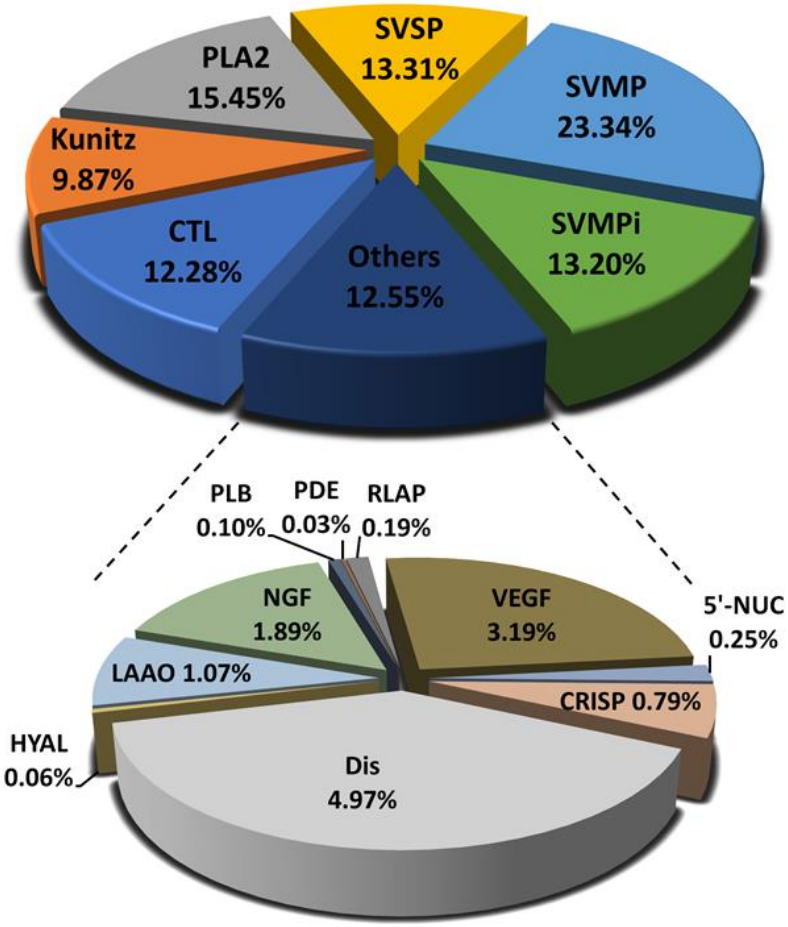


Figure 2: Representation of toxins families in *Vasp* venom gland transcriptome according to their relative expression levels. 5'-NUC, 5'-nucleotidase; CRISP, cysteine rich secretary protein; CTL, C-type lectin; Dis, disintegrin; HYAL, hyaluronidase; LAAO, L-amino acid oxidase; NGF, nerve growth factor; PDE, phosphodiesterase; PLA2, phospholipase A2; PL-B, phospholipase B; RLAP, renin-like aspartic protease; SVMP, snake venom metalloprotease; SVMPi, snake venom metalloprotease inhibitor; SVSP, snake venom serine protease; VEGF, vascular endothelial growth factor.

3.1.2 Enzyme inhibitor toxins

Kunitz peptides have been found in diverse animal venoms such as snake, cone snails, sea anemone and spiders [669–673]. The original function described for kunitz-type peptides was the inhibition of diverse serine proteases, however through natural evolution, they began to develop toxic activities such as ion channel block [673]. Kunitz-type peptides display a conserved scaffold of about 60 amino acids stabilized by three disulfide bridges [657,673]. All the 6 different kunitz type peptides expressed in the transcriptome of *Vasp* representing 9.87% of toxins expression levels (Figure 2) are structurally close to serine protease inhibitor, which suggests that the kunitz-type peptides of *Vasp* are not employed as toxic weapon. From the other 4 different snake venom metalloprotease inhibitors expressed in the transcriptome of *Vasp* representing 13.20% of toxin expression levels (Figure 2), one of them is related to an endogenous tripeptide snake venom metalloprotease inhibitor ~~and could be an interesting candidate to alleviate *Vasp* envenomation~~ [674]~~.~~ described from a number of other Viperinae snake species [674,675].

3.1.3 Neurotoxic toxins

Phospholipases A2 primarily assure digestive functions but also a toxic function, which can manifest in many different ways such as neurotoxicity, myotoxicity, cardiotoxicity and anticoagulant activity [657,676]. Indeed the neurotoxicity of *Vasp* venom is attributed to the presence of three basic monomeric PLA-2, namely ammodytoxins A,B and C, which are able to block the neuromuscular junction at the presynaptic level [677,678]. From the 7 different PLA-2 contigs (including 3 complete sequences, see Figure 3) retrieved from the transcriptome of *Vasp* representing 15.45% of toxin expression levels (Figure 2), it is interesting to note that the most expressed one is closely related to ammodytoxin B, which may explain the neurotoxicity of some *Vasp* bites (Figure 3). From the two other complete sequences, one is 100% identical to ammodytin I2 from *Vasp*, whereas the other shows 99.28% sequence identity with ammodytin I1 from *Vasp* (only one conservative substitution of I>L). The others uncomplete PLA-2 sequences are similar to ammodytin I1, ammodytin L(2) and ammodytin L1, which are not expected to be neurotoxic considering their similarity with ammodytin I2 [679].

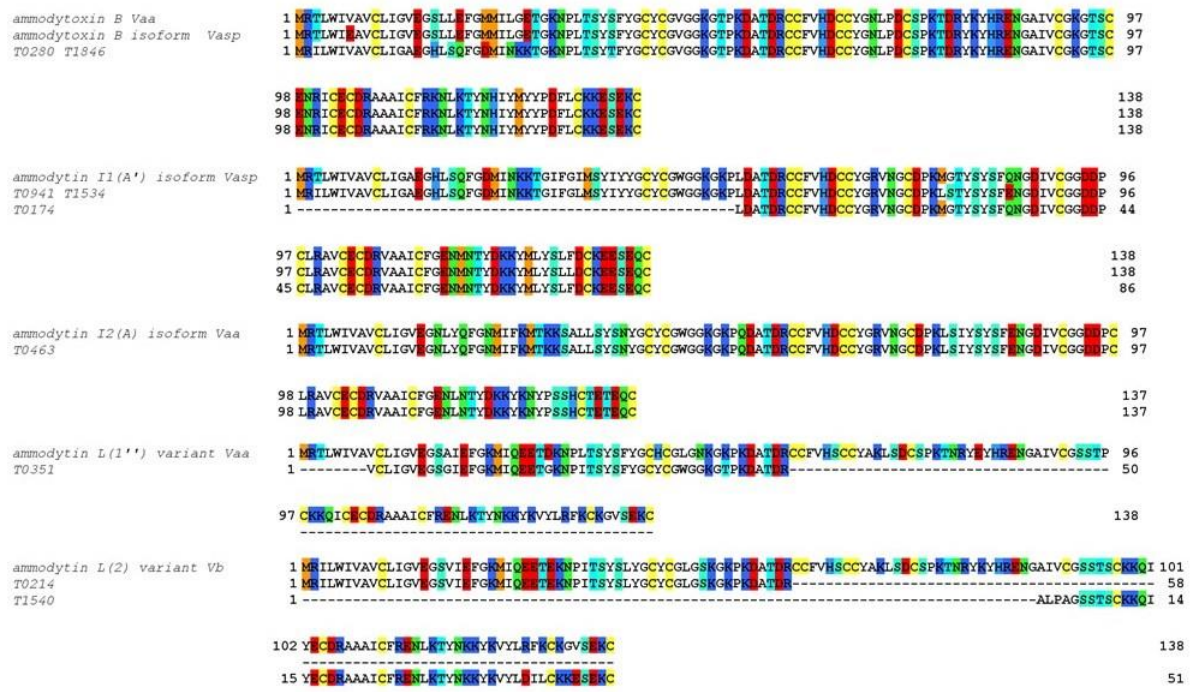


Figure 3: Sequence alignments of PLA2. From the 7 PLA2 related sequence retrieved, 3 are complete and 4 are truncated.

3.1.4 Toxins with other functions

Some other minor components with many roles have been identified in the transcriptome of *Vasp* such as vascular endothelial growth factors (VEGFs) that promote proliferation of vascular endothelial cells and hypotension [680], a renin-like aspartic protease (RLAP) similar to *Echis ocellatus* RLAP, which is highly similar to human and mouse renin, and possibly facilitates the spread of the venom by inducing local hypertension [657]. One hyaluronidase sequence was identified in *Vasp* transcriptome, and this enzyme family is found in wide range of animal venoms where it facilitates the diffusion of other toxins present in the venom by hydrolysis of the hyaluronic acid [681]. Finally, one nerve growth factor (NGF) and one nucleotidase were also found but their putative role in venoms remains to be elucidated.

3.2 Toxins identification from venom proteome analysis

3.2.1 Comparative LC-MS of Vasp venom

The venom of one specimen of *Vasp* was collected sequentially using different methods and analyzed by LC-MS. Overall, the venoms collected after a defensive bite (letting the snake bite), classic (forcing the snake bite) and forced milking (same as classic milking but by stimulating the venom glands) show a very similar total ion chromatogram trace (Figure 4). The profile of each sample can be separated

into 3 major groups, based on the retention time and molecular masse range. The first group (< 20 minutes elution) contains mostly small molecular masses (< 2 kDa), whereas the second group (~20-33 min) contains a majority of masses in the 2-4 kDa range. Yet, about half of the TIC is constituted of masses > 10 kDa, including 3 distinct subgroups. Eluting between 33 and 40 minutes, molecular masses around 13-14 kDa likely correspond to the PLA2 enzymes and CTL. The remaining subgroups (30-60 kDa and 13-30 kDa) may include SVMP (50-60 kDa), PLB (64 kDa), LAAO (57 kDa), RLAP (43 kDa), QC (42 kDa), SP (28-30 kDa), NGF (27 kDa) and CRISP (25 kDa).

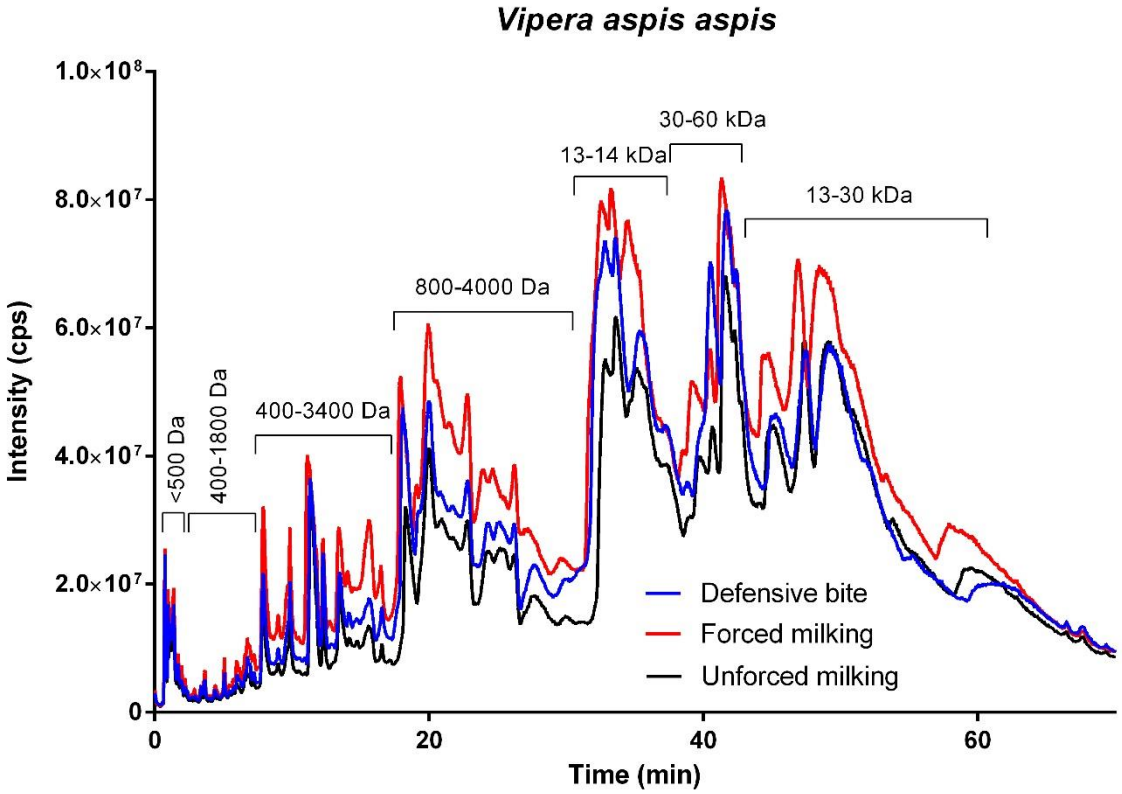


Figure 4: Total Ion Chromatogram profiles of *Vasp* venom.

3.2.2 Bottom up analysis of Vasp venom

To reduce the bias due to intraspecific variations and to allow for a direct correlation between transcriptomic and proteomic data, the crude venom was also milked from the same specimen prior to RNA extraction. The MS/MS spectra obtained after reduction, alkylation and digestion of the crude venom were analyzed by PEAKS studio software using the full assembled transcriptome translated into the six reading frames. Overall, each toxin family identified in our transcriptomic analysis was validated by proteomics (with a minimum of one sequence per family). A total of 64 peptides/proteins were identified in the venom using our assembled transcriptome as reference database (Table 1). Overall, a good correlation is observed between expression levels in transcriptome and venom composition,

although variable half-lives of distinct mRNA and differential regulation of the post transcription machinery can affect this ratio [682]. For instance, PLA2 expression levels are significantly different in the venom gland transcriptome 15% (Figure 2) compared to the proteome ≈30% (Figure 5). Surprisingly, the most expressed PLA2 are similar to ammodytin I1, which is probably not neurotoxic considering its similarity with ammodytin I2 [679]. The reported neurotoxicity of *Vasp* bites could rather be explained by the low level of ammodytoxin B-related toxin, which was the most highly expressed PLA2 sequence in the transcriptome [677,678]. C-type lectins assume diverse functions, hence the many isoforms detected in the transcriptome (24% of the toxin transcriptome according to contig numbers, data not shown), but expression level in the proteome is much lower ≈12% (Figure 5A).

Snake venom serine protease and disintegrin expression levels in both transcriptome and proteome are relatively conserved, respectively representing around 12% and 5% of the expression levels (Figures 2 and 5A). Snake venom metalloproteases are moderately expressed in the venom, representing around 11% (Figure 5A), whereas transcriptome expression levels were about 22% (Figure 2) of all toxins. Interestingly, the toxin families represented by only one transcriptomic sequence were all validated by proteomics, including 5'-nucleotidase, hyaluronidase, L-amino-acid oxidase, NGF, phosphodiesterase, PLB and renin-like aspartic protease, despite some low transcriptomic expression levels.

Interestingly, out of the 64 identified sequences, 7 were not initially annotated as venom proteins in our curated transcriptome.

3.3 Comparison with vipera berus berus venom

Vasp and *Vbb* are the two vipers most commonly encountered in France, in this section we will discuss the difference between the venom of these two species potentially explaining the differences in symptoms following envenomation.

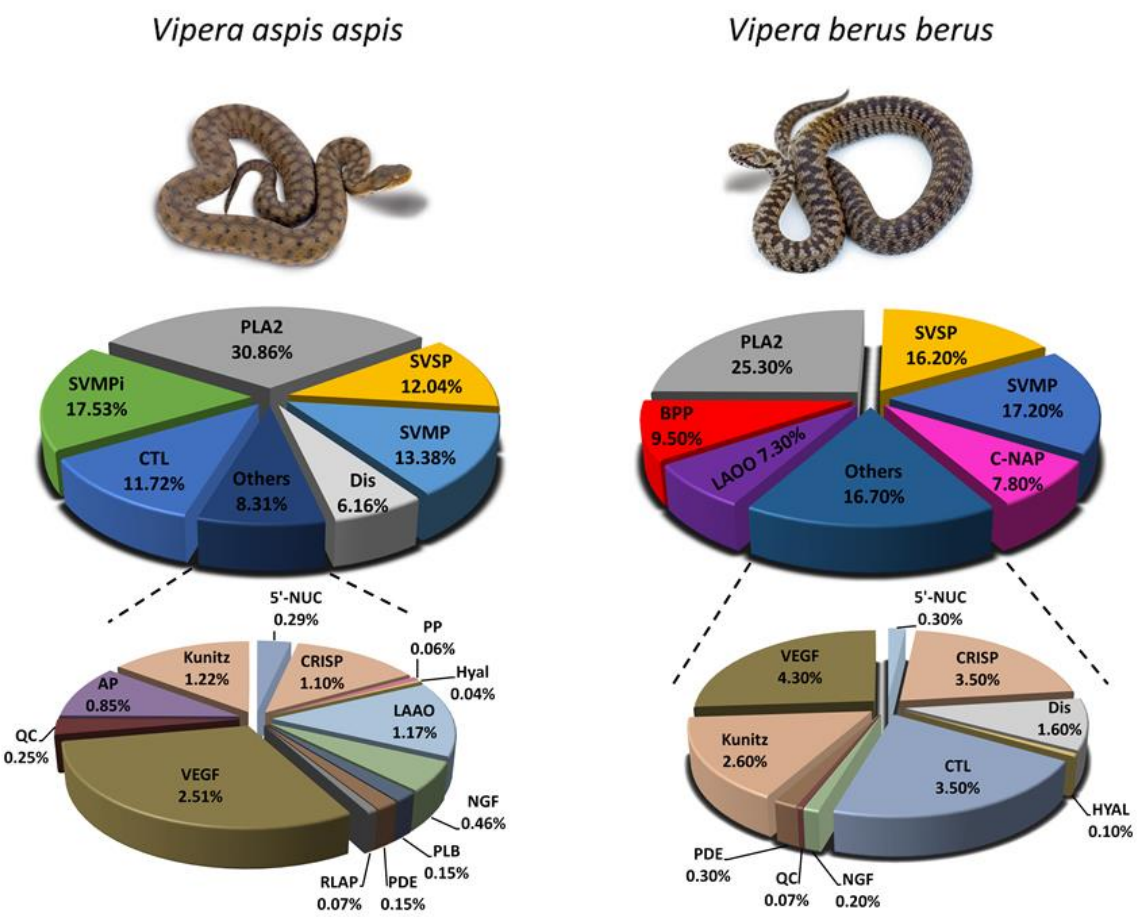


Figure 5: Representation of toxins families in (A) *Vasp* venom and comparison to (B) *Vbb* venom [683] according to their relative expression level in proteome. 5'-NUC, 5'-nucleotidase; CRISP, cysteine rich secretary protein; CTL, C-type lectin; Dis, disintegrin; HYAL, hyaluronidase; LAOO, L-amino acid oxidase; NGF, nerve growth factor; PDE, phosphodiesterase; PLA2, phospholipase A2; PL-B, phospholipase B; RLAP, renin-like aspartic protease; SVMP, snake venom metalloprotease; SVMPi, snake venom metalloprotease inhibitor; SVSP, snake venom serine protease; VEGF, vascular endothelial growth factor.

Vipera berus berus is the most widely distributed viper in Europe causing more bites than any of its congeners [684,685]. The principal local effect of *Vbb* bites are hemorrhage, edema, myonecrosis, bruising and pain [684,686,687], hypotension being the most important sign of systemic envenomation [632,688–690]. However, neurotoxic effects, systemic hemorrhage and coagulopathy caused by *Vbb* envenomation are rare [684,690]. The two most striking differences between compositions of the two venoms are smaller level of snaclec and PLA2. Latinović *et al.* showed that there is no ammodytoxin-like PLA2 in *Vbb* venom [657] unlike *Vasp* venom, which exhibits significant level of ammodytoxin B in

both venom proteome and transcriptome explaining the recently observed neurotoxicity of *Vasp* bites [636,638]. Snaclec are the major inducers of thrombocytopenia, which is clinically manifested as excessive platelet aggregation/agglutination [657], despite low level of C-type lectin in *Vbb* venom compared to *Vasp* venom thrombocytopenia is relatively frequently in patients poisoned by *Vbb* [689]. Thus, CTL from *Vasp* venom along with others perturbing hemostasis toxins actively contribute to hemorrhage and coagulopathy observed in *Vasp* envenomation [631,636,638].

Overall, proteo-transcriptomic *Vasp* venom analysis allowed the identification of 64 proteins. Moreover, the venom composition is consistent to *Vasp* bite effects with hemorrhagic effect probably induced by class III SVMPs, SNACLECs related to hemostasis impairing toxins presumably playing a role in reported thrombocytopenia and an Ammodytoxin B-related PLA2 potentially explaining the neurotoxic effect.

Acknowledgements

This study received financial support from the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, São Paulo Research Foundation, scholarship to FGA, 2016/20641-9 to F.G.A.), National Council for Scientific and Technological Development (CNPq, scholarship to FGA n. 150037/2018-0)

Supporting information

Table 1: Proteins identified in the *Vasp* venom by PEAKS studio software using assembled transcriptome as database. d indicate a deamidation, b a biotinylation and sequence homology is the result of highest score protein matching the uniprot database. Underline indicates sequences that were not initially annotated as venom proteins

Accession	-10lgP	Coverage	#Unique peptide	#Spec	PTM	Avg. Mass	Sequence homology (% identity)
T0941_R_2.5479_L_552_(-2)	384.53	68%	26	1056	Q20d N86d N115d N117d	20230	Ammodytin I1 isoform 1_ Vipera aspis (99.3%)
T0029_R_1.5154_L_2477_(-3)	357.34	43%	30	392	A324R Q409d N413d Q414d N433d N460d	90848	Metalloproteinase of class P-II MPII-1_ Vipera ammodytes ammodytes (88.3%)
T0018_R_0.5786_L_2733_(-2)	353.79	44%	70	419		101771	L-amino-acid oxidase_Daboia russelii (92.9%)
T0204_R_0.0162_L_1300_(-3)	343.79	54%	30	231	V176R	48596	Metalloproteinase_Echis coloratus (76.6%)
T0026_R_0.5265_L_2497_(-3)	336.41	42%	31	292	Q198d Q205d N479d N549d	92452	H3 metalloproteinase 1_Vipera ammodytes ammodytes (96.3%)
T0463_R_0.3263_L_853_(-2)	315.94	39%	23	509	V128A H129Q F178L L182V N183S	30384	Neutral phospholipase A2 ammodytin I2_Vipera ammodytes ammodytes (100%)

T0228_R_1.0199_L_1232_(-3)	311.11	42%	30	536	Q114d Q140d N256d	45425	Group III snake venom metalloproteinase_Echis ocellatus (79.3%)
T0049_R_0.1477_L_2242_(-3)	308.81	35%	19	143	Q322b	83429	Zinc metalloproteinase-disintegrin-like VLAIP-B_Macrovipera lebetina (94.9%)
T0004_R_0.1265_L_3240_(-3)	301.33	32%	35	181	D386N Q556d	119908	Xaa-Pro aminopeptidase 2_Boiga irregularis (88.3%)
T0221_R_1.2515_L_1238_(-3)	297.93	40%	18	209	G246A	45222	Metalloproteinase H4-A_Vipera ammodytes ammodytes (82.4%)
T0280_R_2.4064_L_1122_(-3)	294.85	25%	7	392	Q152d T170S N201d	40586	Basic phospholipase A2 ammodytoxin B_ Vipera ammodytes ammodytes (92.8%)
T0024_R_0.5406_L_2574_(-1)	291.63	17%	14	324	N97d N99d H132R Q294d	95126	Serine protease VLSP-1_Macrovipera lebetina (92.7%)
T0178_R_0.3396_L_1392_(-2)	281.89	40%	19	121	I76V L78V	49084	Cysteine-rich venom protein_Vipera berus (92.1%)
T0594_R_0.2911_L_741_(-3)	279.72	57%	29	195	T130M Q193E N194D	27321	Snaclec A5_Macrovipera lebetina (75.2%)
T0012_R_0.4819_L_2844_(-2)	279.28	14%	16	330	N77d N79d D214N	103148	Chymotrypsin-like protease VLCTLP_Macrovipera lebetina (88.7%)
T1846_R_2.2645_L_300_(-1)	271.20	72%	2	216	N18d R73M D77N I78V	11411	Basic phospholipase A2 ammodytoxin C_Vipera ammodytes ammodytes (98.9%)
T0023_R_0.5727_L_2575_(-3)	270.77	32%	15	137	Q481d	96232	Metalloproteinase F1_Vipera ammodytes ammodytes (92.3%)
T0751_R_1.7657_L_642_(-2)	269.91	38%	15	378	N64d N93d	23048	Disintegrin Dis-1_Vipera ammodytes ammodytes (98.4%)
T0025_R_0.136_L_2508_(-2)	268.43	32%	22	94		91858	Ecto-5'-nucleotidase_Crotalus horridus (91.8%)
T0102_R_1.2094_L_1743_(-3)	265.08	37%	16	483	K163R R178S	62896	Serine proteinase SP-3_Vipera ammodytes ammodytes
T0257_R_1.0899_L_1162_(-2)	263.84	56%	21	108	E327R	44038	Metalloproteinase of class P-III MPIII-1_Vipera ammodytes ammodytes (96.7%)
T0174_R_0.6632_L_1397_(-3)	260.06	15%	2	568	N113d N115d	52230	Acidic phospholipase A2 ammodytin I1_Vipera ammodytes ammodytes (100%)
T0898_R_1.0141_L_571_(-2)	251.83	52%	17	126	N125d	21225	Snaclec B3/B5_Macrovipera lebetina (87.2%)
T1651_R_0.4837_L_368_(-2)	250.22	46%	9	91	A48W	13886	C-type lectin-like protein 3B_Macrovipera lebetina (95.9%)
T0140_R_0.6417_L_1526_(-2)	249.83	33%	21	227	N290d N313d	57136	Metalloprotease P-IIa 1_Protobothrops flavoviridis (82.9%)
T0203_R_0.3692_L_1300_(-1)	241.54	25%	4	239	G165N	45896	C-type lectin-like protein 3A_Macrovipera lebetina (95.6%)
T1491_R_0.6941_L_400_(-2)	238.65	57%	25	294	I15T	14388	Endogenous tripeptide metalloproteinase inhibitor_Vipera ammodytes ammodytes (75%)
T0824_R_1.0948_L_600_(-2)	238.19	43%	6	71		22444	C-type lectin-like protein 3B_Macrovipera lebetina (94.6%)
T0578_R_1.4236_L_750_(-3)	234.90	57%	6	496	Q174d	27511	Snake venom serine protease nikobin_Vipera nikolskii (93.5%)
T0905_R_0.3701_L_568_(-2)	231.15	43%	7	84	Q55d	21354	C-type lectin-like protein 3A_Macrovipera lebetina (89.6%)

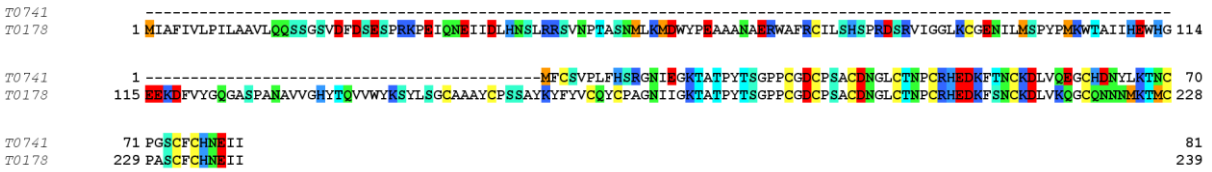
T0051_R_2.2812_L_2212_(-1)	230.66	15%	11	233	I329M	81522	Enzymatically inactive serine proteinase-like protein SPH-1_Vipera ammodytes ammodytes (99.5%)
T0076_R_0.9025_L_1950_(-3)	229.87	14%	16	114		70907	Venom nerve growth factor_Macrovipera lebetina (97.5%)
T0313_R_0.9123_L_1038_(-1)	225.03	20%	19	190	N109d	37625	Disintegrin Dis-1_Vipera ammodytes ammodytes (83.6%)
T0335_R_0.9979_L_999_(-1)	224.19	31%	4	139	F221V Q319d N323d Q324d	37275	Metalloproteinase of class P-II MPII-3_Vipera ammodytes ammodytes (88.8%)
T1850_R_1.405_L_300_(-2)	223.73	52%	2	438	A9G E17Q N71d	10647	Endogenous tripeptide metalloproteinase inhibitor_Vipera ammodytes ammodytes (79.1%)
T0027_R_0.6434_L_2486_(-3)	218.61	7%	7	95	A359Q N364d N375d	93106	Serine protease_Ovophis okinavensis (74.8%)
T0224_R_1.6894_L_1236_(-3)	218.42	24%	18	391		44311	Vammin-1_Vipera ammodytes ammodytes (99.3%)
T0255_R_0.9174_L_1166_(-2)	215.37	43%	4	64	S344R	43565	Metalloproteinase H4-A_Vipera ammodytes ammodytes (87.1%)
T0059_R_0.226_L_2112_(-2)	211.65	19%	9	172	N311d	75766	Serine proteinase SP-4_Vipera ammodytes ammodytes (87.7%)
T0007_R_0.1209_L_3041_(-3)	210.15	16%	6	56		113498	Coagulation factor X-activating enzyme heavy chain_Macrovipera lebetina (87.1%)
T0158_R_0.1249_L_1459_(-3)	202.04	17%	4	133	N798d N81d Q381d	52883	Serine proteinase SP-2_Vipera ammodytes ammodytes (94%)
T0020_R_0.0183_L_2635_(-2)	191.94	23%	15	53		99205	Phosphodiesterase_Macrovipera lebetina (98.6%)
T0609_R_0.1686_L_730_(-2)	182.88	22%	5	65	N168d A171E	27788	Snaclec coagulation factor X-activating enzyme light chain 2_Macrovipera lebetina (85.4%)
T0096_R_0.0534_L_1790_(-1)	181.16	23%	10	37		68579	Phospholipase B-like_Echis coloratus (94.2%)
T0416_R_0.8706_L_900_(-2)	175.95	18%	15	80	L150S K169D K169N	31328	Kunitz-type serine protease inhibitor ki-VN_Vipera nikolskii (91.5%)
T0580_R_0.3735_L_748_(-3)	173.51	38%	9	30	K75R Q93d	27710	C-type lectin snaclec-1_Vipera ammodytes ammodytes (76.9%)
T0458_R_0.5124_L_862_(-2)	172.94	13%	7	38		30869	Metalloproteinase_Echis coloratus (75%)
T1116_R_0.2236_L_498_(-3)	169.59	25%	5	39		19150	Snaclec coagulation factor X-activating enzyme light chain 2_Macrovipera lebetina (84.1%)
T0741_R_0.0859_L_648_(-2)	168.22	31%	3	31		22332	Cysteine-rich secretory protein CRISP-1_Vipera ammodytes ammodytes (90%)
T0619_R_0.0656_L_720_(-1)	163.86	21%	7	20		27357	C-type lectin 1_Bitis gabonica (76.1%)

T0463_R_0.3263_L_853_(-3)	162.80	12%	3	38	K269R	28564	Serine proteinase SP-5_Vipera ammodytes ammodytes (65%)
T0131_R_0.0309_L_1584_(-1)	158.93	23%	11	51		59208	Glutaminyl-peptide cyclotransferases_Daboia russelii (97.8%)
T0441_R_0.0974_L_880_(-1)	157.17	15%	4	59		31359	Metalloproteinase F1_Vipera ammodytes ammodytes (89.5%)
T0917_R_0.5443_L_564_(-2)	155.53	36%	6	29	Q152Y	21395	C-type lectin snaclec-4_Vipera ammodytes ammodytes (95.9%)
T0119_R_0.0149_L_1657_(-2)	151.01	17%	7	22		62789	Aminopeptidase_Bitis rhinoceros (95.8%)
T0642_R_0.1723_L_700_(-3)	143.36	15%	5	16		26095	C-type lectin-like protein 4B_Macrovipera lebetina (96%)
T0729_R_0.0128_L_651_(-2)	143.04	37%	4	28		25126	Aminopeptidase_Bitis rhinoceros (94.9%)
T0189_R_0.1352_L_1346_(-2)	142.16	13%	2	22		47419	Snaclec coagulation factor X-activating enzyme light chain 1_Macrovipera lebetin (85.9%)
T0019_R_0.0185_L_2728_(-1)	136.75	5%	2	19	N786d	97936	Disintegrin lebein-2-alpha_Macrovipera lebetina (90.9%)
T0082_R_0.0305_L_1896_(-1)	132.53	9%	5	10		72376	Hyaluronidase-1_Cerastes cerastes (95.5%)
T0189_R_0.1352_L_1346_(-1)	130.83	5%	5	31		49074	C-type lectin-like protein 3B_Macrovipera lebetina (96.5%)
T0407_R_0.1158_L_911_(-3)	128.98	9%	9	54	N65D	31357	Kunitz-type serine protease inhibitor 1_Vipera ammodytes ammodytes (83.3%)
T0037_R_0.1006_L_2374_(-3)	127.76	7%	5	21		87986	Renin-like aspartic protease_Echis ocellatus (87.7%)
T0232_R_0.0124_L_1228_(-2)	125.56	11%	5	9		46038	Proactivator polypeptide-like_Crotalus adamanteus (92.2%)

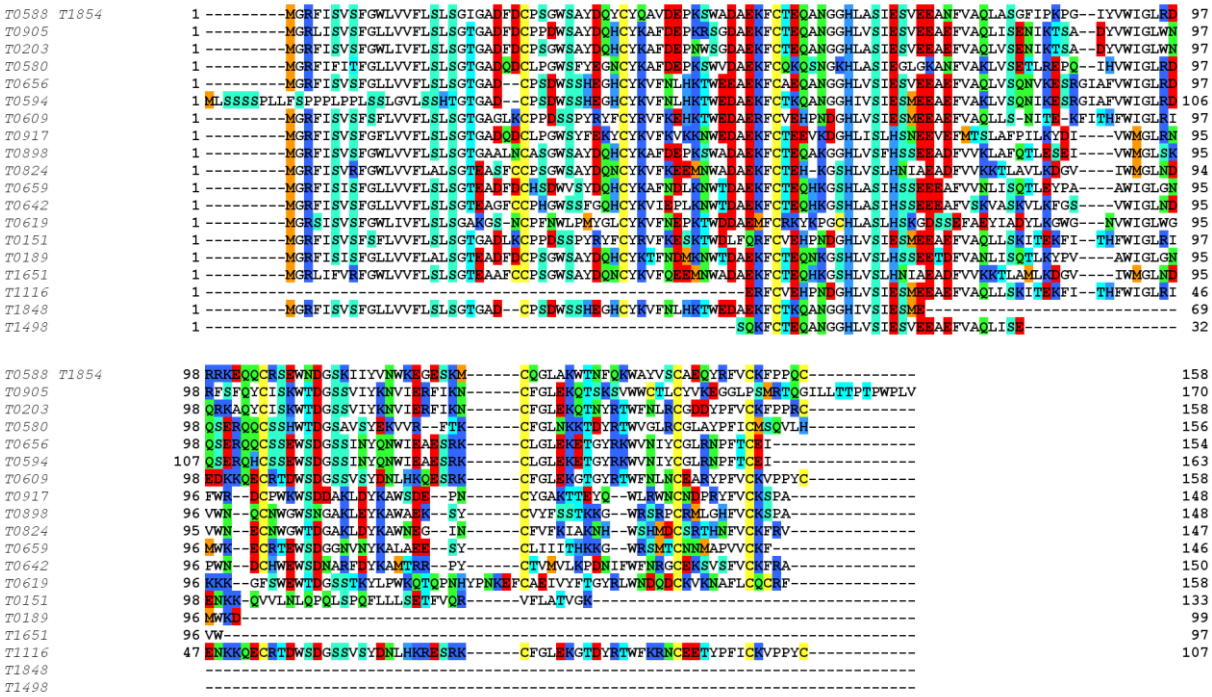
Figure S1: Alignments of the 80 toxin-related sequences retrieved from the *Vipera aspis aspis* venom gland transcriptome.

5’-nucleotidase

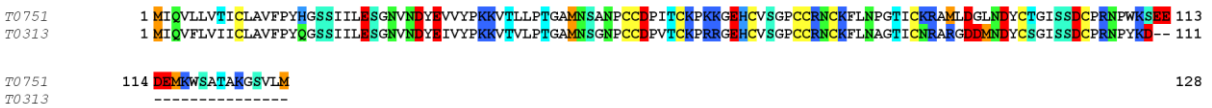
CRISP



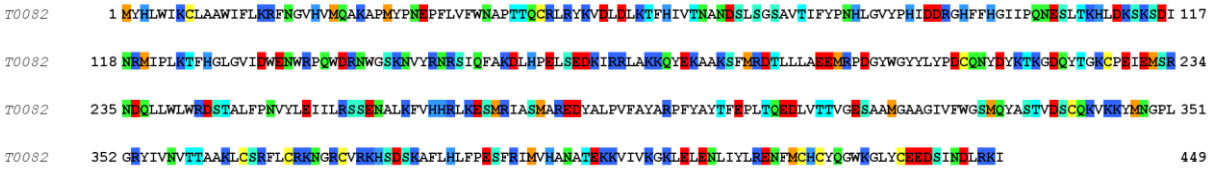
C-type Lectin



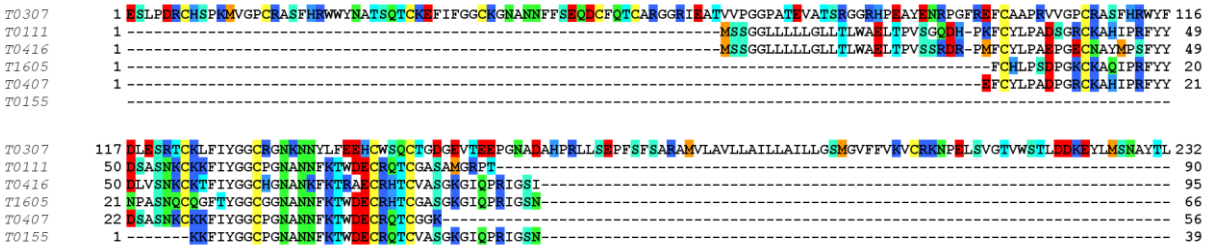
Disintegrin



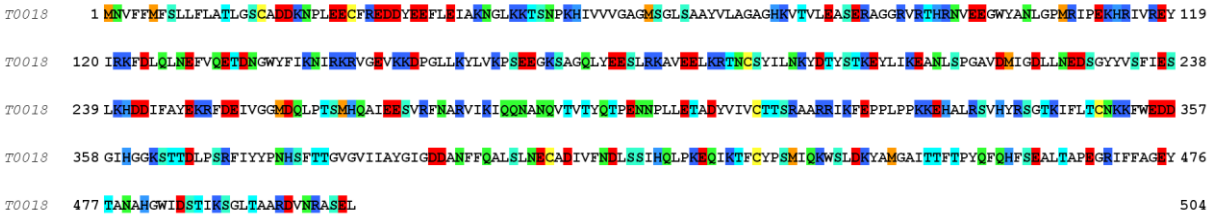
Hyaluronidase



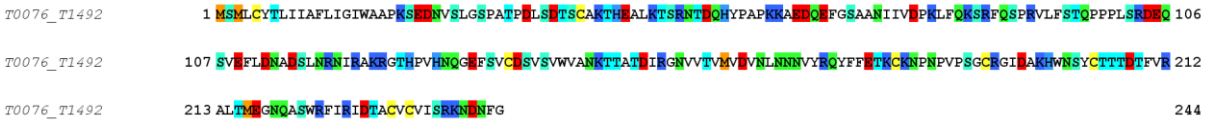
Kunitz



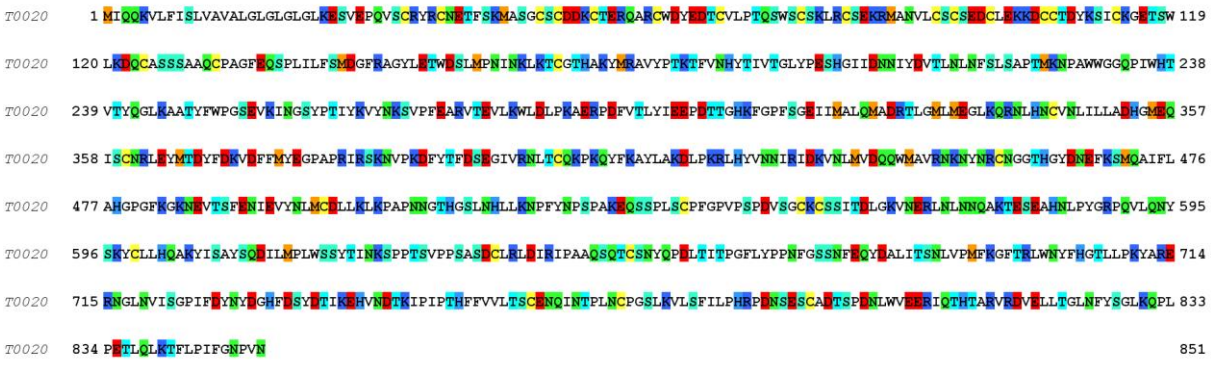
L-amino-acid oxidase



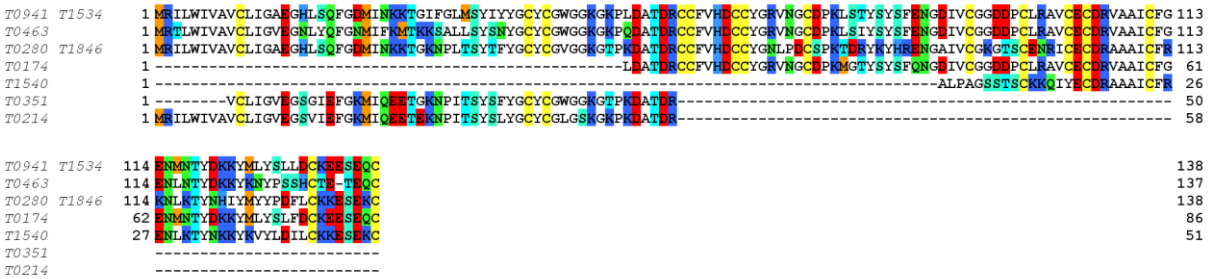
NGF



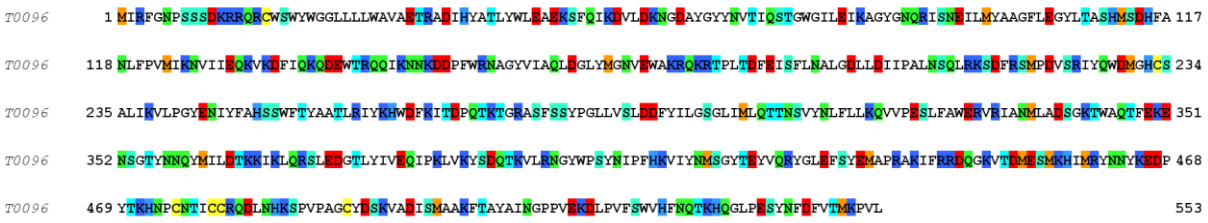
Phosphodiesterase



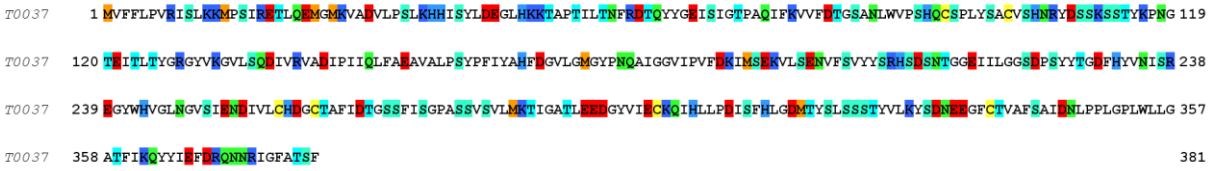
PLA2



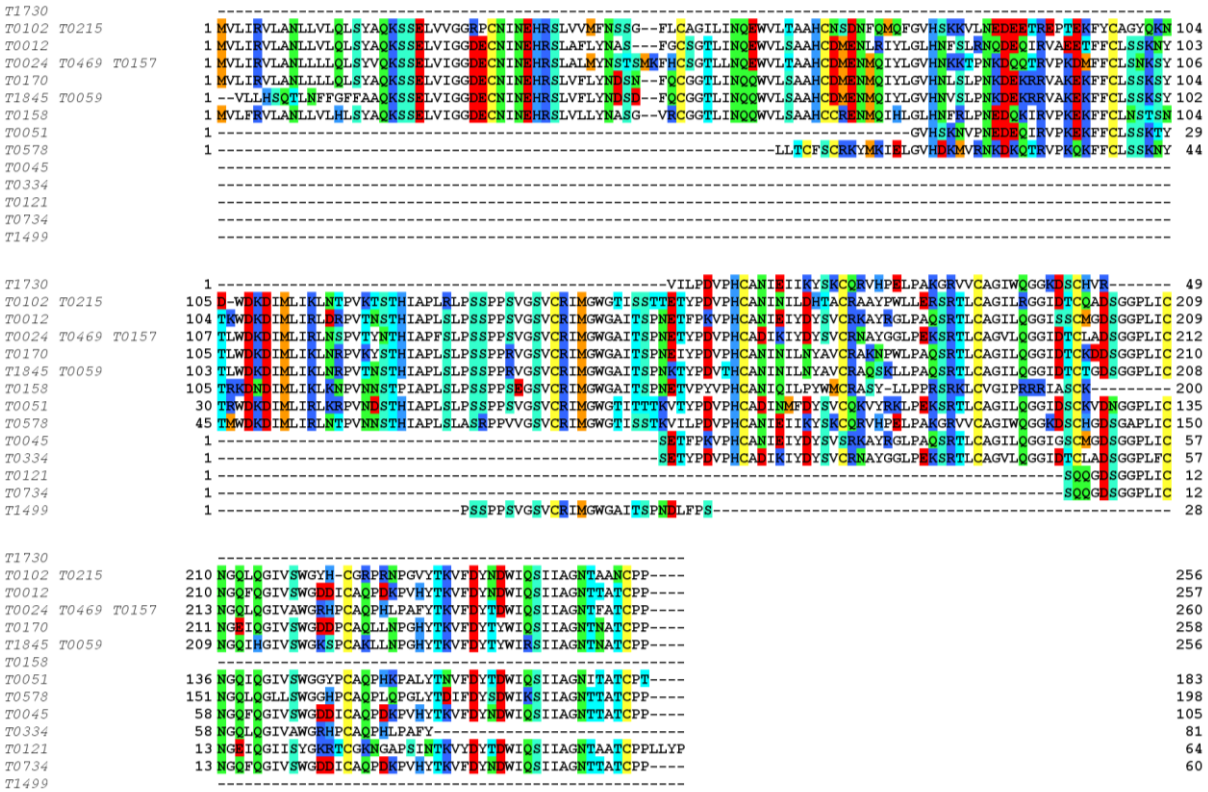
PLB



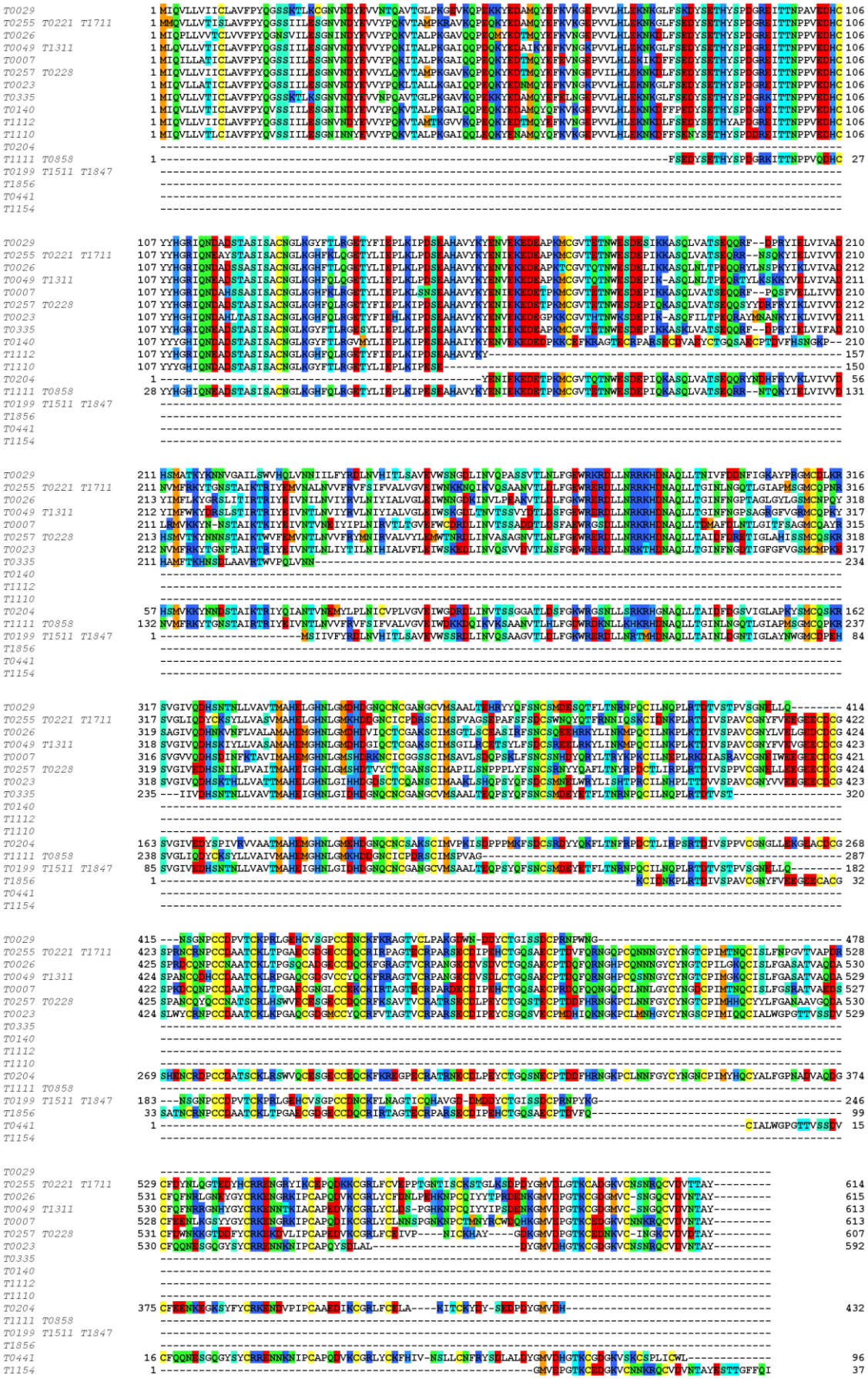
Renin-like aspartic protease



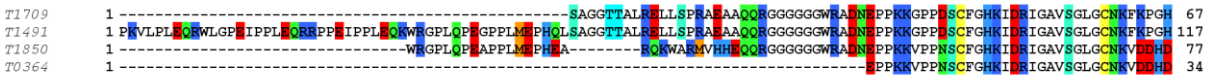
Serine Protease



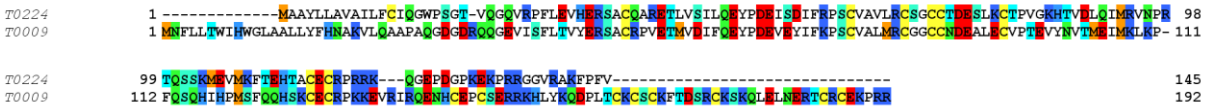
SVMP



SVMP inhibitor



VEGF



Conclusions et perspectives

L'étude protéo-transcriptomique du venin de *Vipera aspis aspsis* a mis en évidence plus d'une cinquantaine de nouvelles séquences. Aucune activité de ces séquences n'a encore été démontrée expérimentalement, cependant celle-ci peut être estimé par homologie avec les séquences existantes et ainsi permettre la conception rationnelle de nouveau composés bioactifs, à l'image du captopril, un inhibiteur de l'enzyme de conversion, qui a été développé à partir de peptides potentialisant la bradykinine isolée du venin de *Bothrops jararaca*.

6. Conclusions et perspectives

D'un point de vue général, la synthèse et la purification d'une dizaine de peptides variant de 9 à 40 résidus et de 0 à 3 ponts disulfures ont pu être effectués. La plupart des toxines synthétisées se sont révélées bioactives ce qui souligne le fort potentiel des peptides de venins comme outils pharmacologiques et a permis la publication de 5 articles scientifiques ainsi qu'a minima 3 autres en préparation.

Le cœur de ce travail de thèse a été la mise en place d'une stratégie de synthèse pour les toxines à 2 ponts disulfures, celles-ci étant majoritairement représentées par la famille des α-conotoxines qui sont des précieux outils pharmacologiques pour l'étude des récepteurs à l'acétylcholine nicotiniques neuronaux. Ces dernières sont classiquement synthétisées en formant le premier pont disulfure par oxydation à l'air des deux résidus cystéines libres, sur des temps relativement longs allant jusqu'à 24 heures ; les deux autres cystéines étant protégées par des groupements protecteurs acm qui sont déprotégés et oxydés à l'iode pour former le deuxième pont disulfure. La méthodologie mise en place est largement inspirée de la littérature exceptée que le premier pont disulfure (entre les cystéines libres) est formé en des temps très courts (5 min) grâce à l'utilisation du DTP. Ce réactif bien que décrit pour la première fois en 1999 comme activateur des groupements thiols pour la formation de ponts disulfures en des temps très courts (5 min), n'avais jusqu'à présent pas été employé pour la formation de multiples ponts disulfures de manière séquentielle. Enfin un gain de temps supplémentaire a été permis par la suppression de l'étape de lyophilisation intermédiaire. L'évolution de la méthodologie de synthèse est décrite dans la Figure ***17***. Cette méthodologie a pu être appliquée aussi bien à des peptides de taille moyenne (≈ 15 résidus) qu'à des peptides de plus grande taille (40 résidus) et des peptides modifiés cycliques du N-ter au C-ter. Elle est rapide et efficace et serait donc particulièrement pertinente dans le cadre de la production d'un grand nombre d'analogues par chimie combinatoire en vue d'un screening sur un récepteur en particulier ou sur une large gamme de récepteurs.

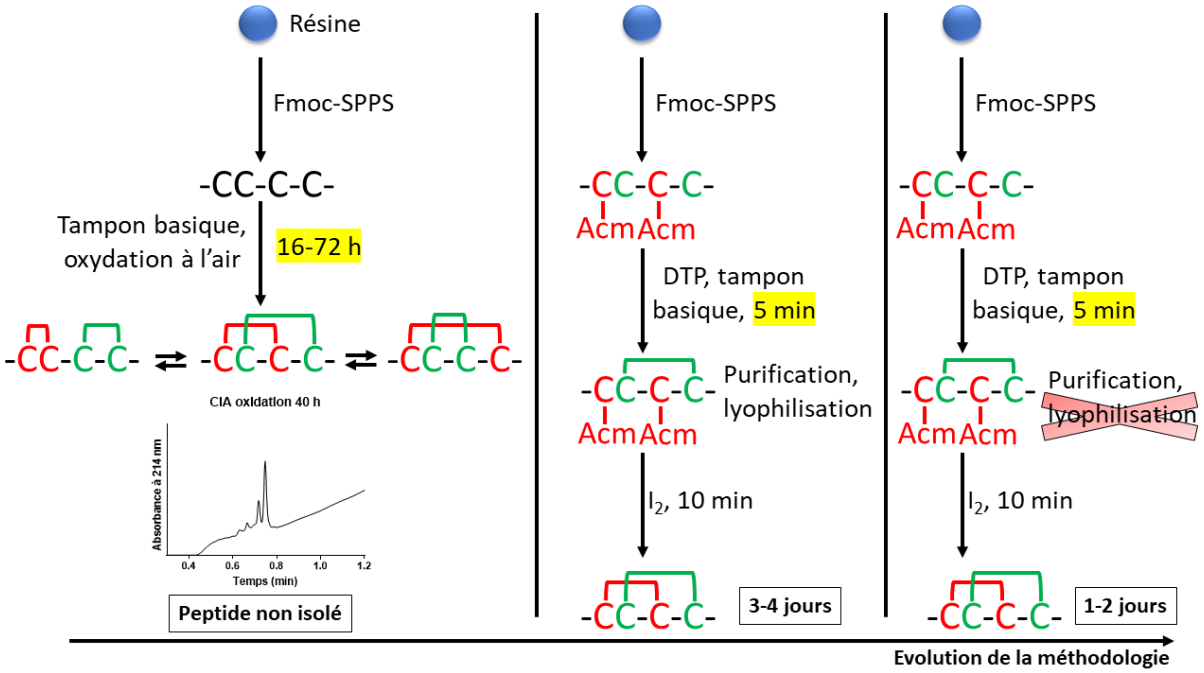


Figure 17: Evolution de la méthodologie de synthèse de peptides à deux ponts disulfures au cours du temps. Les essais infructueux de repliement oxydatif sous controle thermodynamique, notamment sur la séquence CIA, nous a conduit à l'utilisation de méthodologie de formation régiosélective. Les temps de synthèse ont pu être grandement améliorés grâce à l'utilisation du DTP qui permet une formation rapide du premier pont disulfure et la suppression de l'étape de lyophilisation intermédiaire.

La mise en place d'une stratégie de synthèse à 3 ponts s'est révélée plus délicate et a mis en évidence une première limitation du DTP. En effet celui-ci ne semble pas efficace pour la formation de ponts difficiles à former pour des raisons de contraintes stériques. En effet il s'est avéré que l'ordre de formation des ponts est essentiel et cela questionne sérieusement l'intérêt de l'utilisation d'une stratégie régiosélective pour la synthèse de peptides contenant plus de 2 ponts disulfures. En effet en plus de nécessiter de connaitre au préalable l'appariement des ponts disulfures, il est nécessaire de tester plusieurs ordres de formations, de ce point de vue le folding oxydatif sous contrôle thermodynamique apparaît être une meilleure alternative.

Les toxines bioactives malgré leur résistance accrue à la dégradation enzymatique grâce à leurs multiples ponts disulfures sont tout de même sujettes aux problèmes de perméabilité membranaire caractéristique des peptides. Malgré les progrès effectués en chimie médicinale, il n'existe à l'heure actuelle aucune recette magique pour rendre un peptide disponible par voie orale. Il ne fait nul doute que les futures recherches en chimie médicinale seront dirigés en ce sens.

Références

1. Newman, D.J.; Cragg, G.M. Natural Products As Sources of New Drugs over the 30 Years from 1981 to 2010. *J. Nat. Prod.* **2012**, *75*, 311–335.
2. Ray, A. Beyond debacle and debate: developing solutions in drug safety. *Nat Rev Drug Discov* **2009**, *8*, 775–779.
3. Watkins, P.B. Drug Safety Sciences and the Bottleneck in Drug Development. *Clin Pharmacol Ther* **2011**, *89*, 788–790.
4. Drewry, D.H.; Macarron, R. Enhancements of screening collections to address areas of unmet medical need: an industry perspective. *Current Opinion in Chemical Biology* **2010**, *14*, 289–298.
5. Zinzalla, G.; Thurston, D.E. Targeting protein–protein interactions for therapeutic intervention: a challenge for the future. *Future Medicinal Chemistry* **2009**, *1*, 65–93.
6. Bunnage, M.E. Getting pharmaceutical R&D back on target. *Nat Chem Biol* **2011**, *7*, 335–339.
7. Ahn, N.G.; Wang, A.H.-J. Proteomics and genomics: perspectives on drug and target discovery. *Current Opinion in Chemical Biology* **2008**, *12*, 1–3.
8. Kapetanovic, I.M. Computer-aided drug discovery and development (CADDD): In silico-chemico-biological approach. *Chemico-Biological Interactions* **2008**, *171*, 165–176.
9. Schneider, H.-C.; Klabunde, T. Understanding drugs and diseases by systems biology? *Bioorganic & Medicinal Chemistry Letters* **2013**, *23*, 1168–1176.
10. Sun, L. Peptide-Based Drug Development. *Mod Chem Appl* **2013**, *01*.
11. Craik, D.J.; Fairlie, D.P.; Liras, S.; Price, D. The Future of Peptide-based Drugs: Peptides in Drug Development. *Chemical Biology & Drug Design* **2013**, *81*, 136–147.
12. Scannell, J.W.; Blanckley, A.; Boldon, H.; Warrington, B. Diagnosing the decline in pharmaceutical R&D efficiency. *Nat Rev Drug Discov* **2012**, *11*, 191–200.
13. Lau, J.L.; Dunn, M.K. Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Bioorganic & Medicinal Chemistry* **2018**, *26*, 2700–2707.
14. Banting, F.G.; Best, C.H.; Collip, J.B.; Campbell, W.R.; Fletcher, A.A. Pancreatic Extracts in the Treatment of Diabetes Mellitus. *Can Med Assoc J* **1922**, *12*, 141–146.
15. Hench, P.S.; Kendall, E.C.; Slocumb, C.H.; Polley, H.F. Adrenocortical Hormone in Arthritis : Preliminary Report. *Annals of the Rheumatic Diseases* **1949**, *8*, 97–104.
16. Vigneaud, V. du; Ressler, C.; Swan, C.J.M.; Roberts, C.W.; Katsoyannis, P.G.; Gordon, S. THE SYNTHESIS OF AN OCTAPEPTIDE AMIDE WITH THE HORMONAL ACTIVITY OF OXYTOCIN. *J. Am. Chem. Soc.* **1953**, *75*, 4879–4880.
17. Vasopressin Synthesized: Du Vigneaud uses same methods in synthesizing two forms of vasopressin as he used for oxytocin. *Chem. Eng. News* **1956**, *34*, 2754.
18. Vlieghe, P.; Lisowski, V.; Martinez, J.; Khrestchatisky, M. Synthetic therapeutic peptides: science and market. *Drug Discovery Today* **2010**, *15*, 40–56.
19. Sato, A.K.; Viswanathan, M.; Kent, R.B.; Wood, C.R. Therapeutic peptides: technological advances driving peptides into development. *Current Opinion in Biotechnology* **2006**, *17*, 638–642.
20. Giribaldi, J.; Dutertre, S. α-Conotoxins to explore the molecular, physiological and pathophysiological functions of neuronal nicotinic acetylcholine receptors. *Neuroscience Letters* **2018**, *679*, 24–34.
21. Raufman, J.-P. Bioactive peptides from lizard venoms. *Regulatory Peptides* **1996**, *61*, 1–18.
22. Eng, J. Exendin peptides. *Mt. Sinai J. Med.* **1992**, *59*, 147–149.
23. Eng, J.; Kleinman, W.A.; Singh, L.; Singh, G.; Raufman, J.P. Isolation and characterization of exendin-4, an exendin-3 analogue, from Heloderma suspectum venom. Further evidence for an exendin receptor on dispersed acini from guinea pig pancreas. *J. Biol. Chem.* **1992**, *267*, 7402–7405.
24. Meier, J.J. GLP-1 receptor agonists for individualized treatment of type 2 diabetes mellitus. *Nat Rev Endocrinol* **2012**, *8*, 728–742.

25. Malone, J.; Trautmann, M.; Wilhelm, K.; Taylor, K.; Kendall, D.M. Exenatide once weekly for the treatment of type 2 diabetes. *Expert Opin Investig Drugs* **2009**, *18*, 359–367.
26. Zaccardi, F.; Htike, Z.Z.; Webb, D.R.; Khunti, K.; Davies, M.J. Benefits and Harms of Once-Weekly Glucagon-like Peptide-1 Receptor Agonist Treatments: A Systematic Review and Network Meta-analysis. *Ann. Intern. Med.* **2016**, *164*, 102–113.
27. Marso, S.P.; Bain, S.C.; Consoli, A.; Eliaschewitz, F.G.; Jódar, E.; Leiter, L.A.; Lingvay, I.; Rosenstock, J.; Seufert, J.; Warren, M.L.; et al. Semaglutide and Cardiovascular Outcomes in Patients with Type 2 Diabetes. *N Engl J Med* **2016**, *375*, 1834–1844.
28. Marso, S.P.; Daniels, G.H.; Brown-Frandsen, K.; Kristensen, P.; Mann, J.F.E.; Nauck, M.A.; Nissen, S.E.; Pocock, S.; Poulter, N.R.; Ravn, L.S.; et al. Liraglutide and Cardiovascular Outcomes in Type 2 Diabetes. *N Engl J Med* **2016**, *375*, 311–322.
29. Neer, R.M.; Arnaud, C.D.; Zanchetta, J.R.; Prince, R.; Gaich, G.A.; Reginster, J.Y.; Hodsman, A.B.; Eriksen, E.F.; Ish-Shalom, S.; Genant, H.K.; et al. Effect of parathyroid hormone (1-34) on fractures and bone mineral density in postmenopausal women with osteoporosis. *N. Engl. J. Med.* **2001**, *344*, 1434–1441.
30. Camacho, P.M.; Petak, S.M.; Binkley, N.; Clarke, B.L.; Harris, S.T.; Hurley, D.L.; Kleerekoper, M.; Lewiecki, E.M.; Miller, P.D.; Narula, H.S.; et al. AMERICAN ASSOCIATION OF CLINICAL ENDOCRINOLOGISTS AND AMERICAN COLLEGE OF ENDOCRINOLOGY CLINICAL PRACTICE GUIDELINES FOR THE DIAGNOSIS AND TREATMENT OF POSTMENOPAUSAL OSTEOPOROSIS — 2016-- *EXECUTIVE SUMMARY*. *Endocrine Practice* **2016**, *22*, 1111–1118.
31. Kendler, D.L.; Marin, F.; Zerbini, C.A.F.; Russo, L.A.; Greenspan, S.L.; Zikan, V.; Bagur, A.; Malouf-Sierra, J.; Lakatos, P.; Fahrleitner-Pammer, A.; et al. Effects of teriparatide and risedronate on new fractures in post-menopausal women with severe osteoporosis (VERO): a multicentre, double-blind, double-dummy, randomised controlled trial. *The Lancet* **2018**, *391*, 230–240.
32. Hopkins, A.L.; Groom, C.R. The druggable genome. *Nat Rev Drug Discov* **2002**, *1*, 727–730.
33. Stupp, R.; Hegi, M.E.; Gorlia, T.; Erridge, S.C.; Perry, J.; Hong, Y.-K.; Aldape, K.D.; Lhermitte, B.; Pietsch, T.; Grujicic, D.; et al. Cilengitide combined with standard treatment for patients with newly diagnosed glioblastoma with methylated MGMT promoter (CENTRIC EORTC 26071-22072 study): a multicentre, randomised, open-label, phase 3 trial. *The Lancet Oncology* **2014**, *15*, 1100–1108.
34. Chang, Y.S.; Graves, B.; Guerlavais, V.; Tovar, C.; Packman, K.; To, K.-H.; Olson, K.A.; Kesavan, K.; Gangurde, P.; Mukherjee, A.; et al. Stapled α−helical peptide drug development: A potent dual inhibitor of MDM2 and MDMX for p53-dependent cancer therapy. *Proc Natl Acad Sci USA* **2013**, *110*, E3445–E3454.
35. Meerovitch, K.; Torkildsen, G.; Lonsdale, J.; Goldfarb, H.; Lama, T.; Cumberlidge, G.; Ousler III, G.W. Safety and efficacy of MIM D3 ophthalmic solutions in a randomized placebo controlled Phase 2 clinical trial in patients with dry eye. *OPTH* **2013**, 1275.
36. Birk, A.V.; Liu, S.; Soong, Y.; Mills, W.; Singh, P.; Warren, J.D.; Seshan, S.V.; Pardee, J.D.; Szeto, H.H. The Mitochondrial-Targeted Compound SS-31 Re-Energizes Ischemic Mitochondria by Interacting with Cardiolipin. *JASN* **2013**, *24*, 1250–1261.
37. King, G.F. Venoms as a platform for human drugs: translating toxins into therapeutics. *Expert Opinion on Biological Therapy* **2011**, *11*, 1469–1484.
38. Pennington, M.W.; Czerwinski, A.; Norton, R.S. Peptide therapeutics from venom: Current status and potential. *Bioorganic & Medicinal Chemistry* **2018**, *26*, 2738–2758.
39. Bhattacharjee, P.; Bhattacharyya, D. Therapeutic Use of Snake Venom Components: A Voyage from Ancient to Modern India. *MROC* **2014**, *11*, 45–54.
40. Harvey, A.L. Toxins and drug discovery. *Toxicon* **2014**, *92*, 193–200.
41. Oldrati, V.; Arrell, M.; Violette, A.; Perret, F.; Sprüngli, X.; Wolfender, J.-L.; Stöcklin, R. Advances in venomics. *Molecular BioSystems* **2016**, *12*, 3530–3543.
42. Lewis, R.J.; Garcia, M.L. Therapeutic potential of venom peptides. *Nature Reviews Drug Discovery* **2003**, *2*, 790–802.

43. Domon, B. Mass Spectrometry and Protein Analysis. *Science* **2006**, *312*, 212–217.
44. Mardis, E.R. Next-Generation Sequencing Platforms. *Annual Rev. Anal. Chem.* **2013**, *6*, 287–303.
45. Vetter, I. Development and Optimization of FLIPR High Throughput Calcium Assays for Ion Channels and GPCRs. In *Calcium Signaling*; Islam, Md.S., Ed.; Springer Netherlands: Dordrecht, 2012; Vol. 740, pp. 45–82 ISBN 978-94-007-2887-5.
46. Calvete, J.J. Venomics: integrative venom proteomics and beyond. *Biochem. J.* **2017**, *474*, 611–634.
47. Karas, Michael.; Hillenkamp, Franz. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* **1988**, *60*, 2299–2301.
48. Fenn, J.; Mann, M.; Meng, C.; Wong, S.; Whitehouse, C. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989**, *246*, 64–71.
49. Aebersold, R.; Mann, M. Mass spectrometry-based proteomics. **2003**, *422*, 10.
50. Hunt, D.F.; Yates, J.R.; Shabanowitz, J.; Winston, S.; Hauer, C.R. Protein sequencing by tandem mass spectrometry. *Proceedings of the National Academy of Sciences* **1986**, *83*, 6233–6237.
51. Biemann, K. Contributions of mass spectrometry to peptide and protein structure. *Biomed. Environ. Mass Spectrom.* **1988**, *16*, 99–111.
52. Papayannopoulos, I.A. The interpretation of collision-induced dissociation tandem mass spectra of peptides. *Mass Spectrom. Rev.* **1995**, *14*, 49–73.
53. Paizs, B.; Suhai, S. Fragmentation pathways of protonated peptides. *Mass Spectrom. Rev.* **2005**, *24*, 508–548.
54. Yu, Wen.; Vath, J.E.; Huberty, M.C.; Martin, S.A. Identification of the facile gas-phase cleavage of the Asp-Pro and Asp-Xxx peptide bonds in matrix-assisted laser desorption time-of-flight mass spectrometry. *Anal. Chem.* **1993**, *65*, 3015–3023.
55. Zhokhov, S.S.; Kovalyov, S.V.; Samgina, T.Yu.; Lebedev, A.T. An EThcD-Based Method for Discrimination of Leucine and Isoleucine Residues in Tryptic Peptides. *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 1600–1611.
56. Heather, J.M.; Chain, B. The sequence of sequencers: The history of sequencing DNA. *Genomics* **2016**, *107*, 1–8.
57. Torres, A.F.C.; Huang, C.; Chong, C.-M.; Leung, S.W.; Prieto-da-Silva, Á.R.B.; Havt, A.; Quinet, Y.P.; Martins, A.M.C.; Lee, S.M.Y.; Rádis-Baptista, G. Transcriptome Analysis in Venom Gland of the Predatory Giant Ant Dinoponera quadriceps: Insights into the Polypeptide Toxin Arsenal of Hymenopterans. *PLoS ONE* **2014**, *9*, e87556.
58. Calvete, J.J.; Sanz, L.; Angulo, Y.; Lomonte, B.; Gutiérrez, J.M. Venoms, venomics, antivenomics. *FEBS Letters* **2009**, *583*, 1736–1743.
59. Rodríguez de la Vega, R.C.; Schwartz, E.F.; Possani, L.D. Mining on scorpion venom biodiversity. *Toxicon* **2010**, *56*, 1155–1161.
60. Prashanth, J.R.; Lewis, R.J.; Dutertre, S. Towards an integrated venomics approach for accelerated conopeptide discovery. *Toxicon* **2012**, *60*, 470–477.
61. Calvete, J.J. The continuing saga of snake venom disintegrins. *Toxicon* **2013**, *62*, 40–49.
62. Brahma, R.K.; McCleary, R.J.R.; Kini, R.M.; Doley, R. Venom gland transcriptomics for identifying, cataloging, and characterizing venom proteins in snakes. *Toxicon* **2015**, *93*, 1–10.
63. Fox, E.J.; Reid-Bayliss, K.S.; Emond, M.J.; Loeb, L.A. Accuracy of Next Generation Sequencing Platforms. *Next Gener Seq Appl* **2014**, *1*.
64. Violette, A.; Biass, D.; Dutertre, S.; Koua, D.; Piquemal, D.; Pierrat, F.; Stöcklin, R.; Favreau, P. Large-scale discovery of conopeptides and conoproteins in the injectable venom of a fish-hunting cone snail using a combined proteomic and transcriptomic approach. *Journal of Proteomics* **2012**, *75*, 5215–5225.
65. Low, D.H.W.; Sunagar, K.; Undheim, E.A.B.; Ali, S.A.; Alagon, A.C.; Ruder, T.; Jackson, T.N.W.; Pineda Gonzalez, S.; King, G.F.; Jones, A.; et al. Dracula's children: Molecular evolution of vampire bat venom. *Journal of Proteomics* **2013**, *89*, 95–111.

66. Li, R.; Yu, H.; Xue, W.; Yue, Y.; Liu, S.; Xing, R.; Li, P. Jellyfish venomics and venom gland transcriptomics analysis of Stomolophus meleagris to reveal the toxins associated with sting. *Journal of Proteomics* **2014**, *106*, 17–29.
67. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *Journal of Molecular Biology* **1990**, *215*, 403–410.
68. Kaas, Q.; Craik, D. Bioinformatics-Aided Venomics. *Toxins* **2015**, *7*, 2159–2187.
69. Leonardi, A.; Biass, D.; Kordiš, D.; Stöcklin, R.; Favreau, P.; Križaj, I. *Conus consors* Snail Venom Proteomics Proposes Functions, Pathways, and Novel Families Involved in Its Venomic System. *J. Proteome Res.* **2012**, *11*, 5046–5058.
70. Koua, D.; Laht, S.; Kaplinski, L.; Stöcklin, R.; Remm, M.; Favreau, P.; Lisacek, F. Position-specific scoring matrix and hidden Markov model complement each other for the prediction of conopeptide superfamilies. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2013**, *1834*, 717–724.
71. Cornet, V.; Henry, J.; Corre, E.; Le Corguille, G.; Zanuttini, B.; Zatylny-Gaudin, C. Dual role of the cuttlefish salivary proteome in defense and predation. *Journal of Proteomics* **2014**, *108*, 209–222.
72. Wong, E.S.W.; Hardy, M.C.; Wood, D.; Bailey, T.; King, G.F. SVM-Based Prediction of Propeptide Cleavage Sites in Spider Toxins Identifies Toxin Innovation in an Australian Tarantula. *PLoS ONE* **2013**, *8*, e66279.
73. Craig, R.; Beavis, R.C. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* **2004**, *20*, 1466–1467.
74. Zhang, J.; Xin, L.; Shan, B.; Chen, W.; Xie, M.; Yuen, D.; Zhang, W.; Zhang, Z.; Lajoie, G.A.; Ma, B. PEAKS DB: De Novo sequencing assisted database search for sensitive and accurate peptide identification. *Mol Cell Proteomics* **2011**, mcp.M111.010587.
75. Biass, D.; Violette, A.; Hulo, N.; Lisacek, F.; Favreau, P.; Stöcklin, R. Uncovering Intense Protein Diversification in a Cone Snail Venom Gland Using an Integrative Venomics Approach. *J. Proteome Res.* **2015**, *14*, 628–638.
76. Rokyta, D.R.; Wray, K.P.; McGivern, J.J.; Margres, M.J. The transcriptomic and proteomic basis for the evolution of a novel venom phenotype within the Timber Rattlesnake (Crotalus horridus). *Toxicon* **2015**, *98*, 34–48.
77. Sanggaard, K.W.; Dyrlund, T.F.; Thomsen, L.R.; Nielsen, T.A.; Brøndum, L.; Wang, T.; Thøgersen, I.B.; Enghild, J.J. Characterization of the gila monster (Heloderma suspectum suspectum) venom proteome. *Journal of Proteomics* **2015**, *117*, 1–11.
78. Menon, R.; Gasser, R.B.; Mitreva, M.; Ranganathan, S. An analysis of the transcriptome of Teladorsagia circumcincta: its biological and biotechnological implications. *BMC Genomics* **2012**, *13*, S10.
79. Undheim, E.A.B.; Sunagar, K.; Hamilton, B.R.; Jones, A.; Venter, D.J.; Fry, B.G.; King, G.F. Multifunctional warheads: Diversification of the toxin arsenal of centipedes via novel multidomain transcripts. *Journal of Proteomics* **2014**, *102*, 1–10.
80. Jungo, F.; Bougueleret, L.; Xenarios, I.; Poux, S. The UniProtKB/Swiss-Prot Tox-Prot program: A central hub of integrated venom protein data. *Toxicon* **2012**, *60*, 551–557.
81. Oliveira, J.S.; Fuentes-Silva, D.; King, G.F. Development of a rational nomenclature for naming peptide and protein toxins from sea anemones. *Toxicon* **2012**, *60*, 539–550.
82. Akondi, K.B.; Muttenthaler, M.; Dutertre, S.; Kaas, Q.; Craik, D.J.; Lewis, R.J.; Alewood, P.F. Discovery, Synthesis, and Structure–Activity Relationships of Conotoxins. *Chem. Rev.* **2014**, *114*, 5815–5847.
83. King, G.F.; Gentz, M.C.; Escoubas, P.; Nicholson, G.M. A rational nomenclature for naming peptide toxins from spiders and other venomous animals. *Toxicon* **2008**, *52*, 264–276.
84. Becker, S.; Terlau, H. Toxins from cone snails: properties, applications and biotechnological production. *Appl Microbiol Biotechnol* **2008**, *79*, 1–9.
85. Merrifield, R.B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.

86. Amblard, M.; Fehrentz, J.-A.; Martinez, J.; Subra, G. Methods and Protocols of Modern Solid Phase Peptide Synthesis. *Molecular Biotechnology* **2006**, *33*, 239–254.
87. Carpino, L.A.; Han, G.Y. 9-Fluorenylmethoxycarbonyl function, a new base-sensitive amino-protecting group. *Journal of the American Chemical Society* **1970**, *92*, 5748–5749.
88. Walter, G. Production and use of antibodies against synthetic peptides. *Journal of Immunological Methods* **1986**, *88*, 149–161.
89. Dryland, A.; Sheppard, R.C. Peptide synthesis. Part 8. A system for solid-phase synthesis under low pressure continuous flow conditions. *J. Chem. Soc., Perkin Trans. 1* **1986**, 125.
90. *Fmoc solid phase peptide synthesis: a practical approach*; Chan, W.C., White, P.D., Eds.; The practical approach series; Oxford University Press: New York, 2000; ISBN 978-0-19-963725-6.
91. White, P.; Keyte, J.W.; Bailey, K.; Bloomberg, G. Expediting the Fmoc solid phase synthesis of long peptides through the application of dimethyloxazolidine dipeptides. *Journal of Peptide Science* **2004**, *10*, 18–26.
92. Wöhr, T.; Mutter, M. Pseudo-prolines in peptide synthesis: Direct insertion of serine and threonine derived oxazolidines in dipeptides. *Tetrahedron Letters* **1995**, *36*, 3847–3848.
93. Cardona, V.; Eberle, I.; Barthélémy, S.; Beythien, J.; Doerner, B.; Schneeberger, P.; Keyte, J.; White, P.D. Application of Dmb-Dipeptides in the Fmoc SPPS of Difficult and Aspartimide-Prone Sequences. *Int J Pept Res Ther* **2008**, *14*, 285–292.
94. Góngora-Benítez, M.; Tulla-Puche, J.; Albericio, F. Multifaceted Roles of Disulfide Bonds. Peptides as Therapeutics. *Chem. Rev.* **2014**, *114*, 901–926.
95. Grishin, A.A.; Wang, C.-I.A.; Muttenthaler, M.; Alewood, P.F.; Lewis, R.J.; Adams, D.J. α-Conotoxin AuIB Isomers Exhibit Distinct Inhibitory Mechanisms and Differential Sensitivity to Stoichiometry of α3β4 Nicotinic Acetylcholine Receptors. *J. Biol. Chem.* **2010**, *285*, 22254–22263.
96. Tietze, A.A.; Tietze, D.; Ohlenschläger, O.; Leipold, E.; Ullrich, F.; Kühl, T.; Mischo, A.; Buntkowsky, G.; Görlach, M.; Heinemann, S.H.; et al. Structurally Diverse μ-Conotoxin PIIIA Isomers Block Sodium Channel Na_V 1.4. *Angewandte Chemie International Edition* **2012**, *51*, 4058–4061.
97. Gehrmann, J.; Alewood, P.F.; Craik, D.J. Structure determination of the three disulfide bond isomers of α-conotoxin GI: a model for the role of disulfide bonds in structural stability 1 1Edited by P. E. Wright. *Journal of Molecular Biology* **1998**, *278*, 401–415.
98. Han, Y.; Albericio, F.; Barany, G. Occurrence and Minimization of Cysteine Racemization during Stepwise Solid-Phase Peptide Synthesis [1,2]. *J. Org. Chem.* **1997**, *62*, 4307–4312.
99. Grandas, A.; Jorba, X.; Giralt, E.; Pedroso, E. Anchoring of Fmoc-amino acids to hydroxymethyl resins. *International Journal of Peptide and Protein Research* **2009**, *33*, 386–390.
100. Fujiwara, Y.; Akaji, K.; Kiso, Y. Racemization-Free Synthesis of C-Terminal Cysteine-Peptide Using 2-Chlorotrityl Resin. *Chem. Pharm. Bull.* **1994**, *42*, 724–726.
101. *Oxidative Folding of Peptides and Proteins:*; Moroder, L., Buchner, J., Eds.; RSC Biomolecular Sciences; Royal Society of Chemistry: Cambridge, 2008; ISBN 978-0-85404-148-0.
102. Wong, C.T.T.; Taichi, M.; Nishio, H.; Nishiuchi, Y.; Tam, J.P. Optimal Oxidative Folding of the Novel Antimicrobial Cyclotide from *Hedyotis biflora* Requires High Alcohol Concentrations. *Biochemistry* **2011**, *50*, 7275–7283.
103. Góngora-Benítez, M.; Tulla-Puche, J.; Paradís-Bas, M.; Werbitzky, O.; Giraud, M.; Albericio, F. Optimized Fmoc solid-phase synthesis of the cysteine-rich peptide linaclotide. *Biopolymers* **2011**, *96*, 69–80.
104. Steiner, A.M.; Bulaj, G. Optimization of oxidative folding methods for cysteine-rich peptides: a study of conotoxins containing three disulfide bridges. *J. Peptide Sci.* **2011**, *17*, 1–7.
105. Arolas, J.L.; Aviles, F.X.; Chang, J.-Y.; Ventura, S. Folding of small disulfide-rich proteins: clarifying the puzzle. *Trends in Biochemical Sciences* **2006**, *31*, 292–301.
106. Benham, C.J.; Saleet Jafri, M. Disulfide bonding patterns and protein topologies. *Protein Science* **2008**, *2*, 41–54.

107. Reinwarth, M.; Nasu, D.; Kolmar, H.; Avrutina, O. Chemical Synthesis, Backbone Cyclization and Oxidative Folding of Cystine-knot Peptides — Promising Scaffolds for Applications in Drug Design. *Molecules* **2012**, *17*, 12533–12552.
108. Jia, Y.; Liu, H.; Bao, W.; Weng, M.; Chen, W.; Cai, Y.; Zheng, Z.; Zou, G. Functional analysis of propeptide as an intramolecular chaperone for in vivo folding of subtilisin nattokinase. *FEBS Letters* **2010**, *584*, 4789–4796.
109. Kamber, B.; Hartmann, A.; Eisler, K.; Riniker, B.; Rink, H.; Sieber, P.; Rittel, W. The Synthesis of Cystine Peptides by Iodine Oxidation ofS-Trityl-cysteine andS-Acetamidomethyl-cysteine Peptides. *Helv. Chim. Acta* **1980**, *63*, 899–915.
110. Fujii, N.; Otaka, A.; Funakoshi, S.; Bessho, K.; Watanabe, T.; Akaji, K.; Yajima, H. Studies on peptides. CLI. Syntheses of cystine-peptides by oxidation of s-protected cysteine-peptides with thallium(III) trifluoroacetate. *Chem. Pharm. Bull.* **1987**, *35*, 2339–2347.
111. Muttenthaler, M.; Ramos, Y.G.; Feytens, D.; de Araujo, A.D.; Alewood, P.F. p-Nitrobenzyl protection for cysteine and selenocysteine: A more stable alternative to the acetamidomethyl group. *Biopolymers* **2010**, *94*, 423–432.
112. Lamthanh, H.; Roumestand, C.; Deprun, C.; Ménez, A. Side reaction during the deprotection of (S-acetamidomethyl)cysteine in a peptide with a high serine and threonine content. *Int. J. Pept. Protein Res.* **1993**, *41*, 85–95.
113. Nishiuchi, Y.; Sakakibara, S. Primary and secondary structure of conotoxin GI, a neurotoxic tridecapeptide from a marine snail. *FEBS Letters* **1982**, *148*, 260–262.
114. Ponsati, B.; Giralt, E.; Andreu, D. Solid-phase approaches to regiospecific double disulfide formation. Application to a fragment of bovine pituitary peptide. *Tetrahedron* **1990**, *46*, 8255–8266.
115. Monje, V.D.; Haack, J.A.; Naisbitt, S.R.; Miljanich, G.; Ramachandran, J.; Nasdasdi, L.; Olivera, B.M.; Hillyard, D.R.; Gray, W.R. A new Conus peptide ligand for Ca channel subtypes. *Neuropharmacology* **1993**, *32*, 1141–1149.
116. Cuthbertson, A.; Indrevoll, B. A method for the one-pot regioselective formation of the two disulfide bonds of α-conotoxin SI. *Tetrahedron Letters* **2000**, *41*, 3661–3663.
117. Nielsen, J.S.; Buczek, P.; Bulaj, G. Cosolvent-assisted oxidative folding of a bicyclicα-conotoxin ImI. *J. Peptide Sci.* **2004**, *10*, 249–256.
118. Cuthbertson, A.; Indrevoll, B. Regioselective Formation, Using Orthogonal Cysteine Protection, of an α-Conotoxin Dimer Peptide Containing Four Disulfide Bonds. *Org. Lett.* **2003**, *5*, 2955–2957.
119. Otaka, A.; Koide, T.; Shide, A.; Fujii, N. Application of dimethylsulphoxide(DMSO) / trifluoroacetic acid(TFA) oxidation to the synthesis of cystine-containing peptide. *Tetrahedron Letters* **1991**, *32*, 1223–1226.
120. Atherton, E.; Sheppard, R.C.; Ward, P. Peptide synthesis. Part 7. Solid-phase synthesis of conotoxin G1. *J. Chem. Soc., Perkin Trans. 1* **1985**, 2065.
121. Szabó, I.; Schlosser, G.; Hudecz, F.; Mező, G. Disulfide bond rearrangement during regioselective oxidation in PhS(O)Ph/CH3SiCl3 mixture for the synthesis of α-conotoxin GI. *Biopolymers* **2007**, *88*, 20–28.
122. Munson, M.C.; Barany, G. Synthesis of .alpha.-conotoxin SI, a bicyclic tridecapeptide amide with two disulfide bridges: illustration of novel protection schemes and oxidation strategies. *J. Am. Chem. Soc.* **1993**, *115*, 10203–10210.
123. Galanis, A.S.; Albericio, F.; Grøtli, M. Enhanced microwave-assisted method for on-bead disulfide bond formation: Synthesis of α-conotoxin MII. *Biopolymers* **2009**, *92*, 23–34.
124. Hargittai, B.; Barany, G. Controlled syntheses of natural and disulfide-mispaired regioisomers of α-conotoxin SI*: Syntheses of α-conotoxin SI regioisomers. *The Journal of Peptide Research* **1999**, *54*, 468–479.
125. Vassilatis, D.K.; Hohmann, J.G.; Zeng, H.; Li, F.; Ranchalis, J.E.; Mortrud, M.T.; Brown, A.; Rodriguez, S.S.; Weller, J.R.; Wright, A.C.; et al. The G protein-coupled receptor repertoires of human and mouse. *Proceedings of the National Academy of Sciences* **2003**, *100*, 4903–4908.

126. Kini, R.M.; Doley, R. Structure, function and evolution of three-finger toxins: Mini proteins with multiple targets. *Toxicon* **2010**, *56*, 855–867.
127. Caulfield, M.P.; Birdsall, N.J. International Union of Pharmacology. XVII. Classification of muscarinic acetylcholine receptors. *Pharmacol. Rev.* **1998**, *50*, 279–290.
128. Koivula, K.; Rondinelli, S.; Näsman, J. The three-finger toxin MTα is a selective α2B-adrenoceptor antagonist. *Toxicon* **2010**, *56*, 440–447.
129. Quinton, L.; Girard, E.; Maiga, A.; Rekik, M.; Lluel, P.; Masuyer, G.; Larregola, M.; Marquer, C.; Ciolek, J.; Magnin, T.; et al. Isolation and pharmacological characterization of AdTx1, a natural peptide displaying specific insurmountable antagonism of the α1A-adrenoceptor: Specific peptide antagonist for the α1A-adrenoceptor. *British Journal of Pharmacology* **2010**, *159*, 316–325.
130. Rouget, C.; Quinton, L.; Maïga, A.; Gales, C.; Masuyer, G.; Malosse, C.; Chamot-Rooke, J.; Thai, R.; Mourier, G.; De Pauw, E.; et al. Identification of a novel snake peptide toxin displaying high affinity and antagonist behaviour for the α2-adrenoceptors: Specific peptide antagonist for α2-adrenoceptors. *British Journal of Pharmacology* **2010**, *161*, 1361–1374.
131. Ménez, A.; Gatineau, E.; Roumestand, C.; Harvey, A.L.; Mouawad, L.; Gilquin, B.; Toma, F. Do cardiotoxins possess a functional site? Structural and chemical modification studies reveal the functional site of the cardiotoxin from Naja nigricollis. *Biochimie* **1990**, *72*, 575–588.
132. Kornisiuk, E.; Jerusalinsky, D.; Cerveñansky, C.; Harvey, A.L. Binding of muscarinic toxins MTx1 and MTx2 from the venom of the green mamba Dendroaspis angusticeps to cloned human muscarinic cholinoceptors. *Toxicon* **1995**, *33*, 11–18.
133. Harvey, A.L.; Kornisiuk, E.; Bradley, K.N.; Cerveñansky, C.; Durán, R.; Adrover, M.; Sánchez, G.; Jerusalinsky, D. Effects of Muscarinic Toxins MT1 and MT2 from Green Mamba on Different Muscarinic Cholinoceptors. *Neurochemical Research* **2002**, *27*, 1543–1554.
134. Jolkkonen, M.; van Giersbergen, P.L.M.; Hellman, U.; Wernstedt, C.; Karlsson, E. A toxin from the green mamba *Dendroaspis angusticeps* : Amino acid sequence and selectivity for muscarinic m4 receptors. *FEBS Letters* **1994**, *352*, 91–94.
135. Liang, J.S.; Carsi-Gabrenas, J.; Krajewski, J.L.; McCafferty, J.M.; Purkerson, S.L.; Santiago, M.P.; Strauss, W.L.; Valentine, H.H.; Potter, L.T. Anti-muscarinic toxins from Dendroaspis angusticeps. *Toxicon* **1996**, *34*, 1257–1267.
136. Olianas, M.C.; Ingianni, A.; Maullu, C.; Adem, A.; Karlsson, E.; Onali, P. Selectivity profile of muscarinic toxin 3 in functional assays of cloned and native receptors. *J. Pharmacol. Exp. Ther.* **1999**, *288*, 164–170.
137. Karlsson, E.; Jolkkonen, M.; Mulugeta, E.; Onali, P.; Adem, A. Snake toxins with high selectivity for subtypes of muscarinic acetylcholine receptors. *Biochimie* **2000**, *82*, 793–806.
138. Olianas, M.C.; Maullu, C.; Adem, A.; Mulugeta, E.; Karlsson, E.; Onali, P. Inhibition of acetylcholine muscarinic M_1 receptor function by the M_1-selective ligand muscarinic toxin 7 (MT-7). *British Journal of Pharmacology* **2000**, *131*, 447–452.
139. 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 M_1 Muscarinic Receptor. *Mol Pharmacol* **2003**, *63*, 26–35.
140. Kukkonen, A.; Peräkylä, M.; Åkerman, K.E.O.; Näsman, J. Muscarinic Toxin 7 Selectivity Is Dictated by Extracellular Receptor Loops. *J. Biol. Chem.* **2004**, *279*, 50923–50929.
141. Freedman, J.E.; Snyder, S.H. Vipoxin. A protein from Russell's viper venom with high affinity for biogenic amine receptors. *J. Biol. Chem.* **1981**, *256*, 13172–13179.
142. Bevan, P.; Hiestand, P. beta-RTX. A receptor-active protein from Russell's viper (Vipera russelli russelli) venom. *J. Biol. Chem.* **1983**, *258*, 5319–5326.
143. Miyoshi, S.; Tu, A.T. Phospholipase A2fromNaja naja sputatrixVenom Is a Muscarinic Acetylcholine Receptor Inhibitor. *Archives of Biochemistry and Biophysics* **1996**, *328*, 17–25.
144. Miyoshi, S.; Tu, A.T. Muscarinic Acetylcholine Receptor (mAChR) Inhibitor from Snake Venom: Interaction with Subtypes of Human mAChR. *Archives of Biochemistry and Biophysics* **1999**, *369*, 114–118.

145. Takasaki, C.; Tamiya, N.; Bdolah, A.; Wollberg, Z.; Kochva, E. Sarafotoxins S6: several isotoxins from Atractaspis engaddensis (burrowing asp) venom that affect the heart. *Toxicon* **1988**, *26*, 543–548.
146. Kloog, Y.; Ambar, I.; Sokolovsky, M.; Kochva, E.; Wollberg, Z.; Bdolah, A. Sarafotoxin, a novel vasoconstrictor peptide: phosphoinositide hydrolysis in rat heart and brain. *Science* **1988**, *242*, 268–270.
147. Williams, D.L.; Jones, K.L.; Pettibone, D.J.; Lis, E.V.; Clineschmidt, B.V. Sarafotoxin S6c: An agonist which distinguishes between endothelin receptor subtypes. *Biochemical and Biophysical Research Communications* **1991**, *175*, 556–561.
148. Lauer-Fields, J.L.; Cudic, M.; Wei, S.; Mari, F.; Fields, G.B.; Brew, K. Engineered Sarafotoxins as Tissue Inhibitor of Metalloproteinases-like Matrix Metalloproteinase Inhibitors. *J. Biol. Chem.* **2007**, *282*, 26948–26955.
149. Schweitz, H.; Pacaud, P.; Diochot, S.; Moinier, D.; Lazdunski, M. MIT $_1$, a black mamba toxin with a new and highly potent activity on intestinal contraction. *FEBS Letters* **1999**, *461*, 183–188.
150. Mollay, C.; Wechselberger, C.; Mignogna, G.; Negri, L.; Melchiorri, P.; Barra, D.; Kreil, G. Bv8, a small protein from frog skin and its homologue from snake venom induce hyperalgesia in rats. *European Journal of Pharmacology* **1999**, *374*, 189–196.
151. Lin, D.C.-H.; Bullock, C.M.; Ehlert, F.J.; Chen, J.-L.; Tian, H.; Zhou, Q.-Y. Identification and Molecular Characterization of Two Closely Related G Protein-coupled Receptors Activated by Prokineticins/Endocrine Gland Vascular Endothelial Growth Factor. *J. Biol. Chem.* **2002**, *277*, 19276–19280.
152. Masuda, Y.; Takatsu, Y.; Terao, Y.; Kumano, S.; Ishibashi, Y.; Suenaga, M.; Abe, M.; Fukusumi, S.; Watanabe, T.; Shintani, Y.; et al. Isolation and identification of EG-VEGF/prokineticins as cognate ligands for two orphan G-protein-coupled receptors. *Biochemical and Biophysical Research Communications* **2002**, *293*, 396–402.
153. Soga, T.; Matsumoto, S.; Oda, T.; Saito, T.; Hiyama, H.; Takasaki, J.; Kamohara, M.; Ohishi, T.; Matsushime, H.; Furuichi, K. Molecular cloning and characterization of prokineticin receptors. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **2002**, *1579*, 173–179.
154. Negri, L.; Lattanzi, R.; Giannini, E.; Melchiorri, P. Bv8/Prokineticin proteins and their receptors. *Life Sciences* **2007**, *81*, 1103–1116.
155. Bullock, C.M.; Li, J.-D.; Zhou, Q.-Y. Structural Determinants Required for the Bioactivities of Prokineticins and Identification of Prokineticin Receptor Antagonists. *Mol Pharmacol* **2004**, *65*, 582–588.
156. Longenecker, H.E.; Hurlbut, W.P.; Mauro, A.; Clark, A.W. Effects of Black Widow Spider Venom on the Frog Neuromuscular Junction: Effects on End-plate Potential, Miniature End-plate Potential and Nerve Terminal Spike. *Nature* **1970**, *225*, 701–703.
157. Silva, J.-P.; Suckling, J.; Ushkaryov, Y. Penelope's web: using α-latrotoxin to untangle the mysteries of exocytosis. *Journal of Neurochemistry* **2009**, *111*, 275–290.
158. Davletov, B.A.; Shamotienko, O.G.; Lelianova, V.G.; Grishin, E.V.; Ushkaryov, Y.A. Isolation and Biochemical Characterization of a Ca $^{2+}$ -independent α-Latrotoxin-binding Protein. *J. Biol. Chem.* **1996**, *271*, 23239–23245.
159. Krasnoperov, V.G.; Bittner, M.A.; Beavis, R.; Kuang, Y.; Salnikow, K.V.; Chepurny, O.G.; Little, A.R.; Plotnikov, A.N.; Wu, D.; Holz, R.W.; et al. alpha-Latrotoxin stimulates exocytosis by the interaction with a neuronal G-protein-coupled receptor. *Neuron* **1997**, *18*, 925–937.
160. Fredriksson, R.; Lagerström, M.C.; Lundin, L.-G.; Schiöth, H.B. The G-Protein-Coupled Receptors in the Human Genome Form Five Main Families. Phylogenetic Analysis, Paralogon Groups, and Fingerprints. *Mol Pharmacol* **2003**, *63*, 1256–1272.
161. Lang, J. Ca2+-independent insulin exocytosis induced by alpha -latrotoxin requires latrophilin, a G protein-coupled receptor. *The EMBO Journal* **1998**, *17*, 648–657.

162. Ashton, A.C.; Volynski, K.E.; Lelianova, V.G.; Orlova, E.V.; Van Renterghem, C.; Canepari, M.; Seagar, M.; Ushkaryov, Y.A. α-Latrotoxin, Acting via Two Ca $^{2+}$ -dependent Pathways, Triggers Exocytosis of Two Pools of Synaptic Vesicles. *J. Biol. Chem.* **2001**, *276*, 44695–44703.
163. Sharpe, I.A.; Gehrmann, J.; Loughnan, M.L.; Thomas, L.; Adams, D.A.; Atkins, A.; Palant, E.; Craik, D.J.; Adams, D.J.; Alewood, P.F.; et al. Two new classes of conopeptides inhibit the α1-adrenoceptor and noradrenaline transporter. *Nat Neurosci* **2001**, *4*, 902–907.
164. Sharpe, I.A.; Thomas, L.; Loughnan, M.; Motin, L.; Palant, E.; Croker, D.E.; Alewood, D.; Chen, S.; Graham, R.M.; Alewood, P.F.; et al. Allosteric α_1 -Adrenoreceptor Antagonism by the Conopeptide ρ-TIA. *J. Biol. Chem.* **2003**, *278*, 34451–34457.
165. Cheng, L.L.; Stoev, S.; Manning, M.; Derick, S.; Pena, A.; Mimoun, M.B.; Guillon, G. Design of Potent and Selective Agonists for the Human Vasopressin V_{1b} Receptor Based on Modifications of [Deamino-Cys ^{1}]arginine Vasopressin at Position 4. *J. Med. Chem.* **2004**, *47*, 2375–2388.
166. Dutertre, S.; Ulens, C.; Büttner, R.; Fish, A.; van Elk, R.; Kendel, Y.; Hopping, G.; Alewood, P.F.; Schroeder, C.; Nicke, A.; et al. AChBP-targeted α-conotoxin correlates distinct binding orientations with nAChR subtype selectivity. *EMBO J* **2007**, *26*, 3858–3867.
167. Cruz, L.J.; de Santos, V.; Zafaralla, G.C.; Ramilo, C.A.; Zeikus, R.; Gray, W.R.; Olivera, B.M. Invertebrate vasopressin/oxytocin homologs. Characterization of peptides from Conus geographus and Conus straitus venoms. *J. Biol. Chem.* **1987**, *262*, 15821–15824.
168. Craig, A.G.; Norberg, T.; Griffin, D.; Hoeger, C.; Akhtar, M.; Schmidt, K.; Low, W.; Dykert, J.; Richelson, E.; Navarro, V.; et al. Contulakin-G, an *O* -Glycosylated Invertebrate Neurotensin. *J. Biol. Chem.* **1999**, *274*, 13752–13759.
169. Matavel, A.; Cruz, J.S.; Penaforte, C.L.; Araújo, D.A.M.; Kalapothakis, E.; Prado, V.F.; Diniz, C.R.; Cordeiro, M.N.; Beirão, P.S.L. Electrophysiological characterization and molecular identification of the Phoneutria nigriventer peptide toxin PnTx2-6 1. *FEBS Letters* **2002**, *523*, 219–223.
170. Nunes, K.P.; Costa-Gonçalves, A.; Lanza, L.F.; Cortes, S.F.; Cordeiro, M.N.; Richardson, M.; Pimenta, A.M.C.; Webb, R.C.; Leite, R.; De Lima, M.E. Tx2-6 toxin of the Phoneutria nigriventer spider potentiates rat erectile function. *Toxicon* **2008**, *51*, 1197–1206.
171. Andrade, E.; Villanova, F.; Borra, P.; Leite, K.; Troncone, L.; Cortez, I.; Messina, L.; Paranhos, M.; Claro, J.; Srougi, M. Penile erection induced *in vivo* by a purified toxin from the Brazilian spider *Phoneutria nigriventer*. *BJU International* **2008**, *102*, 835–837.
172. Matavel, A.; Fleury, C.; Oliveira, L.C.; Molina, F.; de Lima, M.E.; Cruz, J.S.; Cordeiro, M.N.; Richardson, M.; Ramos, C.H.I.; Beirão, P.S.L. Structure and Activity Analysis of Two Spider Toxins That Alter Sodium Channel Inactivation Kinetics. *Biochemistry* **2009**, *48*, 3078–3088.
173. Silva, C.N.; Nunes, K.P.; Torres, F.S.; Cassoli, J.S.; Santos, D.M.; Almeida, F.D.M.; Matavel, A.; Cruz, J.S.; Santos-Miranda, A.; Nunes, A.D.C.; et al. PnPP-19, a Synthetic and Nontoxic Peptide Designed from a Phoneutria nigriventer Toxin, Potentiates Erectile Function via NO/cGMP. *The Journal of Urology* **2015**, *194*, 1481–1490.
174. Leite, K.R.M.; Andrade, E.; Ramos, A.T.; Magnoli, F.C.; Srougi, M.; Troncone, L.R.P. Phoneutria nigriventer spider toxin Tx2-6 causes priapism and death: A histopathological investigation in mice. *Toxicon* **2012**, *60*, 797–801.
175. Freitas, A.C.N.; Pacheco, D.F.; Machado, M.F.M.; Carmona, A.K.; Duarte, I.D.G.; de Lima, M.E. PnPP-19, a spider toxin peptide, induces peripheral antinociception through opioid and cannabinoid receptors and inhibition of neutral endopeptidase. *Br J Pharmacol* **2016**, *173*, 1491–1501.
176. Freitas, A.C.N.; Silva, G.C.; Pacheco, D.F.; Pimenta, A.M.C.; Lemos, V.S.; Duarte, I.D.G.; de Lima, M.E. The synthetic peptide PnPP-19 induces peripheral antinociception via activation of NO/cGMP/KATP pathway: Role of eNOS and nNOS. *Nitric Oxide* **2017**, *64*, 31–38.
177. Freitas, A.C.N.; Peigneur, S.; Macedo, F.H.P.; Menezes-Filho, J.E.; Millns, P.; Medeiros, L.F.; Arruda, M.A.; Cruz, J.; Holliday, N.D.; Tytgat, J.; et al. The Peptide PnPP-19, a Spider Toxin

Derivative, Activates μ-Opioid Receptors and Modulates Calcium Channels. *Toxins* **2018**, *10*, 43.
178. Lewis, R.J. Conotoxins: Molecular and Therapeutic Targets. In *Marine Toxins as Research Tools*; Fusetani, N., Kem, W., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2009; Vol. 46, pp. 45–65 ISBN 978-3-540-87892-6.
179. Lewis, R.J. Conotoxin Venom Peptide Therapeutics. In *Pharmaceutical Biotechnology*; Guzmán, C.A., Feuerstein, G.Z., Eds.; Springer New York: New York, NY, 2009; Vol. 655, pp. 44–48 ISBN 978-1-4419-1131-5.
180. Kress, H.G.; Simpson, K.H.; Marchettini, P.; Ver Donck, A.; Varrassi, G. Intrathecal therapy: what has changed with the introduction of ziconotide. *Pain Pract* **2009**, *9*, 338–347.
181. Nicke, A.; Wonnacott, S.; Lewis, R.J. alpha-Conotoxins as tools for the elucidation of structure and function of neuronal nicotinic acetylcholine receptor subtypes. *Eur J Biochem* **2004**, *271*, 2305–2319.
182. Lebbe, E.K.M.; Tytgat, J. In the picture: disulfide-poor conopeptides, a class of pharmacologically interesting compounds. *Journal of Venomous Animals and Toxins including Tropical Diseases* **2016**, *22*.
183. Barberis, C.; Mouillac, B.; Durroux, T. Structural bases of vasopressin/oxytocin receptor function. *Journal of Endocrinology* **1998**, *156*, 223–229.
184. Mouillac, B.; Chini, B.; Balestre, M.-N.; Elands, J.; Trumpp-Kallmeyer, S.; Hoflack, J.; Hibert, M.; Jard, S.; Barberis, C. The Binding Site of Neuropeptide Vasopressin V1a Receptor: EVIDENCE FOR A MAJOR LOCALIZATION WITHIN TRANSMEMBRANE REGIONS. *Journal of Biological Chemistry* **1995**, *270*, 25771–25777.
185. Petersson, M. Cardiovascular effects of oxytocin. *Prog. Brain Res.* **2002**, *139*, 281–288.
186. Dutertre, S.; Croker, D.; Daly, N.L.; Andersson, Å.; Muttenthaler, M.; Lumsden, N.G.; Craik, D.J.; Alewood, P.F.; Guillon, G.; Lewis, R.J. Conopressin-T from Conus tulipa Reveals an Antagonist Switch in Vasopressin-like Peptides. *J. Biol. Chem.* **2008**, *283*, 7100–7108.
187. Lewis, R.J.; Dutertre, S.; Vetter, I.; Christie, M.J. Conus venom peptide pharmacology. *Pharmacol. Rev.* **2012**, *64*, 259–298.
188. Möller, C.; Marí, F. A vasopressin/oxytocin-related conopeptide with γ-carboxyglutamate at position 8. *Biochem J* **2007**, *404*, 413–419.
189. Postina, R.; Kojro, E.; Fahrenholz, F. Separate Agonist and Peptide Antagonist Binding Sites of the Oxytocin Receptor Defined by Their Transfer into the V_2 Vasopressin Receptor. *J. Biol. Chem.* **1996**, *271*, 31593–31601.
190. Rodrigo, J.; Pena, A.; Murat, B.; Trueba, M.; Durroux, T.; Guillon, G.; Rognan, D. Mapping the Binding Site of Arginine Vasopressin to V_{1a} and V_{1b} Vasopressin Receptors. *Molecular Endocrinology* **2007**, *21*, 512–523.
191. Liu, L.; Chew, G.; Hawrot, E.; Chi, C.; Wang, C.; Xu, M. Two Potent α3/5 Conotoxins from Piscivorous Conus achatinus. *Acta Biochim Biophys Sin* **2007**, *39*, 438–444.
192. Favreau, P.; Krimm, I.; Le Gall, F.; Bobenrieth, M.J.; Lamthanh, H.; Bouet, F.; Servent, D.; Molgo, J.; Ménez, A.; Letourneux, Y.; et al. Biochemical characterization and nuclear magnetic resonance structure of novel alpha-conotoxins isolated from the venom of Conus consors. *Biochemistry* **1999**, *38*, 6317–6326.
193. Gray, W.R.; Rivier, J.E.; Galyean, R.; Cruz, L.J.; Olivera, B.M. Conotoxin MI. Disulfide bonding and conformational states. *J. Biol. Chem.* **1983**, *258*, 12247–12251.
194. Maux, D.; Enjalbal, C.; Martinez, J.; Aubagnac, J.-L. New example of proline-induced fragmentation in electrospray ionization mass spectrometry of peptides. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1470–1475.
195. Chufán, E.E.; De, M.; Eipper, B.A.; Mains, R.E.; Amzel, L.M. Amidation of Bioactive Peptides: The Structure of the Lyase Domain of the Amidating Enzyme. *Structure* **2009**, *17*, 965–973.
196. Giribaldi, J.; Wilson, D.; Nicke, A.; El Hamdaoui, Y.; Laconde, G.; Faucherre, A.; Moha Ou Maati, H.; Daly, N.; Enjalbal, C.; Dutertre, S. Synthesis, Structure and Biological Activity of CIA and CIB, Two α-Conotoxins from the Predation-Evoked Venom of Conus catus. *Toxins* **2018**, *10*, 222.

197. Jard, S.; Bockaert, J. Stimulus-response coupling in neurohypophysial peptide target cells. *Physiological Reviews* **1975**, *55*, 489–536.
198. Dutertre, S.; Jin, A.-H.; Vetter, I.; Hamilton, B.; Sunagar, K.; Lavergne, V.; Dutertre, V.; Fry, B.G.; Antunes, A.; Venter, D.J.; et al. Evolution of separate predation- and defence-evoked venoms in carnivorous cone snails. *Nat Commun* **2014**, *5*.
199. Prashanth, J.R.; Brust, A.; Jin, A.-H.; Alewood, P.F.; Dutertre, S.; Lewis, R.J. Cone snail venomics: from novel biology to novel therapeutics. *Future Med Chem* **2014**, *6*, 1659–1675.
200. Essack, M.; Bajic, V.B.; Archer, J.A.C. Conotoxins that confer therapeutic possibilities. *Mar Drugs* **2012**, *10*, 1244–1265.
201. Lebbe, E.K.M.; Peigneur, S.; Wijesekara, I.; Tytgat, J. Conotoxins Targeting Nicotinic Acetylcholine Receptors: An Overview. *Mar Drugs* **2014**, *12*, 2970–3004.
202. Dutertre, S.; Nicke, A.; Tsetlin, V.I. Nicotinic acetylcholine receptor inhibitors derived from snake and snail venoms. *Neuropharmacology* **2017**, *127*, 196–223.
203. McIntosh, J.M.; Olivera, B.M.; Cruz, L.J. Conus peptides as probes for ion channels. *Meth. Enzymol.* **1999**, *294*, 605–624.
204. McIntosh, J.M.; Santos, A.D.; Olivera, B.M. Conus peptides targeted to specific nicotinic acetylcholine receptor subtypes. *Annu. Rev. Biochem.* **1999**, *68*, 59–88.
205. Olivera, B.M.; Cruz, L.J. Conotoxins, in retrospect. *Toxicon* **2001**, *39*, 7–14.
206. Gray, W.R.; Olivera, B.M.; Cruz, L.J. Peptide toxins from venomous Conus snails. *Annu. Rev. Biochem.* **1988**, *57*, 665–700.
207. Walker, C.S.; Steel, D.; Jacobsen, R.B.; Lirazan, M.B.; Cruz, L.J.; Hooper, D.; Shetty, R.; DelaCruz, R.C.; Nielsen, J.S.; Zhou, L.M.; et al. The T-superfamily of Conotoxins. *J. Biol. Chem.* **1999**, *274*, 30664–30671.
208. Jin, A.-H.; Daly, N.L.; Nevin, S.T.; Wang, C.-I.A.; Dutertre, S.; Lewis, R.J.; Adams, D.J.; Craik, D.J.; Alewood, P.F. Molecular Engineering of Conotoxins: The Importance of Loop Size to α-Conotoxin Structure and Function. *J. Med. Chem.* **2008**, *51*, 5575–5584.
209. Gotti, C.; Zoli, M.; Clementi, F. Brain nicotinic acetylcholine receptors: native subtypes and their relevance. *Trends Pharmacol. Sci.* **2006**, *27*, 482–491.
210. Zoli, M.; Pistillo, F.; Gotti, C. Diversity of native nicotinic receptor subtypes in mammalian brain. *Neuropharmacology* **2015**, *96*, 302–311.
211. Couturier, S.; Bertrand, D.; Matter, J.-M.; Hernandez, M.-C.; Bertrand, S.; Millar, N.; Valera, S.; Barkas, T.; Ballivet, M. A neuronal nicotinic acetylcholine receptor subunit (α7) is developmentally regulated and forms a homo-oligomeric channel blocked by α-BTX. *Neuron* **1990**, *5*, 847–856.
212. Rubboli, F.; Court, J.A.; Sala, C.; Morris, C.; Chini, B.; Perry, E.; Clementi, F. Distribution of Nicotinic Receptors in the Human Hippocampus and Thalamus. *European Journal of Neuroscience* **1994**, *6*, 1596–1604.
213. Gotti, C.; Clementi, F. Neuronal nicotinic receptors: from structure to pathology. *Progress in Neurobiology* **2004**, *74*, 363–396.
214. Steinlein, O.K.; Bertrand, D. Nicotinic receptor channelopathies and epilepsy. *Pflugers Arch - Eur J Physiol* **2010**, *460*, 495–503.
215. Picciotto, M.R.; Caldarone, B.J.; Brunzell, D.H.; Zachariou, V.; Stevens, T.R.; King, S.L. Neuronal nicotinic acetylcholine receptor subunit knockout mice: physiological and behavioral phenotypes and possible clinical implications. *Pharmacology & Therapeutics* **2001**, *92*, 89–108.
216. Hogg, R.C.; Hopping, G.; Alewood, P.F.; Adams, D.J.; Bertrand, D. α-Conotoxins PnIA and [A10L]PnIA Stabilize Different States of the α7-L247T Nicotinic Acetylcholine Receptor. *J. Biol. Chem.* **2003**, *278*, 26908–26914.
217. Kim, J.-B. Channelopathies. *Korean J Pediatr* **2014**, *57*, 1–18.
218. Ellison, M.; McIntosh, J.M.; Olivera, B.M. α-Conotoxins ImI and ImII SIMILAR α7 NICOTINIC RECEPTOR ANTAGONISTS ACT AT DIFFERENT SITES. *J. Biol. Chem.* **2003**, *278*, 757–764.

219. McIntosh, J.M.; Yoshikami, D.; Mahe, E.; Nielsen, D.B.; Rivier, J.E.; Gray, W.R.; Olivera, B.M. A nicotinic acetylcholine receptor ligand of unique specificity, alpha-conotoxin ImI. *J. Biol. Chem.* **1994**, *269*, 16733–16739.
220. Ellison, M.; Gao, F.; Wang, H.-L.; Sine, S.M.; McIntosh, J.M.; Olivera, B.M. α-Conotoxins ImI and ImII Target Distinct Regions of the Human α7 Nicotinic Acetylcholine Receptor and Distinguish Human Nicotinic Receptor Subtypes. *Biochemistry* **2004**, *43*, 16019–16026.
221. Armishaw, C.; Jensen, A.A.; Balle, T.; Clark, R.J.; Harpsøe, K.; Skonberg, C.; Liljefors, T.; Strømgaard, K. Rational Design of α-Conotoxin Analogues Targeting α7 Nicotinic Acetylcholine Receptors. *J Biol Chem* **2009**, *284*, 9498–9512.
222. Armishaw, C.J.; Singh, N.; Medina-Franco, J.L.; Clark, R.J.; Scott, K.C.M.; Houghten, R.A.; Jensen, A.A. A Synthetic Combinatorial Strategy for Developing α-Conotoxin Analogs as Potent α7 Nicotinic Acetylcholine Receptor Antagonists. *J Biol Chem* **2010**, *285*, 1809–1821.
223. Whiteaker, P.; Christensen, S.; Yoshikami, D.; Dowell, C.; Watkins, M.; Gulyas, J.; Rivier, J.; Olivera, B.M.; McIntosh, J.M. Discovery, Synthesis, and Structure Activity of a Highly Selective α7 Nicotinic Acetylcholine Receptor Antagonist. *Biochemistry* **2007**, *46*, 6628–6638.
224. Terpinskaya, T.I.; Osipov, A.V.; Kuznetsova, T.E.; Ryzhkovskaya, E.L.; Ulaschik, V.S.; Ivanov, I.A.; Tsetlin, V.I.; Utkin, Y.N. α-conotoxins revealed different roles of nicotinic cholinergic receptor subtypes in oncogenesis of Ehrlich tumor and in the associated inflammation. *Dokl. Biochem. Biophys.* **2015**, *463*, 216–219.
225. Whiteaker, P.; Marks, M.J.; Christensen, S.; Dowell, C.; Collins, A.C.; McIntosh, J.M. Synthesis and Characterization of ^{125}I-α-Conotoxin ArIB[V11L;V16A], a Selective α7 Nicotinic Acetylcholine Receptor Antagonist. *J Pharmacol Exp Ther* **2008**, *325*, 910–919.
226. Jin, A.-H.; Vetter, I.; Dutertre, S.; Abraham, N.; Emidio, N.B.; Inserra, M.; Murali, S.S.; Christie, M.J.; Alewood, P.F.; Lewis, R.J. MrIC, a Novel α-Conotoxin Agonist in the Presence of PNU at Endogenous α7 Nicotinic Acetylcholine Receptors. *Biochemistry* **2014**, *53*, 1–3.
227. Wang, S.; Zhao, C.; Liu, Z.; Wang, X.; Liu, N.; Du, W.; Dai, Q. Structural and Functional Characterization of a Novel α-Conotoxin Mr1.7 from Conus marmoreus Targeting Neuronal nAChR α3β2, α9α10 and α6/α3β2β3 Subtypes. *Mar Drugs* **2015**, *13*, 3259–3275.
228. Luo, S.; Zhangsun, D.; Schroeder, C.I.; Zhu, X.; Hu, Y.; Wu, Y.; Weltzin, M.M.; Eberhard, S.; Kaas, Q.; Craik, D.J.; et al. A novel α4/7-conotoxin LvIA from Conus lividus that selectively blocks α3β2 vs. α6/α3β2β3 nicotinic acetylcholine receptors. *FASEB J* **2014**, *28*, 1842–1853.
229. Wu, Y.; Zhangsun, D.; Zhu, X.; Kaas, Q.; Zhangsun, M.; Harvey, P.J.; Craik, D.J.; McIntosh, J.M.; Luo, S. α-Conotoxin [S9A]TxID Potently discriminates between α3β4 and α6/α3β4 Nicotinic Acetylcholine Receptors. *J. Med. Chem.* **2017**.
230. Banerjee, J.; Yongye, A.B.; Chang, Y.-P.; Gyanda, R.; Medina-Franco, J.L.; Armishaw, C.J. Design and synthesis of α-conotoxin GID analogues as selective α4β2 nicotinic acetylcholine receptor antagonists. *Biopolymers* **2014**, *102*, 78–87.
231. Azam, L.; Maskos, U.; Changeux, J.-P.; Dowell, C.D.; Christensen, S.; De Biasi, M.; McIntosh, J.M. α-Conotoxin BuIA[T5A;P6O]: a novel ligand that discriminates between α6ß4 and α6ß2 nicotinic acetylcholine receptors and blocks nicotine-stimulated norepinephrine release. *FASEB J.* **2010**, *24*, 5113–5123.
232. Azam, L.; Yoshikami, D.; McIntosh, J.M. Amino Acid Residues That Confer High Selectivity of the α6 Nicotinic Acetylcholine Receptor Subunit to α-Conotoxin MII[S4A,E11A,L15A]. *J. Biol. Chem.* **2008**, *283*, 11625–11632.
233. McIntosh, J.M.; Azam, L.; Staheli, S.; Dowell, C.; Lindstrom, J.M.; Kuryatov, A.; Garrett, J.E.; Marks, M.J.; Whiteaker, P. Analogs of α-Conotoxin MII Are Selective for α6-Containing Nicotinic Acetylcholine Receptors. *Mol Pharmacol* **2004**, *65*, 944–952.
234. Luo, S.; Zhangsun, D.; Wu, Y.; Zhu, X.; Hu, Y.; McIntyre, M.; Christensen, S.; Akcan, M.; Craik, D.J.; McIntosh, J.M. Characterization of a Novel α-Conotoxin from Conus textile That Selectively Targets α6/α3β2β3 Nicotinic Acetylcholine Receptors. *J. Biol. Chem.* **2013**, *288*, 894–902.

235. Hone, A.J.; Ruiz, M.; Scadden, M.; Christensen, S.; Gajewiak, J.; Azam, L.; McIntosh, J.M. Positional Scanning Mutagenesis of α-Conotoxin PeIA Identifies Critical Residues That Confer Potency and Selectivity for α6/α3β2β3 and α3β2 Nicotinic Acetylcholine Receptors. *J. Biol. Chem.* **2013**, *288*, 25428–25439.
236. Christensen, S.B.; Bandyopadhyay, P.K.; Olivera, B.M.; McIntosh, J.M. αS-Conotoxin GVIIIB Potently and Selectively Blocks α9α10 Nicotinic Acetylcholine Receptors. *Biochem Pharmacol* **2015**, *96*, 349–356.
237. Romero, H.K.; Christensen, S.B.; Di Cesare Mannelli, L.; Gajewiak, J.; Ramachandra, R.; Elmslie, K.S.; Vetter, D.E.; Ghelardini, C.; Iadonato, S.P.; Mercado, J.L.; et al. Inhibition of α9α10 nicotinic acetylcholine receptors prevents chemotherapy-induced neuropathic pain. *Proc Natl Acad Sci U S A* **2017**, *114*, E1825–E1832.
238. Spies, M.; Lips, K.S.; Kurzen, H.; Kummer, W.; Haberberger, R.V. Nicotinic acetylcholine receptors containing subunits alpha3 and alpha5 in rat nociceptive dorsal root ganglion neurons. *J. Mol. Neurosci.* **2006**, *30*, 55–56.
239. Marks, M.J.; Smith, K.W.; Collins, A.C. Differential Agonist Inhibition Identifies Multiple Epibatidine Binding Sites in Mouse Brain. *J Pharmacol Exp Ther* **1998**, *285*, 377–386.
240. Pang, X.; Liu, L.; Ngolab, J.; Zhao-Shea, R.; McIntosh, J.M.; Gardner, P.D.; Tapper, A.R. Habenula cholinergic neurons regulate anxiety during nicotine withdrawal via nicotinic acetylcholine receptors. *Neuropharmacology* **2016**, *107*, 294–304.
241. Hikosaka, O. The habenula: from stress evasion to value-based decision-making. *Nat Rev Neurosci* **2010**, *11*, 503–513.
242. Salas, R.; Sturm, R.; Boulter, J.; Biasi, M.D. Nicotinic Receptors in the Habenulo-Interpeduncular System Are Necessary for Nicotine Withdrawal in Mice. *J. Neurosci.* **2009**, *29*, 3014–3018.
243. McIntosh, J.M.; Absalom, N.; Chebib, M.; Elgoyhen, A.B.; Vincler, M. Alpha9 nicotinic acetylcholine receptors and the treatment of pain. *Biochem. Pharmacol.* **2009**, *78*, 693–702.
244. Callaghan, B.; Haythornthwaite, A.; Berecki, G.; Clark, R.J.; Craik, D.J.; Adams, D.J. Analgesic alpha-conotoxins Vc1.1 and Rg1A inhibit N-type calcium channels in rat sensory neurons via GABAB receptor activation. *J. Neurosci.* **2008**, *28*, 10943–10951.
245. Zhu, X.; Bi, J.; Yu, J.; Li, X.; Zhang, Y.; Zhangsun, D.; Luo, S. Recombinant Expression and Characterization of α-Conotoxin LvIA in Escherichia coli. *Marine Drugs* **2016**, *14*, 11.
246. Dutertre, S.; Nicke, A.; Lewis, R.J. β2 Subunit Contribution to 4/7 α-Conotoxin Binding to the Nicotinic Acetylcholine Receptor. *J. Biol. Chem.* **2005**, *280*, 30460–30468.
247. Kompella, S.N.; Cuny, H.; Hung, A.; Adams, D.J. Molecular Basis for Differential Sensitivity of *α*-Conotoxin RegIIA at Rat and Human Neuronal Nicotinic Acetylcholine Receptors. *Mol Pharmacol* **2015**, *88*, 993–1001.
248. Everhart, D.; Cartier, G.E.; Malhotra, A.; Gomes, A.V.; McIntosh, J.M.; Luetje, C.W. Determinants of Potency on α-Conotoxin MII, a Peptide Antagonist of Neuronal Nicotinic Receptors. *Biochemistry* **2004**, *43*, 2732–2737.
249. Napier, I.A.; Klimis, H.; Rycroft, B.K.; Jin, A.H.; Alewood, P.F.; Motin, L.; Adams, D.J.; Christie, M.J. Intrathecal α-conotoxins Vc1.1, AuIB and MII acting on distinct nicotinic receptor subtypes reverse signs of neuropathic pain. *Neuropharmacology* **2012**, *62*, 2202–2207.
250. Klimis, H.; Adams, D.J.; Callaghan, B.; Nevin, S.; Alewood, P.F.; Vaughan, C.W.; Mozar, C.A.; Christie, M.J. a novel mechanism of inhibition of high-voltage activated calcium channels by α-conotoxins contributes to relief of nerve injury-induced neuropathic pain. *Pain* **2011**, *152*, 259–266.
251. Everhart, D.; Reiller, E.; Mirzoian, A.; McIntosh, J.M.; Malhotra, A.; Luetje, C.W. Identification of Residues That Confer α-Conotoxin-PnIA Sensitivity on the α3 Subunit of Neuronal Nicotinic Acetylcholine Receptors. *J Pharmacol Exp Ther* **2003**, *306*, 664–670.
252. Lin, B.; Xu, M.; Zhu, X.; Wu, Y.; Liu, X.; Zhangsun, D.; Hu, Y.; Xiang, S.-H.; Kasheverov, I.E.; Tsetlin, V.I.; et al. From crystal structure of α-conotoxin GIC in complex with *Ac*-AChBP to

molecular determinants of its high selectivity for α3β2 nAChR. *Scientific Reports* **2016**, *6*, srep22349.
253. Luo, S.; Akondi, K.B.; Zhangsun, D.; Wu, Y.; Zhu, X.; Hu, Y.; Christensen, S.; Dowell, C.; Daly, N.L.; Craik, D.J.; et al. Atypical α-Conotoxin LtIA from Conus litteratus Targets a Novel Microsite of the α3β2 Nicotinic Receptor. *J. Biol. Chem.* **2010**, *285*, 12355–12366.
254. Gotti, C.; Moretti, M.; Gaimarri, A.; Zanardi, A.; Clementi, F.; Zoli, M. Heterogeneity and complexity of native brain nicotinic receptors. *Biochemical Pharmacology* **2007**, *74*, 1102–1111.
255. Hernandez, S.C.; Vicini, S.; Xiao, Y.; Dávila-García, M.I.; Yasuda, R.P.; Wolfe, B.B.; Kellar, K.J. The Nicotinic Receptor in the Rat Pineal Gland Is an α3β4 Subtype. *Mol Pharmacol* **2004**, *66*, 978–987.
256. Turner, J.R.; Kellar, K.J. Nicotinic Cholinergic Receptors in the Rat Cerebellum: Multiple Heteromeric Subtypes. *J. Neurosci.* **2005**, *25*, 9258–9265.
257. Grady, S.R.; Moretti, M.; Zoli, M.; Marks, M.J.; Zanardi, A.; Pucci, L.; Clementi, F.; Gotti, C. Rodent habenulo-interpeduncular pathway expresses a large variety of uncommon nAChR subtypes, but only the α3β4* and α3β3β4* subtypes mediate acetylcholine release. *J Neurosci* **2009**, *29*, 2272–2282.
258. Jensen, A.A.; Frølund, B.; Liljefors, T.; Krogsgaard-Larsen, P. Neuronal nicotinic acetylcholine receptors: structural revelations, target identifications, and therapeutic inspirations. *J. Med. Chem.* **2005**, *48*, 4705–4745.
259. Changeux, J.-P. Nicotine addiction and nicotinic receptors: lessons from genetically modified mice. *Nat Rev Neurosci* **2010**, *11*, 389–401.
260. Luo, S.; Kulak, J.M.; Cartier, G.E.; Jacobsen, R.B.; Yoshikami, D.; Olivera, B.M.; McIntosh, J.M. α-Conotoxin AuIB Selectively Blocks α3β4 Nicotinic Acetylcholine Receptors and Nicotine-Evoked Norepinephrine Release. *J. Neurosci.* **1998**, *18*, 8571–8579.
261. Dutton, J.L.; Bansal, P.S.; Hogg, R.C.; Adams, D.J.; Alewood, P.F.; Craik, D.J. A New Level of Conotoxin Diversity, a Non-native Disulfide Bond Connectivity in α-Conotoxin AuIB Reduces Structural Definition but Increases Biological Activity. *J. Biol. Chem.* **2002**, *277*, 48849–48857.
262. Grishin, A.A.; Cuny, H.; Hung, A.; Clark, R.J.; Brust, A.; Akondi, K.; Alewood, P.F.; Craik, D.J.; Adams, D.J. Identifying Key Amino Acid Residues That Affect α-Conotoxin AuIB Inhibition of α3β4 Nicotinic Acetylcholine Receptors. *J. Biol. Chem.* **2013**, *288*, 34428–34442.
263. Cuny, H.; Kompella, S.N.; Tae, H.-S.; Yu, R.; Adams, D.J. Key Structural Determinants in the Agonist Binding Loops of Human β2 and β4 Nicotinic Acetylcholine Receptor Subunits Contribute to α3β4 Subtype Selectivity of α-Conotoxins. *J. Biol. Chem.* **2016**, *291*, 23779–23792.
264. Franco, A.; Kompella, S.N.; Akondi, K.B.; Melaun, C.; Daly, N.L.; Luetje, C.W.; Alewood, P.F.; Craik, D.J.; Adams, D.J.; Marí, F. RegIIA: An α4/7-conotoxin from the venom of Conus regius that potently blocks α3β4 nAChRs. *Biochemical Pharmacology* **2012**, *83*, 419–426.
265. Kompella, S.N.; Hung, A.; Clark, R.J.; Marí, F.; Adams, D.J. Alanine Scan of α-Conotoxin RegIIA Reveals a Selective α3β4 Nicotinic Acetylcholine Receptor Antagonist. *J Biol Chem* **2015**, *290*, 1039–1048.
266. Luo, S.; Zhangsun, D.; Zhu, X.; Wu, Y.; Hu, Y.; Christensen, S.; Harvey, P.J.; Akcan, M.; Craik, D.J.; McIntosh, J.M. Characterization of a Novel Alpha-Conotoxin TxID from Conus textile that Potently Blocks rat Alpha3beta4 Nicotinic Acetylcholine Receptors. *J Med Chem* **2013**, *56*, 9655–9663.
267. Chang, Y.-P.; Banerjee, J.; Dowell, C.; Wu, J.; Gyanda, R.; Houghten, R.A.; Toll, L.; McIntosh, J.M.; Armishaw, C.J. Discovery of a Potent and Selective α3β4 Nicotinic Acetylcholine Receptor Antagonist from an α-Conotoxin Synthetic Combinatorial Library. *J. Med. Chem.* **2014**, *57*, 3511–3521.
268. Gotti, C.; Moretti, M.; Clementi, F.; Riganti, L.; McIntosh, J.M.; Collins, A.C.; Marks, M.J.; Whiteaker, P. Expression of Nigrostriatal α6-Containing Nicotinic Acetylcholine Receptors Is

Selectively Reduced, but Not Eliminated, by β3 Subunit Gene Deletion. *Mol Pharmacol* **2005**, *67*, 2007–2015.
269. Grady, S.R.; Salminen, O.; McIntosh, J.M.; Marks, M.J.; Collins, A.C. Mouse Striatal Dopamine Nerve Terminals Express α4α5β2 and Two Stoichiometric Forms of α4β2*-Nicotinic Acetylcholine Receptors. *J Mol Neurosci* **2010**, *40*, 91–95.
270. Wonnacott, S. Presynaptic nicotinic ACh receptors. *Trends Neurosci.* **1997**, *20*, 92–98.
271. Flores, C.M.; Rogers, S.W.; Pabreza, L.A.; Wolfe, B.B.; Kellar, K.J. A subtype of nicotinic cholinergic receptor in rat brain is composed of alpha 4 and beta 2 subunits and is up-regulated by chronic nicotine treatment. *Mol Pharmacol* **1992**, *41*, 31–37.
272. Rueter, L.E.; Donnelly-Roberts, D.L.; Curzon, P.; Briggs, C.A.; Anderson, D.J.; Bitner, R.S. A-85380: a pharmacological probe for the preclinical and clinical investigation of the alphabeta neuronal nicotinic acetylcholine receptor. *CNS Drug Rev* **2006**, *12*, 100–112.
273. Ebbert, J.O. Emerging drugs for the treatment of tobacco dependence. *Expert Opinion on Emerging Drugs* **2009**, *14*, 23–32.
274. Taly, A.; Corringer, P.-J.; Guedin, D.; Lestage, P.; Changeux, J.-P. Nicotinic receptors: allosteric transitions and therapeutic targets in the nervous system. *Nat Rev Drug Discov* **2009**, *8*, 733–750.
275. Millard, E.L.; Nevin, S.T.; Loughnan, M.L.; Nicke, A.; Clark, R.J.; Alewood, P.F.; Lewis, R.J.; Adams, D.J.; Craik, D.J.; Daly, N.L. Inhibition of Neuronal Nicotinic Acetylcholine Receptor Subtypes by α-Conotoxin GID and Analogues. *J. Biol. Chem.* **2009**, *284*, 4944–4951.
276. Nicke, A.; Loughnan, M.L.; Millard, E.L.; Alewood, P.F.; Adams, D.J.; Daly, N.L.; Craik, D.J.; Lewis, R.J. Isolation, Structure, and Activity of GID, a Novel α4/7-Conotoxin with an Extended N-terminal Sequence. *J. Biol. Chem.* **2003**, *278*, 3137–3144.
277. Suresh, A.; Hung, A. Molecular simulation study of the unbinding of α-conotoxin [Υ4E]GID at the α7 and α4β2 neuronal nicotinic acetylcholine receptors. *J. Mol. Graph. Model.* **2016**, *70*, 109–121.
278. Wang, S.; Du, T.; Liu, Z.; Wang, S.; Wu, Y.; Ding, J.; Jiang, L.; Dai, Q. Characterization of a T-superfamily conotoxin TxVC from Conus textile that selectively targets neuronal nAChR subtypes. *Biochemical and Biophysical Research Communications* **2014**, *454*, 151–156.
279. Beissner, M.; Dutertre, S.; Schemm, R.; Danker, T.; Sporning, A.; Grubmüller, H.; Nicke, A. Efficient Binding of 4/7 α-Conotoxins to Nicotinic α4β2 Receptors Is Prevented by Arg185 and Pro195 in the α4 Subunit. *Mol Pharmacol* **2012**, *82*, 711–718.
280. Liu, J.; McGlinn, A.M.; Fernandes, A.; Milam, A.H.; Strang, C.E.; Andison, M.E.; Lindstrom, J.M.; Keyser, K.T.; Stone, R.A. Nicotinic Acetylcholine Receptor Subunits in Rhesus Monkey Retina. *Invest Ophthalmol Vis Sci* **2009**, *50*, 1408–1415.
281. Hone, A.J.; Scadden, M.; Gajewiak, J.; Christensen, S.; Lindstrom, J.; McIntosh, J.M. α-Conotoxin PeIA[S9H,V10A,E14N] Potently and Selectively Blocks α6β2β3 versus α6β4 Nicotinic Acetylcholine Receptors. *Mol Pharmacol* **2012**, *82*, 972–982.
282. Mackey, E.D.W.; Engle, S.E.; Kim, M.R.; O'Neill, H.C.; Wageman, C.R.; Patzlaff, N.E.; Wang, Y.; Grady, S.R.; McIntosh, J.M.; Marks, M.J.; et al. α6* Nicotinic Acetylcholine Receptor Expression and Function in a Visual Salience Circuit. *J. Neurosci.* **2012**, *32*, 10226–10237.
283. Klink, R.; d'Exaerde, A. de K.; Zoli, M.; Changeux, J.-P. Molecular and Physiological Diversity of Nicotinic Acetylcholine Receptors in the Midbrain Dopaminergic Nuclei. *J. Neurosci.* **2001**, *21*, 1452–1463.
284. Azam, L.; Winzer-Serhan, U.H.; Chen, Y.; Leslie, F.M. Expression of neuronal nicotinic acetylcholine receptor subunit mRNAs within midbrain dopamine neurons. *J. Comp. Neurol.* **2002**, *444*, 260–274.
285. Champtiaux, N.; Gotti, C.; Cordero-Erausquin, M.; David, D.J.; Przybylski, C.; Léna, C.; Clementi, F.; Moretti, M.; Rossi, F.M.; Novère, N.L.; et al. Subunit Composition of Functional Nicotinic Receptors in Dopaminergic Neurons Investigated with Knock-Out Mice. *J. Neurosci.* **2003**, *23*, 7820–7829.

286. Pons, S.; Fattore, L.; Cossu, G.; Tolu, S.; Porcu, E.; McIntosh, J.M.; Changeux, J.P.; Maskos, U.; Fratta, W. Crucial Role of α4 and α6 Nicotinic Acetylcholine Receptor Subunits from Ventral Tegmental Area in Systemic Nicotine Self-Administration. *J. Neurosci.* **2008**, *28*, 12318–12327.
287. Yang, K.; Jin, G.; Wu, J. Mysterious α6-containing nAChRs: function, pharmacology, and pathophysiology. *Acta Pharmacol Sin* **2009**, *30*, 740–751.
288. Quik, M.; McIntosh, J.M. Striatal α6* Nicotinic Acetylcholine Receptors: Potential Targets for Parkinson's Disease Therapy. *J Pharmacol Exp Ther* **2006**, *316*, 481–489.
289. Kim, H.-W.; McIntosh, J.M. α6 nAChR subunit residues that confer α-conotoxin BuIA selectivity. *FASEB J* **2012**, *26*, 4102–4110.
290. Hone, A.J.; McIntosh, J.M.; Azam, L.; Lindstrom, J.; Lucero, L.; Whiteaker, P.; Passas, J.; Blazquez, J.; Albillos, A. -Conotoxins Identify the 3 4* Subtype as the Predominant Nicotinic Acetylcholine Receptor Expressed in Human Adrenal Chromaffin Cells. *Molecular Pharmacology* **2015**, *88*, 881–893.
291. Hone, A.J.; Meyer, E.L.; McIntyre, M.; McIntosh, J.M. Nicotinic acetylcholine receptors in dorsal root ganglion neurons include the α6β4* subtype. *FASEB J* **2012**, *26*, 917–926.
292. Dowell, C.; Olivera, B.M.; Garrett, J.E.; Staheli, S.T.; Watkins, M.; Kuryatov, A.; Yoshikami, D.; Lindstrom, J.M.; McIntosh, J.M. α-Conotoxin PIA Is Selective for α6 Subunit-Containing Nicotinic Acetylcholine Receptors. *J. Neurosci.* **2003**, *23*, 8445–8452.
293. Pucci, L.; Grazioso, G.; Dallanoce, C.; Rizzi, L.; De Micheli, C.; Clementi, F.; Bertrand, S.; Bertrand, D.; Longhi, R.; De Amici, M.; et al. Engineering of α-conotoxin MII-derived peptides with increased selectivity for native α6β2* nicotinic acetylcholine receptors. *FASEB J.* **2011**, *25*, 3775–3789.
294. Clark, R.J.; Fischer, H.; Dempster, L.; Daly, N.L.; Rosengren, K.J.; Nevin, S.T.; Meunier, F.A.; Adams, D.J.; Craik, D.J. Engineering stable peptide toxins by means of backbone cyclization: Stabilization of the α-conotoxin MII. *Proc Natl Acad Sci U S A* **2005**, *102*, 13767–13772.
295. Ray, M.; Bohr, I.; McIntosh, J.M.; Ballard, C.; McKeith, I.; Chalon, S.; Guilloteau, D.; Perry, R.; Perry, E.; Court, J.A.; et al. Involvement of alpha6/alpha3 neuronal nicotinic acetylcholine receptors in neuropsychiatric features of Dementia with Lewy bodies: [(125)I]-alpha-conotoxin MII binding in the thalamus and striatum. *Neurosci. Lett.* **2004**, *372*, 220–225.
296. Cui, C.; Booker, T.K.; Allen, R.S.; Grady, S.R.; Whiteaker, P.; Marks, M.J.; Salminen, O.; Tritto, T.; Butt, C.M.; Allen, W.R.; et al. The β3 Nicotinic Receptor Subunit: A Component of α-Conotoxin MII-Binding Nicotinic Acetylcholine Receptors that Modulate Dopamine Release and Related Behaviors. *J. Neurosci.* **2003**, *23*, 11045–11053.
297. Quik, M.; Polonskaya, Y.; Kulak, J.M.; McIntosh, J.M. Vulnerability of 125I-α-Conotoxin MII Binding Sites to Nigrostriatal Damage in Monkey. *J. Neurosci.* **2001**, *21*, 5494–5500.
298. Whiteaker, P.; McIntosh, J.M.; Luo, S.; Collins, A.C.; Marks, M.J. 125I-α-Conotoxin MII Identifies a Novel Nicotinic Acetylcholine Receptor Population in Mouse Brain. *Mol Pharmacol* **2000**, *57*, 913–925.
299. Bibevski, S.; Zhou, Y.; McIntosh, J.M.; Zigmond, R.E.; Dunlap, M.E. Functional nicotinic acetylcholine receptors that mediate ganglionic transmission in cardiac parasympathetic neurons. *J. Neurosci.* **2000**, *20*, 5076–5082.
300. Mao, D.; Yasuda, R.P.; Fan, H.; Wolfe, B.B.; Kellar, K.J. Heterogeneity of Nicotinic Cholinergic Receptors in Rat Superior Cervical and Nodose Ganglia. *Molecular Pharmacology* **2006**, *70*, 1693–1699.
301. David, R.; Ciuraszkiewicz, A.; Simeone, X.; Orr-Urtreger, A.; Papke, R.L.; McIntosh, J.M.; Huck, S.; Scholze, P. Biochemical and functional properties of distinct nicotinic acetylcholine receptors in the superior cervical ganglion of mice with targeted deletions of nAChR subunit genes. *European Journal of Neuroscience* **2010**, *31*, 978–993.
302. Plazas, P.V.; Katz, E.; Gomez-Casati, M.E.; Bouzat, C.; Elgoyhen, A.B. Stoichiometry of the α9α10 Nicotinic Cholinergic Receptor. *J. Neurosci.* **2005**, *25*, 10905–10912.
303. Lips, K.S.; Pfeil, U.; Kummer, W. Coexpression of alpha 9 and alpha 10 nicotinic acetylcholine receptors in rat dorsal root ganglion neurons. *Neuroscience* **2002**, *115*, 1–5.

304. Haberberger, R.V.; Bernardini, N.; Kress, M.; Hartmann, P.; Lips, K.S.; Kummer, W. Nicotinic acetylcholine receptor subtypes in nociceptive dorsal root ganglion neurons of the adult rat. *Auton Neurosci* **2004**, *113*, 32–42.
305. Callaghan, B.; Adams, D.J. Analgesic α-conotoxins Vc1.1 and RgIA inhibit N-type calcium channels in sensory neurons of α9 nicotinic receptor knockout mice. *Channels (Austin)* **2010**, *4*, 51–54.
306. Elgoyhen, A.B.; Johnson, D.S.; Boulter, J.; Vetter, D.E.; Heinemann, S. Alpha 9: an acetylcholine receptor with novel pharmacological properties expressed in rat cochlear hair cells. *Cell* **1994**, *79*, 705–715.
307. Elgoyhen, A.B.; Vetter, D.E.; Katz, E.; Rothlin, C.V.; Heinemann, S.F.; Boulter, J. alpha10: a determinant of nicotinic cholinergic receptor function in mammalian vestibular and cochlear mechanosensory hair cells. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 3501–3506.
308. Vetter, D.E.; Liberman, M.C.; Mann, J.; Barhanin, J.; Boulter, J.; Brown, M.C.; Saffiote-Kolman, J.; Heinemann, S.F.; Elgoyhen, A.B. Role of alpha9 nicotinic ACh receptor subunits in the development and function of cochlear efferent innervation. *Neuron* **1999**, *23*, 93–103.
309. Koval, L.; Lykhmus, O.; Zhmak, M.; Khruschov, A.; Tsetlin, V.; Magrini, E.; Viola, A.; Chernyavsky, A.; Qian, J.; Grando, S.; et al. Differential involvement of α4β2, α7 and α9α10 nicotinic acetylcholine receptors in B lymphocyte activation in vitro. *Int. J. Biochem. Cell Biol.* **2011**, *43*, 516–524.
310. Peng, H.; Ferris, R.L.; Matthews, T.; Hiel, H.; Lopez-Albaitero, A.; Lustig, L.R. Characterization of the human nicotinic acetylcholine receptor subunit alpha (alpha) 9 (CHRNA9) and alpha (alpha) 10 (CHRNA10) in lymphocytes. *Life Sci.* **2004**, *76*, 263–280.
311. Satkunanathan, N.; Livett, B.; Gayler, K.; Sandall, D.; Down, J.; Khalil, Z. Alpha-conotoxin Vc1.1 alleviates neuropathic pain and accelerates functional recovery of injured neurones. *Brain Res.* **2005**, *1059*, 149–158.
312. Wu, C.-H.; Lee, C.-H.; Ho, Y.-S. Nicotinic Acetylcholine Receptor-Based Blockade: Applications of Molecular Targets for Cancer Therapy. *Clin Cancer Res* **2011**, *17*, 3533–3541.
313. Mohammadi, S.; Christie, M.J. α9-nicotinic acetylcholine receptors contribute to the maintenance of chronic mechanical hyperalgesia, but not thermal or mechanical allodynia. *Mol Pain* **2014**, *10*, 64.
314. McIntosh, J.M.; Plazas, P.V.; Watkins, M.; Gomez-Casati, M.E.; Olivera, B.M.; Elgoyhen, A.B. A Novel α-Conotoxin, PeIA, Cloned from Conus pergrandis, Discriminates between Rat α9α10 and α7 Nicotinic Cholinergic Receptors. *J. Biol. Chem.* **2005**, *280*, 30107–30112.
315. Xu, S.; Zhang, T.; Kompella, S.N.; Yan, M.; Lu, A.; Wang, Y.; Shao, X.; Chi, C.; Adams, D.J.; Ding, J.; et al. Conotoxin αD-GeXXA utilizes a novel strategy to antagonize nicotinic acetylcholine receptors. *Scientific Reports* **2015**, *5*, srep14261.
316. Yang, L.; Tae, H.-S.; Fan, Z.; Shao, X.; Xu, S.; Zhao, S.; Adams, D.; Wang, C. A Novel Lid-Covering Peptide Inhibitor of Nicotinic Acetylcholine Receptors Derived from αD-Conotoxin GeXXA. *Marine Drugs* **2017**, *15*, 164.
317. Luo, S.; Zhangsun, D.; Harvey, P.J.; Kaas, Q.; Wu, Y.; Zhu, X.; Hu, Y.; Li, X.; Tsetlin, V.I.; Christensen, S.; et al. Cloning, synthesis, and characterization of αO-conotoxin GeXIVA, a potent α9α10 nicotinic acetylcholine receptor antagonist. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, E4026-4035.
318. Li, X.; Hu, Y.; Wu, Y.; Huang, Y.; Yu, S.; Ding, Q.; Zhangsun, D.; Luo, S. Anti-hypersensitive effect of intramuscular administration of αO-conotoxin GeXIVA[1,2] and GeXIVA[1,4] in rats of neuropathic pain. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **2016**, *66*, 112–119.
319. Ellison, M.; Haberlandt, C.; Gomez-Casati, M.E.; Watkins, M.; Elgoyhen, A.B.; McIntosh, J.M.; Olivera, B.M. Alpha-RgIA: a novel conotoxin that specifically and potently blocks the alpha9alpha10 nAChR. *Biochemistry* **2006**, *45*, 1511–1517.
320. Ellison, M.; Feng, Z.-P.; Park, A.J.; Zhang, X.; Olivera, B.M.; McIntosh, J.M.; Norton, R.S. Alpha-RgIA, a novel conotoxin that blocks the alpha9alpha10 nAChR: structure and identification of key receptor-binding residues. *J. Mol. Biol.* **2008**, *377*, 1216–1227.

321. Azam, L.; McIntosh, J.M. Molecular basis for the differential sensitivity of rat and human α9α10 nAChRs to α-conotoxin RgIA. *J Neurochem* **2012**, *122*, 1137–1144.
322. Azam, L.; Papakyriakou, A.; Zouridakis, M.; Giastas, P.; Tzartos, S.J.; McIntosh, J.M. Molecular interaction of α-conotoxin RgIA with the rat α9α10 nicotinic acetylcholine receptor. *Mol. Pharmacol.* **2015**, *87*, 855–864.
323. Halai, R.; Callaghan, B.; Daly, N.L.; Clark, R.J.; Adams, D.J.; Craik, D.J. Effects of cyclization on stability, structure, and activity of α-conotoxin RgIA at the α9α10 nicotinic acetylcholine receptor and GABA(B) receptor. *J. Med. Chem.* **2011**, *54*, 6984–6992.
324. Chhabra, S.; Belgi, A.; Bartels, P.; van Lierop, B.J.; Robinson, S.D.; Kompella, S.N.; Hung, A.; Callaghan, B.P.; Adams, D.J.; Robinson, A.J.; et al. Dicarba analogues of α-conotoxin RgIA. Structure, stability, and activity at potential pain targets. *J. Med. Chem.* **2014**, *57*, 9933–9944.
325. Vincler, M.; Wittenauer, S.; Parker, R.; Ellison, M.; Olivera, B.M.; McIntosh, J.M. Molecular mechanism for analgesia involving specific antagonism of alpha9alpha10 nicotinic acetylcholine receptors. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 17880–17884.
326. Pacini, A.; Micheli, L.; Maresca, M.; Branca, J.J.V.; McIntosh, J.M.; Ghelardini, C.; Di Cesare Mannelli, L. The α9α10 nicotinic receptor antagonist α-conotoxin RgIA prevents neuropathic pain induced by oxaliplatin treatment. *Exp. Neurol.* **2016**, *282*, 37–48.
327. Di Cesare Mannelli, L.; Cinci, L.; Micheli, L.; Zanardelli, M.; Pacini, A.; McIntosh, J.M.; Ghelardini, C. α-conotoxin RgIA protects against the development of nerve injury-induced chronic pain and prevents both neuronal and glial derangement. *Pain* **2014**, *155*, 1986–1995.
328. Christensen, S.B.; Hone, A.J.; Roux, I.; Kniazeff, J.; Pin, J.-P.; Upert, G.; Servent, D.; Glowatzki, E.; McIntosh, J.M. RgIA4 Potently Blocks Mouse α9α10 nAChRs and Provides Long Lasting Protection against Oxaliplatin-Induced Cold Allodynia. *Front Cell Neurosci* **2017**, *11*, 219.
329. Colomer, C.; Olivos-Oré, L.A.; Vincent, A.; McIntosh, J.M.; Artalejo, A.R.; Guérineau, N.C. Functional characterization of alpha9-containing cholinergic nicotinic receptors in the rat adrenal medulla: implication in stress-induced functional plasticity. *J. Neurosci.* **2010**, *30*, 6732–6742.
330. Halai, R.; Clark, R.J.; Nevin, S.T.; Jensen, J.E.; Adams, D.J.; Craik, D.J. Scanning mutagenesis of alpha-conotoxin Vc1.1 reveals residues crucial for activity at the alpha9alpha10 nicotinic acetylcholine receptor. *J. Biol. Chem.* **2009**, *284*, 20275–20284.
331. Yu, R.; Kompella, S.N.; Adams, D.J.; Craik, D.J.; Kaas, Q. Determination of the α-conotoxin Vc1.1 binding site on the α9α10 nicotinic acetylcholine receptor. *J. Med. Chem.* **2013**, *56*, 3557–3567.
332. Sandall, D.W.; Satkunanathan, N.; Keays, D.A.; Polidano, M.A.; Liping, X.; Pham, V.; Down, J.G.; Khalil, Z.; Livett, B.G.; Gayler, K.R. A Novel α-Conotoxin Identified by Gene Sequencing Is Active in Suppressing the Vascular Response to Selective Stimulation of Sensory Nerves in Vivo. *Biochemistry* **2003**, *42*, 6904–6911.
333. Clark, R.J.; Jensen, J.; Nevin, S.T.; Callaghan, B.P.; Adams, D.J.; Craik, D.J. The Engineering of an Orally Active Conotoxin for the Treatment of Neuropathic Pain. *Angewandte Chemie International Edition* **2010**, *49*, 6545–6548.
334. Wu, J.; Liu, Q.; Tang, P.; Mikkelsen, J.D.; Shen, J.; Whiteaker, P.; Yakel, J.L. Heteromeric α7β2 Nicotinic Acetylcholine Receptors in the Brain. *Trends Pharmacol Sci* **2016**, *37*, 562–574.
335. Moretti, M.; Zoli, M.; George, A.A.; Lukas, R.J.; Pistillo, F.; Maskos, U.; Whiteaker, P.; Gotti, C. The novel α7β2-nicotinic acetylcholine receptor subtype is expressed in mouse and human basal forebrain: biochemical and pharmacological characterization. *Mol. Pharmacol.* **2014**, *86*, 306–317.
336. Liu, Q.; Huang, Y.; Shen, J.; Steffensen, S.; Wu, J. Functional α7β2 nicotinic acetylcholine receptors expressed in hippocampal interneurons exhibit high sensitivity to pathological level of amyloid β peptides. *BMC Neurosci* **2012**, *13*, 155.
337. Liu, Q.; Huang, Y.; Xue, F.; Simard, A.; DeChon, J.; Li, G.; Zhang, J.; Lucero, L.; Wang, M.; Sierks, M.; et al. A novel nicotinic acetylcholine receptor subtype in basal forebrain cholinergic neurons with high sensitivity to amyloid peptides. *J. Neurosci.* **2009**, *29*, 918–929.

338. Azam, L.; Winzer-Serhan, U.; Leslie, F.M. Co-expression of alpha7 and beta2 nicotinic acetylcholine receptor subunit mRNAs within rat brain cholinergic neurons. *Neuroscience* **2003**, *119*, 965–977.
339. Sudweeks, S.N.; Yakel, J.L. Functional and molecular characterization of neuronal nicotinic ACh receptors in rat CA1 hippocampal neurons. *J. Physiol. (Lond.)* **2000**, *527 Pt 3*, 515–528.
340. Thomsen, M.S.; Zwart, R.; Ursu, D.; Jensen, M.M.; Pinborg, L.H.; Gilmour, G.; Wu, J.; Sher, E.; Mikkelsen, J.D. α7 and β2 Nicotinic Acetylcholine Receptor Subunits Form Heteromeric Receptor Complexes that Are Expressed in the Human Cortex and Display Distinct Pharmacological Properties. *PLoS ONE* **2015**, *10*, e0130572.
341. Himaya, S.W.A.; Jin, A.-H.; Dutertre, S.; Giacomotto, J.; Mohialdeen, H.; Vetter, I.; Alewood, P.F.; Lewis, R.J. Comparative Venomics Reveals the Complex Prey Capture Strategy of the Piscivorous Cone Snail Conus catus. *J. Proteome Res.* **2015**, *14*, 4372–4381.
342. Puillandre, N.; Duda, T.F.; Meyer, C.; Olivera, B.M.; Bouchet, P. One, four or 100 genera? A new classification of the cone snails. *J Molluscan Stud* **2015**, *81*, 1–23.
343. Sharman, J.L.; Benson, H.E.; Pawson, A.J.; Lukito, V.; Mpamhanga, C.P.; Bombail, V.; Davenport, A.P.; Peters, J.A.; Spedding, M.; Harmar, A.J.; et al. IUPHAR-DB: updated database content and new features. *Nucleic Acids Res.* **2013**, *41*, D1083-1088.
344. Vetter, I.; Lewis, R.J. Therapeutic potential of cone snail venom peptides (conopeptides). *Curr Top Med Chem* **2012**, *12*, 1546–1552.
345. Becker, S.; Terlau, H. Toxins from cone snails: properties, applications and biotechnological production. *Appl Microbiol Biotechnol* **2008**, *79*, 1–9.
346. Brady, R.; Baell, J.; Norton, R. Strategies for the Development of Conotoxins as New Therapeutic Leads. *Marine Drugs* **2013**, *11*, 2293–2313.
347. Pennington, M.W.; Czerwinski, A.; Norton, R.S. Peptide therapeutics from venom: Current status and potential. *Bioorganic & Medicinal Chemistry* **2017**.
348. Dutton, J.L.; Craik, D.J. alpha-Conotoxins: nicotinic acetylcholine receptor antagonists as pharmacological tools and potential drug leads. *Curr. Med. Chem.* **2001**, *8*, 327–344.
349. Cooper, E.; Couturier, S.; Ballivet, M. Pentameric structure and subunit stoichiometry of a neuronal nicotinic acetylcholine receptor. *Nature* **1991**, *350*, 235–238.
350. Twede, V.D.; Miljanich, G.; Olivera, B.M.; Bulaj, G. Neuroprotective and cardioprotective conopeptides: An emerging class of drug leads. *Curr Opin Drug Discov Devel* **2009**, *12*, 231–239.
351. Exley, R.; Clements, M.A.; Hartung, H.; McIntosh, J.M.; Cragg, S.J. Alpha6-containing nicotinic acetylcholine receptors dominate the nicotine control of dopamine neurotransmission in nucleus accumbens. *Neuropsychopharmacology* **2008**, *33*, 2158–2166.
352. Garza, A.; Huang, L.Z.; Son, J.-H.; Winzer-Serhan, U.H. Expression of nicotinic acetylcholine receptors and subunit messenger RNAs in the enteric nervous system of the neonatal rat. *Neuroscience* **2009**, *158*, 1521–1529.
353. Zdanowski, R.; Krzyżowska, M.; Ujazdowska, D.; Lewicka, A.; Lewicki, S. Role of α7 nicotinic receptor in the immune system and intracellular signaling pathways. *Central European Journal of Immunology* **2015**, *3*, 373–379.
354. Spence, I.; Gillessen, D.; Gregson, R.P.; Quinn, R.J. Characterization of the neurotoxic constituents of Conus geographus (L) venom. *Life Sciences* **1977**, *21*, 1759–1769.
355. Dutertre, S.; Jin, A.-H.; Alewood, P.F.; Lewis, R.J. Intraspecific variations in Conus geographus defence-evoked venom and estimation of the human lethal dose. *Toxicon* **2014**, *91*, 135–144.
356. Gyanda, R.; Banerjee, J.; Chang, Y.-P.; Phillips, A.M.; Toll, L.; Armishaw, C.J. Oxidative folding and preparation of α-conotoxins for use in high-throughput structure–activity relationship studies. *J. Pept. Sci.* **2013**, *19*, 16–24.
357. Maruyama, K.; Nagasawa, H.; Suzuki, A. 2,2′-Bispyridyl disulfide rapidly induces intramolecular disulfide bonds in peptides. *Peptides* **1999**, *20*, 881–884.
358. Güntert, P. Automated NMR structure calculation with CYANA. *Methods Mol. Biol.* **2004**, *278*, 353–378.

359. Jacobsen, R.B.; DelaCruz, R.G.; Grose, J.H.; McIntosh, J.M.; Yoshikami, D.; Olivera, B.M. Critical Residues Influence the Affinity and Selectivity of α-Conotoxin MI for Nicotinic Acetylcholine Receptors. *Biochemistry* **1999**, *38*, 13310–13315.
360. Groebe, D.R.; Gray, W.R.; Abramson, S.N. Determinants Involved in the Affinity of α-Conotoxins GI and SI for the Muscle Subtype of Nicotinic Acetylcholine Receptors †. *Biochemistry* **1997**, *36*, 6469–6474.
361. Cartier, G.E.; Yoshikami, D.; Gray, W.R.; Luo, S.; Olivera, B.M.; McIntosh, J.M. A New -Conotoxin Which Targets 32 Nicotinic Acetylcholine Receptors. *J. Biol. Chem.* **1996**, *271*, 7522–7528.
362. Price-Carter, M.; Gray, W.R.; Goldenberg, D.P. Folding of ω-Conotoxins. 1. Efficient Disulfide-Coupled Folding of Mature Sequences in Vitro. *Biochemistry* **1996**, *35*, 15537–15546.
363. Maslennikov, I.V.; Sobol, A.G.; Gladky, K.V.; Lugovskoy, A.A.; Ostrovsky, A.G.; Tsetlin, V.I.; Ivanov, V.T.; Arseniev, A.S. Two distinct structures of α-conotoxin GI in aqueous solution. *European Journal of Biochemistry* **1998**, *254*, 238–247.
364. Gouda, H.; Yamazaki, K.; Hasegawa, J.; Kobayashi, Y.; Nishiuchi, Y.; Sakakibara, S.; Hirono, S. Solution structure of alpha-conotoxin MI determined by 1H-NMR spectroscopy and molecular dynamics simulation with the explicit solvent water. *Biochim. Biophys. Acta* **1997**, *1343*, 327–334.
365. Terlau, H.; Shon, K.J.; Grilley, M.; Stocker, M.; Stühmer, W.; Olivera, B.M. Strategy for rapid immobilization of prey by a fish-hunting marine snail. *Nature* **1996**, *381*, 148–151.
366. Jakubowski, J.A. Intraspecific variation of venom injected by fish-hunting Conus snails. *Journal of Experimental Biology* **2005**, *208*, 2873–2883.
367. Dutertre, S.; Biass, D.; Stöcklin, R.; Favreau, P. Dramatic intraspecimen variations within the injected venom of Conus consors: An unsuspected contribution to venom diversity. *Toxicon* **2010**, *55*, 1453–1462.
368. Bren, N.; Sine, S.M. Hydrophobic Pairwise Interactions Stabilize α-Conotoxin MI in the Muscle Acetylcholine Receptor Binding Site. *J. Biol. Chem.* **2000**, *275*, 12692–12700.
369. Pucci, L.; Grazioso, G.; Dallanoce, C.; Rizzi, L.; De Micheli, C.; Clementi, F.; Bertrand, S.; Bertrand, D.; Longhi, R.; De Amici, M.; et al. Engineering of α-conotoxin MII-derived peptides with increased selectivity for native α6β2* nicotinic acetylcholine receptors. *FASEB J.* **2011**, *25*, 3775–3789.
370. Chun, J.B.S.; Baker, M.R.; Kim, D.H.; LeRoy, M.; Toribo, P.; Bingham, J.-P. Cone snail milked venom dynamics – A quantitative study of Conus purpurascens. *Toxicon* **2012**, *60*, 83–94.
371. Kohn, A.J. PISCIVOROUS GASTROPODS OF THE GENUS CONUS*. *Proc Natl Acad Sci U S A* **1956**, *42*, 168–171.
372. Kelley, W.P.; Schulz, J.R.; Jakubowski, J.A.; Gilly, W.F.; Sweedler, J.V. Two Toxins from Conus striatus that Individually Induce Tetanic Paralysis. *Biochemistry* **2006**, *45*, 14212–14222.
373. Cragg, G.M.; Newman, D.J. Natural products: A continuing source of novel drug leads. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2013**, *1830*, 3670–3695.
374. Puillandre, N.; Duda, T.F.; Meyer, C.; Olivera, B.M.; Bouchet, P. One, four or 100 genera? A new classification of the cone snails. *J Molluscan Stud* **2015**, *81*, 1–23.
375. Davis, J.; Jones, A.; Lewis, R.J. Remarkable inter- and intra-species complexity of conotoxins revealed by LC/MS. *Peptides* **2009**, *30*, 1222–1227.
376. Albuquerque, E.X.; Pereira, E.F.R.; Alkondon, M.; Rogers, S.W. Mammalian Nicotinic Acetylcholine Receptors: From Structure to Function. *Physiological Reviews* **2009**, *89*, 73–120.
377. Ladner, R.C.; Sato, A.K.; Gorzelany, J.; de Souza, M. Phage display-derived peptides as therapeutic alternatives to antibodies. *Drug Discovery Today* **2004**, *9*, 525–529.
378. Byk, G.; Halle, D.; Zeltser, I.; Bitan, G.; Selinger, Z.; Gilon, C. Synthesis and Biological Activity of NK-1 Selective, N-Backbone Cyclic Analogs of the C-Terminal Hexapeptide of Substance P. *J. Med. Chem.* **1996**, *39*, 3174–3178.
379. Hess, S.; Linde, Y.; Ovadia, O.; Safrai, E.; Shalev, D.E.; Swed, A.; Halbfinger, E.; Lapidot, T.; Winkler, I.; Gabinet, Y.; et al. Backbone Cyclic Peptidomimetic Melanocortin-4 Receptor

Agonist as a Novel Orally Administrated Drug Lead for Treating Obesity. *J. Med. Chem.* **2008**, *51*, 1026–1034.
380. Linde, Y.; Ovadia, O.; Safrai, E.; Xiang, Z.; Portillo, F.P.; Shalev, D.E.; Haskell-Luevano, C.; Hoffman, A.; Gilon, C. Structure-activity relationship and metabolic stability studies of backbone cyclization and N-methylation of melanocortin peptides. *Biopolymers* **2008**, *90*, 671–682.
381. Ovadia, O.; Linde, Y.; Haskell-Luevano, C.; Dirain, M.L.; Sheynis, T.; Jelinek, R.; Gilon, C.; Hoffman, A. The effect of backbone cyclization on PK/PD properties of bioactive peptide-peptoid hybrids: The melanocortin agonist paradigm. *Bioorganic & Medicinal Chemistry* **2010**, *18*, 580–589.
382. Armishaw, C.J.; Jensen, A.A.; Balle, L.D.; Scott, K.C.M.; Sørensen, L.; Strømgaard, K. Improving the Stability of α-Conotoxin AuIB Through N-to-C Cyclization: The Effect of Linker Length on Stability and Activity at Nicotinic Acetylcholine Receptors. *Antioxidants & Redox Signaling* **2011**, *14*, 65–76.
383. Lovelace, E.S.; Armishaw, C.J.; Colgrave, M.L.; Wahlstrom, M.E.; Alewood, P.F.; Daly, N.L.; Craik, D.J. Cyclic MrIA: A Stable and Potent Cyclic Conotoxin with a Novel Topological Fold that Targets the Norepinephrine Transporter. *J. Med. Chem.* **2006**, *49*, 6561–6568.
384. Koradi, R.; Billeter, M.; Wüthrich, K. MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* **1996**, *14*, 51–55, 29–32.
385. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera?A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612.
386. Barlos, K.; Chatzi, O.; Gatos, D.; Stavropoulos, G. 2-Chlorotrityl chloride resin. Studies on anchoring of Fmoc-amino acids and peptide cleavage. *Int. J. Pept. Protein Res.* **1991**, *37*, 513–520.
387. Cheneval, O.; Schroeder, C.I.; Durek, T.; Walsh, P.; Huang, Y.-H.; Liras, S.; Price, D.A.; Craik, D.J. Fmoc-Based Synthesis of Disulfide-Rich Cyclic Peptides. *J. Org. Chem.* **2014**, *79*, 5538–5544.
388. Wüthrich, K. *NMR of proteins and nucleic acids*; The George Fisher Baker non-resident lectureship in chemistry at Cornell University; Wiley: New York, 1986; ISBN 978-0-471-82893-8.
389. Shen, Y.; Delaglio, F.; Cornilescu, G.; Bax, A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* **2009**, *44*, 213–223.
390. Jimenez, E.C.; Olivera, B.M.; Teichert, R.W. αC-Conotoxin PrXA: A New Family of Nicotinic Acetylcholine Receptor Antagonists †. *Biochemistry* **2007**, *46*, 8717–8724.
391. Molgo, J. Effects of Aminopyridines on Neuromuscular Transmission. In *Aminopyridines and Similarly Acting Drugs: Effects on Nerves, Muscles and Synapses*; Elsevier, 1982; pp. 95–116 ISBN 978-0-08-028000-4.
392. Molgó, J.; Lemeignan, M.; Guerrero, S. Facilitatory effects of 4-aminopyridine on strontium-mediated evoked and delayed transmitter release from motor nerve terminals. *European Journal of Pharmacology* **1982**, *84*, 1–7.
393. Sanders, D.B.; Juel, V.C.; Harati, Y.; Smith, A.G.; Peltier, A.C.; Marburger, T.; Lou, J.; Pascuzzi, R.M.; Richman, D.P.; Xie, T.; et al. 3,4-diaminopyridine base effectively treats the weakness of Lambert-Eaton myasthenia. *Muscle Nerve* **2018**, *57*, 561–568.
394. Jonsson, M.; Gurley, D.; Dabrowski, M.; Larsson, O.; Johnson, E.C.; Eriksson, L.I. Distinct Pharmacologic Properties of Neuromuscular Blocking Agents on Human Neuronal Nicotinic Acetylcholine Receptors: A Possible Explanation for the Train-of-four Fade. *Anesthesiology* **2006**, *105*, 521–533.
395. Fagerlund, M.J.; Eriksson, L.I. Current concepts in neuromuscular transmission. *British Journal of Anaesthesia* **2009**, *103*, 108–114.
396. Ning, J.; Li, R.; Ren, J.; Zhangsun, D.; Zhu, X.; Wu, Y.; Luo, S. Alanine-Scanning Mutagenesis of α-Conotoxin GI Reveals the Residues Crucial for Activity at the Muscle Acetylcholine Receptor. *Marine Drugs* **2018**, *16*, 507.

397. Tajima, T.; Amaya, J.; Katayama, K.; Koizumi, T. Difference of train-of-four fade induced by nondepolarizing neuromuscular blocking drugs: a theoretical consideration on the underlying mechanisms. *J Anesth* **1995**, *9*, 333–337.
398. Nagashima, M.; Yasuhara, S.; Martyn, J.A.J. Train-of-Four and Tetanic Fade Are Not Always a Prejunctional Phenomenon as Evaluated by Toxins Having Highly Specific Pre- and Postjunctional Actions: *Anesthesia & Analgesia* **2013**, *116*, 994–1000.
399. Daly, N.L.; Callaghan, B.; Clark, R.J.; Nevin, S.T.; Adams, D.J.; Craik, D.J. Structure and Activity of α-Conotoxin PeIA at Nicotinic Acetylcholine Receptor Subtypes and GABA $_B$ Receptor-coupled N-type Calcium Channels. *J. Biol. Chem.* **2011**, *286*, 10233–10237.
400. You; Li; Xiong; Zhu; Zhangsun; Zhu; Luo α-Conotoxin TxIB: A Uniquely Selective Ligand for α6/α3β2β3 Nicotinic Acetylcholine Receptor Attenuates Nicotine-Induced Conditioned Place Preference in Mice. *Marine Drugs* **2019**, *17*, 490.
401. Kaas, Q.; Yu, R.; Jin, A.-H.; Dutertre, S.; Craik, D.J. ConoServer: updated content, knowledge, and discovery tools in the conopeptide database. *Nucleic Acids Research* **2012**, *40*, D325–D330.
402. Rigby, A.C.; Lucas-Meunier, E.; Kalume, D.E.; Czerwiec, E.; Hambe, B.; Dahlqvist, I.; Fossier, P.; Baux, G.; Roepstorff, P.; Baleja, J.D.; et al. A conotoxin from Conus textile with unusual posttranslational modifications reduces presynaptic Ca2+ influx. *Proceedings of the National Academy of Sciences* **1999**, *96*, 5758–5763.
403. Elliger, C.A.; Richmond, T.A.; Lebaric, Z.N.; Pierce, N.T.; Sweedler, J.V.; Gilly, W.F. Diversity of conotoxin types from Conus californicus reflects a diversity of prey types and a novel evolutionary history. *Toxicon* **2011**, *57*, 311–322.
404. Bush, K.A.; Stenflo, J.; Roth, D.A.; Czerwiec, E.; Harrist, A.; Begley, G.S.; Furie, B.C.; Furie, B. Hydrophobic Amino Acids Define the Carboxylation Recognition Site in the Precursor of the γ-Carboxyglutamic-Acid-Containing Conotoxin ε-TxIX from the Marine Cone Snail *Conus textile* †. *Biochemistry* **1999**, *38*, 14660–14666.
405. Liu, J.; Wu, Q.; Pi, C.; Zhao, Y.; Zhou, M.; Wang, L.; Chen, S.; Xu, A. Isolation and characterization of a T-superfamily conotoxin from Conus litteratus with targeting tetrodotoxin-sensitive sodium channels. *Peptides* **2007**, *28*, 2313–2319.
406. Luo, S.; Zhangsun, D.; Wu, Y.; Zhu, X.; Xie, L.; Hu, Y.; Zhang, J.; Zhao, X. Identification and Molecular Diversity of T-superfamily Conotoxins from Conus lividus and Conus litteratus. *Chem Biol Drug Design* **2006**, *68*, 97–106.
407. McIntosh, J.M.; Dowell, C.; Watkins, M.; Garrett, J.E.; Yoshikami, D.; Olivera, B.M. α-Conotoxin GIC from Conus geographus, a Novel Peptide Antagonist of Nicotinic Acetylcholine Receptors. *J. Biol. Chem.* **2002**, *277*, 33610–33615.
408. El Hamdaoui, Y.; Wu, X.; Clark, R.J.; Giribaldi, J.; Anangi, R.; Craik, D.J.; King, G.F.; Dutertre, S.; Kaas, Q.; Herzig, V.; et al. Periplasmic Expression of 4/7 α-Conotoxin TxIA Analogs in E. coli Favors Ribbon Isomer Formation – Suggestion of a Binding Mode at the α7 nAChR. *Front. Pharmacol.* **2019**, *10*, 577.
409. Dineley, K.T.; Pandya, A.A.; Yakel, J.L. Nicotinic ACh receptors as therapeutic targets in CNS disorders. *Trends in Pharmacological Sciences* **2015**, *36*, 96–108.
410. Lombardo, S.; Maskos, U. Role of the nicotinic acetylcholine receptor in Alzheimer's disease pathology and treatment. *Neuropharmacology* **2015**, *96*, 255–262.
411. Mohamed, T.S.; Jayakar, S.S.; Hamouda, A.K. Orthosteric and Allosteric Ligands of Nicotinic Acetylcholine Receptors for Smoking Cessation. *Front. Mol. Neurosci.* **2015**, *8*.
412. Hone, A.J.; Talley, T.T.; Bobango, J.; Huidobro Melo, C.; Hararah, F.; Gajewiak, J.B.; Christensen, S.B.; Harvey, P.J.; Craik, D.J.; McIntosh, J.M. Molecular determinants of α-conotoxin potency for inhibition of human and rat α6β4 nicotinic acetylcholine receptors. *Journal of Biological Chemistry* **2018**, jbc.RA118.005649.
413. Kaas, Q.; Westermann, J.-C.; Craik, D.J. Conopeptide characterization and classifications: An analysis using ConoServer. *Toxicon* **2010**, *55*, 1491–1509.

414. Antoniewicz, M.R. Methods and advances in metabolic flux analysis: a mini-review. *J Ind Microbiol Biotechnol* **2015**, *42*, 317–325.
415. Klint, J.K.; Senff, S.; Saez, N.J.; Seshadri, R.; Lau, H.Y.; Bende, N.S.; Undheim, E.A.B.; Rash, L.D.; Mobli, M.; King, G.F. Production of Recombinant Disulfide-Rich Venom Peptides for Structural and Functional Analysis via Expression in the Periplasm of E. coli. *PLoS ONE* **2013**, *8*, e63865.
416. Anangi, R.; Rash, L.D.; Mobli, M.; King, G.F. Functional Expression in Escherichia coli of the Disulfide-Rich Sea Anemone Peptide APETx2, a Potent Blocker of Acid-Sensing Ion Channel 3. *Marine Drugs* **2012**, *10*, 1605–1618.
417. Gloor, S.; Pongs, O.; Schmalzing, G. A vector for the synthesis of cRNAs encoding Myc epitope-tagged proteins in xenopus laevis oocytes. *Gene* **1995**, *160*, 213–217.
418. Hwang, T.L.; Shaka, A.J. Water Suppression That Works. Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients. *Journal of Magnetic Resonance, Series A* **1995**, *112*, 275–279.
419. Braunschweiler, L.; Ernst, R.R. Coherence transfer by isotropic mixing: Application to proton correlation spectroscopy. *Journal of Magnetic Resonance (1969)* **1983**, *53*, 521–528.
420. Jeener, J.; Meier, B.H.; Bachmann, P.; Ernst, R.R. Investigation of exchange processes by two-dimensional NMR spectroscopy. *The Journal of Chemical Physics* **1979**, *71*, 4546–4553.
421. Palmer, A.G.; Cavanagh, J.; Wright, P.E.; Rance, M. Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy. *Journal of Magnetic Resonance (1969)* **1991**, *93*, 151–170.
422. Vranken, W.F.; Boucher, W.; Stevens, T.J.; Fogh, R.H.; Pajon, A.; Llinas, M.; Ulrich, E.L.; Markley, J.L.; Ionides, J.; Laue, E.D. The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins* **2005**, *59*, 687–696.
423. Wishart, D.S.; Bigam, C.G.; Holm, A.; Hodges, R.S.; Sykes, B.D. 1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J Biomol NMR* **1995**, *5*, 67–81.
424. Shen, Y.; Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol NMR* **2013**, *56*, 227–241.
425. Nederveen, A.J.; Doreleijers, J.F.; Vranken, W.; Miller, Z.; Spronk, C.A.E.M.; Nabuurs, S.B.; Güntert, P.; Livny, M.; Markley, J.L.; Nilges, M.; et al. RECOORD: A recalculated coordinate database of 500+ proteins from the PDB using restraints from the BioMagResBank. *Proteins* **2005**, *59*, 662–672.
426. Brünger, A.T.; Adams, P.D.; Clore, G.M.; DeLano, W.L.; Gros, P.; Grosse-Kunstleve, R.W.; Jiang, J.S.; Kuszewski, J.; Nilges, M.; Pannu, N.S.; et al. Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination. *Acta Crystallogr D Biol Crystallogr* **1998**, *54*, 905–921.
427. Dominguez, C.; Boelens, R.; Bonvin, A.M.J.J. HADDOCK: A Protein−Protein Docking Approach Based on Biochemical or Biophysical Information. *J. Am. Chem. Soc.* **2003**, *125*, 1731–1737.
428. Linge, J.P.; Nilges, M. [No title found]. *Journal of Biomolecular NMR* **1999**, *13*, 51–59.
429. Williams, C.J.; Headd, J.J.; Moriarty, N.W.; Prisant, M.G.; Videau, L.L.; Deis, L.N.; Verma, V.; Keedy, D.A.; Hintze, B.J.; Chen, V.B.; et al. MolProbity: More and better reference data for improved all-atom structure validation: PROTEIN SCIENCE.ORG. *Protein Science* **2018**, *27*, 293–315.
430. Leffler, A.E.; Kuryatov, A.; Zebroski, H.A.; Powell, S.R.; Filipenko, P.; Hussein, A.K.; Gorson, J.; Heizmann, A.; Lyskov, S.; Tsien, R.W.; et al. Discovery of peptide ligands through docking and virtual screening at nicotinic acetylcholine receptor homology models. *Proc Natl Acad Sci USA* **2017**, *114*, E8100–E8109.
431. Šali, A.; Blundell, T.L. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *Journal of Molecular Biology* **1993**, *234*, 779–815.
432. Walsh, R.M.; Roh, S.-H.; Gharpure, A.; Morales-Perez, C.L.; Teng, J.; Hibbs, R.E. Structural principles of distinct assemblies of the human α4β2 nicotinic receptor. *Nature* **2018**, *557*, 261–265.

433. Shen, M.; Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein Sci.* **2006**, *15*, 2507–2524.
434. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J.L.; Dror, R.O.; Shaw, D.E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **2010**, NA-NA.
435. Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics* **2007**, *126*, 014101.
436. Miller, B.R.; McGee, T.D.; Swails, J.M.; Homeyer, N.; Gohlke, H.; Roitberg, A.E. *MMPBSA.py* : An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314–3321.
437. Lyskov, S.; Chou, F.-C.; Conchúir, S.Ó.; Der, B.S.; Drew, K.; Kuroda, D.; Xu, J.; Weitzner, B.D.; Renfrew, P.D.; Sripakdeevong, P.; et al. Serverification of Molecular Modeling Applications: The Rosetta Online Server That Includes Everyone (ROSIE). *PLoS ONE* **2013**, *8*, e63906.
438. Tyka, M.D.; Keedy, D.A.; André, I.; DiMaio, F.; Song, Y.; Richardson, D.C.; Richardson, J.S.; Baker, D. Alternate States of Proteins Revealed by Detailed Energy Landscape Mapping. *Journal of Molecular Biology* **2011**, *405*, 607–618.
439. Raveh, B.; London, N.; Schueler-Furman, O. Sub-angstrom modeling of complexes between flexible peptides and globular proteins. *Proteins* **2010**, NA-NA.
440. Leaver-Fay, A.; O'Meara, M.J.; Tyka, M.; Jacak, R.; Song, Y.; Kellogg, E.H.; Thompson, J.; Davis, I.W.; Pache, R.A.; Lyskov, S.; et al. Scientific Benchmarks for Guiding Macromolecular Energy Function Improvement. In *Methods in Enzymology*; Elsevier, 2013; Vol. 523, pp. 109–143 ISBN 978-0-12-394292-0.
441. Azam, L.; McIntosh, J.M. Alpha-conotoxins as pharmacological probes of nicotinic acetylcholine receptors. *Acta Pharmacol Sin* **2009**, *30*, 771–783.
442. Kapust, R.B.; Tözsér, J.; Copeland, T.D.; Waugh, D.S. The P1′ specificity of tobacco etch virus protease. *Biochemical and Biophysical Research Communications* **2002**, *294*, 949–955.
443. Loughnan, M.L.; Nicke, A.; Jones, A.; Adams, D.J.; Alewood, P.F.; Lewis, R.J. Chemical and Functional Identification and Characterization of Novel Sulfated α-Conotoxins from the Cone Snail *Conus a nemone*. *Journal of Medicinal Chemistry* **2004**, *47*, 1234–1241.
444. Carstens, B.B.; Swedberg, J.; Berecki, G.; Adams, D.J.; Craik, D.J.; Clark, R.J. Effects of linker sequence modifications on the structure, stability, and biological activity of a cyclic α-conotoxin. *Biopolymers* **2016**, *106*, 864–875.
445. Schubert, M.; Labudde, D.; Oschkinat, H.; Schmieder, P. [No title found]. *Journal of Biomolecular NMR* **2002**, *24*, 149–154.
446. Wishart, D.S.; Sykes, B.D.; Richards, F.M. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *Journal of Molecular Biology* **1991**, *222*, 311–333.
447. Wu, Y.; Wu, X.; Yu, J.; Zhu, X.; Zhangsun, D.; Luo, S. Influence of Disulfide Connectivity on Structure and Bioactivity of α-Conotoxin TxIA. *Molecules* **2014**, *19*, 966–979.
448. Kasheverov, I.E.; Zhmak, M.N.; Fish, A.; Rucktooa, P.; Khruschov, A.Yu.; Osipov, A.V.; Ziganshin, R.H.; D'hoedt, D.; Bertrand, D.; Sixma, T.K.; et al. Interaction of α-conotoxin ImII and its analogs with nicotinic receptors and acetylcholine-binding proteins: additional binding sites on *Torpedo* receptor. *Journal of Neurochemistry* **2009**, *111*, 934–944.
449. Wu, X.; Tae, H.-S.; Huang, Y.-H.; Adams, D.J.; Craik, D.J.; Kaas, Q. Stoichiometry dependent inhibition of rat α3β4 nicotinic acetylcholine receptor by the ribbon isomer of α-conotoxin AuIB. *Biochem. Pharmacol.* **2018**.
450. Ahmed, M.C.; Papaleo, E.; Lindorff-Larsen, K. How well do force fields capture the strength of salt bridges in proteins? *PeerJ* **2018**, *6*, e4967.
451. Singer, H.M.; Erhardt, M.; Steiner, A.M.; Zhang, M.-M.; Yoshikami, D.; Bulaj, G.; Olivera, B.M.; Hughes, K.T. Selective Purification of Recombinant Neuroactive Peptides Using the Flagellar Type III Secretion System. *mBio* **2012**, *3*, e00115-12.

452. Yu, J.; Zhu, X.; Yang, Y.; Luo, S.; Zhangsun, D. Expression in *Escherichia coli* of fusion protein comprising α-conotoxin TxIB and preservation of selectivity to nicotinic acetylcholine receptors in the purified product. *Chemical Biology & Drug Design* **2018**, *91*, 349–358.
453. Kang, T.S.; Vivekanandan, S.; Jois, S.D.S.; Kini, R.M. Effect of C-Terminal Amidation on Folding and Disulfide-Pairing of α-Conotoxin ImI. *Angew. Chem. Int. Ed.* **2005**, *44*, 6333–6337.
454. Ul-Hasan, S.; Burgess, D.M.; Gajewiak, J.; Li, Q.; Hu, H.; Yandell, M.; Olivera, B.M.; Bandyopadhyay, P.K. Characterization of the peptidylglycine α-amidating monooxygenase (PAM) from the venom ducts of neogastropods, Conus bullatus and Conus geographus. *Toxicon* **2013**, *74*, 215–224.
455. Zieliński, M.; Wójtowicz-Krawiec, A.; Mikiewicz, D.; Kęsik-Brodacka, M.; Cecuda-Adamczewska, V.; Marciniak-Rusek, A.; Sokołowska, I.; Łukasiewicz, N.; Gurba, L.; Odrowąż-Sypniewski, M.; et al. Expression of recombinant human bifunctional peptidylglycine α-amidating monooxygenase in CHO cells and its use for insulin analogue modification. *Protein Expression and Purification* **2016**, *119*, 102–109.
456. Jin, A.-H.; Brandstaetter, H.; Nevin, S.T.; Tan, C.C.; Clark, R.J.; Adams, D.J.; Alewood, P.F.; Craik, D.J.; Daly, N.L. Structure of alpha-conotoxin BuIA: influences of disulfide connectivity on structural dynamics. *BMC Struct. Biol.* **2007**, *7*, 28.
457. Carstens, B.B.; Berecki, G.; Daniel, J.T.; Lee, H.S.; Jackson, K.A.V.; Tae, H.-S.; Sadeghi, M.; Castro, J.; O'Donnell, T.; Deiteren, A.; et al. Structure-Activity Studies of Cysteine-Rich α-Conotoxins that Inhibit High-Voltage-Activated Calcium Channels via $GABA_B$ Receptor Activation Reveal a Minimal Functional Motif. *Angew. Chem. Int. Ed.* **2016**, *55*, 4692–4696.
458. McIntosh, J.M.; Corpuz, G.O.; Layer, R.T.; Garrett, J.E.; Wagstaff, J.D.; Bulaj, G.; Vyazovkina, A.; Yoshikami, D.; Cruz, L.J.; Olivera, B.M. Isolation and Characterization of a Novel *Conus* Peptide with Apparent Antinociceptive Activity. *J. Biol. Chem.* **2000**, *275*, 32391–32397.
459. Safavi-Hemami, H.; Gorasia, D.G.; Steiner, A.M.; Williamson, N.A.; Karas, J.A.; Gajewiak, J.; Olivera, B.M.; Bulaj, G.; Purcell, A.W. Modulation of Conotoxin Structure and Function Is Achieved through a Multienzyme Complex in the Venom Glands of Cone Snails. *J. Biol. Chem.* **2012**, *287*, 34288–34303.
460. Yu, R.; Craik, D.J.; Kaas, Q. Blockade of Neuronal α7-nAChR by α-Conotoxin ImI Explained by Computational Scanning and Energy Calculations. *PLoS Comput Biol* **2011**, *7*, e1002011.
461. Yu, R.; Kaas, Q.; Craik, D.J. Delineation of the Unbinding Pathway of α-Conotoxin ImI from the α7 Nicotinic Acetylcholine Receptor. *J. Phys. Chem. B* **2012**, *116*, 6097–6105.
462. Livett, B.G.; Gayler, K.R.; Khalil, Z. Drugs from the Sea: Conopeptides as Potential Therapeutics. *Current Medicinal Chemistry* **2004**, *11*, 1715–1723.
463. M'Kadmi, C.; Cabral, A.; Barrile, F.; Giribaldi, J.; Cantel, S.; Damian, M.; Mary, S.; Denoyelle, S.; Dutertre, S.; Péraldi-Roux, S.; et al. N-terminal Liver-expressed antimicrobial peptide 2 (LEAP2) region exhibits inverse agonist activity toward the ghrelin receptor. *Journal of Medicinal Chemistry* **2018**, *62*, 965–973.
464. Kojima, M.; Hosoda, H.; Date, Y.; Nakazato, M.; Matsuo, H.; Kangawa, K. Ghrelin is a growth-hormone-releasing acylated peptide from stomach. *Nature* **1999**, *402*, 656–660.
465. Müller, T.D.; Nogueiras, R.; Andermann, M.L.; Andrews, Z.B.; Anker, S.D.; Argente, J.; Batterham, R.L.; Benoit, S.C.; Bowers, C.Y.; Broglio, F.; et al. Ghrelin. *Molecular Metabolism* **2015**, *4*, 437–460.
466. Howard, A.D.; Feighner, S.D.; Cully, D.F.; Arena, J.P.; Liberator, P.A.; Rosenblum, C.I.; Hamelin, M.; Hreniuk, D.L.; Palyha, O.C.; Anderson, J.; et al. A Receptor in Pituitary and Hypothalamus That Functions in Growth Hormone Release. *Science* **1996**, *273*, 974–977.
467. Holst, B.; Cygankiewicz, A.; Jensen, T.H.; Ankersen, M.; Schwartz, T.W. High Constitutive Signaling of the Ghrelin Receptor—Identification of a Potent Inverse Agonist. *Molecular Endocrinology* **2003**, *17*, 2201–2210.
468. Damian, M.; Marie, J.; Leyris, J.-P.; Fehrentz, J.-A.; Verdié, P.; Martinez, J.; Banères, J.-L.; Mary, S. High Constitutive Activity Is an Intrinsic Feature of Ghrelin Receptor Protein: A STUDY WITH

A FUNCTIONAL MONOMERIC GHS-R1a RECEPTOR RECONSTITUTED IN LIPID DISCS. *J. Biol. Chem.* **2012**, *287*, 3630–3641.

469. Pantel, J. Loss of constitutive activity of the growth hormone secretagogue receptor in familial short stature. *Journal of Clinical Investigation* **2006**, *116*, 760–768.
470. Fernandez, G.; Cabral, A.; Andreoli, M.F.; Labarthe, A.; M'Kadmi, C.; Ramos, J.G.; Marie, J.; Fehrentz, J.-A.; Epelbaum, J.; Tolle, V.; et al. Evidence Supporting a Role for Constitutive Ghrelin Receptor Signaling in Fasting-Induced Hyperphagia in Male Mice. *Endocrinology* **2018**, *159*, 1021–1034.
471. Uchida, A.; Zigman, J.M.; Perelló, M. Ghrelin and eating behavior: evidence and insights from genetically-modified mouse models. *Front. Neurosci.* **2013**, *7*.
472. Krause, A.; Sillard, R.; Kleemeier, B.; Klüver, E.; Maronde, E.; Conejo-García, J.R.; Forssmann, W.G.; Schulz-Knappe, P.; Nehls, M.C.; Wattler, F.; et al. Isolation and biochemical characterization of LEAP-2, a novel blood peptide expressed in the liver. *Protein Sci* **2003**, *12*, 143–152.
473. Henriques, S.T.; Tan, C.C.; Craik, D.J.; Clark, R.J. Structural and Functional Analysis of Human Liver-Expressed Antimicrobial Peptide 2. *ChemBioChem* **2010**, *11*, 2148–2157.
474. Thiébaud, P.; Garbay, B.; Auguste, P.; Sénéchal, C.L.; Maciejewska, Z.; Fédou, S.; Gauthereau, X.; Costaglioli, P.; Thézé, N. Overexpression of Leap2 impairs Xenopus embryonic development and modulates FGF and activin signals. *Peptides* **2016**, *83*, 21–28.
475. Ge, X.; Yang, H.; Bednarek, M.A.; Galon-Tilleman, H.; Chen, P.; Chen, M.; Lichtman, J.S.; Wang, Y.; Dalmas, O.; Yin, Y.; et al. LEAP2 Is an Endogenous Antagonist of the Ghrelin Receptor. *Cell Metab.* **2017**.
476. Leyris, J.P.; Roux, T.; Trinquet, E.; Verdie, P.; Fehrentz, J.A.; Oueslati, N.; Douzon, S.; Bourrier, E.; Lamarque, L.; Gagne, D.; et al. Homogeneous time-resolved fluorescence-based assay to screen for ligands targeting the growth hormone secretagogue receptor type 1a. *Anal Biochem* **2011**, *408*, 253–262.
477. Damian, M.; Marie, J.; Leyris, J.-P.; Fehrentz, J.-A.; Verdie, P.; Martinez, J.; Baneres, J.-L.; Mary, S. High Constitutive Activity Is an Intrinsic Feature of Ghrelin Receptor Protein A STUDY WITH A FUNCTIONAL MONOMERIC GHS-R1a RECEPTOR RECONSTITUTED IN LIPID DISCS. *Journal of Biological Chemistry* **2012**, *287*, 3630–3641.
478. Damian, M.; Mary, S.; Maingot, M.; M'Kadmi, C.; Gagne, D.; Leyris, J.-P.; Denoyelle, S.; Gaibelet, G.; Gavara, L.; Costa, M.G. de S.; et al. Ghrelin receptor conformational dynamics regulate the transition from a preassembled to an active receptor:Gq complex. *Proceedings of the National Academy of Sciences of the United States of America* **2015**, *112*, 1601–1606.
479. Kozasa, T. Purification of G protein subunits from Sf9 insect cells using hexahistidine-tagged alpha and beta gamma subunits. *Methods in molecular biology (Clifton, N.J.)* **2004**, *237*, 21–38.
480. McEwen, D.P.; Gee, K.R.; Kang, H.C.; Neubig, R.R. Fluorescent BODIPY-GTP analogs: Real-time measurement of nucleotide binding to G proteins. *Analytical Biochemistry* **2001**, *291*, 109–117.
481. Granier, S.; Kim, S.; Shafer, A.M.; Ratnala, V.R.P.; Fung, J.J.; Zare, R.N.; Kobilka, B. Structure and Conformational Changes in the C-terminal Domain of the β_2-Adrenoceptor: INSIGHTS FROM FLUORESCENCE RESONANCE ENERGY TRANSFER STUDIES. *Journal of Biological Chemistry* **2007**, *282*, 13895–13905.
482. Leyris, J.-P.; Roux, T.; Trinquet, E.; Verdié, P.; Fehrentz, J.-A.; Oueslati, N.; Douzon, S.; Bourrier, E.; Lamarque, L.; Gagne, D.; et al. Homogeneous time-resolved fluorescence-based assay to screen for ligands targeting the growth hormone secretagogue receptor type 1a. *Analytical Biochemistry* **2011**, *408*, 253–262.
483. Els, S.; Schild, E.; Petersen, P.S.; Kilian, T.-M.; Mokrosinski, J.; Frimurer, T.M.; Chollet, C.; Schwartz, T.W.; Holst, B.; Beck-Sickinger, A.G. An Aromatic Region To Induce a Switch between Agonism and Inverse Agonism at the Ghrelin Receptor. *J. Med. Chem.* **2012**, *55*, 7437–7449.

484. M'Kadmi, C.; Leyris, J.-P.; Onfroy, L.; Galés, C.; Saulière, A.; Gagne, D.; Damian, M.; Mary, S.; Maingot, M.; Denoyelle, S.; et al. Agonism, Antagonism, and Inverse Agonism Bias at the Ghrelin Receptor Signaling. *J. Biol. Chem.* **2015**, *290*, 27021–27039.
485. Bdioui, S.; Verdi, J.; Pierre, N.; Trinquet, E.; Roux, T.; Kenakin, T. Equilibrium Assays Are Required to Accurately Characterize the Activity Profiles of Drugs Modulating Gq-Protein-Coupled Receptors. *Mol Pharmacol* **2018**, *94*, 992–1006.
486. Holst, B.; Holliday, N.D.; Bach, A.; Elling, C.E.; Cox, H.M.; Schwartz, T.W. Common Structural Basis for Constitutive Activity of the Ghrelin Receptor Family. *J. Biol. Chem.* **2004**, *279*, 53806–53817.
487. DiGruccio, M.R.; Mawla, A.M.; Donaldson, C.J.; Noguchi, G.M.; Vaughan, J.; Cowing-Zitron, C.; van der Meulen, T.; Huising, M.O. Comprehensive alpha, beta and delta cell transcriptomes reveal that ghrelin selectively activates delta cells and promotes somatostatin release from pancreatic islets. *Molecular Metabolism* **2016**, *5*, 449–458.
488. Mary, S.; Damian, M.; Louet, M.; Floquet, N.; Fehrentz, J.-A.; Marie, J.; Martinez, J.; Baneres, J.-L. Ligands and signaling proteins govern the conformational landscape explored by a G protein-coupled receptor. *Proceedings of the National Academy of Sciences* **2012**, *109*, 8304–8309.
489. Damian, M.; Mary, S.; Maingot, M.; M'Kadmi, C.; Gagne, D.; Leyris, J.-P.; Denoyelle, S.; Gaibelet, G.; Gavara, L.; Garcia de Souza Costa, M.; et al. Ghrelin receptor conformational dynamics regulate the transition from a preassembled to an active receptor:Gq complex. *Proc Natl Acad Sci USA* **2015**, *112*, 1601–1606.
490. Adriaenssens, A.E.; Svendsen, B.; Lam, B.Y.H.; Yeo, G.S.H.; Holst, J.J.; Reimann, F.; Gribble, F.M. Transcriptomic profiling of pancreatic alpha, beta and delta cell populations identifies delta cells as a principal target for ghrelin in mouse islets. *Diabetologia* **2016**, *59*, 2156–2165.
491. Yanagi, S.; Sato, T.; Kangawa, K.; Nakazato, M. The Homeostatic Force of Ghrelin. *Cell Metabolism* **2018**, *27*, 786–804.
492. Cabral, A.; Valdivia, S.; Fernandez, G.; Reynaldo, M.; Perello, M. Divergent Neuronal Circuitries Underlying Acute Orexigenic Effects of Peripheral or Central Ghrelin: Critical Role of Brain Accessibility. *J Neuroendocrinol* **2014**, *26*, 542–554.
493. Nijenhuis, W.A.J.; Oosterom, J.; Adan, R.A.H. AgRP(83–132) Acts as an Inverse Agonist on the Human-Melanocortin-4 Receptor. *Molecular Endocrinology* **2001**, *15*, 164–171.
494. Barrile, F.; M'Kadmi, C.; De Francesco, P.N.; Cabral, A.; García Romero, G.; Mustafá, E.R.; Cantel, S.; Damian, M.; Mary, S.; Denoyelle, S.; et al. Development of a novel fluorescent ligand of growth hormone secretagogue receptor based on the N-Terminal Leap2 region. *Molecular and Cellular Endocrinology* **2019**, *498*, 110573.
495. Cornejo, M.P.; Castrogiovanni, D.; Schiöth, H.B.; Reynaldo, M.; Marie, J.; Fehrentz, J.; Perello, M. Growth hormone secretagogue receptor signaling affects high fat intake independently of plasma levels of ghrelin and leap2, in a 4-day binge eating model. *J Neuroendocrinol* **2019**, jne.12785.
496. Brik, A.; Laps, S.; Sun, H.; Kamnesky, G. Palladium Mediated Direct Disulfide Bond Formation in Proteins Containing S-Acetamidomethyl-Cysteine under Aqueous Conditions. *Angewandte Chemie International Edition* **2019**.
497. Kent, S.B.H. Chemical Synthesis of Peptides and Proteins. *Annu. Rev. Biochem.* **1988**, *57*, 957–989.
498. Dawson, P.; Muir, T.; Clark-Lewis, I.; Kent, S. Synthesis of proteins by native chemical ligation. *Science* **1994**, *266*, 776–779.
499. Fang, G.-M.; Li, Y.-M.; Shen, F.; Huang, Y.-C.; Li, J.-B.; Lin, Y.; Cui, H.-K.; Liu, L. Protein Chemical Synthesis by Ligation of Peptide Hydrazides. *Angew. Chem. Int. Ed.* **2011**, *50*, 7645–7649.
500. Johnson, E.C.B.; Kent, S.B.H. Insights into the Mechanism and Catalysis of the Native Chemical Ligation Reaction. *J. Am. Chem. Soc.* **2006**, *128*, 6640–6646.

501. Flood, D.T.; Hintzen, J.C.J.; Bird, M.J.; Cistrone, P.A.; Chen, J.S.; Dawson, P.E. Leveraging the Knorr Pyrazole Synthesis for the Facile Generation of Thioester Surrogates for use in NCL. *Angewandte Chemie International Edition* **2018**.
502. Drukewitz, S.H.; Fuhrmann, N.; Undheim, E.A.B.; Blanke, A.; Giribaldi, J.; Mary, R.; Laconde, G.; Dutertre, S.; von Reumont, B.M. A Dipteran's Novel Sucker Punch: Evolution of Arthropod Atypical Venom with a Neurotoxic Component in Robber Flies (Asilidae, Diptera). *Toxins* **2018**, *10*, 29.
503. Fry, B.G.; Roelants, K.; Champagne, D.E.; Scheib, H.; Tyndall, J.D.A.; King, G.F.; Nevalainen, T.J.; Norman, J.A.; Lewis, R.J.; Norton, R.S.; et al. The Toxicogenomic Multiverse: Convergent Recruitment of Proteins Into Animal Venoms. *Annual Review of Genomics and Human Genetics* **2009**, *10*, 483–511.
504. Casewell, N.R.; Wüster, W.; Vonk, F.J.; Harrison, R.A.; Fry, B.G. Complex cocktails: the evolutionary novelty of venoms. *Trends in Ecology & Evolution* **2013**, *28*, 219–229.
505. von Reumont, B.M.; Campbell, L.I.; Richter, S.; Hering, L.; Sykes, D.; Hetmank, J.; Jenner, R.A.; Bleidorn, C. A Polychaete's Powerful Punch: Venom Gland Transcriptomics of Glycera Reveals a Complex Cocktail of Toxin Homologs. *Genome Biology and Evolution* **2014**, *6*, 2406–2423.
506. von Reumont, B.; Campbell, L.; Jenner, R. Quo Vadis Venomics? A Roadmap to Neglected Venomous Invertebrates. *Toxins* **2014**, *6*, 3488–3551.
507. Hargreaves, A.D.; Swain, M.T.; Hegarty, M.J.; Logan, D.W.; Mulley, J.F. Restriction and Recruitment--Gene Duplication and the Origin and Evolution of Snake Venom Toxins. *Genome Biology and Evolution* **2014**, *6*, 2088–2095.
508. Martinson, E.O.; Mrinalini; Kelkar, Y.D.; Chang, C.H.; Werren, J.H. The Evolution of Venom by Co-option of Single-Copy Genes. *Current Biology* **2017**.
509. Wong, E.S.W.; Papenfuss, A.T.; Whittington, C.M.; Warren, W.C.; Belov, K. A Limited Role for Gene Duplications in the Evolution of Platypus Venom. *Molecular Biology and Evolution* **2011**, *29*, 167–177.
510. Vonk, F.J.; Casewell, N.R.; Henkel, C. V; Heimberg, A.M.; Jansen, H.J.; McCleary, R.J.R.; Kerkkamp, H.M.E.; Vos, R.A.; Guerreiro, I.; Calvete, J.J.; et al. The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. *Proceedings of the National Academy of Sciences* **2013**, *110*, 20651–20656.
511. Undheim, E.A.B.; Mobli, M.; King, G.F. Toxin structures as evolutionary tools: Using conserved 3D folds to study the evolution of rapidly evolving peptides. *BioEssays* **2016**, *38*, 539–548.
512. Thomas, G.W.C.; Hahn, M.W.; Hahn, Y. The effects of increasing the number of taxa on inferences of molecular convergence. *Genome Biology and Evolution* **2017**, evw306.
513. Casewell, N.R.; Visser, J.C.; Baumann, K.; Dobson, J.; Han, H.; Kuruppu, S.; Morgan, M.; Romilio, A.; Weisbecker, V.; Mardon, K.; et al. The Evolution of Fangs, Venom, and Mimicry Systems in Blenny Fishes. *Current Biology* **2017**, *27*, 1184–1191.
514. Undheim, E.; Fry, B.; King, G. Centipede Venom: Recent Discoveries and Current State of Knowledge. *Toxins* **2015**, *7*, 679–704.
515. von Reumont, B.M.; Blanke, A.; Richter, S.; Alvarez, F.; Bleidorn, C.; Jenner, R.A. The First Venomous Crustacean Revealed by Transcriptomics and Functional Morphology: Remipede Venom Glands Express a Unique Toxin Cocktail Dominated by Enzymes and a Neurotoxin. *Molecular Biology and Evolution* **2014**, *31*, 48–58.
516. Walker, A.A.; Madio, B.; Jin, J.; Undheim, E.A.B.; Fry, B.G.; King, G.F. Melt With This Kiss: Paralyzing and Liquefying Venom of The Assassin Bug Pristhesancus plagipennis (Hemiptera: Reduviidae). *Molecular & Cellular Proteomics* **2017**, *16*, 552–566.
517. Yeates, D.K.; Wiegmann, B.M.; Courtney, G.W.; Meier, R.; Lambkin, C.; Pape, T. Phylogeny and systematics of Diptera: Two decades of progress and prospects. *Zootaxa* **2007**.
518. Wiegmann, B.M.; Trautwein, M.D.; Winkler, I.S.; Barr, N.B.; Kim, J.-W.; Lambkin, C.; Bertone, M.A.; Cassel, B.K.; Bayless, K.M.; Heimberg, A.M.; et al. Episodic radiations in the fly tree of life. *Proceedings of the National Academy of Sciences* **2011**.

519. Dikow, T. A phylogenetic hypothesis for Asilidae based on a total evidence analysis of morphological and DNA sequence data (Insecta: Diptera: Brachycera: Asiloidea). *Organisms Diversity and Evolution* **2009**.
520. Kahan, D. The toxic effect of the bite and the proteolytic activity of the saliva and stomach contents of the robber flies (Diptera Asilidae). *Israel Journal of Zoology* **1964**, *13*, 47–57.
521. Melin, D.E. *Contributions to the Knowledge of the Biology, Metamorphosis and Distribution of the Swedish Asilids in Relation to the Whole Family of Asilids*; Zoologiska bidrag från Uppsala; Almqvist & Wiksells, 1923;
522. Hobby, B.M. Rhodesian Asilidae (Diptera) and Their Prey Collected by Mr C. F. M. Swynnerton. *The Journal of Animal Ecology* **1935**, *4*, 90.
523. Dennis, D.S.; Lavigne, R.J. Hymenoptera as prey of robber flies (Diptera: Asilidae) with new prey records. *Journal of the Entomological Research Society* **2007**.
524. Dennis, D.S.; Lavigne, R.J.; Dennis, J.G. Hemiptera (heteroptera/homoptera) as prey of robber flies (diptera: Asilidae) with unpublished records. *Journal of the Entomological Research Society* **2010**.
525. Dennis, D.S.; Lavigne, R.J.; Dennis, J.G. Spiders (Araneae) as prey of robber flies (Diptera: Asilidae). *Journal of the Entomological Research Society* 2012.
526. Joern, A.; Rudd, N.T. Impact of predation by the robber fly Proctacanthus milbertii (Diptera: Asilidae) on grasshopper (Orthoptera: Acrididae) populations. *Oecologia* **1982**.
527. Rabinovich, M.; Corley, J.C. An important new predator of honey bees the robber fly Mallophora ruficauda Wiedemann (Diptera, Asilidae) in Argentina. *American Bee Journal* **1997**.
528. Owsley, W.B. The Comparative Morphology of Internal Structures of the Asilidae (Diptera). *Annals of the Entomological Society of America* **1946**, *39*, 33–68.
529. Wardill, T.J.; Fabian, S.T.; Pettigrew, A.C.; Stavenga, D.G.; Nordström, K.; Gonzalez-Bellido, P.T. A Novel Interception Strategy in a Miniature Robber Fly with Extreme Visual Acuity. *Current Biology* **2017**.
530. Whitfield, F.G.S. The Relation between the Feeding-habits and the Structure of the Month-parts in the Asilidae (Diptera). *Proceedings of the Zoological Society of London* **1925**, *95*, 599–638.
531. Musso, J.J.; Garnier, R.; Legier, F. Comparison of toxicity of venom of some asilids (Diptera Brachycera) on locusts. *Annales De La Societe Entomologique De France* **1978**, *14*, 177–184.
532. King, G.F. Venoms as a platform for human drugs: translating toxins into therapeutics. *Expert Opinion on Biological Therapy* **2011**, *11*, 1469–1484.
533. King, G.F.; Hardy, M.C. Spider-Venom Peptides: Structure, Pharmacology, and Potential for Control of Insect Pests. *Annual Review of Entomology* **2013**, *58*, 475–496.
534. Geller-Grimm, F. Photographic atlas and identification key to the robber flies of Germany (Diptera:Asilidae) Available online: http://www.robberflies.info/keyger/start.html (accessed on Oct 1, 2017).
535. A quality control tool for high throughput sequence data. Available online: http://www.bioinformatics.babraham.ac.uk/projects/ (accessed on Oct 1, 2017).
536. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120.
537. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* **2011**, *29*, 644–652.
538. TransDecoder (Find Coding Regions Within Transcripts) Available online: https://github.com/TransDecoder (accessed on Oct 1, 2017).
539. UniProt: a hub for protein information. *Nucleic Acids Research* **2014**, *43*, D204–D212.
540. HMMER: biosequence analysis using profile hidden Markov models Available online: http://hmmer.org/ (accessed on Oct 1, 2017).

541. Finn, R.D.; Coggill, P.; Eberhardt, R.Y.; Eddy, S.R.; Mistry, J.; Mitchell, A.L.; Potter, S.C.; Punta, M.; Qureshi, M.; Sangrador-Vegas, A.; et al. The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Research* **2016**.
542. Hoffmann, S.; Otto, C.; Kurtz, S.; Sharma, C.M.; Khaitovich, P.; Vogel, J.; Stadler, P.F.; Hackermüller, J. Fast mapping of short sequences with mismatches, insertions and deletions using index structures. *PLoS Computational Biology* **2009**.
543. Bray, N.L.; Pimentel, H.; Melsted, P.; Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology* **2016**.
544. Brekhman, V.; Malik, A.; Haas, B.; Sher, N.; Lotan, T. Transcriptome profiling of the dynamic life cycle of the scypohozoan jellyfish Aurelia aurita. *BMC Genomics* **2015**, *16*, 74.
545. Lewis Ames, C.; Ryan, J.F.; Bely, A.E.; Cartwright, P.; Collins, A.G. A new transcriptome and transcriptome profiling of adult and larval tissue in the box jellyfish Alatina alata: an emerging model for studying venom, vision and sex. *BMC Genomics* **2016**, *17*.
546. Jungo, F.; Bougueleret, L.; Xenarios, I.; Poux, S. The UniProtKB/Swiss-Prot Tox-Prot program: A central hub of integrated venom protein data. *Toxicon* **2012**.
547. Jones, P.; Binns, D.; Chang, H.-Y.; Fraser, M.; Li, W.; McAnulla, C.; McWilliam, H.; Maslen, J.; Mitchell, A.; Nuka, G.; et al. InterProScan 5: genome-scale protein function classification. *Bioinformatics* **2014**, *30*, 1236–1240.
548. Petersen, T.N.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* **2011**.
549. Li, W.; Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **2006**, *22*, 1658–1659.
550. Zheng, J.-S.; Tang, S.; Qi, Y.-K.; Wang, Z.-P.; Liu, L. Chemical synthesis of proteins using peptide hydrazides as thioester surrogates. *Nat. Protocols* **2013**.
551. Yushkevich, P.A.; Piven, J.; Hazlett, H.C.; Smith, R.G.; Ho, S.; Gee, J.C.; Gerig, G. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage* **2006**.
552. Blender. Available online: blender.org (accessed on Oct 1, 2017).
553. Otto, C.; Stadler, P.F.; Hoffmann, S. Lacking alignments? The next-generation sequencing mapper segemehl revisited. *Bioinformatics* **2014**.
554. Undheim, E.A.B.; Jones, A.; Clauser, K.R.; Holland, J.W.; Pineda, S.S.; King, G.F.; Fry, B.G. Clawing through Evolution: Toxin Diversification and Convergence in the Ancient Lineage Chilopoda (Centipedes). *Molecular Biology and Evolution* **2014**, *31*, 2124–2148.
555. Magalhães, G.S.; Lopes-Ferreira, M.; Junqueira-de-Azevedo, I.L.M.; Spencer, P.J.; Araújo, M.S.; Portaro, F.C. V; Ma, L.; Valente, R.H.; Juliano, L.; Fox, J.W.; et al. Natterins, a new class of proteins with kininogenase activity characterized from Thalassophryne nattereri fish venom. *Biochimie* **2005**, *87*, 687–699.
556. Tamura, S.; Yamakawa, M.; Shiomi, K. Purification, characterization and cDNA cloning of two natterin-like toxins from the skin secretion of oriental catfish Plotosus lineatus. *Toxicon* **2011**, *58*, 430–438.
557. Lopes-Ferreira, M.; Grund, L.Z.; Lima, C. Thalassophryne nattereri fish venom: from the envenoming to the understanding of the immune system. *The journal of venomous animals and toxins including tropical diseases* **2014**.
558. Pallaghy, P.K.; Norton, R.S.; Nielsen, K.J.; Craik, D.J. A common structural motif incorporating a cystine knot and a triple-stranded β-sheet in toxic and inhibitory polypeptides. *Protein Science* **1994**.
559. Wang, X. hong; Smith, R.; Fletcher, J.I.; Wilson, H.; Wood, C.J.; Howden, M.E.H.; King, G.F. Structure-function studies of ω-atracotoxin, a potent antagonist of insect voltage-gated calcium channels. *European Journal of Biochemistry* **1999**.
560. Fletcher, J.I.; Smith, R.; O'Donoghue, S.I.; Nilges, M.; Connor, M.; Howden, M.E.; Christie, M.J.; King, G.F. The structure of a novel insecticidal neurotoxin, omega-atracotoxin-HV1, from the venom of an Australian funnel web spider. *Nature structural biology* **1997**.

561. Madio, B.; Undheim, E.A.B.; King, G.F. Revisiting venom of the sea anemone Stichodactyla haddoni: Omics techniques reveal the complete toxin arsenal of a well-studied sea anemone genus. *Journal of Proteomics* **2017**.
562. von Reumont, B.; Undheim, E.; Jauss, R.-T.; Jenner, R. Venomics of Remipede Crustaceans Reveals Novel Peptide Diversity and Illuminates the Venom's Biological Role. *Toxins* **2017**, *9*, 234.
563. Musso, J.J. Observations sur le comortement alimentaires l'anatomie e l'histologie des galndes salivaires de deux asilides mediterraneens. *Annales De La Societe Entomologique De France* **1968**, *4*, 245–255.
564. Undheim, E.A.B.; Mobli, M.; King, G.F. Toxin structures as evolutionary tools: Using conserved 3D folds to study the evolution of rapidly evolving peptides. *BioEssays* **2016**, *38*, 539–548.
565. Cohen, A.C. Extra-Oral Digestion in Predaceous Terrestrial Arthropoda. *Annual Review of Entomology* **1995**, *40*, 85–103.
566. Cohen, A.C. Solid-to-Liquid Feeding: The Inside(s) Story of Extra-Oral Digestion in Predaceous Arthropoda. *American Entomologist* **1998**, *44*, 103–117.
567. D.M., C.; B.L., L.; R.C., C.; C.N., M. Dengue-2 Alters Salivary Gland Protein Expression In Infected Aedes Aegypti Mosquitoes. *American Journal of Tropical Medicine and Hygiene* 2010.
568. Almeras, L.; Fontaine, A.; Belghazi, M.; Bourdon, S.; Boucomont-Chapeaublanc, E.; Orlandi-Pradines, E.; Baragatti, M.; Corre-Catelin, N.; Reiter, P.; Pradines, B.; et al. Salivary Gland Protein Repertoire from Aedes aegypti Mosquitoes. *Vector-Borne and Zoonotic Diseases* **2010**.
569. Wongkamchai, S.; Khongtak, P.; Leemingsawat, S.; Komalamisra, N.; Junsong, N.; Kulthanan, K.; Wisuthsarewong, W.; Boitano, J.J. Comparative identification of protein profiles and major allergens of saliva, salivary gland and whole body extracts of mosquito species in Thailand. *Asian Pacific Journal of Allergy and Immunology* **2010**.
570. Ma, D.; Li, Y.; Dong, J.; An, S.; Wang, Y.; Liu, C.; Yang, X.; Yang, H.; Xu, X.; Lin, D.; et al. Purification and characterization of two new allergens from the salivary glands of the horsefly, Tabanusyao. *Allergy* **2010**.
571. Volfova, V.; Tothova, V.; Volf, P. Hyaluronidase activity in the salivary glands of tabanid flies. *Insect Biochemistry and Molecular Biology* **2016**.
572. Xu, X.; Yang, H.; Ma, D.; Wu, J.; Wang, Y.; Song, Y.; Wang, X.; Lu, Y.; Yang, J.; Lai, R. Toward an Understanding of the Molecular Mechanism for Successful Blood Feeding by Coupling Proteomics Analysis with Pharmacological Testing of Horsefly Salivary Glands. *Molecular & Cellular Proteomics* **2008**.
573. Takác, P.; Nunn, M. a; Mészáros, J.; Pechánová, O.; Vrbjar, N.; Vlasáková, P.; Kozánek, M.; Kazimírová, M.; Hart, G.; Nuttall, P. a; et al. Vasotab, a vasoactive peptide from horse fly Hybomitra bimaculata (Diptera, Tabanidae) salivary glands. *The Journal of experimental biology* **2006**.
574. Misof, B.; Liu, S.; Meusemann, K.; Peters, R.S.; Flouri, T.; Beutel, R.G.; Niehuis, O.; Petersen, M. Phylogenomics resolves the timing and pattern of insect evolution. *Science* **2014**.
575. Sunagar, K.; Moran, Y. The Rise and Fall of an Evolutionary Innovation: Contrasting Strategies of Venom Evolution in Ancient and Young Animals. *PLoS Genetics* **2015**.
576. Catterall, W.A. Voltage-Gated Calcium Channels. *Cold Spring Harbor Perspectives in Biology* **2011**, *3*, a003947–a003947.
577. Bourinet, E.; Alloui, A.; Monteil, A.; Barrère, C.; Couette, B.; Poirot, O.; Pages, A.; McRory, J.; Snutch, T.P.; Eschalier, A.; et al. Silencing of the Cav3.2 T-type calcium channel gene in sensory neurons demonstrates its major role in nociception. *EMBO J* **2005**, *24*, 315–324.
578. Todorovic, S.M.; Jevtovic-Todorovic, V. The role of T-type calcium channels in peripheral and central pain processing. *CNS Neurol Disord Drug Targets* **2006**, *5*, 639–653.
579. Flatters, S.J.L.; Bennett, G.J. Ethosuximide reverses paclitaxel- and vincristine-induced painful peripheral neuropathy: *Pain* **2004**, *109*, 150–161.
580. Dogrul, A.; Gardell, L.R.; Ossipov, M.H.; Tulunay, C.F.; Lai, J.; Porreca, F. Reversal of experimental neuropathic pain by T-type calcium channel blockers: *Pain* **2003**, *105*, 159–168.

581. Klint, J.K.; Senff, S.; Rupasinghe, D.B.; Er, S.Y.; Herzig, V.; Nicholson, G.M.; King, G.F. Spider-venom peptides that target voltage-gated sodium channels: Pharmacological tools and potential therapeutic leads. *Toxicon* **2012**, *60*, 478–491.
582. Dos Santos, R.G.; Van Renterghem, C.; Martin-Moutot, N.; Mansuelle, P.; Cordeiro, M.N.; Diniz, C.R.; Mori, Y.; De Lima, M.E.; Seagar, M. *Phoneutria nigriventer* ω-Phonetoxin IIA Blocks the Ca_v2 Family of Calcium Channels and Interacts with ω-Conotoxin-binding Sites. *J. Biol. Chem.* **2002**, *277*, 13856–13862.
583. Wang, M.; Guan, X.; Liang, S. The cross channel activities of spider neurotoxin huwentoxin-I on rat dorsal root ganglion neurons. *Biochemical and Biophysical Research Communications* **2007**, *357*, 579–583.
584. Liang, S.; Zhang, D.; Pan, X.; Chen, Q.; Zhou, P. Properties and amino acid sequence of huwentoxin-I, a neurotoxin purified from the venom of the Chinese bird spider Selenocosmia Huwena. *Toxicon* **1993**, *31*, 969–978.
585. Jiang, L.; Peng, L.; Chen, J.; Zhang, Y.; Xiong, X.; Liang, S. Molecular diversification based on analysis of expressed sequence tags from the venom glands of the Chinese bird spider Ornithoctonus huwena,. *Toxicon* **2008**, *51*, 1479–1489.
586. Liu, Z.; Dai, J.; Dai, L.; Deng, M.; Hu, Z.; Hu, W.; Liang, S. Function and Solution Structure of Huwentoxin-X, a Specific Blocker of N-type Calcium Channels, from the Chinese Bird Spider *Ornithoctonus huwena*. *J. Biol. Chem.* **2006**, *281*, 8628–8635.
587. Bindokas, V.P.; Adams, M.E. ?-Aga-I: A presynaptic calcium channel antagonist from venom of the funnel web spider,Agelenopsis aperta. *J. Neurobiol.* **1989**, *20*, 171–188.
588. Adams, E.; Bindokas, P.; Venema, J. o-Agatoxins: Novel Calcium Channel Antagonists of Two Subtypes from Funnel Web Spider (Agelenopsis uperta) Venom. 8.
589. Ertel, E.A.; Warren, V.A.; Adams, M.E.; Griffin, P.R.; Cohen, C.J.; Smith, M.M. Type III .omega.-Agatoxins: A Family of Probes for Similar Binding Sites on L- and N-Type Calcium Channels. *Biochemistry* **1994**, *33*, 5098–5108.
590. Yan, L.; Adams, M.E. The Spider Toxin ω-Aga IIIA Defines a High Affinity Site on Neuronal High Voltage-activated Calcium Channels. *J. Biol. Chem.* **2000**, *275*, 21309–21316.
591. Leão, R.M.; Cruz, J.S.; Diniz, C.R.; Cordeiro, M.N.; Beirão, P.S.L. Inhibition of neuronal high-voltage activated calcium channels by the ω-Phoneutria nigriventer Tx3-3 peptide toxin. *Neuropharmacology* **2000**, *39*, 1756–1767.
592. Cordeiro, M. do N.; de Figueiredo, S.G.; Valentim, A. do C.; Diniz, C.R.; von Eickstedt, V.R.D.; Gilroy, J.; Richardson, M. Purification and amino acid sequences of six Tx3 type neurotoxins from the venom of the Brazilian 'armed' spider Phoneutria Nigriventer (keys.). *Toxicon* **1993**, *31*, 35–42.
593. Vieira, L.B.; Kushmerick, C.; Hildebrand, M.E.; Garcia, E.; Stea, A.; Cordeiro, M.N.; Richardson, M.; Gomez, M.V.; Snutch, T.P. Inhibition of High Voltage-Activated Calcium Channels by Spider Toxin PnTx3-6. *J Pharmacol Exp Ther* **2005**, *314*, 1370–1377.
594. Vieira, L.B.; Pimenta, A.M.C.; Richardson, M.; Bemquerer, M.P.; Reis, H.J.; Cruz, J.S.; Gomez, M.V.; Santoro, M.M.; Ferreira-de-Oliveira, R.; Figueiredo, S.G.; et al. Leftward Shift in the Voltage-Dependence for Ca2+ Currents Activation Induced by a New Toxin from Phoneutria reidyi (Aranae, Ctenidae) Venom. *Cell Mol Neurobiol* **2007**, *27*, 129–146.
595. Richardson, M.; Pimenta, A.M.C.; Bemquerer, M.P.; Santoro, M.M.; Beirao, P.S.L.; Lima, M.E.; Figueiredo, S.G.; Bloch, C.; Vasconcelos, E.A.R.; Campos, F.A.P.; et al. Comparison of the partial proteomes of the venoms of Brazilian spiders of the genus Phoneutria. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* **2006**, *142*, 173–187.
596. Newcomb, R.; Palma, A.; Fox, J.; Gaur, S.; Lau, K.; Chung, D.; Cong, R.; Bell, J.R.; Horne, B. SNX-325, a novel calcium antagonist from the spider Segestria florentina. *Biochemistry* **1995**, *34*, 8341–8347.
597. Heinke, B.; Balzer, E.; Sandkuhler, J. Pre- and postsynaptic contributions of voltage-dependent Ca2+ channels to nociceptive transmission in rat spinal lamina I neurons. *Eur J Neurosci* **2004**, *19*, 103–111.

598. Rycroft, B.K.; Vikman, K.S.; Christie, M.J. Inflammation reduces the contribution of N-type calcium channels to primary afferent synaptic transmission onto NK1 receptor-positive lamina I neurons in the rat dorsal horn: N-type calcium channel function in the dorsal horn during inflammation. *The Journal of Physiology* **2007**, *580*, 883–894.
599. Mochizuki, M.; Tsuda, S.; Tanimura, K.; Nishiuchi, Y. Regioselective Formation of Multiple Disulfide Bonds with the Aid of Postsynthetic S-Tritylation. *Org. Lett.* **2015**, *17*, 2202–2205.
600. Oldrati, V.; Koua, D.; Allard, P.-M.; Hulo, N.; Arrell, M.; Nentwig, W.; Lisacek, F.; Wolfender, J.-L.; Kuhn-Nentwig, L.; Stöcklin, R. Peptidomic and transcriptomic profiling of four distinct spider venoms. *PLoS ONE* **2017**, *12*, e0172966.
601. Deuis, J.R.; Dekan, Z.; Wingerd, J.S.; Smith, J.J.; Munasinghe, N.R.; Bhola, R.F.; Imlach, W.L.; Herzig, V.; Armstrong, D.A.; Rosengren, K.J.; et al. Pharmacological characterisation of the highly NaV1.7 selective spider venom peptide Pn3a. *Sci Rep* **2017**, *7*.
602. Takeuchi, K.; Park, E.J.; Lee, C.W.; Kim, J.I.; Takahashi, H.; Swartz, K.J.; Shimada, I. Solution Structure of ω-Grammotoxin SIA, A Gating Modifier of P/Q and N-type Ca2+ Channel. *Journal of Molecular Biology* **2002**, *321*, 517–526.
603. Vetter, I.; Dekan, Z.; Knapp, O.; Adams, D.J.; Alewood, P.F.; Lewis, R.J. Isolation, characterization and total regioselective synthesis of the novel μO-conotoxin MfVIA from Conus magnificus that targets voltage-gated sodium channels. *Biochemical Pharmacology* **2012**, *84*, 540–548.
604. Eliasen, R.; Andresen, T.L.; Conde-Frieboes, K.W. Handling a tricycle: Orthogonal versus random oxidation of the tricyclic inhibitor cystine knotted peptide gurmarin. *Peptides* **2012**, *37*, 144–149.
605. Ji, J.A.; Zhang, B.; Cheng, W.; Wang, Y.J. Methionine, tryptophan, and histidine oxidation in a model protein, PTH: Mechanisms and stabilization. *Journal of Pharmaceutical Sciences* **2009**, *98*, 4485–4500.
606. Simat, T.J.; Steinhart, H. Oxidation of Free Tryptophan and Tryptophan Residues in Peptides and Proteins. *J. Agric. Food Chem.* **1998**, *46*, 490–498.
607. Lee, M.G.; Rogers, C.M. Degradation of tryptophan in aqueous solution. *J Parenter Sci Technol* **1988**, *42*, 20–22.
608. Harris, K.M.; Flemer, S.; Hondal, R.J. Studies on Deprotection of Cysteine and Selenocysteine Side-Chain Protecting Groups. *J Pept Sci* **2007**, *13*, 81–93.
609. Castañeda, O.; Sotolongo, V.; Amor, A.M.; Stöcklin, R.; Anderson, A.J.; Harvey, A.L.; Engström, Å.; Wernstedt, C.; Karlsson, E. Characterization of a potassium channel toxin from the Caribbean sea anemone Stichodactyla helianthus. *Toxicon* **1995**, *33*, 603–613.
610. Pennington, M.W.; Byrnes, M.E.; Zaydenberg, I.; Khaytin, I.; De Chastonay, J.; Krafte, D.S.; Hill, R.; Mahnir, V.M.; Volberg, W.. A.; Gorczyca, W.; et al. Chemical synthesis and characterization of ShK toxin: a potent potassium channel inhibitor from a sea anemone. *International Journal of Peptide and Protein Research* **1996**, *46*, 354–358.
611. Pohl, J.; Hubalek, F.; Byrnes, M.E.; Nielsen, K.R.; Woods, A.; Pennington, M.W. Assignment of the three disulfide bonds in ShK toxin: A potent potassium channel inhibitor from the sea anemone Stichodactyla helianthus. *Letters in Peptide Science* **1995**, *1*, 291–297.
612. Tudor, J.E.; Pallaghy, P.K.; Pennington, M.W.; Norton, R.S. Solution structure of ShK toxin, a novel potassium channel inhibitor from a sea anemone. *Nat. Struct. Biol.* **1996**, *3*, 317–320.
613. Beeton, C. Targeting Effector Memory T Cells with a Selective Peptide Inhibitor of Kv1.3 Channels for Therapy of Autoimmune Diseases. *Molecular Pharmacology* **2005**, *67*, 1369–1381.
614. Beeton, C.; Wulff, H.; Standifer, N.E.; Azam, P.; Mullen, K.M.; Pennington, M.W.; Kolski-Andreaco, A.; Wei, E.; Grino, A.; Counts, D.R.; et al. Kv1.3 channels are a therapeutic target for T cell-mediated autoimmune diseases. *Proceedings of the National Academy of Sciences* **2006**, *103*, 17414–17419.

615. Pennington, M.W.; Beeton, C.; Galea, C.A.; Smith, B.J.; Chi, V.; Monaghan, K.P.; Garcia, A.; Rangaraju, S.; Giuffrida, A.; Plank, D.; et al. Engineering a Stable and Selective Peptide Blocker of the Kv1.3 Channel in T Lymphocytes. *Molecular Pharmacology* **2009**, *75*, 762–773.
616. Tarcha, E.J.; Olsen, C.M.; Probst, P.; Peckham, D.; Muñoz-Elías, E.J.; Kruger, J.G.; Iadonato, S.P. Safety and pharmacodynamics of dalazatide, a Kv1.3 channel inhibitor, in the treatment of plaque psoriasis: A randomized phase 1b trial. *PLoS ONE* **2017**, *12*, e0180762.
617. Chi, V.; Pennington, M.W.; Norton, R.S.; Tarcha, E.J.; Londono, L.M.; Sims-Fahey, B.; Upadhyay, S.K.; Lakey, J.T.; Iadonato, S.; Wulff, H.; et al. Development of a sea anemone toxin as an immunomodulator for therapy of autoimmune diseases. *Toxicon* **2012**, *59*, 529–546.
618. Land, J.; Lintermans, L.L.; Stegeman, C.A.; Muñoz-Elías, E.J.; Tarcha, E.J.; Iadonato, S.P.; Heeringa, P.; Rutgers, A.; Abdulahad, W.H. Kv1.3 Channel Blockade Modulates the Effector Function of B Cells in Granulomatosis with Polyangiitis. *Front. Immunol.* **2017**, *8*, 1205.
619. J. Prentis, P.; Pavasovic, A.; S. Norton, R. Sea Anemones: Quiet Achievers in the Field of Peptide Toxins. *Toxins* **2018**, *10*, 36.
620. Gerdol, M.; Cervelli, M.; Mariottini, P.; Oliverio, M.; Dutertre, S.; Modica, M. A Recurrent Motif: Diversity and Evolution of ShKT Domain Containing Proteins in the Vampire Snail Cumia reticulata. *Toxins* **2019**, *11*, 106.
621. Dang, B.; Shen, R.; Kubota, T.; Mandal, K.; Bezanilla, F.; Roux, B.; Kent, S.B.H. Inversion of the Side-Chain Stereochemistry of Indvidual Thr or Ile Residues in a Protein Molecule: Impact on the Folding, Stability, and Structure of the ShK Toxin. *Angew. Chem. Int. Ed.* **2017**, *56*, 3324–3328.
622. Dang, B.; Kubota, T.; Mandal, K.; Bezanilla, F.; Kent, S.B.H. Native Chemical Ligation at Asx-Cys, Glx-Cys: Chemical Synthesis and High-Resolution X-ray Structure of ShK Toxin by Racemic Protein Crystallography. *Journal of the American Chemical Society* **2013**, *135*, 11911–11919.
623. Ma, B.; Zhang, K.; Hendrie, C.; Liang, C.; Li, M.; Doherty-Kirby, A.; Lajoie, G. PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Communications in Mass Spectrometry* **2003**, *17*, 2337–2342.
624. HUPO Test Sample Working Group; Bell, A.W.; Deutsch, E.W.; Au, C.E.; Kearney, R.E.; Beavis, R.; Sechi, S.; Nilsson, T.; Bergeron, J.J.M. A HUPO test sample study reveals common problems in mass spectrometry–based proteomics. *Nat Methods* **2009**, *6*, 423–430.
625. Kapp, E.A.; Schütz, F.; Connolly, L.M.; Chakel, J.A.; Meza, J.E.; Miller, C.A.; Fenyo, D.; Eng, J.K.; Adkins, J.N.; Omenn, G.S.; et al. An evaluation, comparison, and accurate benchmarking of several publicly available MS/MS search algorithms: Sensitivity and specificity analysis. *Proteomics* **2005**, *5*, 3475–3490.
626. Elias, J.E.; Gygi, S.P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* **2007**, *4*, 207–214.
627. Käll, L.; Storey, J.D.; MacCoss, M.J.; Noble, W.S. Assigning Significance to Peptides Identified by Tandem Mass Spectrometry Using Decoy Databases. *J. Proteome Res.* **2008**, *7*, 29–34.
628. Brosch, M.; Yu, L.; Hubbard, T.; Choudhary, J. Accurate and Sensitive Peptide Identification with Mascot Percolator. *J. Proteome Res.* **2009**, *8*, 3176–3181.
629. Jollivet, V.; Hamel, J.F.; de Haro, L.; Labadie, M.; Sapori, J.M.; Cordier, L.; Villa, A.; Nisse, P.; Puskarczyk, E.; Berthelon, L.; et al. European viper envenomation recorded by French poison control centers: A clinical assessment and management study. *Toxicon* **2015**, *108*, 97–103.
630. Audebert, F.; Sorkine, M.; Bon, C. Envenoming by viper bites in France: clinical gradation and biological quantification by ELISA. *Toxicon* **1992**, *30*, 599–609.
631. De Haro, L.; Glaizal, M.; Tichadou, L.; Blanc-Brisset, I.; Hayek-Lanthois, M. Asp Viper (Vipera aspis) Envenomation: Experience of the Marseille Poison Centre from 1996 to 2008. *Toxins* **2009**, *1*, 100–112.
632. Calderón, L.; Lomonte, B.; Gutiérrez, J.M.; Tarkowski, A.; Hanson, L.A. Biological and biochemical activities of Vipera berus (European viper) venom. *Toxicon* **1993**, *31*, 743–753.

633. Samel, M.; Siigur, J. Isolation and characterization of hemorrhagic metalloproteinase from Vipera berus berus (common viper) venom. *Comp. Biochem. Physiol. C, Comp. Pharmacol. Toxicol.* **1990**, *97*, 209–214.
634. Komori, Y.; Sugihara, H. Purification and physiological study of a hypotensive factor from the venom of Vipera aspis aspis (aspic viper). *Toxicon* **1990**, *28*, 359–369.
635. Komori, Y.; Nikai, T.; Sugihara, H. Isolation and characterization of procoagulant from the venom of Vipera aspis aspis. *Int. J. Biochem.* **1993**, *25*, 761–767.
636. Haro, L. de; Robbe-Vincent, A.; Saliou, B.; Valli, M.; Bon, C.; Choumet, V. Unusual neurotoxic envenomations byVipera aspis aspis snakes in France. *Hum Exp Toxicol* **2002**, *21*, 137–145.
637. Mancheva, I.; Kleinschmidt, T.; Aleksiev, B.; Braunitzer, G. Sequence homology between phospholipase and its inhibitor in snake venom. The primary structure of phospholipase A2 of vipoxin from the venom of the Bulgarian viper (Vipera ammodytes ammodytes, Serpentes). *Biol. Chem. Hoppe-Seyler* **1987**, *368*, 343–352.
638. Jan, V.; Maroun, R. c; Robbe-Vincent, A.; De Haro, L.; Choumet, V. Toxicity evolution of Vipera aspis aspis venom: identification and molecular modeling of a novel phospholipase A2 heterodimer neurotoxin 1. *FEBS Letters* **2002**, *527*, 263–268.
639. Ferquel, E.; de Haro, L.; Jan, V.; Guillemin, I.; Jourdain, S.; Teynié, A.; d'Alayer, J.; Choumet, V. Reappraisal of Vipera aspis Venom Neurotoxicity. *PLoS ONE* **2007**, *2*.
640. Zanetti, G.; Duregotti, E.; Locatelli, C.A.; Giampreti, A.; Lonati, D.; Rossetto, O.; Pirazzini, M. Variability in venom composition of European viper subspecies limits the cross-effectiveness of antivenoms. *Scientific Reports* **2018**, *8*, 9818.
641. Lamb, T.; Haro, L. de; Lonati, D.; Brvar, M.; Eddleston, M. Antivenom for European Vipera species envenoming. *Clinical Toxicology* **2017**, *55*, 557–568.
642. Guiavarch, M.; Médus, M.; Tichadou, L.; Glaizal, M.; de Haro, L. Efficacité variable de l'antivenin Viperfav® pour traiter les envenimations vipérines avec neurotoxicité. *La Presse Médicale* **2011**, *40*, 654–656.
643. Calvete, J.J.; Lomonte, B. A bright future for integrative venomics. *Toxicon* **2015**, *107*, 159–162.
644. Pla, D.; Petras, D.; Saviola, A.J.; Modahl, C.M.; Sanz, L.; Pérez, A.; Juárez, E.; Frietze, S.; Dorrestein, P.C.; Mackessy, S.P.; et al. Transcriptomics-guided bottom-up and top-down venomics of neonate and adult specimens of the arboreal rear-fanged Brown Treesnake, Boiga irregularis, from Guam. *Journal of Proteomics* **2018**, *174*, 71–84.
645. Amorim, F.G.; Costa, T.R.; Baiwir, D.; De Pauw, E.; Quinton, L.; Sampaio, S.V. Proteopeptidomic, Functional and Immunoreactivity Characterization of Bothrops moojeni Snake Venom: Influence of Snake Gender on Venom Composition. *Toxins* **2018**, *10*, 177.
646. Degueldre, M.; Verdenaud, M.; Legarda, G.; Minambres, R.; Zuniga, S.; Leblanc, M.; Gilles, N.; Ducancel, F.; De Pauw, E.; Quinton, L. Diversity in sequences, post-translational modifications and expected pharmacological activities of toxins from four Conus species revealed by the combination of cutting-edge proteomics, transcriptomics and bioinformatics. *Toxicon* **2017**, *130*, 116–125.
647. Han, Y.; Ma, B.; Zhang, K. SPIDER: software for protein identification from sequence tags with de novo sequencing error. *J Bioinform Comput Biol* **2005**, *3*, 697–716.
648. Zybailov, B.; Mosley, A.L.; Sardiu, M.E.; Coleman, M.K.; Florens, L.; Washburn, M.P. Statistical Analysis of Membrane Proteome Expression Changes in *Saccharomyces c erevisiae*. *Journal of Proteome Research* **2006**, *5*, 2339–2347.
649. Lochner, A.; Giannone, R.J.; Keller, M.; Antranikian, G.; Graham, D.E.; Hettich, R.L. Label-free Quantitative Proteomics for the Extremely Thermophilic Bacterium *Caldicellulosiruptor obsidiansis* Reveal Distinct Abundance Patterns upon Growth on Cellobiose, Crystalline Cellulose, and Switchgrass. *Journal of Proteome Research* **2011**, *10*, 5302–5314.
650. Al Shweiki, M.R.; Mönchgesang, S.; Majovsky, P.; Thieme, D.; Trutschel, D.; Hoehenwarter, W. Assessment of Label-Free Quantification in Discovery Proteomics and Impact of Technological

Factors and Natural Variability of Protein Abundance. *Journal of Proteome Research* **2017**, *16*, 1410–1424.
651. Hoopmann, M.R.; Winget, J.M.; Mendoza, L.; Moritz, R.L. StPeter: Seamless Label-Free Quantification with the Trans-Proteomic Pipeline. *Journal of Proteome Research* **2018**, *17*, 1314–1320.
652. Sajevic, T.; Leonardi, A.; Križaj, I. An overview of hemostatically active components of *Vipera ammodytes ammodytes* venom. *Toxin Reviews* **2014**, *33*, 33–36.
653. Sanchez, E.F.; Flores-Ortiz, R.J.; Alvarenga, V.G.; Eble, J.A. Direct Fibrinolytic Snake Venom Metalloproteinases Affecting Hemostasis: Structural, Biochemical Features and Therapeutic Potential. *Toxins (Basel)* **2017**, *9*.
654. Gutiérrez, J.M.; Escalante, T.; Rucavado, A.; Herrera, C. Hemorrhage Caused by Snake Venom Metalloproteinases: A Journey of Discovery and Understanding. *Toxins (Basel)* **2016**, *8*.
655. *Venomous reptiles and their toxins: evolution, pathophysiology, and biodiscovery*; Fry, B., Ed.; Oxford University Press: New York, NY, 2015; ISBN 978-0-19-930939-9.
656. Sajevic, T.; Leonardi, A.; Kovačič, L.; Lang-Balija, M.; Kurtović, T.; Pungerčar, J.; Halassy, B.; Trampuš-Bakija, A.; Križaj, I. VaH3, one of the principal hemorrhagins in Vipera ammodytes ammodytes venom, is a homodimeric P-IIIc metalloproteinase. *Biochimie* **2013**, *95*, 1158–1170.
657. Latinović, Z.; Leonardi, A.; Šribar, J.; Sajevic, T.; Žužek, M.C.; Frangež, R.; Halassy, B.; Trampuš-Bakija, A.; Pungerčar, J.; Križaj, I. Venomics of Vipera berus berus to explain differences in pathology elicited by Vipera ammodytes ammodytes envenomation: Therapeutic implications. *Journal of Proteomics* **2016**, *146*, 34–47.
658. Serrano, S.M.T. The long road of research on snake venom serine proteinases. *Toxicon* **2013**, *62*, 19–26.
659. Clemetson, K.J. Snaclecs (snake C-type lectins) that inhibit or activate platelets by binding to receptors. *Toxicon* **2010**, *56*, 1236–1246.
660. Ogawa, T.; Chijiwa, T.; Oda-Ueda, N.; Ohno, M. Molecular diversity and accelerated evolution of C-type lectin-like proteins from snake venom. *Toxicon* **2005**, *45*, 1–14.
661. Crocker, P.R.; Paulson, J.C.; Varki, A. Siglecs and their roles in the immune system. *Nature Reviews Immunology* **2007**, *7*, 255–266.
662. van den Berg, L.M.; Gringhuis, S.I.; Geijtenbeek, T.B.H. An evolutionary perspective on C-type lectins in infection and immunity: C-type lectins in infection and immunity. *Annals of the New York Academy of Sciences* **2012**, *1253*, 149–158.
663. Calvete, J.J.; Marcinkiewicz, C.; Monleón, D.; Esteve, V.; Celda, B.; Juárez, P.; Sanz, L. Snake venom disintegrins: evolution of structure and function. *Toxicon* **2005**, *45*, 1063–1074.
664. Calvete, J.J.; Moreno-Murciano, M.P.; Theakston, R.D.G.; Kisiel, D.G.; Marcinkiewicz, C. Snake venom disintegrins: novel dimeric disintegrins and structural diversification by disulphide bond engineering. *Biochem J* **2003**, *372*, 725–734.
665. Chakrabarty, D.; Chanda, C. Snake Venom Disintegrins. In *Snake Venoms*; Gopalakrishnakone, P., Inagaki, H., Mukherjee, A.K., Rahmy, T.R., Vogel, C.-W., Eds.; Springer Netherlands: Dordrecht, 2015; pp. 1–11 ISBN 978-94-007-6648-8.
666. Chen, H.-S.; Wang, Y.-M.; Huang, W.-T.; Huang, K.-F.; Tsai, I.-H. Cloning, characterization and mutagenesis of Russell's viper venom l-amino acid oxidase: Insights into its catalytic mechanism. *Biochimie* **2012**, *94*, 335–344.
667. Yamazaki, Y.; Morita, T. Structure and function of snake venom cysteine-rich secretory proteins. *Toxicon* **2004**, *44*, 227–231.
668. Sunagar, K.; Johnson, W.E.; O'Brien, S.J.; Vasconcelos, V.; Antunes, A. Evolution of CRISPs Associated with Toxicoferan-Reptilian Venom and Mammalian Reproduction. *Mol Biol Evol* **2012**, *29*, 1807–1822.
669. Schweitz, H.; Bruhn, T.; Guillemare, E.; Moinier, D.; Lancelin, J.M.; Béress, L.; Lazdunski, M. Kalicludines and kaliseptine. Two different classes of sea anemone toxins for voltage sensitive K+ channels. *J. Biol. Chem.* **1995**, *270*, 25121–25126.

670. Dy, C.Y.; Buczek, P.; Imperial, J.S.; Bulaj, G.; Horvath, M.P. Structure of conkunitzin-S1, a neurotoxin and Kunitz-fold disulfide variant from cone snail. *Acta Crystallogr. D Biol. Crystallogr.* **2006**, *62*, 980–990.
671. Sasaki, S.D.; Azzolini, S.S.; Hirata, I.Y.; Andreotti, R.; Tanaka, A.S. Boophilus microplus tick larvae, a rich source of Kunitz type serine proteinase inhibitors. *Biochimie* **2004**, *86*, 643–649.
672. Strydom, D.J. Protease inhibitors as snake venom toxins. *Nature New Biol.* **1973**, *243*, 88–89.
673. Yuan, C.-H.; He, Q.-Y.; Peng, K.; Diao, J.-B.; Jiang, L.-P.; Tang, X.; Liang, S.-P. Discovery of a Distinct Superfamily of Kunitz-Type Toxin (KTT) from Tarantulas. *PLoS One* **2008**, *3*.
674. Yee, K.T.; Pitts, M.; Tongyoo, P.; Rojnuckarin, P.; Wilkinson, M.C. Snake Venom Metalloproteinases and Their Peptide Inhibitors from Myanmar Russell's Viper Venom. *Toxins (Basel)* **2016**, *9*.
675. Wagstaff, S.C.; Favreau, P.; Cheneval, O.; Laing, G.D.; Wilkinson, M.C.; Miller, R.L.; Stöcklin, R.; Harrison, R.A. Molecular characterisation of endogenous snake venom metalloproteinase inhibitors. *Biochemical and Biophysical Research Communications* **2008**, *365*, 650–656.
676. Manjunatha Kini, R. Excitement ahead: structure, function and mechanism of snake venom phospholipase A2 enzymes. *Toxicon* **2003**, *42*, 827–840.
677. Georgieva, D.; Risch, M.; Kardas, A.; Buck, F.; von Bergen, M.; Betzel, C. Comparative Analysis of the Venom Proteomes of *Vipera ammodytes ammodytes* and *Vipera ammodytes meridionalis*. *Journal of Proteome Research* **2008**, *7*, 866–886.
678. Pungercar, J.; Krizaj, I.; Liang, N.S.; Gubensek, F. An aromatic, but not a basic, residue is involved in the toxicity of group-II phospholipase A2 neurotoxins. *Biochem J* **1999**, *341*, 139–145.
679. Krizaj, I.; Liang, N.-S.; Pungercar, J.; Strukelj, B.; Ritonja, A.; Gubensek, F. Amino acid and cDNA sequences of a neutral phospholipase A2 from the long-nosed viper (Vipera ammodytes ammodytes) venom. *European Journal of Biochemistry* **1992**, *204*, 1057–1062.
680. Yamazaki, Y.; Takani, K.; Atoda, H.; Morita, T. Snake Venom Vascular Endothelial Growth Factors (VEGFs) Exhibit Potent Activity through Their Specific Recognition of KDR (VEGF Receptor 2). *J. Biol. Chem.* **2003**, *278*, 51985–51988.
681. Sannaningaiah, D.; Subbaiah, G.K.; Kempaiah, K. Pharmacology of spider venom toxins. *Toxin Reviews* **2014**, *33*, 206–220.
682. Haider, S.; Pal, R. Integrated Analysis of Transcriptomic and Proteomic Data. *Curr Genomics* **2013**, *14*, 91–110.
683. Al-Shekhadat, R.I.; Lopushanskaya, K.S.; Segura, Á.; Gutiérrez, J.M.; Calvete, J.J.; Pla, D. Vipera berus berus Venom from Russia: Venomics, Bioactivities and Preclinical Assessment of Microgen Antivenom. *Toxins* **2019**, *11*, 90.
684. Reading, C.J. Incidence, pathology, and treatment of adder (Vipera berus L.) bites in man. *Emergency Medicine Journal* **1996**, *13*, 346–351.
685. Chippaux, J.-P. Epidemiology of snakebites in Europe: A systematic review of the literature. *Toxicon* **2012**, *59*, 86–99.
686. Valenta, J.; Stach, Z.; Stříteský, M.; Michálek, P. Common Viper Bites in the Czech Republic - Epidemiological and Clinical Aspects during 15 Year Period (1999–2013). *Prague Medical Report* **2014**, *115*, 120–127.
687. Garkowski, A.; Czupryna, P.; Zajkowska, A.; Pancewicz, S.; Moniuszko, A.; Kondrusik, M.; Grygorczuk, S.; Gołębicki, P.; Letmanowski, M.; Zajkowska, J. Vipera berus bites in Eastern Poland – a retrospective analysis of 15 case studies. *Annals of Agricultural and Environmental Medicine* **2012**, *19*, 6.
688. Malina, T.; Krecsak, L.; Warrell, D.A. Neurotoxicity and hypertension following European adder (Vipera berus berus) bites in Hungary: case report and review. *QJM* **2008**, *101*, 801–806.
689. Magdalan, J.; Trocha, M.; Merwid-Ląd, A.; Sozański, T.; Zawadzki, M. Vipera berus Bites in the Region of Southwest Poland—A Clinical Analysis of 26 Cases. *Wilderness & Environmental Medicine* **2010**, *21*, 114–119.

690. Warrell, D.A. Treatment of bites by adders and exotic venomous snakes. *BMJ* **2005**, *331*, 1244–1247.

Résumé/Abstract

Titre : Synthèse de peptides bioactifs inspirés des venins

Les venins animaux ont évolué pour la défense et pour faciliter la prédation. Ainsi, les toxines qui les composent ont souvent pour but d'incapaciter la proie ou le prédateur, notamment en ciblant des fonctions physiologiques clés telles que l'hémostase, la jonction neuromusculaire ou la perception de la douleur et sont généralement de nature peptidique. Ces peptides riches en ponts disulfures provenant des venins présentent une structure tridimensionnelle contrainte et une stabilité plasmatique accrue comparé aux peptides linéaires. La conservation et la ressemblance des récepteurs de la proie/du prédateur avec les récepteurs humains expliquent que ces peptides de venins soit une source unique de composés originaux pour le design d'outils pharmacologiques et de molécules thérapeutiques. Bien que la plupart des toxines typiques modulent les canaux ioniques assurant la transmission de l'information nerveuse, nous avons d'abord commencé par étudier le potentiel des toxines linéaires et à 1 pont sur des cibles plus rarement explorées tels que les récepteurs couplés aux protéines G (chapitre 1). Cependant, la majorité des toxines étant composées de 2 ou 3 ponts disulfures, nous avons ensuite opté pour la mise en place d'une stratégie robuste de synthèse de peptides à deux ponts. Ainsi, la synthèse de plusieurs α-conotoxines (chapitre 2) a permis leur caractérisation structurale par RMN et pharmacologique sur les récepteurs nicotiniques. Une extension de ce travail sur une séquence prometteuse ayant une activité duale sur les récepteurs musculaire et neuronaux a été effectuée par cyclisation N-C. Enfin, ce travail a été étendu à la synthèse de toxines plus longue ciblant d'autres canaux ioniques et comportant 3 ponts disulfures (chapitre 3). Pour finir, la synthèse de nouvelles toxines découvertes dans des banques de données issues de divers transcriptomes de glandes à venin a été initiée, et nous avons réalisé l'analyse protéo-transcriptomique d'un venin de serpent pour l'identification de nouvelles toxines (chapitre 4).

Mots-clés : Venins, Toxines, Ponts disulfures, Synthèse, Caractérisation structurale, Caractérisation pharmacologique, Outils pharmacologiques

---

Title: Synthesis of bioactive peptides from venoms

Animal venoms have evolved for defence and to facilitate predation. The toxins they contain are often intended to incapacitate the prey or predator, by targeting key physiological functions such as hemostasis, neuromuscular junction or pain perception and are generally peptides. These disulfide-rich venom peptides have a constrained three-dimensional structure and increased plasma stability compared to linear peptides. The conservation and similarity of prey/predator receptors to human receptors makes venom peptides a unique source of lead compounds for the design of pharmacological tools and therapeutic compounds. Although most typical toxins modulate ion channels that transmit nerve information, we first studied the potential of linear and 1-bridge toxins on more rarely explored targets such as G-protein coupled receptors (Chapter 1). However, since most toxins are composed of 2 or 3 disulfide bridges, we then opted to set up a robust strategy for synthesis of toxins with two disulfide bridges. Thus, the synthesis of several α-conotoxins (chapter 2) allowed their structural characterization by NMR and pharmacological characterizaton on nicotinic receptors. An extension of this work on a promising sequence with dual activity on muscle and neural receptors was performed by N-C cyclization. This work was extended to the synthesis of longer toxins targeting other ion channels and containing 3 disulfide bridges (chapter 3). Finally, the synthesis of new toxins discovered in databases from various venom gland transcriptomes was initiated, and we performed proteo-transcriptomic analysis of snake venom for the identification of new toxins (Chapter 4).

Keywords: Venoms, Toxins, Disulfide bridges, Synthesis, Structural characterization, Pharmacological characterization, Pharmacological tools