Des toxines animales comme ligands innovants des récepteurs aux mélanocortines

Steve Reynaud

▶ **To cite this version:**

Steve Reynaud. Des toxines animales comme ligands innovants des récepteurs aux mélanocortines. Pharmacologie. Université Paris Saclay (COmUE), 2019. Français. NNT : 2019SACLS596 . tel-03510297

HAL Id: tel-03510297
https://theses.hal.science/tel-03510297

Submitted on 4 Jan 2022

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

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

NNT : 2019SACLS596

Thèse de doctorat

Des toxines animales comme ligands innovants des récepteurs aux mélanocortines

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université Paris-Sud

École doctorale n°569
Innovation thérapeutique : du fondamental à l'appliqué

Spécialité de doctorat:
Sciences pharmacologiques

Thèse présentée et soutenue à Gif-sur-Yvette, le 19/12/2019, par

Steve Reynaud

Composition du Jury :

Philippe Minard Professeur, Université Paris-Sud (I2BC)	Président
Jean-Louis Banères Directeur de Recherche, CNRS (IBMM)	Rapporteur
Sonia Cantel Maitre de conférences, Université de Montpellier 2 (IBMM)	Rapporteur
Michel Bouvier Professeur, Université de Montréal (IRIC)	Examinateur
Loïc Quinton Professeur, Université de Liège (MSLab)	Examinateur
Nicolas Gilles Chercheur, CEA (SIMOPRO)	Directeur de thèse

À ma famille,

Remerciements

À mon directeur de thèse

Je tiens tout d'abord à remercier mon directeur de thèse, Dr Nicolas Gilles, qui, de par ses conseils et son soutien, m'a appris la persévérance et l'optimisme nécessaires au travail de chercheur mais également un large panel de connaissances qui m'ont permis de m'ouvrir dans des domaines scientifiques extrêmement variés. Merci Nicolas.

Aux membres du jury

Je remercie bien sûr tous les membres du jury : Dr Florine Cavelier, Dr Sonia Cantel, Dr Jean-Louis Banères, mes rapporteurs, pour leur immense gentillesse et leur avis scientifiques qu'ils m'ont donnés quitte à traverser plusieurs orages et des situations exeptionnelles ; mes examinateurs, Prof Michel Bouvier, Prof Philippe Minard et Prof Loïc Quinton, d'avoir accepté d'apporter leur précieuse expertise sur mon travail de thèse.

À une équipe pluridisciplinaire

Merci au Dr Denis Servent de m'avoir accueilli dans son équipe pluridisciplinaire où chacun peut apprendre quelque chose à l'autre et dans un domaine aussi passionnant que la toxinologie. Un grand merci également au Dr Gilles Mourier, toujours bienveillant, pour son aide en chimie peptidique ; au Dr Pascal Kessler pour sa bonne humeur et de m'avoir trouvé un surnom que personne d'autre n'aurait pu trouver ; au Dr Philippe Robin, pour son aide en biologie moléculaire et en culture cellulaire, toujours disponible pour répondre à mes questions ; au Dr Muriel Amar pour son aide en test de liaison ; au Dr Jordi Molgó, preuve d'une incroyable passion de la recherche, et dont la dissection d'un nerf d'une souris restera l'un des moments les plus marquant de ma thèse ; enfin un énorme merci au Dr Évelyne Benoit pour son extrême gentillesse, sa disponibilité, le partage de ses connaissances,… Une équipe ne peut bien fonctionner sans une aide administrative, donc merci également à Sylviane Barbe et Évelyne Rabouille pour leur précieuse aide.

À un service à l'écoute

D'autres membres du service m'ont également apporté leur aide. Je pense notamment au Dr Alain Lecoq, à Sandrine Morel-Altmeyer et à Fabrice Beau qui ont répondu avec beaucoup de patience à mes questions sur la gestion des risques radioactifs et chimiques. Un grand merci également au Dr Robert « Bob » Thai et Stéven Dubois pour les nombreuses analyses que je leur ai demandées et qui m'ont beaucoup aidé dans l'avancement de mon projet de thèse. Merci aussi au Dr Oscar Pereira Ramos pour les scripts qui m'ont fait gagner un temps monstrueux dans l'analyse des données.

Aux collaborateurs

Ce travail de thèse n'aurait jamais pu être mené à bien sans de précieux collaborateurs. Ainsi un grand merci aux Dr Hervé Meudal et Dr Céline Landon pour leurs explications claires sur la RMN et leur contribution dans mon premier article scientifique. Merci également aux Dr Annette G. Beck-Sickinger et Dr Karin Mörl pour les tests de fonction et pour avoir accepté d'être ma première collaboration internationale. Merci au Dr Ricardo C. Rodríguez de la Vega avec qui je partage une passion pour l'Évolution. Merci aux Dr Steve

Peigneur et Dr Jan Tytgat pour les tests en électrophysiologie. Et enfin, un grand merci à Suli-Anne Laurin, future docteure et à son directeur de thèse, Dr Michel Bouvier, pour leurs réponses à mes questions, qui ont débloqué mon projet alors qu'il stagnait. Parmi les collaborateurs, il y a aussi des stagiaires : Michael Susset, compagnon de pontage dirigé ; Florian Blondel, destructeur de pont disulfure ; Marine Ghazarian, experte du marquage ; Mélanie Montagna, reine du test de liaison et de la SAR ; Steven Ascoët, partenaire de cinétiques interminables, bon courage pour la suite futur docteur. Merci à vous tous. J'espère que vous avez autant appris avec moi que j'ai appris avec vous.

Aux collègues-amis

Trois ans, c'est assez long pour connaitre un grand nombre de personnes qui deviennent progressivement des amis. Merci Justyna, tes conseils, ton soutien ont été sans faille. Merci Nicolò. Merci pour ton aide, ton partage de connaissances. Les étudiants qui seront avec toi auront beaucoup de chance. Merci Margot. Toujours souriante et de bonne humeur, parler avec toi de tout et de rien étaient un vrai plaisir. Anne-Cécile, merci pour ton soutien. Maintenant, je te passe le flambeau. Bon courage future docteure. Merci Lidia. Tu as toujours été là. Tu es la définition du mot persévérance. À l'occasion, on se refera des randonnées. Merci à mes partenaires de galère, Tânia, Marcela, Laura et Livia pour votre soutien et pour m'avoir montré que c'était possible d'y arriver. Un grand merci à toi aussi, Daad, pour ton précieux soutien lors de ma première année de thèse. Un énorme merci à Pavel, Alexandra et Pierrick pour votre soutien. Manger et faire des pauses avec vous me faisaient oublier le stress de la thèse et me redonnait du courage. Merci aussi à Mihn et Fanny. Merci à mes compagnons de bureau, Manon, Julia, Christelle, Hugo et Marion pour des discussions scientifiques et un peu moins sérieuses. Un merci spécial à Manon. Tu es la première personne que j'ai vraiment connue au laboratoire. On pouvait vraiment parler de tout. Que de bons moments passés au bureau et à l'extérieur. Un autre merci spécial à Julia. Travailler avec toi était un pur plaisir même dans les galères. Vous avez su toutes les deux me soutenir dans les pires moments de ma thèse. Merci encore.

À ma famille

Remercier sa famille à la fin d'une étape comme la thèse nécessiterait plusieurs pages. Je préfère vous dire, ici, un simple merci et vous montrer ma gratitude directement en face. Merci à mes parents, mon frère, Jonathan, Anaïs et à mes oncles et mes tantes qui ont quitté le Sud de la France pour venir assister à ma soutenance de thèse. Parce que parfois, des amis deviennent comme la famille. Merci Mara, Benjamin, Jérémy, Océane, Théo, Lilly. Même si vous étiez loin, votre soutien était présent. Merci également à Dounia, Myriam et Michel Zamiati, des personnes formidables au grand cœur qui m'ont accueilli et m'ont aidé à surmonter une épreuve. Je vous serai éternellement reconnaissant. Et enfin un merci infini à Héloïse. Tu as su me supporter dans tous les sens du terme jusqu'à la fin malgré les épreuves et mes deux mains gauches. Merci.

Liste des abréviations

AC	Adénlyate cyclase
ACE	Enzyme de conversion de l'angiotensine (*Angiotensin Converting Enzyme*)
Acm	Acétamidométhyl
ACN	Acétonitrile
ACTH	Hormone corticotrope (*Adreno CorticoTrohic Hormone*)
AGL	Acides gras libres
AgRP	Agouti-related peptide
AMPc	Adénosine monophosphate cyclique
ARC	Noyau arqué (*Arcuate Nucleus*)
$ARC^{AgRP/NPY}$	Neurones du noyau arqué exprimant les peptides AgRP et NPY
ARC^{POMC}	Neurones du noyau arqué exprimant POMC
ASIC	Canal ionique sensible à l'acidité (*Acid sensing ion channel*)
ASIP	Agouti-signaling peptide
ATP	Adénosine triphosphate
BHE	Barrière hématoencéphalique
BLA	Amygdale basolatérale (*Basolateral Amygdala*)
BNST	Noyau du lit de la stria terminalis (*Bed Nucleus of the Stria Terminalis*)
BPTI	Inhibiteur pancréatique bovin de la trypsine (*Bovine Pancreatic Trypsin Inhibitor*)
BRET	Transfert d'énergie par résonnance de bioluminescence (*Bioluminescence Resonance Energy Transfer*)
CART	Transcrit régulé par la cocaïne et l'amphétamine (*Cocaine and Amphetamine Regulated Transcript)*
Ca_v	Canaux calciques dépendant du potentiel
CC	Cystine
CCK	Cholécystokinine
CCK-AR	Récepteur A à CCK (*CCK-A* receptor)
CGRP	Peptide relié au gène de la calcitonine (*Calcitonin-Gene Related Peptide*)
CEA	Commissariat à l'Énergie Atomique et aux Énergies Alternatives
CHO	Cellules d'ovaires de hamster (*Chinese Hamster Ovary cells*)
CLIP	*Corticotropin-like Intermediate Peptide*
CSαβ	*Cysteine-stabilized-α-helix and β-sheet*
D1R	Récepteur D1 à la dopamine
DA	Dopamine
DAG	Diacylglycérol
DC	Dichroïsme circulaire
DCM	Dichlorométhane
DEA	Dépenses énergétiques adaptatives
DER	Dépenses énergétiques au repos
DET	Dépenses énergétiques totales
DIA	Dépenses énergétiques induites par l'activité physique
DMV	Noyau dorso-médian du nerf vague
dpm	Désintégrations par minute

ECL	Boucle extracellulaire (*Extracellular Loop*)
EC50	Concentration d'agoniste nécessaire pour obtenir 50 % de la réponse maximale
ERK	Kinase régulée par un signal extracellulaire (*Extracellular Signalling Kinase)*
FDA	Food and Drug Administration
Fmoc	Fluorénylméthoxycarbonyle
FRET	Transfert d'énergie par résonnance de Förster (*Förster Resonance Energy Transfer*)
GABA	Acide γ-aminobutyrique
GDP	Guanosine 5'-diphosphate
GEF	Facteur d'échange nucléotidique (*Guanine-nucleotide-Exchange Factor*)
GHSR	Récepteur à la ghréline (*Growth hormone secretagogue receptor*)
GSSG	Disulfure de glutathion
GLP-1	*Gucagon-like peptide 1*
Glu	Glucose
Glut	Glutamate
GlutR	Récepteur au glutamate
GN	Ganglion inférieur du nerf vague
GRK	Kinase spécifique des RCPGs (*G-protein-coupled Receptor Kinase*)
GTP	Guanosine 5'-triphosphate
HCTU	O-(1H-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HEK	Cellules de rein d'embryon humain (*Human Embryonic kidney*)
HTRF	Fluorescence résolue en temps homogène (*Homogeneous Time Resolved Fluorescence*)
HTS	Criblage à haut débit (*High Throughput Screening*)
ICK	Inhibitor Cystine Knot
ICL	Boucle intracellulaire (*Intracellular Loop*)
icv	Injection intracérébroventriculaire
IC50	Concentration de ligand inhibant 50 % de la liaison du traceur
IMC	Indice de masse corporelle
IML	Noyau intermédiolatéral de la moelle épinière
InsCtx	Cortex insulaire
IP	Intrapéritonéal
IP_3	Inositol TRISphosphate
RI	Récepteur à l'insuline
JNK3	c-Jun N-terminal kinases 3
Kir	Canal potassique rectifiant entrant (*Potassium Inward Rectifier channel*)
k_{on}	Constante cinétique d'association
k_{off}	Constante cinétique de dissociation
K_p	Constante de protection
KTT	Toxines du type Kunitz
KTx	Toxine de scorpion active sur les canaux potassiques
λ-KTx	Famille de toxines ICK de scorpion actives sur les canaux potassiques
K_v	Canaux potassiques dépendant du potentiel
LH	Hypothalamus latéral (*Lateral Hypothalamus*)
LHb	Habenula latéral (*Lateral Habenula*)

LPBN	Noyau parabrachial latéral (*Lateral Parabracial Nucleus*)
mAChR	Récepteur muscarinique à l'acétylcholine
MAP	*Mitogen-Activated Protein*
MC4R	Sous-type 4 des récepteurs aux mélanocortines (*Melanocortin-4 receptor*)
MCR	Récepteur aux mélanocortines (*Melanocortin receptor*)
MCH	Hormone concentrant la mélanine (*Melanin Concentrating Hormone*)
MCHR	Récepteur à MCH (*Melanin Concentrating Hormone Receptor*)
ME	Moelle épinière
MRAP	Protéine accessoire des récepteurs aux mélanocortines (*Melanocortin Receptor Accessory Protein*)
MSH	Hormone mélanotrope (*Melanocyte-stimulating hormone*)
NAc	Noyau accumbens.
nAChR	Récepteur nicotinique à l'acétylcholine
NaSpTx	Toxines d'araignées actives sur les canaux sodiques dépendant du potentiel
Na_v	Canal sodique dépendant du potentiel
NDP-α-MSH	[NLe4,D-Phe7]α-MSH
NER	Réparation par excision de nucléotides (*Nuclotide Excision Repair*)
NMP	N-méthylpyrrolidone
NPY	Neuropeptide Y
NPY2R	Récepteur 2 des récepteurs au neuropeptide 2
NTS	Noyau du tractus solitaire
OB-R	Récepteur à la leptine
OMS	Organisation Mondiale de la Santé
Orx	Orexine
OxR	Récepteur à l'orexine
OXT	Ocytocine
pI	Point isoélectrique
PIP_2	Phosphatidylinositol 4,5-bisphosphate
PKA	Protéine kinase A
PKC	Protéine kinase C
PLA2	Phospholipase A2
PLC	Phospholipase C
POMC	Pro-opiomélanocortine
POR	Cortex postrhinal (*Postrhinal cortex*)
PTM	Modification post-traductionnelle (*Post-Translational Modification*)
PVN	Noyau paraventriculaire (*Paraventricular nucleus*)
PVT	Thalamus paraventriculaire (Paraventricular thalamus)
PYY	Peptide YY
RCPG	Récepteur couplé aux protéines G
RGS	Molécules régulatrices de la signalisation des protéines G (*Regulator of G-protein signaling*)
SAR	Relation stucture-activité (*Structure-Activity Relationship*)
seq	séquence
Sim1	Facteur de transcription Single-minded 1
SNC	Système nerveux central
SNP	Système nerveux périphérique
SPA	Test de scintillation de proximité (*Scintillation Proximity Assay*)

SPPS	Synthèse peptidique en phase solide (*Solid-Phase Peptide Synthesis*)
3FT	Toxine à trois doigts (*Three-Finger Toxin*)
$t_{1/2}$	Temps de demi-vie
TFA	Acide trifluoroactéique (*Trifluoroacetic acid*)
TM	Transmembranaire
TRPV1	Récepteur 1 vanilloïde à potentiel transitoire (Transient Receptor Potential Vanilloide 1
UV	Unltraviolet
Vgat	Transporteur vésiculaire du GABA
Vglut2	Type 2 du transporteur vésiculaire du glutamate (Vglut2)
VMH	Noyau ventomédial de l'hypothalamus
VPL	Noyau ventral postéro-latéral du thalamus (*Ventral Posterolateral Thalamus*)
VPM	Noyau ventral postéro-médial du thalamus (*Ventral Posteromedial Thalamus*)
VTA	Aire tegmentale ventrale (*Ventral Tegmental Area*)

Liste des Figures

Figure 1. Structure schématique générale d'un récepteur couplé aux protéines G (RCPG) en 3D (A) et en 2D (B)..... 25
Figure 2. Cycle d'activation et de désensibilisation d'un RCPG..... 27
Figure 3. Génération des différentes mélanocortines à partir du clivage de POMC..... 32
Figure 4. Circuits de la faim induite par un déficit énergétique.. 35
Figure 5. Voie neuronale illustrant le processus de traitement sensoriel de signaux alimentaires dépendant de la sensation de faim..... 38
Figure 6. Circuits du renforcement positif de la prise alimentaire par l'hypothalamus latéral (LH). 39
Figure 7. Circuits de la satiation (effet à court terme)..... 43
Figure 8. Circuits de la satiété (effet à long terme). 47
Figure 9. Rôle du sous-type 4 des récepteurs aux mélanocortines (MC4R) dans les dépenses énergétiques via son action sur le tissu adipeux brun..... 52
Figure 10. Représentation schématique d'une BHE peu perméable (à gauche) et d'une BHE perméable de l'éminence médiane (à droite) 54
Figure 11. Variants perte de fonction ou gain de fonction du MC4R 58
Figure 12. Développement à partir de l'hormone mélanotropes des principaux agonistes peptidiques MC4R sur le marché ou en phase clinique III 61
Figure 13. La cachexie cancéreuse, un syndrome multi-organes 64
Figure 14. Diversité des principaux systèmes d'injection des animaux venimeux 67
Figure 15. Arbre représentant les grands groupes d'animaux venimeux 68
Figure 16. Mécanismes à l'origine de l'évolution des venins 70
Figure 17. Convergence évolutive des activités des toxines animales 71
Figure 18. Représentation schématique de la structure canonique d'une toxine ICK 73
Figure 19. Séquences consensus des familles de NaSpTx..... 74
Figure 20. Caractéristiques structurales des toxines Kunitz 76
Figure 21. Caractéristiques structurales des toxines à 3 doigts de serpent 78
Figure 22. Caractéristiques structurales des α-Ktx de scorpion 81
Figure 23. Comparaison de deux stratégies pour l'étude des venins : bioguidage (à gauche), Venomics (à droite) 87
Figure 24. Consortium Venomics 88
Figure 25. Statistical representation of the sequence banks built from omics analysis (upper panel) and produced Venomics bank (lower panel)**.** 91
Figure 26. Statistical representation of the hits following Venomics bank screening on hMC4R 97
Figure 27. Analytical HPLC chromatogram of the final purified products of N-TRTX-Preg1a and N-BUTX-Ptr1a 104
Figure 28. Pharmacological characterization of N-TRTX-Preg1a and N-BUTX-Ptr1a..... 105
Figure 29. Structural characterization of N-TRTX-Preg1a and N-BUTX-Ptr1a 107
Figure 30. Oxydation aléatoire de N-CTX-Can1a 112
Figure 31. Stratégie de pontage dirigé partiel de N-CTX-Can1a (exemple de la forme Can1a-2) 113

Figure 32. Profil en HPLC analytique des différents produits de N-CTX-Can1a à deux ponts disulfure obtenus après 3 jours d'oxydation aléatoire 113
Figure 33. Attribution des pics obtenus en oxydation aléatoire à l'aide des formes oxydées après pontage dirigé de N-CTX-Can1a. 114
Figure 34. Mécanisme d'inversion de Cys et Arg dans la séquence de N-CTX-Can1a 115
Figure 35. Oxydation aléatoire de N-CTX-Mil1a 118
Figure 36. Stratégie de pontage dirigé de N-CTX-Mil1a 119
Figure 37. Attribution des pics obtenus en oxydation aléatoire à l'aide des formes oxydées après pontage dirigé de N-CTX-Mil1a 120
Figure 38. Domaine EGF-like (Prosite PDOC00021) 122
Figure 39. Analyse LC-MS (ESI-IT) du produit obtenu après oxydation aléatoire de la toxine N-CTX-Ltg1a 124
Figure 40. Attribution des pics obtenus en oxydation aléatoire (en bleu) de N-CTX-Ltg1a à l'aide des formes oxydées après pontage dirigé 124
Figure 41. Séquences des trois formes à un pont disulfure issues de N-CTX-Ltg1a 125
Figure 42. Comparaison des structures de la L-glutamine (A) et d'une cystéine alkylée à l'iodoacétamide (B) dans une chaine peptidique 126
Figure 43. Analyse en dichroisme circulaire de la toxine N-CTX-Ltg1a et de son dérivé, la Hunger-Toxin (HT1-0). 127
Figure 44. Structure de 6-Azido-L-Lysine 128
Figure 45. Analyse qualité de HT1-0 marqué à la fluorescéine (HT1-42) au niveau de K21. 131
Figure 46. Analyse qualité de HT1-0 marqué à la rhodamine au niveau de K21 (HT1-43) 131
Figure 47. Marquage radioactif à l'iode 125 de HT1-49 133
Figure 48. Cinétique d'association de ^{125}I-HT1-49 en scintillation liquide 134
Figure 49. Exemple de cinétique d'association et de dissociation de ^{125}I-HT1-49 en scintillation de proximité (SPA) 134
Figure 50. Visualisation de la liaison de HT1-0 marqué à la rhodamine sur le récepteur hMC4R-GFP 135
Figure 51. Cinétique d'activation de différentes voies de signalisation du hMC4R en présence (•) ou en absence (○) de HT1-0 (1 µM) (A) ou d'α-MSH (2 µM) (B) 137
Figure 52. Étude de la fonction antagoniste de HT1-0 sur les voies G_s, G_{15} et β-arrestine 2 de hMC4R activées par α-MSH ou NDP-α-MSH 138
Figure 53. Inhibition de la production d'AMPc (voie G_s) de hMC4R stimulée par NDP-α-MSH en présence de concentrations croissantes de HT1-49 139
Figure 54. Dissociation lente de ^{125}I-NDP-α-MSH du complexe avec hMC4R et conséquence sur l'affinité de HT1-0 en compétition. 142
Figure 55. Principe de la synthèse chimique en phase solide. 154
Figure 56. Exemple de la production en pontage dirigé d'un peptide à 2 ponts disulfure ... 155
Figure 57. Protection des cystéines du peptide réduit par traitement à l'iodoacétamide ... 157
Figure 58. Principe du test de Kaiser 158
Figure 59. Structure des fluorophores utilisés. 159
Figure 60. Spectres d'émission d'accepteurs vert ou rouge et de donneur HTFR (cryptate de Tb^{2+} ou d'Eu^{3+}) 163
Figure 61. Principe de la fluorescence résolue en temps homogène (HTRF). 164
Figure 62. Principe de la quantification de l'AMPc produit par les cellules par compétition de liaison à l'anticorps anti-AMPc. 165

Liste des Tables

Table 1. Vue d'ensemble des différents sous-types des récepteurs aux mélanocortines 32
Table 2. Motifs de cystéines des toxines actives sur les canaux potassiques K_v 82
Table 3. Selected taxa for sourcing 90
Table 4. Cysteine pattern diversity with an even number of cysteine identified in Omics bank 91
Table 5. Cysteine pattern diversity with an odd number of cysteine identified in Omics bank 94
Table 6. Hit rate following the Venomics bank screening on different GPCRs..... 96
Table 7. Cystein pattern diversity of MC4R hits..... 98
Table 8. Toxines animales sélectionnées 102
Table 9. Inhibition constant values of N-TRTX-Preg1a, N-BUTX-Ptr1a on melanocortin receptors 104
Table 10. Comparaison de N-CTX-Can1a avec les séquences les plus proches de la banque Venomics 110
Table 11. Constantes d'inhibition (Ki) moyennes des différentes formes oxydées de N-CTX-Can1a sur hMC4R 114
Table 12. Comparaison de N-CTX-Mil1a avec les séquences les plus proches de la base UniProtKB/SwissProt (A) et de la banque Venomics (B) 117
Table 13. Constantes d'inhibition (Ki) moyennes des différentes formes oxydées produites de N-CTX-Mil1a sur les récepteurs aux mélanocortines 118
Table 14. Comparaison de N-CTX-Ltg1a avec les séquences les plus proches de la banque Venomics 123
Table 15. Constantes d'inhibition (Ki) moyennes des différentes formes produites de N-CTX-Ltg1a et de ses variants sur les récepteurs aux mélanocortines aboutissant à la Hunger Toxin (HT1-0) 125
Table 16. Relation structure-affinité de la Hunger-Toxin HT1-0 sur hMC4R..... 129
Table 17. Statistiques issues de la régression de Shild des courbes d'activation de différentes voies de signalisation de hMC4R par α-MSH et NDP-α-MSH en présence de concentrations croissantes de HT1-0 139
Table 18. Groupements protecteurs des chaines latérales des acides aminés en synthèse chimique en phase solide Fmoc 153

Liste des Annexes

Annexe 1. Article « A Venomics approach coupled to High-Throughput toxin production strategies for the identification of the first melanocortin toxins" soumis au *Journal of Chemical Chemistry* 196
Annexe 2. Informations supplémentaires de l'article « A Venomics approach coupled to High-Throughput toxin production strategies for the identification of the first melanocortin toxins" soumis au Journal of Chemical Chemistry 223
Annexe 3. Criblage de conditions d'oxydation de N-CTX-Can1a 231
Annexe 4. Rendements des différentes étapes de production des différentes formes de N-CTX-Can1a, N-CTX-Mil1a et N-CTX-Ltg1a produites en pontage dirigé 236
Annexe 5. Spectres de masse MALDI-TOF des différentes formes oxydées en pontage dirigé de N-CTX-Can1a..... 237
Annexe 6. Oxydation aléatoire de P014846 238
Annexe 7. LC-MS (ESI-IT) des formes oxydées «Bead », «Ribbon » et « Globular » de N-CTX-Mil1a..... 239

Table des matières

Remerciements 5
Liste des abréviations 7
Liste des Figures 11
Liste des Tables 13
Liste des Annexes 14
INTRODUCTION 19
Contexte 19
Objectifs et démarche 20
SYNTHÈSE BIBLIOGRAPHIQUE 23
1. Récepteurs couplés aux protéines G 24
1.1 Structure générale 24
1.2 Classification GRAFS 25
1.3 Cycle d'activation et désensibilisation 26
1.4 Signalisation indépendante des protéines G 30
1.5 Sélectivité fonctionnelle 30
2. Les récepteurs aux mélanocortines 31
2.1 Caractéristiques structurales 31
2.2 Une régulation pharmacologique fine 31
2.3 Diversité de fonctions physiologiques 33
3. L'homéostasie énergétique 34
3.1 La faim, cette sensation désagréable 34
3.2 Renforcement positif de la prise alimentaire : implication de l'hypothalamus latéral (LH) 39
3.3 Satiation 43
3.4 Satiété 48
3.5 Contrôle des dépenses énergétiques 50
4. MC4R comme cible thérapeutique : une signalisation complexe dans le système nerveux central 53
4.1 Passage de la barrière hématoencéphalique (BHE) 53
4.2 Activité constitutive 55
4.3 Protéines accessoires 55
4.4 Sélectivité fonctionnelle 56
5. État des lieux des principaux troubles de l'homéostasie énergétique 57
5.1 Obésité 57
5.2 Cachexie cancéreuse 62

6. Diversité et intérêt thérapeutique des toxines animales 66
6.1 Venin *versus* Poison 66
6.2 Histoire évolutive des venins 67
6.3 Convergence et diversité des motifs de cystéines 72
6.4 Applications thérapeutiques 83
PROJET VENOMICS 85
1. Stratégie 86
2. Partenaires 88
3. Résultats 89
3.1 A large venomous animal collection 89
3.2 A huge sequence toxin diversity in animal venoms 90
3.3 From sequences to a high throughput screening compatible toxin bank 95
3.4 A rich source of GPCR ligands 96
3.5 First animal toxins on melanocortin receptors 97
4. Discussion 99
5. Conclusion 100
RÉSULTATS et DISCUSSION 101
1. Choix des toxines à étudier 102
2. Les deux premières toxines actives sur les récepteurs aux mélanocortines 103
2.1 Contexte 103
2.2 Résultats 103
2.3 Discussion 108
2.4 Conclusion et précisions 109
3. Deux conotoxines difficiles à caractériser 110
3.1 N-CTX-Can1a 110
3.2 N-CTX-Mil1a 117
4. D'une conotoxine à un potentiel lead sur hMC4R 122
4.1 Originalité de séquence 122
4.2 D'un peptide réticulé à un peptide linéaire 123
4.3 Étude la liaison de HT1-0 sur les récepteurs aux mélanocortines 127
4.4 Fonction 136
4.5 Discussion 139
CONCLUSION et PERSPECTIVES 147
Des toxines originales en tout point 147
HT1-0 : un potentiel lead contre la cachexie cancéreuse ? 148
MATÉRIELS et MÉTHODES 151
1. Production des peptides 152

1.1 Matériel 152
1.2 Synthèse chimique en phase solide Fluorénylméthoxycarbonyle (Fmoc) 152
1.3 Analyse LC/MS du brut de synthèse 155
1.4 Purification 155
1.5 Formation des ponts disulfure 155
1.6 Évaluation de la qualité du peptide produit 156
1.7 Détermination de la concentration en peptide produit 156
2. Modifications chimiques du peptide 157
2.1 Alkylation des cystéines du peptide réduit 157
2.2 Marquage 157
3. Dichroïsme circulaire 160
3.1 Principe 160
3.2 Mode opératoire 160
4. Caractérisation pharmacologique 161
4.1 Tests de liaisons 161
4.2 Tests de fonction 164
5. Obtention de cellules exprimant le récepteur hMC4R marqué à la GFP 167
5.1 Amplification du plasmide codant pour hMC4R-GFP 167
5.2 Transfection de cellules CHO K1 167
6. Visualisation de la liaison de HT1-0 marqué à la rhodamine sur le récepteur hMC4R-GFP 168
BIBLIOGRAPHIE 169
ANNEXES 195

INTRODUCTION

"It is not the strongest of the species that survives, not the most intelligent that survives. It is the one that is the most adaptable to change" — Charles Darwin

Contexte

D'après Darwin, la survie d'une espèce dépend de la capacité des individus à s'adapter aux changements de son environnement. Que ce soit une bactérie ou une espèce plus complexe comme l'Homme, cette adaptation passe universellement par trouver une source d'énergie suffisante pour couvrir les dépenses énergétiques nécessaires à la vie de l'organisme. Ces dépenses sont liées bien sûr aux fonctions physiologiques de base mais également à des besoins supplémentaires coûteux comme la recherche de nourriture, la fuite des prédateurs et la reproduction. Le maintien de la balance entre ces dépenses et les apports énergétiques est un moteur important de l'évolution des espèces. Un dysfonctionnement chez un individu des systèmes de l'homéostasie énergétique lui confère directement un désavantage vis-à-vis des autres individus. Avec le développement des sociétés occidentales, l'espèce humaine est désormais confrontée à deux problèmes d'ordre alimentaire : d'un côté, une épidémie mondiale d'obésité liée à une disponibilité calorique excédentaire et de l'autre, des états de dénutrition soit liés aux pressions socioculturelles comme l'anorexie nerveuse soit à la recrudescence du nombre de cancers avec la cachexie cancéreuse. La cachexie cancéreuse est définie comme un syndrome métabolique complexe associé à la tumeur et aux traitements anti-cancéreux et caractérisé par une perte de la masse musculaire avec ou sans perte de masse graisseuse. Elle contribue ainsi fortement à l'augmentation de la morbidité et de la mortalité des patients cancéreux (Braun and Marks, 2010). Mieux comprendre les différents systèmes impliqués dans l'homéostasie énergétique est ainsi nécessaire pour pouvoir cibler et développer de nouveaux médicaments contre les troubles de l'homéostasie énergétique.

Le maintien de cette homéostasie implique la détection au niveau périphérique (tissu adipeux blanc, estomac, intestin) de signaux hormonaux ou mécaniques. Ces informations seront ensuite traitées par différentes régions du cerveau permettant l'initiation des sensations de faim ou de satiété. Chaque région présente des systèmes de régulation complexes influant sur l'activité des autres régions. Par exemple, le système limbique joue un rôle clé dans le caractère hédonique de la prise alimentaire et influe sur l'activité d'une autre région majeure, l'hypothalamus.

L'hypothalamus, situé au niveau du système nerveux central, présente différents noyaux dont les plus importants dans l'homéostasie énergétique sont le noyau arqué (ARC) et le noyau paraventriculaire (PVN). Ce derniet est le principal site d'expression du sous-type 4 des récepteurs aux mélanocortines (MC4R), un récepteur couplé aux protéines G (RCPG). Des mutations perte de fonction du MC4R entrainent une hyperphagie aboutissant à une obésité morbide. C'est la première cause d'obésité monogénique chez l'Homme et représente jusqu'à 6% des cas d'obésité précoce ou sévère (https://www.orpha.net). L'activité du récepteur est finement régulée à la fois par des agonistes et des antagonistes endogènes. Les agonistes, les mélanocortines, partagent tous le même pharmacophore HFRW responsable de l'effet anorexigène. L'Agouti-related peptide (AgRP) diminue l'activité du MC4R et a ainsi un effet orexigène. Des molécules développées à partir des ligands

endogènes pourraient donc être des candidats médicaments dans le traitement de l'obésité génétique avec des agonistes ou des états de dénutrition avec des antagonistes.

Les premiers efforts des industries pharmaceutiques se sont portés sur la découverte de nouveaux agonistes, le marché lié à l'obésité étant sans doute plus prometteur que celui de la cachexie ou de l'anorexie. Cela a ainsi permis de développer de nouveaux agonistes de nature peptidique ou des petites molécules plus sélectifs que le ligand endogène de départ. Néanmoins, une seule molécule, le Setmélanotide®, est actuellement en développement clinique pour des pathologies associées à l'obésité comme la déficience en récepteurs de la principale hormone de la satiété, la leptine. Du côté des antagonistes, les travaux se concentrent essentiellement sur l'ingénierie de l'AgRP pour accroitre son efficacité et sa sélectivité. Néanmoins, aucune molécule antagoniste du MC4R n'est pour l'instant en phase clinique.

Les résultats des travaux sur les ligands endogènes se sont ainsi avérés décevants. L'inconvénient de ne travailler que sur les ligands endogènes est la pauvreté structurale des outils développés. Il est donc nécessaire de découvrir de nouvelles structures actives sur les récepteurs aux mélanocortines (MCRs) afin de trouver de nouvelles stratégies de modulation du MC4R. Une solution est d'explorer une source connue pour sa diversité structurale et son potentiel thérapeutique, à savoir les venins.

Néanmoins, l'essentiel de la biodiversité venimeuse est représenté par les Arthropodes terrestres, des organismes de seulement quelques centimètres. Or, les méthodes classiques d'étude par bioguidage[1] requièrent de grandes quantités de venins, chose très difficile à obtenir de petites espèces dont la glande venimeuse fait seulement quelques millimètres. L'exploration de cette ressource naturelle est donc extrêment complexe et incomplète. C'est dans ce contexte qu'a été mené de 2011 à 2015 un grand projet européen, le projet Venomics, coordonné par Nicolas Gilles du Commissariat à l'Énergie Atomique et aux Énergies Alternatives (CEA). L'objectif était de construire la première banque de toxines animales synthétiques retranscrivant au mieux la diversité moléculaire des venins. Cette banque devait pouvoir être facilement criblée sur des cibles thérapeutiques. Pour cela, des méthodes omiques ont permis de séquencer les toxines produites par près de 200 espèce venimeuses. Ensuite, l'utilisation de stratégies de production optimisées a abouti à la production de la banque Venomics. Cette dernière a enfin été criblée sur différents récepteurs couplés aux protéines G impliqués dans diverses pathologies sans traitement actuel efficace. L'une de ces cibles était le sous-type 4 des récepteurs aux mélanocortines. C'est dans ce contexte que je suis arrivé dans le laboratoire « Toxines, récepteurs et canaux ioniques » du CEA de Saclay d'abord en tant que stagiaire ingénieur puis en tant que doctorant.

Objectifs et démarche

Le premier objectif du projet de thèse était d'analyser d'un point de vue statistique les différentes données générées par le projet Venomics en vue d'une future publication.

Le deuxième, purement moléculaire, visait en la caractérisation pharmacologique des premières toxines animales actives sur les récepteurs aux mélanocortines. Pour cela, la première étape est de sélectionner parmi les hits issus du criblage primaire, les toxines à étudier principalement sur la base de leur originalité structurale. Une fois la sélection faite,

[1] La stratégie de bioguidage consiste en des cycles de fractionnement et de test sur une cible moléculaire des différentes fractions générées jusqu'à la purification finale de la molécule responsable de l'activité observée.

les toxines doivent être caractérisées d'un point de vue pharmacologique avec l'étude de leur affinité, leur fonction et leur sélectivité.

Enfin, le dernier objectif était à plus long terme avec une visée thérapeutique via l'obtention d'un potentiel lead dans le traitement des troubles de la prise alimentaire. Afin de maximiser ses chances de poursuite dans le processus de développement thérapeutique, le lead doit être affin et sélectif de sa cible moléculaire mais aussi facile à produire. L'idée est ainsi de partir du meilleur hit issu du criblage, c'est-à-dire celui présentant la meilleure affinité pour hMC4R. Afin de convertir cette molécule en potentiel lead, le premier travail a d'abord consisté en l'amélioration de la production du peptide. Les toxines étant des peptides riches en ponts disulfure, leur production peut s'avérer complexe. Ainsi, dans un premier temps, l'importance de ces ponts dans l'activité de la toxine sur les récepteurs MCRs est étudiée. Une fois, la production optimisée, la seconde tache est d'étudier de façon approfondie la liaison de la forme simplifiée sur hMC4R. Pour cela, des études de relation structure-activité (SAR) du peptide sur le récepteur sont menées. Il s'agit principalement d'observer l'effet d'une variation de la séquence en acides aminés sur l'affinité du peptide. Ces informations permettent de dégager des régions pouvant être modifiées sans altérer la liaison du peptide sur hMC4R. Pour étudier les paramètres cinétiques du peptide, le type de modification à faire est son marquage radioactif et/ou fluorescent. Ces stratégies de marquage présentant chacune des avantages et des inconvénients, elles sont toutes les deux explorées dans mon projet de thèse. Enfin, il est important d'identifier l'application thérapeutique potentielle de la molécule (obésité ou cachexie cancéreuse). La fonction du peptide sur différentes voies de signalisation du MC4R est donc étudiée *in vitro*. S'il s'agit d'un agoniste, le peptide ciblera l'obésité. Dans le cas d'un antagoniste, l'application sera la cachexie cancéreuse.

SYNTHÈSE BIBLIOGRAPHIQUE

1. Récepteurs couplés aux protéines G

Les RCPGs sont retrouvés dans presque tous les organismes Eucaryotes. Chez les mammifères, ils forment l'une des plus grandes et des plus diversifiées familles de récepteurs (Fredriksson, 2005). Jusqu'à présent, plus de 800 gènes codant pour des RCPGs (~400 récepteurs olfactifs) ont été identifiés dans le génome humain (Hanlon and Andrew, 2015). Ils permettent à la cellule de détecter et de répondre à un stimulus extracellulaire par la liaison d'un ligand activateur, l'agoniste. La nature chimique de ce dernier est extrêmement variée : peptide, protéine, petite molécule, ion, photon, lipide (Hauser et al., 2017). La transduction du signal se fait ensuite par une cascade de signalisations par l'intermédiaire ou non d'une protéine G hétérotrimérique. Régulant ainsi des processus physiologiques très variés, ces récepteurs sont impliqués dans de nombreuses maladies cardiovasculaires, métaboliques, neurodégénératives, psychiatriques, infectieuses et cancéreuses (Duc et al., 2015). Pour cette raison, ils font l'objet d'une attention particulière de la part des industries pharmaceutiques. En 2017, 475 médicaments, soit 34% de ceux approuvés par la Food and Drug Administration (FDA), ciblent 108 RCPGs non olfactifs. En outre, 321 autres molécules sont en phase clinique dont 20% ciblent de nouveaux RCPGs (Hauser et al., 2017). Malgré ces avancées, certains récepteurs restent sans ligand exogène disponible sans compter la centaine de RCPG orphelins c'est-à-dire pour lesquels aucun ligand endogène n'est connu (Jacobson, 2015).

1.1 Structure générale

Un RCPG possède typiquement 7 hélices α transmembranaires (TM1 à TM7) composées de 25 à 35 résidus avec une extrémité N-terminale extracellulaire et une extrémité C-terminale cytoplasmique (**Figure 1**) (Nikiforovich et al., 2007). L'extrémité C-terminale peut être palmitoylée après la courte hélice 8 à la sortie du TM7. Cette PTM est un ancrage lipidique dans la membrane à l'origine d'une quatrième boucle intracellulaire ICL4 (Palczewski et al., 2000). Les 7 TMs sont reliées par 3 boucles extracellulaires (ECL1, ECL2, ECL3) et 3 boucles intracellulaires (ICL1, ICL2, ICL3). ECL1 ou la fin extracellulaire de TM3 et ECL2 présentent chacune une cystéine pouvant former un pont disulfure entre les deux boucles.

Plusieurs classifications des RCPGs existent, la longueur des extrémités N-terminale et C-terminale constituant un critère majeur. La plus utilisée est le système A, B, C, D, E et F (Attwood and Findlay, 1994). Ce système couvre l'ensemble des RCPGs des vertébrés et des invertébrés. Néanmoins, certaines classes n'existent pas chez l'Homme, à l'exemple de la classe D incluant des récepteurs aux phéromones d'insectes et de levures. Un autre système basé sur une analyse phylogénétique n'intégrant que les RCPGs trouvés dans le génome humain a été proposée en 2003 : classification GRAFS pour Glutamate (G, 15 membres), Rhodopsin (R, 701), Adhésion (A,24), Frizzled (F,24), Sécrétine (S,15) (Fredriksson et al., 2003). C'est cette dernière qui sera utilisée dans cette partie. Une correspondance avec le système A-F sera également faite.

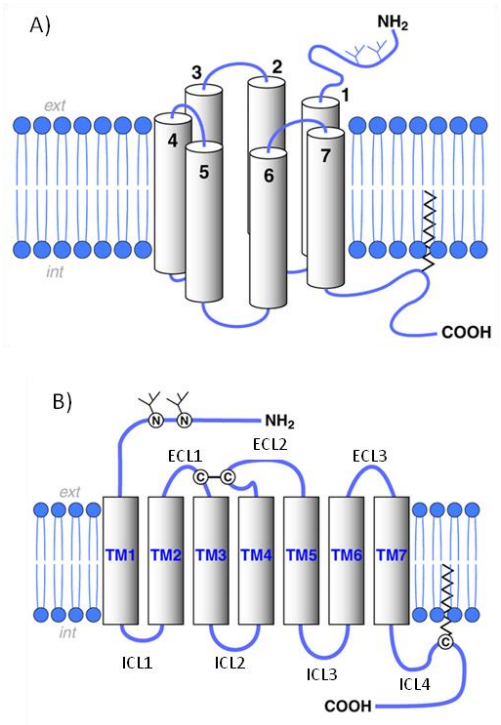


Figure 1. Structure schématique générale d'un récepteur couplé aux protéines G (RCPG) en 3D (A) et en 2D (B). Le récepteur possède 7 hélices α transmembranaires (TM) reliées par trois boucles extracellulaires (ECL) et trois boucles intracellulaires (ICL). ECL1 ou la fin extracellulaire de TM3 et ECL2 sont reliées par un pont disulfure. L'extrémité amino-terminale est extracellulaire tandis que l'extrémité carboxy-terminale est cytoplasmique. Un ancrage lipidique de cette dernière peut être observé et former une quatrième boucle intracellulaire ICL4.

1.2 Classification GRAFS

- **Récepteurs Glutamate (classe C)**

Cette famille regroupe notamment des récepteurs métabotropiques au glutamate et des récepteurs GABA, un récepteur sensible au calcium et des récepteurs du goût. Le ligand des récepteurs métabotropiques, le glutamate, se fixe au niveau d'un domaine de 280 à 580 résidus de l'extrémité N-terminale. Cette extrémité formerait deux lobes séparés par une cavité dans laquelle se fixerait le glutamate. La fixation de ce dernier entrainerait alors la fermeture des lobes autour du ligand comme une plante carnivore d'où le nom en anglais de « Venus fly trap » (Hanlon and Andrew, 2015). La dimérisation du récepteur est également un élément clé pour son activation (Pin et al., 2003).

- **Récepteurs Rhodopsine (classe A)**

C'est la plus grande famille de RCPGs. Hormis quelques exceptions, tous les membres de cette famille partagent certaines caractéristiques : motif NSxxNPxxY du TM7 et motif E/DRY à la limite entre TM3 et ICL2. Le ligand se fixe généralement au niveau des TMs mais la liaison peut également se faire au niveau de l'extrémité N-terminale comme c'est le cas pour les récepteurs aux glycoprotéines (LH, FSH, TSH et LG) (Fredriksson et al., 2003).

Les récepteurs Rhodopsines sont subdivisées en 4 groupes :

- Groupe α

Le groupe α rassemble des récepteurs aux prostaglandines, aux amines, aux opsines, à la mélatonine et les récepteurs MECA. Les récepteurs aux amines sont les plus représentés avec notamment les récepteurs à la sérotonine, à la dopamine, à l'histidine et aux adrénalines. Les récepteurs aux opsines sont connus pour répondre uniquement à un stimulus lumineux. Enfin, les récepteurs MECA comportent les récepteurs aux mélanocortines (MCRs), les RCPGs de la différenciation endothéliale, les récepteurs aux cannabinoïdes et les récepteurs à l'adénosine. Ces récepteurs lient des ligands aux structures

très différentes : par exemple des peptides pour les MCRs ou des lipides pour les RCPGs de la différenciation endothéliale.

- Groupe β

Tous les ligands connus de ce groupe sont de nature peptidique. Il regroupe notamment les récepteurs à la cholécystokinine (CCK) et au neuropeptide Y (NPY).

- Groupe γ

Sont présents dans ce groupe d'une part, les récepteurs à la somatostatine et ceux aux opioïdes, les récepteurs à l'hormone concentrant la mélanine (MCH) ainsi que les récepteurs aux chimiokines et les récepteurs à l'angiotensine ATR1 et ATR2.

- Groupe δ

Ce groupe comporte quatre sous-groupes : les récepteurs MAS liant l'angiotensin 1-7, les récepteurs aux glycoprotéines, les récepteurs aux purines, et les récepteurs olfactifs estimés à 460. Le ligand peut ainsi être de différente nature (glycoprotéique ou nucléotidique par exemples ou être une molécule volatile en ce concerne les récepteurs olfactifs). La majorité des récepteurs olfactifs sont orphelins.

- **Récepteurs Adhésion (classe B)**

Les récepteurs de cette famille contiennent souvent les domaines cadherine ou intégrine avec une extrémité N-terminale plus longue que la famille Sécrétine. Les ligands reconnus sont notamment des composants de la matrice extracellulaire comme le collagène (Luo et al., 2011).

- **Récepteurs Frizzled/Taste2 (classe F)**

Certaines caractéristiques sont propres à cette famille : motif IFL au niveau de TM2, motif SFLL au niveau de TM5 et SxKTL au niveau de TM7 (Fredriksson et al., 2003). Les récepteurs Frizzled jouent un rôle majeur dans le développement embryonnaire en contrôlant par exemple la prolifération et la différenciation cellulaire ainsi que l'apoptose via la reconnaissance de la glycoprotéine Wnt (Fredriksson et al., 2003; Wang et al., 2006). Ils sont au nombre de 10 chez l'Homme. Les récepteurs Taste2 ont quant à eux la particularité d'avoir une extrémité N-terminale trop courte pour lier un ligand. À ce jour, ils restent majoritairement orphelins. Ils jouent cependant un rôle dans la perception du goût.

- **Récepteurs Sécrétine (classe B)**

Le domaine N-terminal extracellulaire plus grand que dans la famille Rhodopsine est crucial pour la reconnaissance et la liaison des ligands. Ces derniers sont typiquement des peptides et des hormones comme la sécrétine, le glucagon, la calcitonine ou encore l'hormone parathyroïde.

1.3 Cycle d'activation et désensibilisation

L'activation du récepteur par son agoniste endogène passe par différentes étapes au cours desquelles les changements conformationnels successifs permettent la transduction du signal perçu par la cellule (**Figure 2**). Afin de décrire les positions importantes dans ce processus, la nomenclature de Ballesteros et Weinstein sera utilisée (Ballesteros and Weinstein, 1995). Selon celle-ci, deux numéros sont attribués à un résidu. Le premier correspond au TM auquel il appartient. Le second correspond à la position du résidu par rapport au résidu le plus conservé auquel on attribue le numéro 50.

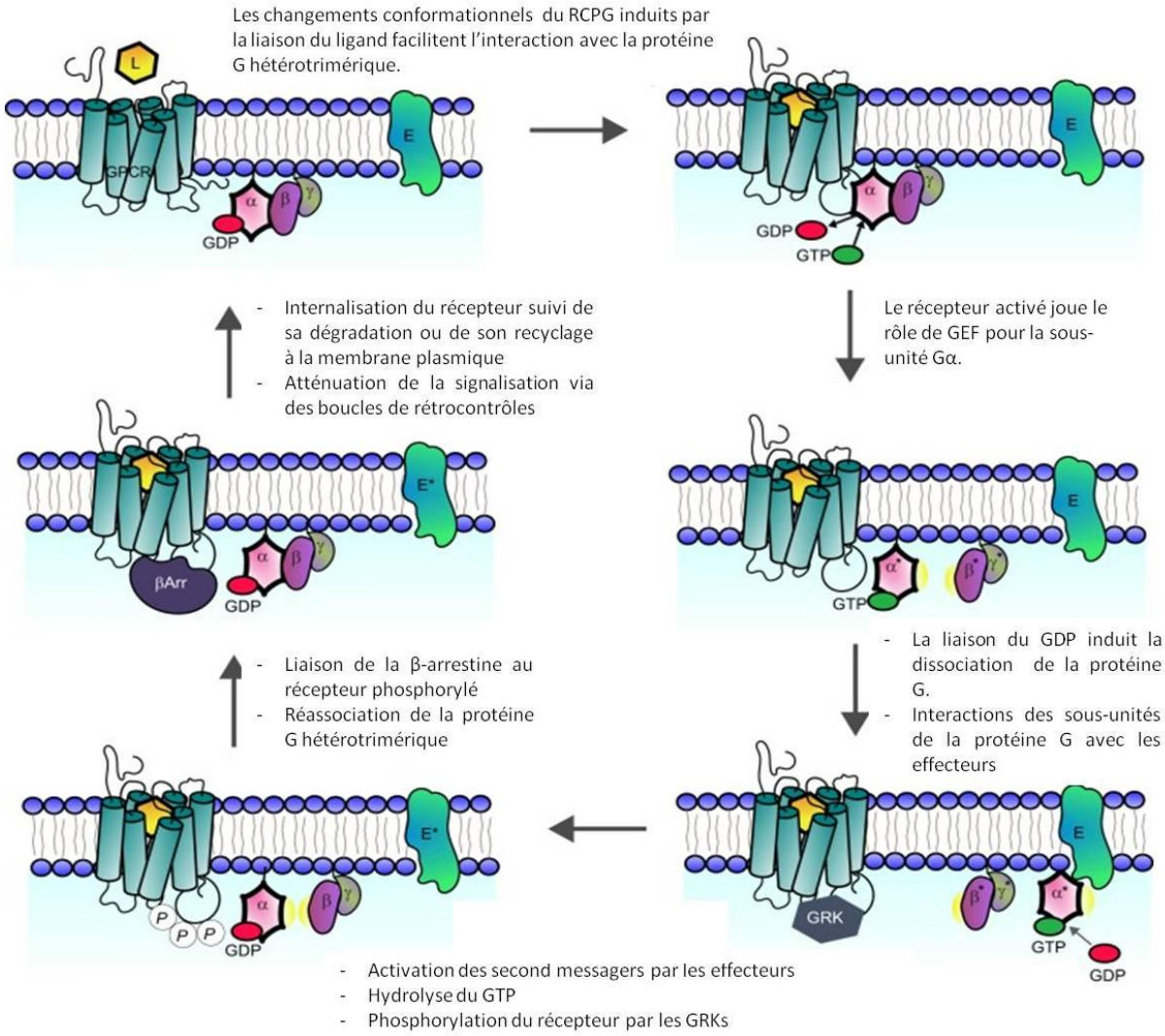


Figure 2. Cycle d'activation et de désensibilisation d'un RCPG. La liaison d'un agoniste L sur le récepteur entraine l'activation de celui-ci via des changements conformationnels impliquant les domaines transmembranaires. La protéine G formée de ses trois sous-unités est alors recrutée. Ce recrutement induit l'échange de GDP en GTP au niveau de la sous-unité Gα et la dissociation de l'hétérotrimère. Les différentes sous-unités peuvent activer différents effecteurs E et par ce biais la production de second messager à l'origine de différentes cascades de signalisation. La désensibilisation du récepteur passe par la phosphorylation de résidus Ser et Thr en parallèle de la réassiociation de l'hétérotrimère via l'hydrolyse de GTP en GDP au niveau de Gα. La phosphorylation du récepteur permet son internalisation par endocytose via un manteau de clathrine. Suivant le motif de phosphorylation, différentes conformations de β-arrestine (βArr) peuvent se fixer déterminant le devenir du récepteur, à savoir son recyclage à la membrane ou son ubiquitination et sa dégradation. GEF : Facteur d'échange nucléotidique (*Guanine-nucleotide-Exchange Factor*). (Adapté de Hanlon and Andrew, 2015)

- **Etat inactivé**

Avant la liaison d'un agoniste, le récepteur est stabilisé à l'état inactivé par différentes interactions. Dans la famille Rhodopsine, le résidu basique Arg$^{3.50}$ du motif hautement conservé E/DRY forme un pont salin avec le résidu acide Asp ou Glu$^{3.49}$. En outre, le pont disulfure formé entre une cystéine d'ECL2 et Cys$^{3.25}$ au-dessus du TM3 participe

également à la stabilité de l'ensemble.

- **Liaison de l'agoniste**

L'agoniste endogène d'un récepteur se lie au niveau d'une zone propre au couple récepteur-agoniste, autrement appelée le site orthostérique. Ce dernier implique les TMs et parfois l'extrémité N-terminale. À la suite de la fixation du ligand, plusieurs changements conformationnels permettent la transition d'un état inactivé à un état activé. Les réarrangements s'effectuent principalement au niveau des hélices transmembranaires. De concert avec TM5, la partie intracellulaire de TM6 subit une rotation et s'écarte du reste des TMs créant une ouverture bordée de TM3, TM5 et TM7 (Latorraca et al., 2017). Ces mouvements sont accompagnés et stabilisés par des microarrangements comme au niveau de $Tyr^{7.53}$du motif NPxxY. La chaine latérale de cette position pointe vers TM1, TM2 et TM7 à l'état inactivé. À la suite de la liaison de l'agoniste, elle change de conformation pour pointer vers le cœur des TMs formant des interactions avec les chaines latérales de différents résidus de TM3 et TM6 (Katritch et al., 2013). En ce qui concerne la zone de fixation de l'agoniste, les modifications structurales sont variées témoignant de la grande diversité de ligands reconnus par les RCPGs (Manglik and Kruse, 2017).

La conséquence de l'ensemble des mouvements décrits ci-dessus est l'exposition de résidus-clés dans la reconnaissance et la liaison de la protéine G.

- **Recrutement de la protéine G**

Une protéine G est un complexe hétérotrimérique formé des sous-unités Gα, Gβ et Gγ. La sous-unité Gα comporte deux domaines, le domaine Ras-like avec un feuillet β et cinq hélices α et un domaine en hélices α. Le domaine Ras-like possède une activité GTPase capable d'hydrolyser la guanosine 5'-triphosphate (GTP) en guanosine 5'-diphosphate (GDP). Le site de liaison du nucléotide est situé entre les deux domaines. En outre, la palmitoylation de l'extrémité N-terminale permet l'ancrage de la sous-unité dans la membrane (Duc et al., 2015). La sous-unité Gβ présente quant à elle une hélice α N-terminale et sept feuillets β dont quatre brins antiparallèles. Elle forme un dimère en superhélice (coilled coil) avec la sous-unité Gγ. Tout comme la sous-unité Gα, le complexe Gβγ est également ancré à la membrane via la cystéine du motif CAAx porté par la sous-unité Gγ (Milligan and Kostenis, 2006). Le complexe se lie à la sous-unité Gα via le domaine Ras-like de cette dernière. L'hétérotrimère ainsi formé peut se lier à un récepteur à l'état inactif ou dériver dans la membrane. Cependant si le récepteur est activé, celui-ci joue alors le rôle de facteur d'échange nucléotidique (GEF) catalysant l'échange du GDP lié à la sous-unité Gα inactive en GTP. Cet état récepteur activé en interaction avec une protéine G est très transitoire. En effet, le GTP est en large excès dans le cytoplasme par rapport au GDP. La liaison du GTP est ainsi extrêmement rapide. Par la suite, cet échange engendre à son tour des changements conformationnels de Gα à l'origine de sa dissociation du complexe Gβγ. La sous-unité Gα ainsi activée peut interagir avec un effecteur et initier toute une cascade de signalisation pouvant faire intervenir un second messager. De nombreuses protéines peuvent jouer le rôle d'effecteur comme des enzymes, des phospholipases, des canaux ioniques, des protéines d'adhésion,…

- **Signalisation**

On dénombre chez l'Homme 21 sous-unités Gα différentes, 6 Gβ et 12 Gγ. Les différentes combinaisons possibles concourent à une grande diversité de réponses intracellulaires. Cependant on distingue quatre grandes classes de protéine G suivant la similarité des sous-unités Gα. Ces dernières sont classiquement divisées en quatre groupes associés à un effecteur particulier. Les deux premiers groupes de Gα composés d'une part de $G\alpha_s$ et d'autre part de $G\alpha_i$ et $G\alpha_o$ agissent tous deux sur l'adénlyate cyclase (AC) (Baltoumas et al., 2013). Cette enzyme membranaire joue le rôle d'effecteur en convertissant l'adénosine triphosphate (ATP) intracellulaire en adénosine monophosphate cyclique (AMPc) (Milligan and Kostenis, 2006). Une élévation des niveaux intracellulaires en AMPc permettra l'activation de la protéine kinase A (PKA) et la cascade d'activation associée. $G\alpha_s$ active AC tandis que les deux autres inhibent son activité. Le troisième groupe avec $G\alpha_{q/11}$ active la phospholipase C (PLC). Cette dernière clive le phosphatidylinositol 4,5-bisphosphate (PIP_2) membranaire en inositol triphosphate (IP_3) et en diacylglycérol (DAG). L'IP_3 reconnait un récepteur spécifique du réticulum endosplasmique permettant l'activation de canaux calciques et le relargage de calcium Ca^{2+} dans le cytosol. Les fortes concentrations cytoplasmiques en Ca^{2+} concourent avec le DAG à l'activation de la protéine kinase C (PKC) mais aussi à la formation du complexe calcium-calmoduline. Celui-ci peut entrainer l'activation d'autres enzymes comme des phosphodiestérases ou la phospholipase A2. Le dernier groupe, $G\alpha_{12/13}$, stimule les GEFs de Rho activant à leur tour la protéine Rho.

Le complexe Gβγ est également capable d'activer différents effecteurs y compris certains ciblés par la sous-unité Gα comme PLC, AC ou la voie des Mitogen-Activated Protein (MAP) Kinases par des mécanismes faisant intervenir Src (Dupré et al., 2009; Khan et al., 2013). En outre, il peut être à l'origine de l'ouverture de canaux ioniques comme les canaux calciques Ca_v1, 2 et 3 ou le canal potassique Kir3 du cœur.

- **Désensibilisation**

L'inactivation de la sous-unité Gα et la reformation du complexe hétérotrimérique nécessite l'hydrolyse du GTP en GDP. L'activité GTPase intrinsèque de la sous-unité Gα étant faible, la réaction d'hydrolyse est accélérée par les molécules régulatrices de la signalisation des protéines G (RGS) (Hanlon and Andrew, 2015). En parallèle, des protéines kinases spécifiques des RCPGs (GRKs) phosphorylent des résidus Ser ou Thr de l'extrémité C-terminale ou d'ICL3 du récepteur. Cette phosphorylation permet la liaison de la β-arrestine et ainsi la désensibilisation du récepteur. Cette dernière passe par activation de phosphodiestéases et de diacylglycérol kinases à l'origine de la dégradation des second messagers et par le recrutement de la clathrine et de son adaptateur AP-2 aboutissant à l'internalisation du complexe récepteur β-arrestine dans les endosomes (Baillie et al., 2003). Le récepteur est ensuite soit dégradé dans les lysosomes à la suite de l'ubiquitination du récepteur soit recyclé à la membrane plasmique. Un nouveau cycle d'activation/inactivation peut alors débuter après déphosphorylation du récepteur.

Cependant, la phosphorylation du récepteur par les GRKs peut engendrer d'autres voies de signalisation indépendantes des protéines G et impliquant différentes conformations des β-arrestines.

1.4 Signalisation indépendante des protéines G

Suivant le motif de phosphorylation généré par les GRKs, la β-arrestine liée au récepteur peut montrer une diversité de conformations à l'origine de ses différentes fonctions : désensibilisation, internalisation et transduction du signal. Il a par exemple été montré que cette liaison pouvait activer la voie des MAP Kinases via la signalisation des kinases régulées par un signal extracellulaire (ERK) de manière totalement indépendante des protéines G pour certains RCPGs (Lotta et al., 2019). Dans ce cas, l'activation d'ERK est soutenue et localisée dans les endosomes alors que celle médiée par les protéines G est transitoire, localisée dans le noyau et promeut la transcription de certains gènes (Eichel et al., 2016). Un autre exemple est la voie c-Jun N-terminal kinases 3 (JNK3) importante pour le cycle cellulaire (Jean-Charles et al., 2017).

La transduction du signal par les RCPGs est ainsi bimodale avec une composante dépendante des protéines G et une autre indépendante des protéines G mais médiée par les β-arrestines nécessitant la phosphorylation par les GRKs. En outre, il est maintenant admis qu'un récepteur donné peut activer différents types de sous-unités Gα. Une ou plusieurs de ces voies peuvent être préférentiellement activées suivant le ligand lié. On parle alors de signalisation biaisée ou de sélectivité fonctionnelle.

1.5 Sélectivité fonctionnelle

Différentes explications de l'existence de la sélectivité fonctionnelle ont été proposées. Un ligand biaisé pourrait entraîner des changements conformationnels différents de ceux induits par un ligand non biaisé appelé alors agoniste équilibré. Ces changements seraient à l'origine d'un ensemble d'états conformationnels du récepteur capables de reconnaitre et de lier différents partenaires intracellulaires et induire différentes voies de signalisation (Strachan et al., 2014). Outre des changements conformationnels différentiels du récepteur, un ligand biaisé pourrait induire différents motifs de phosphorylations du récepteur par les GRKs, déterminant comme on l'a vu précédemment le devenir du récepteur et les voies activées par la suite par la β-arrestine. Par ces mécanismes, la liaison d'un ligand biaisé peut favoriser certains évènements et en atténuer d'autres aboutissants à une voie de signalisation préférentielle. Les cinétiques de signalisation peuvent également être modifiées, amenant à un biais apparent dans le temps (Klein Herenbrink et al., 2016). Sachant que chaque voie de signalisation d'un récepteur peut être associée à une fonction différente, la découverte de ligand biaisé pourrait être à l'origine de médicaments plus efficaces.

En résumé, la régulation de l'activité d'un RCPG est beaucoup plus complexe que l'équation : un RCPG + un ligand = une voie de signalisation / un effet physiologique. Cette complexité permet une régulation fine de l'activité du récepteur et de la fonction physiologique associée. Ces fonctions peuvent être diverses au sein d'une même famille de récepteur. C'est le cas de la famille des récepteurs aux mélanocortines (MCRs).

2. Les récepteurs aux mélanocortines

2.1 Caractéristiques structurales

Au total, cinq sous-types de récepteurs aux mélanocortines humains ont été découverts à ce jour. Ils partagent entre 39% et 56% d'identité de séquence (Fredriksson et al., 2003). Ensemble, ils forment en longueur les plus petits RCPGs de la classe Rhodopsine. Ils possèdent ainsi plusieurs caractéristiques de cette classe avec notamment des sites potentiels de N-glycosylation près de l'extrémité N-terminale et un site supposé de palmitoylation au niveau de l'extrémité C-terminale bien que les sites exacts de ces modifications n'aient pas encore été identifiés. Plusieurs sites de phosphorylations de l'extrémité C-terminale sont également présents (Kumar et al., 2009). Tous les MCRs possèdent le motif DRY à la jonction du TM3. Cependant, certaines caractéristiques sont propres à cette famille. Par exemple, le pont disulfure entre le haut de TM3 et EL2 est manquant mais un autre pont est formé au niveau d'EL3 (Tao, 2010). Les boucles intracellulaires et extracellulaires sont particulièrement courtes comparées aux autres récepteurs de la classe Rhodopsine. En outre, $N^{7.49}$ du motif NPxxY et le résidu $P^{5.50}$ hautement conservés à travers la classe Rhodopsine sont substitués respectivement par un aspartate (D) et une méthionine (M). Une autre particularité est leur régulation à la fois par des agonistes et des antagonistes endogènes.

2.2 Une régulation pharmacologique fine

- **Des agonistes et des antagonistes endogènes**

D'un point de vue pharmacologique, les MCRs, à l'exception de MC2R, peuvent tous être reconnus par les mêmes ligands bien que les affinités diffèrent suivant le sous-type de récepteur considéré (**Table 1**). Les agonistes, les mélanocortines, sont produits essentiellement au niveau central dans l'hypophyse antérieure (Padilla et al., 2010), le noyau arqué de l'hypothalamus (ARC) et le noyau du tractus solitaire (NTS) mai aussi en périphérie dans les mélanocytes (Chakraborty et al., 1996). Elles sont toutes issues de clivages successifs du précurseur pro-opiomélanocortine (POMC) (**Figure 3**). Historiquement, elles ont été décrites comme activant principalement la voie G_s entrainant ainsi une augmentation des niveaux intracellulaires d'AMPc (Gantz and Fong, 2003). Cependant, des études plus récentes montrent que les MCRs peuvent être aussi couplés à d'autres protéines G comme $G_{q/11}$, $G_{i/o}$ ou couplés à la β-arrestine (Büch et al., 2009; Chai et al., 2007; Rodrigues et al., 2012). Comme vu précédemment, les mélanocortines sont issues du clivage de POMC. Elles partagent toutes le même pharmacophore HFRW, séquence minimale pour activer le récepteur (Hruby et al., 1987; Singh et al., 2010). Cette séquence montre une structure en β-turn (Haslach et al., 2009). En ce qui concerne les antagonistes endogènes, chez la souris, ils correspondent aux protéines Agouti et Agouti-related peptide (AgRP). La protéine Agouti est exprimée majoritairement au niveau de la peau. De par son affinité pour différents MCRs, des mutations dominantes du gène causent des phénotypes pléiotropiques chez les souris comme l'obésité, la résistance à l'insuline, le pelage jaune et une prédisposition pour les tumeurs (McNulty et al., 2005). L'équivalent chez les humains est la protéine ASIP (Agouti-signaling protein) qui est beaucoup moins bien caractérisée que son équivalente murine. Néanmoins, elle jouerait également un rôle dans la pigmentation. Enfin, le peptide AgRP est

surtout exprimé dans les glandes surrénales et l'hypothalamus (Haskell-Luevano and Monck, 2001) où il est principalement sécrété sous la forme clivée AgRP(83-132) (Creemers et al., 2006). Il diminuerait le nombre de récepteur exprimé à la surface des cellules potentiellement par la voie β-arrestine (Breit et al., 2006). Et de par son activité agoniste inverse et/ou son activité d'agoniste biaisé G_i, il diminuerait les niveaux intracellulaires d'AMPc (Büch et al., 2009; Haskell-Luevano and Monck, 2001). Les protéines Agouti et AgRP possèdent toutes les deux un domaine C-terminal avec cinq ponts disulfure intégrant un motif Inhibitor Cystine Knot (ICK). Ce domaine expliquerait à lui seul l'effet antagoniste des deux molécules avec un pharmacophore formé du tripeptide Arg-Phe-Phe (Tota et al., 1999).

Table 1. Vue d'ensemble des différents sous-types des récepteurs aux mélanocortines. ACTH : hormone corticotrope ; AgRP : Agouti-related peptide ; MSH : hormone mélanotrope ; SNC : système nerveux central ; SNP : système nerveux périphérique. (Adapté de Jackson et al., 2015)

Sous-type	Ligand endogène (par affinité décroissante)	Antagoniste/agoniste inverse endogène	Principal site d'expression	Fonction principale
MC1R	α-MSH = β-MSH = ACTH > γ-MSH	Agouti	Mélanocytes	Pigmentation, inflammation
MC2R	ACTH	Agouti	Cortex des glandes surrénales	Production des glucocorticoïdes
MC3R	γ-MSH = α-MSH = β-MSH = ACTH	Agouti, AgRP	SNC (hypothalamus), tractus gastro-intestinal, reins	Homéostasie, énergétique, inflammation, anticipation de la prise alimentaire
MC4R	α-MSH = β-MSH = ACTH > γ-MSH	AgRP, Agouti	SNC (hypothalamus)	Homéostasie, énergétique, thermogenèse, contrôle de l'appétit, fonction érectile
MC5R	α-MSH = β-MSH = ACTH > γ-MSH	Agouti	Glandes exocrines	Fonction exocrine, régulation de l'activité des glandes sébacées, lipolyse

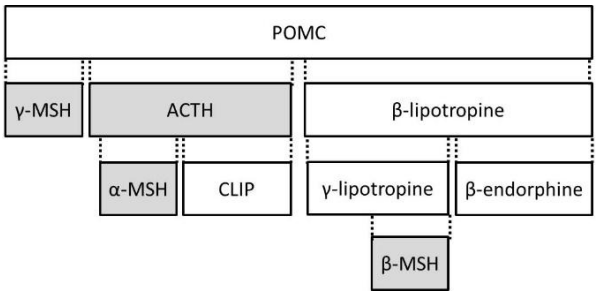


Figure 3. Génération des différentes mélanocortines à partir du clivage de POMC. Le gène POMC code pour un précurseur de 32 kDa avec trois régions : la séquence ACTH (positions 138-176) à l'origine de l'α-MSH (13 aa) , la séquence de la lipotropine β-LPH (179-267) en C terminal clivée pour donner la γ-LPH (55 aa), la bêta-endorphine (30 aa), et la séquence en N terminal contenant celle de la γ-MSH (10 aa). (Adapté de Catania et al., 2004) ACTH : hormone corticotrope ; CLIP : Corticotropin-like Intermediate Peptide ; MSH : hormone mélanotrope ; NPY : neuropeptide Y ; POMC : pro-opiomélanocortine.

- **Rôle des protéines accessoires des recepteurs aux mélanocortines (MRAPs)**

Les MRAPs sont des protéines transmembranaires qui interagissent et régulent la signalisation et le transport de différents RCPGs dont les MCRs. On distingue deux membres de cette famille : MRAP1 et MRAP2. Les deux interagissent *in vitro* avec l'ensemble des MCRs (Chan et al., 2009). Cependant le rôle et la relevance de ces interactions *in vivo* restent encore à déterminer. MRAP1 est connue pour s'exprimer majoritairement au niveau de la glande surrénale où elle joue un rôle essentiel dans la liaison des ligands, la signalisation et le transport à la membrane de MC2R (Metherell et al., 2005). MRAP2 est quant à elle principalement exprimée au niveau du SNC où elle joue un rôle dans le maintien de l'homéostasie énergétique en modulant notamment le transport et l'activité de MC4R (Asai et al., 2013).

L'action conjointe des agonistes et des antagonistes endogènes associés aux MRAPs permet une régulation fine des différentes fonctions physiologiques des MCRs.

2.3 Diversité de fonctions physiologiques

Majoritairement retrouvé au niveau de la peau et plus partculièrement dans les mélanocytes, MC1R joue un rôle majeur dans la prévention du mélanome. Son activation permet la synthèse d'eumélanine (pigment noir) aux dépens de celle de phéomélanine (pigment jaune) protègeant ainsi les mélanocytes de dommages causés par les rayons ultraviolets (UVs) (Kobayashi et al., 1998). De plus, il permet la réparation par excision de nucléotides (NER) et élimine de ce fait de l'ADN génomique les photoproduits induits par les rayons UVs comme les nucléotides endommagés (Kadekaro et al., 2010).

MC2R joue un rôle dans le contrôle de la stéroïdogenèse par les glandes surrénales (Chida et al., 2007). Contrairement aux autres MCRs, il n'est activé que par l'homone corticotrope (ACTH).

MC5R est exprimé dans de nombreux organes et tissus avec un rôle dans la sécrétion au niveau des glandes sébacées et glandes exocrines (Chen et al., 1997). Dans les cellules du muscle squelettique, MC5R permettrait la régulation de l'oxydation des acides gras favorisant ainsi la thermogenèse (An et al., 2007; Enriori et al., 2016). Associé à cette fonction, il favoriserait également la lipolyse dans les adipocytes (Rodrigues et al., 2013). De ce point de vue, MC5R jouerait ainsi un rôle dans l'homéostasie énergétique. Néanmoins, son rôle reste largement méconnu comparé à celui des deux principaux MCRs impliqués dans cette fonction.

MC3R et MC4R sont exprimés principalement au niveau du système nerveux central, notamment au niveau hypothalamique. Pendant longtemps, le rôle exact de MC3R n'était pas connu. Récemment, il a été montré qu'il jouerait un rôle de rhéostat en réponse à des états de nutrition exceptionnels tels qu'une restriction alimentaire ou au contraire lors d'un régime riche en calorie (Ghamari-Langroudi et al., 2018). Mais le principal protagoniste du maintien de l'homéostasie énergétique reste le MC4R comme on va le voir dans la partie suivante.

3. L'homéostasie énergétique

Tout individu a besoin d'énergie pour survivre. Chaque jour, il doit trouver un apport énergétique suffisant pour couvrir les dépenses liées non seulement à son métabolisme de base mais aussi celles liées à ses activités quotidiennes. L'état énergétique subit alors des fluctuations quotidiennes avec des périodes de repas et de jeûne comme pendant le sommeil. Afin d'assurer une stabilité du poids et de la masse graisseuse traduisant un équilibre global de la balance énergétique, des systèmes de régulation performants ont été sélectionnés au cours de l'Évolution. Ces systèmes s'intègrent dans un processus appelé « homéostasie énergétique ». Son maintien dépend de la détection d'informations en périphérie et de leur traitement au niveau central engendrant ainsi une réponse adaptée à l'état énergétique de l'individu (Gautron et al., 2015). C'est ainsi un processus dynamique avec plusieurs phases. La première est la sensation de faim liée à un déficit énergétique perçu au niveau hypothalamique. Elle augmente la motivation à rechercher de la nourriture. La deuxième phase est le renforcement de cette motivation pour continuer de consommer la nourriture. Puis vient la satiation c'est-à-dire la fin du repas une fois que la nourriture consommée est supposée suffisante pour couvrir les besoins énergétiques de l'organisme. C'est un processus rapide, à court terme à différencier de la satiété. Ce dernier état est nécessaire pour éviter une sensation de faim constante et donc la multiplication des repas.

3.1 La faim, cette sensation désagréable

Pour beaucoup, la sensation de faim se définit comme une sensation désagréable liée au désir ou l'envie de manger. Elle va conditionner différents types d'alimentation : l'alimentation spontanée et quotidienne programmée (potentiellement entraînée par l'apport anticipé des oscillateurs circadiens prédictifs), l'alimentation engendrée par l'apparition soudaine d'une source alimentaire hautement appétante et calorique qui pourrait aider à se défendre contre un déficit énergétique futur et l'alimentation suivant une restriction alimentaire entraînée par des signaux de rétroaction indiquant un bilan énergétique négatif (**Figure 4**). C'est ce dernier type qui sera principalement traitée ici.

- **Signal périphérique**

Au niveau périphérique, la sensation de faim est initiée par une hormone: la ghréline produite par les cellules endocrines de l'estomac. À ce jour, elle représente le seul facteur orexigène présent dans la circulation systémique. Pour être active, elle doit être acylée par la ghréline O-acyltransférase (Gutierrez et al., 2008). Elle va augmenter la sécrétion d'acides gastriques et la motilité de l'estomac (Masuda et al., 2000) à l'origine des fameux « gargouillis » lorsqu'on a faim. Mais l'effet sur l'appétit est permis essentiellement par l'activation du récepteur Growth hormone secretagogue (GHSR), un RCPG abondant à la surface des neurones AgRP/NPY du noyau arqué ($ARC^{AgRP/NPY}$) (Kamegai et al., 2000).

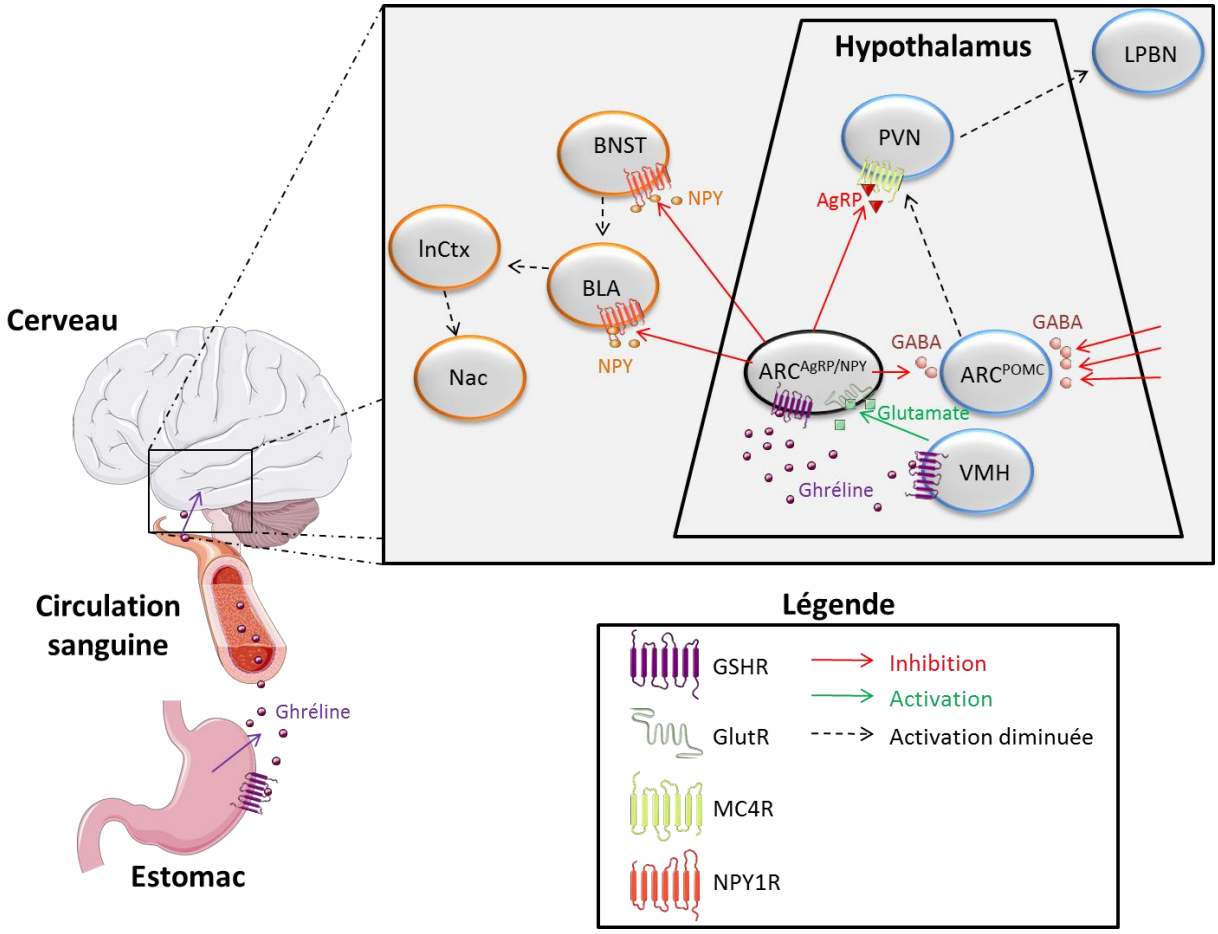


Figure 4. Circuits de la faim induite par un déficit énergétique. En réponse à un déficit énergétique, l'estomac produit la ghréline qui atteint l'hypothalamus via la circulation générale. On distingue alors des effets à long terme sur l'homéostasie énergétique (régions impliquées en bleu), via l'inhibition de la satiété, des effets à court terme sur la motivation à rechercher de la nourriture (régions impliquées en orange). Concernant les effets à long terme, la ghréline, en activant les neurones du noyau ventro-médial de l'hypothalamus (VMH), favorise la création de synapses glutamatergiques entre ces neurones et les neurones $ARC^{AgRP/NPY}$. Ces derniers inhibent l'activité des neurones du noyau paraventriculaire de l'hypothalamus (PVN) directement via la libération de l'antagoniste du sous-type 4 des récepteurs aux mélanocortines (MC4R), l'Agouti-related peptide (AgRP). En parallèle, les neurones $ARC^{AgRP/NPY}$ libèrent du GABA inhibant l'activité des neurones ARC^{POMC} producteurs des mélanocortines, agonistes du MC4R. Il en résulte une forte inhibition de l'activité du PVN qui se répercute sur le centre de la satiété, le noyau parabrachial latéral (LPBN). Les effets sur la motivation à rechercher de la nourriture passent majoritairement par la libération du neuropeptide Y (NPY). Celui-ci reconnait son récepteur au niveau du noyau du lit de la stria terminalis (BNST) et de l'amygdale basolatérale (BLA) diminuant l'activité de celle-ci et donc la peur. L'inhibition qui en résulte du circuit de la récompense, dont le noyau accumbens (Nac) fait parti, pourrait expliquer la valence négative associée à la sensation de faim. GABA : acide γ-aminobutyrique ; GSHR : récepteur à la ghréline ; GlutR : récepteur au glutamate ; NPY1R : récepteur 1 au neuropeptide Y ; POMC : pro-opiomélanocortine.

- **Rôle des neurones ARC$^{AgRP/NPY}$ à long terme dans l'inhibition de la satiété**

Le premier rôle décrit des neurones ARC$^{AgRP/NPY}$ est un rôle à long terme dans l'homéostasie énergétique via l'inhibition des systèmes anorexigènes de la satiété via la production de l'Agouti-related peptide (AgRP) agissant au niveau du noyau paraventriculaire de l'hypothalamus (PVN) (Chai et al., 2003; Cowley et al., 1999). Ce noyau est caractérisé d'un point de vue ontologique par le facteur de transcription Single-minded 1 (Sim1) (Kublaoui et al., 2008) dont une sous-population de neurones du PVN (PVN^{MC4R}) exprime fortement le sous-type 4 des récepteurs aux mélanocortines (MC4R) (Collet et al., 2017). Il reçoit de nombreuses projections des neurones ARC$^{AgRP/NPY}$ et ARC^{POMC} du noyau arqué (Cowley et al., 1999). Dans le cas d'un déficit énergétique, la ghréline entraine l'activation des neurones ARC$^{AgRP/NPY}$ non seulement via la reconnaissance de son récepteur GSHR mais aussi en favorisant l'apposition de nouvelles synapses glutamatergiques excitatrices (Pinto et al., 2004). Ces dernières proviennent essentiellement du noyau ventromédial de l'hypothalamus et du noyau paraventriculaire de l'hypothalamus permettant une boucle de régulation dans le contrôle des comportements de la prise alimentaire (Krashes et al., 2014). Les neurones ARC$^{AgRP/NPY}$ ainsi activés vont agir au niveau du noyau paraventriculaire de l'hypothalamus via deux processus (**Figure 4**). Le premier est direct avec la libération d'AgRP au niveau de ce noyau diminuant ainsi directement l'activité du MC4R. Le second processus est indirect. Les neurones ARC$^{AgRP/NPY}$ libèrent au niveau du noyau arqué le neurotransmetteur inhibiteur GABA qui diminue l'activité des neurones ARC^{POMC} en association avec la création de nouvelles synapses GABAergiques issues de populations neuronales autres que ARC$^{AgRP/NPY}$ (Yi et al., 2013). Pour Atasoy et al., cette régulation intra-noyau arqué médiée par GABA serait plutôt liée à une réponse rapide et non à un rôle à long terme (Atasoy et al., 2012). Tout ceci concourt à la diminution de l'activité des neurones ARC^{POMC} et donc du noyau paraventriculaire de l'hypothalamus. Or, ce dernier projette vers le noyau parabrachial latéral, un noyau clé de la satiété. Selon certains auteurs, les neurones ARC$^{AgRP/NPY}$ projetteraient directement vers ce noyau (Atasoy et al., 2012). Mais dans ce cas, ces projections n'activeraient pas directement la prise alimentaire mais limiteraient le malaise viscéral résultant de l'ablation des neurones ARC$^{AgRP/NPY}$.

Néanmoins, l'activation des neurones ARC$^{AgRP/NPY}$ via la ghréline stimule rapidement la prise alimentaire. Cela a notamment été démontré par des activations optogénétiques (Aponte et al., 2011) ou chimogénétiques (Krashes et al., 2011) de ces neurones. Dans ces études, le comportement alimentaire vorace est induit rapidement même à des moments où la souris mange normalement peu comme pendant la journée. Or la libération d'AgRP n'a un effet orexigène qu'à partir de plusieurs heures (Krashes et al., 2013). L'effet rapide observé fait intervenir d'autres molécules libérées par les neurones ARC$^{AgRP/NPY}$, GABA et le neuropeptide Y (NPY) (Atasoy et al., 2012; Krashes et al., 2013). GABA est le principal neurotransmetteur inhibiteur du système nerveux central chez les mammifères et les oiseaux et NPY un neuropeptide inhibiteur non réticulé de 36 résidus reconnaissant un RCPG de la classe Rhodopsine. Ces deux molécules vont être à l'origine de différents mécanismes aboutissant à l'augmentation de la motivation de l'individu à rechercher de la nourriture.

- **Rôle des neurones ARC$^{AgRP/NPY}$ à court terme sur la motivation à rechercher de la nourriture**

Le premier mécanisme est associé à la valence négative de la sensation de faim c'est-à-dire son caractère désagréable. En effet, l'injection intracrânienne de NPY provoque une

aversion gustative similaire à celle observée à la suite de l'activation optogénétique des neurones $ARC^{AgRP/NPY}$ (Atasoy et al., 2012). Le second mécanisme est l'inhibition de la peur. NPY est autant connu pour ses propriétés orexigènes que pour ses propriétés anxiolytiques (Comeras et al., 2019; Verma et al., 2016). Ce rôle fait sens dans un contexte de déficit énergétique. Un individu affamé doit choisir entre rester dans un endroit sûr et mourir de faim ou sortir à la recherche de nourriture au risque de rencontrer des situations dangereuses. Le troisième est l'augmentation de la valeur de récompense que représente la nourriture. L'animal fera ainsi plus d'efforts pour en obtenir une petite quantité même si elle est moins appétante (Chen et al., 2016; Denis et al., 2015).

L'aspect émotionnel est donc très important dans la sensation de faim et implique l'ensemble du système limbique (**Figure 4**). Ce système est le siège des émotions et est principalement formé de l'hippocampe et du fornix (mémoire), de l'amygdale (agressivité et peur), du gyrus cingulaire et de l'hypothalamus. Ainsi les neurones $ARC^{AgRP/NPY}$ agissent sur ces différentes régions. Ils vont également moduler l'activité d'autres zones cérébrales incluant le thalamus paraventriculaire et le noyau du lit de la stria terminalis, tous deux modulant l'activité de l'amygdale et le cortex insulaire important dans l'intéroception[2]. Bien que des auteurs attribuent une fonction à ces différentes projections (Atasoy et al., 2012), l'interconnexion de ces régions fait qu'elles sont le plus souvent impliquées dans différents rôles des neurones $ARC^{AgRP/NPY}$ dans un contexte de déficit énergétique.

Les circuits expliquant la valence négative portée par l'activation des neurones $ARC^{AgRP/NPY}$ ne sont pour l'instant pas bien décrits mais résultent sûrement de la modulation des autres parties du système limbique comme l'hippocampe pour l'apprentissage de la signification de cette sensation et la solution de trouver de la nourriture rapidement et du circuit de la récompense qui augmente la valeur de la nourriture. Les neurones $ARC^{AgRP/NPY}$ forment des synapses inhibitrices sur une sous-population de neurones du thalamus paraventriculaire qui projette vers l'amygdale communiquant à son tour avec le cortex insulaire. Or le dialogue entre ces deux dernières zones est essentiel pour coder et reconnaitre la valeur intéroceptive des signaux alimentaires (Livneh et al., 2017) et donc promouvoir la recherche de nourriture (**Figure 5**). Ce rôle dans la motivation implique notamment des projections du cortex insulaire vers le striatum ventral et le noyau accumbens appartenant au circuit de la récompense (Parkes and Balleine, 2013). L'inhibition de cette zone et du cortex insulaire serait peut-être à l'origine de la sensation désagréable de la faim. De plus, l'effet sur l'amygdale est renforcé non seulement par des projections inhibitrices des neurones $ARC^{AgRP/NPY}$ vers le noyau du lit de la stria terminalis normalement activateur de l'amygdale mais aussi via une inhibition directe de cette zone avec la libération de NPY concourant à un effet anxiolytique (Pape and Pare, 2010).

Néanmoins il apparait paradoxal qu'une population de neurones qui stimule l'appétit soit aussi liée à un comportement d'évitement. Chez des souris avec un accès alimentaire restreint, la présentation de nourriture ou de stimuli sensoriels prédisant l'arrivée de nourriture cause une diminution rapide de l'activité de la majorité des neurones $ARC^{AgRP/NPY}$ en quelques secondes avant même le début du repas (Betley et al., 2015; Chen et al., 2015; Mandelblat-Cerf et al., 2015). Manger serait donc un comportement dont l'apprentissage antérieur permettrait de réduire la valence négative liée à l'activation des neurones $ARC^{AgRP/NPY}$. La sensation désagréable de la faim s'atténuant, la motivation à chercher de la

[2] L'intéroception est la capacité à évaluer de manière exacte son activité physiologique (par exemple son rythme cardiaque).

nourriture diminue et permet ainsi d'initier la consommation de la nourriture trouvée. Il est important de noter que la seule perception d'une source de nourriture est suffisante pour induire cette chute rapide d'activité des neurones ARC$^{AgRP/NPY}$. Cela suggère donc que la régulation de ces neurones n'est pas uniquement sous le contrôle de facteurs circulants comme la ghréline mais dépend également de signaux sensoriels liés à la perception de l'environnement (Chen et al., 2015). D'ailleurs, Andermann et Lowell suggèrent que cette diminution d'activité des neurones ARC$^{AgRP/NPY}$ serait induite par des signaux alimentaires et serait proportionnelle à la quantité de nourriture qui va être consommée, prévenant ainsi une surconsommation (Andermann and Lowell, 2017). Par exemple, si la quantité de nourriture nouvellement disponible dépasse les besoins énergétiques, alors l'activité des neurones ARC$^{AgRP/NPY}$ chuterait rapidement à un niveau bas mais non nul (Mandelblat-Cerf et al., 2015). Pour maintenir la prise alimentaire initiée par les neurones ARC$^{AgRP/NPY}$, d'autres circuits neuronaux doivent rentrer en jeux comme ceux de l'hypothalamus latéral (**Figure 6**).

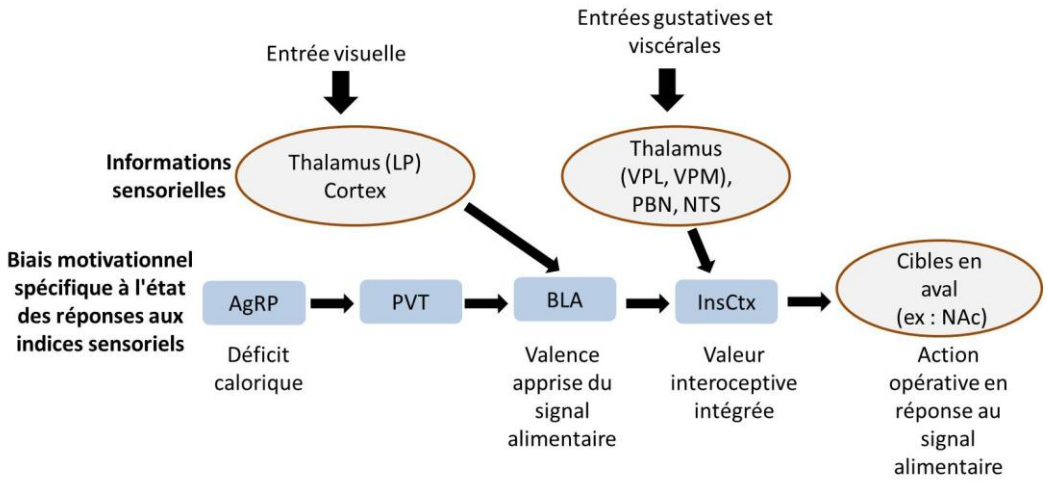


Figure 5. Voie neuronale illustrant le processus de traitement sensoriel de signaux alimentaires dépendant de la sensation de faim. L'amygdale latérale (BLA) et le cortex insulaire (InsCtx) permettent l'évaluation de la valence et de la saillance[3] de signaux alimentaires déjà rencontrés ainsi que la valeur intéroceptive espérée répondant à la question : Vais-je me sentir bien en mangeant et après avoir mangé cette nourriture ? Si la réponse est positive, les circuits moteurs sont recrutés par l'amygdale latérale et le cortex insulaire. La réponse dépend de la sensation de faim initiée par les neurones ARC$^{AgRP/NPY}$ via son action sur le thalamus paraventriuclaire (PVT) et de la saillance des signaux alimentaires déjà rencontrés et évaluée par l'amygdale latérale (BLA). L'estimation dans InsCtx des conséquences intéroceptives est probablement modulée par l'expérience acquise par les entrées gustatives et viscérales. LP : noyau lateral postérieur du thalamus ; POR : cortex postrhinal ; VPL : noyau ventral postéro-latéral du thalamus ; VPM : noyau ventral postéro-médial du thalamus ; PBN : noyau parabrachial ; NTS : noyau du tractus solitaire ; NAc : noyau accumbens. (Adapté de Andermann and Lowell, 2017)

[3] La saillance d'une chose est le fait qu'elle retienne l'attention par rapport aux autres choses présente dans le même environnement.

3.2 Renforcement positif de la prise alimentaire : implication de l'hypothalamus latéral (LH)

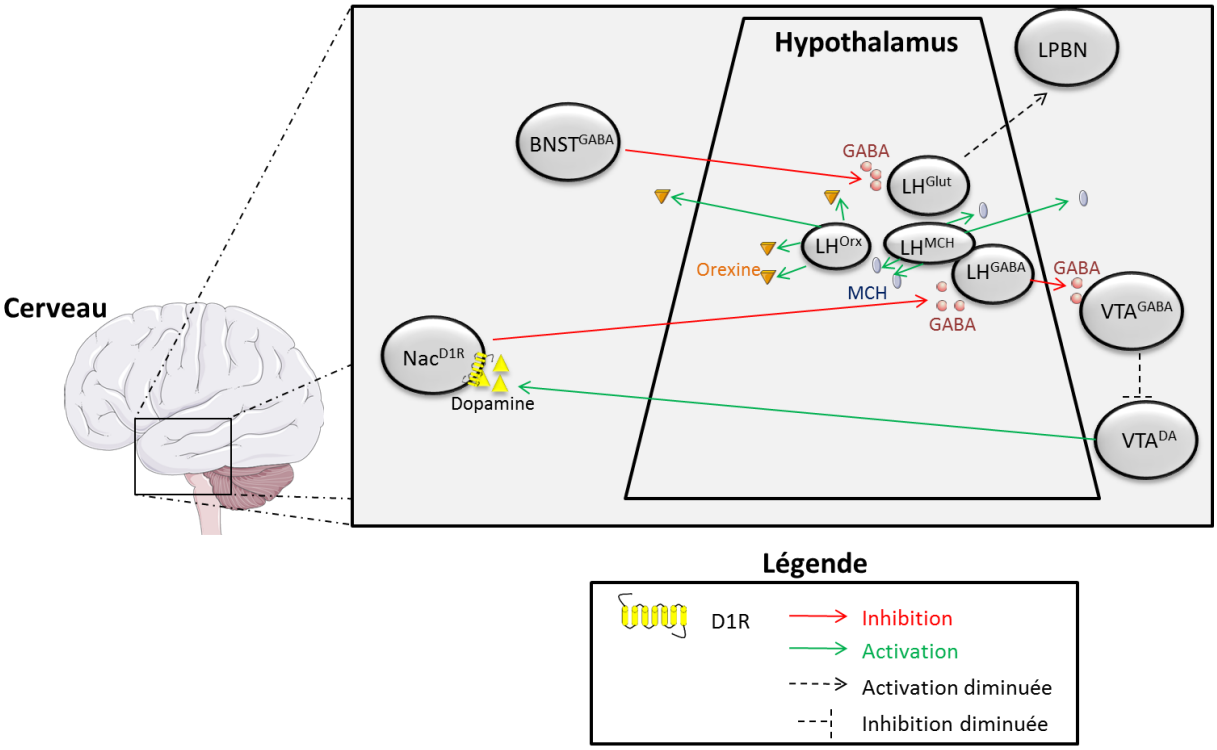


Figure 6. Circuits du renforcement positif de la prise alimentaire par l'hypothalamus latéral (LH). Les neurones GABAergiques du noyau du lit de la stria terminalis (BNST^{GABA}) inhibent préférentiellement les neurones excitateurs glutamatergiques LH^{Glut} diminuant l'activation du centre de la satiété, le noyau parabrachial latéral (LPBN). Lorsqu'ils sont activés, les neurones à orexine (LH^{Orx}) et les neurones à l'hormone concentrant la mélanine (LH^{MCH}) libèrent ces deux peptides et activent d'autres régions cérébrales du circuit de la récompense y compris d'autres populations neuronales du LH augmentant ainsi la motivation et la prise alimentaire. En parallèle, les neurones GABAergiques LH^{GABA} inhibent les neurones GABergiques de l'aire tegmentale ventrale (VTA^{GABA}) levant l'inhibition des neurones dopaminergiques de cette aire (VTA^{DA}). La libération de dopamine induit une valence positive de la prise alimentaire entrainant un renforcement de celle-ci. Cette libération est néanmoins soumise à un rétro-contrôle négatif. La dopamine active les récepteurs dopaminergiques D1 (D1R) exprimés à la surface du noyau accumbens (Nac). La stimulation de ces neurones NAc^{D1R} augmente alors l'inhibition GABAergique des neurones LH^{GABA} favorisant la diminution de la motivation et concourant ainsi à la terminaison du repas. DA : dopamine ; GABA : acide γ-aminobutyrique ; Glut : Glutamate.

Les premières études utilisant des stimulations électriques de LH chez des rongeurs montrent qu'une forte activation de cette zone entraine des comportements alimentaires voraces (Delgado and Anand, 1953) tandis que des lésions réduisent la prise alimentaire (Anand and Brobeck, 1951). Cependant, les manipulations de LH montrent également une disparité des comportements induits par son activation électrique. Certains individus auront des comportements alimentaires tandis que d'autres en auront des différents comme boire, attaquer ou encore une activité sexuelle (Sternson and Eiselt, 2017). Ces résultats

contrastent avec ceux de la photostimulation des neurones $ARC^{AgRP/NPY}$ qui induisent uniquement une réponse alimentaire (Aponte et al., 2011). En outre, d'autres études montrent que les animaux exécutent de façon fiable des réponses instrumentales, comme l'appui sur un levier, pour fournir une auto-stimulation électrique de LH dans les mêmes sites qui provoquent des comportements alimentaires (Margues and Olds, 1962; Olds and Milner, 1954). Ces résultats suggèrent ainsi que le rôle de LH et des circuits neuronaux associés est principalement lié à un renforcement positif de la valeur de récompense. Dans un contexte de prise alimentaire, la récompense est réprésentée par la nourriture.

Le LH englobe un grand nombre de types cellulaires avec différentes modalités de signalisation, y compris divers neurotransmetteurs et neuropeptides. Ces derniers pourraient aider à mieux comprendre le rôle du LH dans le contrôle de l'homéostasie énergétique.

- **Les différentes populations de neurones de l'hypothalamus latéral**

Les critères de différenciation des populations neuronales de LH sont moléculaires et basés soit en fonction des neurotransmetteurs libérés, soit en fonction des neuropeptides produits.

  - Les neurotransmetteurs de LH

Les marqueurs des neurones glutamatergiques (type 2 du transporteur vésiculaire du glutamate Vglut2) et des neurones GABAergiques (transporteur vésiculaire du GABA Vgat) sont fortement exprimés au niveau du LH suggérant la présence de sous-populations de neurones excitateurs et de neurones inhibiteurs (Ziegler et al., 2002). L'activation optogénétique des neurones LH^{Vgat} stimule la prise alimentaire et le comportement d'auto-stimulation suggérant un rôle dans la valeur gratifiante de la nourriture (Jennings et al., 2015). Ces effets sont étonnamment proches de ceux décrits précédemment lors de l'activation du LH. L'ablation génétique des neurones LH^{Vgat} confirme d'ailleurs ces résultats avec des effets opposés. Les mêmes auteurs ont aussi observés que certains groupes de neurones LH^{Vgat} sont actifs pendant la phase de recherche de nourriture tandis que d'autres sont excités pendant la consommation de la nourriture (Jennings et al., 2015).

  - Les neuropeptides de LH

Plusieurs sous-populations de neurones de LH expriment des neuropeptides qui reconnaissent des RCPGs. Certains peptides sont orexigènes comme MCH, l'orexine (Orx) aussi appelée hypocrétine et la galanine tandis que d'autres, la neurotensine et le transcrit régulé par la cocaïne et l'amphétamine (CART) sont anorexigènes (Berthoud and Münzberg, 2011). Mais les deux qui ont reçu le plus de considération sont MCH et l'orexine d'une part pour leur rôle dans l'augmentation de l'appétit et d'autre part parce qu'ils sont essentiellement produits au niveau du LH contrairement aux autres qui sont largement distribués au sein du cerveau. À noter que bien que les sous-populations LH^{MCH} et LH^{Orx} soient bien distinctes l'une de l'autre au niveau de l'hypothalamus latéral (Stuber and Wise, 2016), elles ne correspondraient pas forcément à des populations différentes de LH^{Vglut2}. En effet, certains neurones produisant MCH ou Orx expriment le marqueur Vglut2 (Elias et al., 1998; Rosin et al., 2003). Par contre, les neurones LH^{Vgat} ne sont colocalisés ni avec LH^{MCH} ni avec LH^{Orx} (Jennings et al., 2015) même si certains neurones LH^{MCH} expriment d'autres marqueurs GABAergiques comme GAD67 (Jego et al., 2013).

Le neuropeptide MCH est un peptide cyclique de 19 résidus. Son injection

intracerébroventriculaire (icv) augmente la prise alimentaire et le poids chez les rongeurs (Qu et al., 1996). En outre, la surexpression de MCH entraine une hyperphagie et une obésité (Ludwig et al., 2001) tandis que les souris sans récepteur à MCH ou n'exprimant pas MCH sont hypophagiques et maigres (Alon and Friedman, 2006; Shimada et al., 1998). Cet effet sur la prise alimentaire pourrait être lié à un renforcement positif de la nourriture. En effet, l'activation optogénétique des neurones LH^{MCH} favorise le conditionnement à une saveur (González et al., 2016) tandis que cet effet est bloqué si les neurones LH^{MCH} sont retirés (Alon and Friedman, 2006). De façon intéressante, l'activation de ces neurones favorise le sommeil paradoxal montrant un rôle également dans l'éveil (Jego et al., 2013). Ces deux rôles sont médiés par la reconnaissance de MCH avec un RCPG de la classe Rhodopsine. Deux isoformes du récepteur ont été découverts et sont reconnus avec la même affinité nanomolaire par MCH. MCHR1 est retrouvé chez tous les vertébrés. Il peut être couplé avec les protéines G_i, G_o et G_q (Hawes et al., 2000). Mais la contribution de chacune de ces voies de signalisation aux effets *in vivo* n'a pas encore été élucidée. Le deuxième récepteur, MCHR2, a été retrouvé notamment chez les primates et les chiens mais pas chez les rongeurs rendant l'étude de sa fonction *in vivo* difficile (Tan et al., 2002). Ainsi l'ensemble des études *in vivo* sur les rongeurs ne s'intéressent qu'au rôle de MCHR1. Contrairement à MCHR1, la signalisation de MCHR2 serait restreinte à G_q (Sailer et al., 2001). Malgré ces différences, les deux récepteurs présentent tous deux une large répartition de leur expression au sein du cerveau chez les primates, avec une expression privilégiée au niveau du LH.

Les neurones LH^{Orx} coexpriment l'orexine-A et la dynorphine. L'orexine-A est un peptide de 30 résidus possédant 2 ponts disulfure. Il reconnait deux récepteurs de la classe Rhodopsine (OxR1 et OxR2) qui présentent une large distribution dans le cerveau, principalement au niveau de LH. Les neurones exprimant l'orexine-A sont connus pour leur implication dans la prise alimentaire et les comportements liés à la récompense. L'injection du peptide dans le ventricule latéral augmente la prise alimentaire (Sakurai, 1999) et les souris n'exprimant pas ce peptide sont hypophagiques (Chemelli et al., 1999). Cependant, ces neurones ont surtout été décrits pour leur rôle dans l'éveil. Une ablation génétique de ces neurones entraine une narcolepsie des souris (Chemelli et al., 1999) tandis que leur activation optogénétique augmente l'attention (Adamantidis et al., 2007). Le rôle dans l'éveil est donc opposé à celui des neurones LH^{MCH}. Les deux populations de neurones déchargent d'ailleurs de manière réciproque dans le cycle éveil-sommeil (Hassani et al., 2009). Un lien entre éveil et prise alimentaire pourrait ainsi exister. Une interprétation possible serait que les phénotypes alimentaires associés aux neurones LH^{Orx} découleraient en réalité de leur rôle dans l'éveil. Ces phénotypes correspondraient donc à des comportements naturels qui apparaissent normalement dans certains stades d'éveil de l'animal. Par exemple, un animal mange plus s'il est éveillé.

Au final, LH participe aux phases de recherche et de consommation de la nourriture. L'activité du LH module comme les neurones $ARC^{AgRP/NPY}$ non seulement la motivation de l'individu mais renforce également la prise alimentaire. Ce rôle dual fait intervenir des circuits neuronaux similaires à ceux des $ARC^{AgRP/NPY}$.

- **Circuits neuronaux**

Le LH reçoit de nombreuses afférences excitatrices et inhibitrices provenant de différentes régions corticales et sous-corticales comme le noyau du lit de la stria terminalis

et les régions du système limbique y compris l'amygdale, le noyau parabrachial latéral et les noyaux arqué et paraventriculaire de l'hypothalamus. Des neurotransmetteurs comme la dopamine, la sérotonine et la noradrénaline sont également libérés dans le LH et vont jouer un rôle dans la modulation de ses différentes fonctions dans la motivation et les comportements alimentaires.

Le rôle dans la motivation et la valeur de récompense de la nourriture passent essentiellement par d'autres régions du système limbique. Les neurones GABAergiques du noyau du lit de la stria terminalis (BNST) émettent des projections vers LH et plus particulièrement vers les neurones glutamatergiques. L'activation optogénétique de ce circuit entraine rapidement un comportement alimentaire vorace vers la nourriture la plus appétante et la plus calorique (Jennings et al., 2013).

En aval de LH, les régions cibles dépendent des différentes populations neuronales composant LH. L'une de ces cibles est l'aire tegmentale ventrale (VTA). Le circuit LH-VTA est important pour la motivation et le renforcement de la valeur de la récompense via le conditionnement des réponses en fonction du contexte et le contrôle de la recherche compulsive de récompense. Par exemple, les neurones LH^{GABA} innervent préférentiellement les neurones GABAergiques de VTA inhibant ainsi leur activité (Nieh et al., 2015). En fonction de l'environnement ou du contexte, les projections LH^{GABA}-VTA sont capables de gérer différents comportements de motivation. En présence de nourriture, la stimulation de ce circuit promeut la prise alimentaire (Nieh et al., 2015) alors qu'en présence d'un nouvel objet ou d'un nouvel élément social, elle promeut cette fois-ci respectivement l'interaction et l'investigation (Nieh et al., 2016). La stimulation des neurones LH^{GABA} projetant vers le VTA augmenterait ainsi la motivation de l'animal à remplir une tache, mais entraine également la levée d'inhibition des neurones de VTA libérant de la dopamine (VTA^{DA}). Le résultat est une activation du noyau accumbens (Nac) inhibiteur des neurones LH^{GABA} formant ainsi un rétrocontrôle négatif à l'origine d'une diminution de la motivation (Stuber and Wise, 2016).

En conclusion, les circuits de LH modulent la motivation liée à une récompense, nourriture ou non, et sont étroitement connectés au circuit de la récompense impliquant l'aire tegmentale ventrale. Ils permettent de maintenir les comportements alimentaires initiés par les neurones $ARC^{AgRP/NPY}$ en renforçant la valeur de récompense de la nourriture. Cependant, au contraire d'autres circuits neuronaux orexigènes comme le circuit des neurones $ARC^{AgRP/NPY}$, les circuits de LH ne contrôlent pas la prise alimentaire essentiellement en fonction des besoins énergétiques. Pour terminer la phase de repas, la baisse de motivation initiée par l'activation du noyau accumbens n'est pas suffisante. D'autres circuits doivent être impliqués et modulés en fonction des signaux périphériques perçus lors du repas (**Figure 7**).

3.3 Satiation

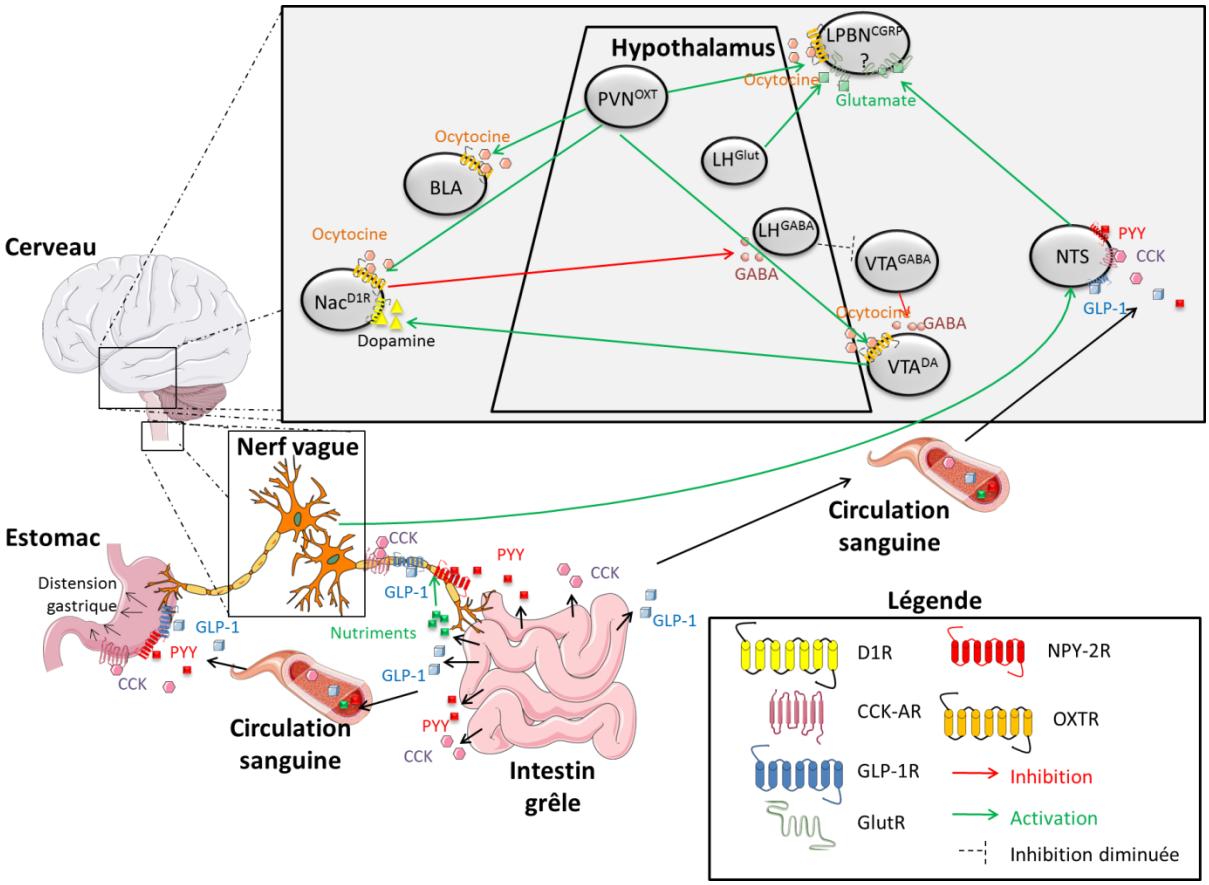


Figure 7. Circuits de la satiation (effet à court terme). La nourriture ingérée entraine une distension au niveau de l'estomac et la libération de nutriments et de peptides intestinaux (GLP-1, CCK et PYY) dans la circulation générale. Au niveau de l'estomac et de l'intestin, la motilité est accrue favorisant dans une boucle positive la production des peptides intestinaux. En outre, ces facteurs mécaniques et humoraux sont également perçus par les afférences vagales ou directement par le noyau du tractus solitaire (NTS) concourant à l'activité de ce dernier. Ensuite, le NTS stimule à son tour le centre de la satiété, le noyau parabrachial latéral LPBN en ciblant potentiellement la population neuronale $LPBN^{CGRP}$ porteuse d'une valence négative entrainant de ce fait la fin du repas. Cet effet est renforcé par les neurones glutamatergiques de l'hypothalamus latéral (LH^{Glut}) et par le noyau paraventriculaire de l'hypothalamus libérant l'ocytocine (PVN^{OXT}). En parallèle, les neurones PVN^{OXT} diminueraient la motivation à manger en activant d'autres régions du système limbique comme l'amygdale (BLA) et le noyau accumbens (Nac). CCK : cholécystokinine ; CCK-AR : récepteur A à CCK ; CGRP : peptide relié au gène de la calcitonine ; D1R : récepteur D1 à la dopamine ; DA : dopamine ; GLP-1 : gucagon-like peptide 1 ; GLP-1R : récepteur au GLP-1 ; Glut : glutamate ; GlutR : récepteur au glutamate ; NPY2R : récepteur 2 au neuropeptide Y ; OXTR : récepteur à l'ocytocine.

Lors d'un repas, la nourriture ingérée doit couvrir les besoins énergétiques immédiats. Cette évaluation nécessite dans un premier temps l'émission de signaux périphériques en fonction de la quantité et de la composition de la nourriture ingérée, et dans un second temps l'intégration au niveau central de ces signaux pour stopper le repas et entrainer la satiation.

- **Signaux périphériques**

Les principaux signaux de la satiation sont hormonaux et originaires du tractus gastro-intestinal. On a vu précédemment le rôle de la ghréline dans la sensation de faim. Lors d'un repas, sa sécrétion est stoppée (Steinert et al., 2017). D'autres hormones vont alors être sécrétées pour diminuer l'appétit en jouant notamment sur la taille et la fréquence des repas (Kim et al., 2018; Steinert et al., 2016). L'une d'entre elles est l'amyline, une hormone sécrétée par les cellules β-pancréatiques lors du repas en l'espace de quelques minutes (Lutz, 2013). Elle reconnait un RCPG au niveau de l'aire postrema, un organe localisé à proximité du noyau du tractus solitaire et dépourvue de barrière hématoencéphalique. Cette reconnaissance entraine une diminution de la prise alimentaire et un ralentissement de la vidange gastrique ainsi que des sécrétions digestives (Young, 2012). Cependant, les hormones de la satiation les plus étudiées sont les peptides intestinaux Glucagon-like peptide 1 (GLP-1), le peptide YY (PYY) et la cholécystokinine (CCK) reconnaissant tous des RCPGs au niveau central et périphérique. En ce qui concerne GLP-1, deux formes avec la même activité circulent dans le sang, GLP-1(7–37) et GLP-1(7-36NH_2), la dernière étant majoritaire chez l'Homme (Steinert et al., 2016). En plus d'être sécrété au niveau intestinal, GLP-1 est aussi produit par certains neurones du noyau du tractus solitaire. Jusqu'à présent, on ne sait pas dans quelle mesure ces deux sources périphérique et centrale sont redondantes et si elles agissent sur différentes cibles et ont des fonctions différentes. Néanmoins, dans les deux cas, le peptide reconnait un RCPG (GLP-1R) exprimé dans le tractus gastro-intestinal mais aussi au niveau d'autres organes comme le pancréas, certaines régions du cerveau comme l'hypothalamus et les afférences du nerf vagal. PYY est présent sous deux formes circulantes : la forme sécrétée PYY(1-36) et la forme active PYY(3-36) résultant du clivage des résidus tyrosine et proline à son extrémité N-terminale notamment dans le sang, le foie, et les cellules endothéliales des capillaires. PYY(3-36) active un RCPG, majoritairement NPY2R retrouvé dans les mêmes régions que GLP-1R (Eysselein et al., 1990). Enfin, la principale forme circulante de CCK est un peptide de 58 résidus (Eysselein et al., 1990). Elle reconnait deux RCPGs mais celui vraisemblablement impliqué dans l'effet anorexigène est l'isoforme A (CCK-AR) (Steinert et al., 2016). Son activation peut se faire via un mode d'action local au niveau de la partie proximale de l'intestin grêle (Cox, 1998) ou endocrine au niveau du muscle pylorique (Moran et al., 1988) et du noyau du tractus solitaire (Reidelberger et al., 2004). L'importance respective de ces deux modes d'action dans l'homéostasie énergétique n'a pas encore été déterminée. Cependant, l'essentiel des articles et des revues se focalisent sur l'effet neuroendocrine au niveau du noyau du tractus solitaire. Ces différents peptides anorexigènes sont tous sécrétés par des cellules spécialisées de l'intestin, les cellules entéroendocrines. Suivant leur localisation, ces cellules ne sécrètent pas les mêmes peptides : celles de la partie proximale cosécrètent GLP-1 et CCK tandis que celles de la région distale sécrètent GLP-1 et PYY (Kim et al., 2018). Leur sécrétion est initiée via la reconnaissance des nutriments ingérés par des récepteurs exprimés à la surface des cellules entéroendocrines. Ces récepteurs regroupent de nombreux RCPGs comme des récepteurs au goût et des récepteurs aux acides gras libres comme GPR40 pour la sécrétion de CCK ou GPR119 pour celle de GLP-1 (Steinert et al., 2016). La composition du repas et notamment la richesse du régime en graisse joue ainsi sur ces sécrétions. GLP-1 et PYY sont par exemple cosécrétés presque immédiatement (quelques minutes) après l'ingestion de glucose (Psichas et al., 2015). Par contre, si le régime est à la fois protéique et lipidique, la sécrétion de GLP-1 est alors ralentie et moins soutenue. Enfin, si un mélange de carbohydrates, de protéines et de lipides est consommé, la sécrétion est biphasique avec un

pic à 15min et à 2h après le repas (Steinert et al., 2017). Or, la disponibilité des nutriments dans l'intestin dépend de la vidange gastrique. Manger augmente le volume gastrique. Cette augmentation est détectée par des mécanorécepteurs au niveau de l'estomac qui vont permettre l'accommodation gastrique, une série de réflexes évitant une pression intragastrique trop importante aboutissant à la vidange de l'estomac. Mais ce processus est également régulé par l'activation des récepteurs aux nutriments de l'intestin via la stimulation de la sécrétion des peptides intestinaux créant ainsi une boucle de rétrocontrôle.

Ces signaux périphériques vont être traités majoritairement par le noyau du tractus solitaire intimement lié au système nerveux autonome pour terminer la phase de repas.

- **Circuits neuronaux**

Le noyau du tractus solitaire reçoit des afférences vagales (Ritter, 2004). Celles-ci innervent les cryptes intestinales et les villosités de l'intestin, l'estomac, le pancréas, et la veine porte transportant le sang des intestins vers le foie (Berthoud, 2004; Berthoud et al., 2004). Elles reçoivent ainsi non seulement des informations sur le volume et mais aussi sur la nature de la nourriture consommée. Le volume ingéré serait détecté par des neurones sensibles à la distension mécanique de l'estomac et de l'intestin (Williams et al., 2016). Mais les nutriments pourraient agir directement sur les afférences vagales via la présence de transporteurs et de récepteurs spécifiques comme le RCPG GPR40 activé par les acides gras à longue chaine (Darling et al., 2014; Li et al., 2016a). Ils vont aussi avoir une action indirecte en stimulant la production des peptides intestinaux GLP-1 ou CCK qui activent leurs récepteurs respectifs. Outre une activation via des afférences neuronales, le noyau du tractus solitaire peut directement être modulé par les peptides intestinaux. Par exemples, l'injection dans cette zone de GLP-1 supprime la prise alimentaire (Grill et al., 2002; Hayes et al., 2016) et l'administration périphérique de PYY(3-36) active les neurones de cette zone et est associée à un effet anorexigène causé vraisemblablement par l'induction d'une aversion gustative (Halatchev and Cone, 2005; Shen et al., 2017). Par contre, l'administration périphérique de CCK ne conduit pas à un changement net de la prise alimentaire puisqu'elle réduit la taille des repas mais augmente leur fréquence (West et al., 1984). CCK apparait donc comme un peptide essentiel dans la terminaison des repas et donc dans la satiation. Le résultat final de tous ces mécanismes est une forte activation du noyau du tractus solitaire. Ce dernier émet des projections glutamatergiques excitatrices vers les neurones du noyau parabrachial latéral exprimant le peptide lié au gène de la calcitonine (CGRP). Lorsqu'elle est activée, cette population est vectrice d'une valence négative d'où la fin du repas (Carter et al., 2015). En outre, leur activation supprime également la prise alimentaire dans des conditions où il est défavorable de manger, par exemple lors d'une maladie (Carter et al., 2013).

Au niveau de l'hypothalamus latéral, les neurones glutamatergiques LH^{Glut} sont activés et stimulent le noyau parabrachial latéral (Jennings et al., 2013; Stamatakis et al., 2016). Outre cet effet sur l'appétit, la stimulation du noyau accumbens diminue la motivation à manger via l'inhibition des neurones GABAergiques LH^{GABA} (O'Connor et al., 2015; Stratford and Kelley, 1999).

Un autre système plus controversée est celui lié à l'ocytocine, un peptide produit dans le noyau paraventriculaire de l'hypothalamus. Des expériences chez des rongeurs, des primates non humains et des humains démontrent un effet anorexigène de l'ocytocine presque immédiat et soutenu avec des doses répétées (Atasoy et al., 2012; Lawson et al., 2015). Cet effet est expliqué par la reconnaissance d'un RCPG au niveau du noyau du tractus

solitaire (NTS) (Lawson, 2017) dont l'activation accroit la sensibilité du NTS aux peptides intestinaux (Atasoy et al., 2012). Outre le noyau du tractus solitaire, les neurones à ocytocine du noyau paraventriculaire de l'hypothalamus projettent également vers l'aire tegmentale ventrale (Mullis et al., 2013) et le noyau accumbens (Herisson et al., 2016), deux régions du circuit limbique contrecarrant les effets sur la motivation induits par les neurones GABAergiques de l'hypothalamus latéral.

Une fois le repas terminé, des mécanismes de régulation à long terme doivent permettre d'éviter une augmentation de la fréquence des repas. Ces mécanismes sont ceux de la satiété et font le lien entre les réserves d'énergie sous forme de graisse et l'hypothalamus (**Figure 8**).

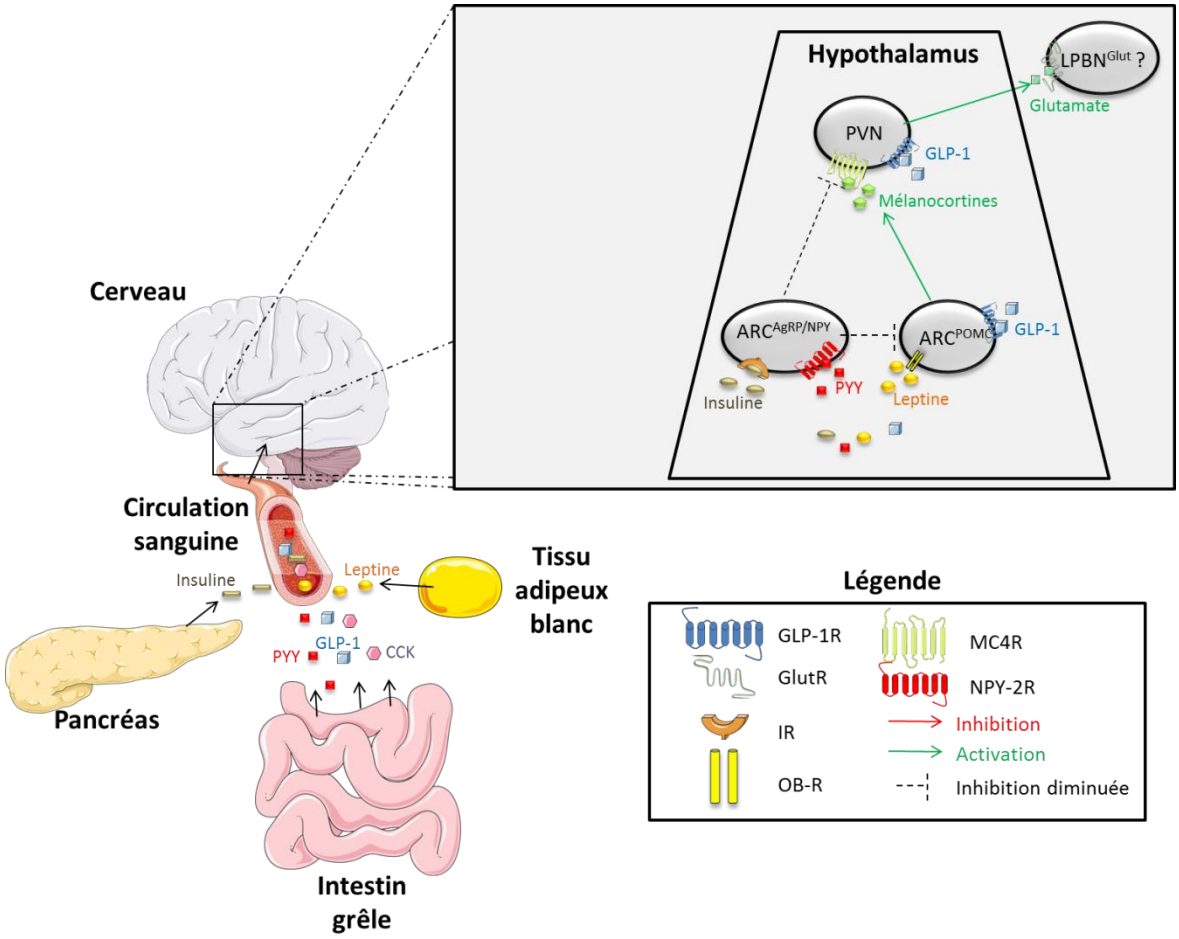


Figure 8. Circuits de la satiété (effet à long terme). En réponse à l'ingestion de nutriments, des signaux périphériques sont émis sous la forme de sécrétions de différents peptides intestinaux comme le Glucagon-like peptide 1 (GLP-1), le peptide YY (PYY) et la cholécystokinine (CCK). En parallèle, le tissu adipeux blanc et le pancréas sécrètent respectivement dans la circulation générale la leptine et l'insuline en fonction des stocks énergétiques notamment sous forme de graisse. En atteignant le noyau arqué de l'hypothalamus (ARC), ces différents peptides activent les neurones ARC^{POMC} et diminuent l'activité des neurones $ARC^{AgRP/NPY}$. L'augmentation de la production des mélanocortines, agonistes du soust-type 4 des récepteurs aux mélanocortines (MC4R), associée à la diminution de celle de l'antagoniste Agouti-related peptide (AgRP) entrainent une forte activation du MC4R au niveau du noyau paraventriculaire (PVN). Ces neurones émettent ensuite des projections excitatrices porteuses d'une valence positive vers le centre de la satiété, le noyau parabrachial latéral (LPBN) et notamment une population neuronale supposée glutamatergique à l'origine de la satiété. GLP-1R : récepteur au GLP-1 ; Glut : glutamate ; GlutR : récepteur au glutamate ; IR : récepteur à l'insuline ; NPY : neuopeptide Y ; NPY2R : récepteur 2 au neuropeptide Y ; OB-R : récepteur à la leptine ; POMC : pro-opiomélanocortine.

3.4 Satiété

- **Signaux périphériques**

Les deux principales hormones de la satiété sont la leptine et l'insuline. Les niveaux circulant de ces deux hormones sont directement proportionnels à l'adiposité de l'individu (Kim et al., 2018; Steinert et al., 2016; Zhao et al., 2003). La leptine sécrétée par les adipocytes reconnait un récepteur (OB-R) de la famille des récepteurs des cytokines exprimé notamment au niveau du noyau arqué de l'hypothalamus (Stephens et al., 1995). L'insuline est quant à elle d'origine pancréatique et joue un rôle crucial dans la régulation de la glycémie. Son récepteur à activité tyrosine kinase est aussi fortement exprimé au niveau hypothalamique (Corp et al., 1986; Havrankova et al., 1978).

- **Circuits neuronaux**

Au niveau de l'hypothalamus, l'insuline et la leptine vont agir en stimulant l'activité des neurones ARC^{POMC} via la reconnaissance de leur récepteur respectif et en diminuant celle des neurones $ARC^{AgRP/NPY}$ soit en entrainant une hyperpolarisation des neurones dans le cas de l'insuline, soit en réduisant leurs afférences glutamatergiques excitatrices dans le cas de la leptine (Cowley et al., 2001; Shen et al., 2017; Yang et al., 2011). Cette réduction serait sans doute médiée par la libération de β-endorphine par les neurones ARC^{POMC} (Yang et al., 2011). Cependant, une autre étude montre que cette libération de β-endorphine est promue par l'activation d'un RCPG, le récepteur 1 aux cannabinoïdes exprimé au niveau des neurones ARC^{POMC} et entraine une augmentation plutôt qu'une diminution de la prise alimentaire (Koch et al., 2015). Cela apporte une explication possible à l'hyperphagie induite par les cannabinoïdes (Waterson and Horvath, 2015). L'activation des neurones ARC^{POMC} résulte également de la diminution de leur inhibition GABAergiques par les neurones $ARC^{AgRP/NPY}$ associée aux faibles niveaux de ghréline circulante. Ainsi par différents mécanismes, dans le cadre d'une balance énergétique positive, les neurones $ARC^{AgRP/NPY}$ sont inhibés alors que les neurones ARC^{POMC} sont fortement stimulés. Ces derniers vont ainsi accroitre leur expression de pro-opiomélanocortine (POMC), le précurseur des mélanocortines et de β-endorphine. Les mélanocortines (principalement α-MSH) vont ensuite reconnaitre le MC4R au niveau du noyau paraventriculaire de l'hypothalamus (PVN) et causer la satiété mais n'auront dans ce cas aucun effet sur les dépenses énergétiques (Balthasar et al., 2005). Contrairement aux neurones $ARC^{AgRP/NPY}$ dont l'activation entraine un comportement alimentaire vorace peu importe si l'animal est nourri ou affamé (Krashes et al., 2011), les neurones ARC^{POMC} doivent être stimulés chroniquement pour observer l'effet anorexigène (Zhan et al., 2013). Deux hypothèses peuvent expliquer cette différence. La première serait liée aux neuropeptides produits. Au contraire des neurones ARC^{POMC}, les neurones $ARC^{AgRP/NPY}$ sont impliqués dans d'autres systèmes (via NPY et GABA) que celui des mélanocortines. Le système hypothalamique des mélanocortines apparait donc bien central dans le maintien de l'homéostasie énergétique à long terme. La seconde explication concerne l'organisation synaptique entre les neurones du noyau arqué (ARC) et ceux du noyau paraventriculaire (PVN). Les neurones ARC^{POMC} montrent par exemple des connexions synaptiques avec le noyau paraventriculaire relativement faibles comparées à celles des neurones $ARC^{AgRP/NPY}$ (Atasoy et al., 2014).

Les peptides intestinaux PYY et GLP-1 joueraient également un rôle dans l'activation des neurones ARC^{POMC} (Kim et al., 2018). Néanmoins, une étude montre que l'activation de GLP-1R au niveau de ces neurones est suffisante mais pas nécessaire pour expliquer l'effet

anorexigène de GLP-1. Par exemple l'injection du peptide dans le noyau arqué diminue la prise alimentaire mais la délétion du récepteur au niveau des neurones ARC^{POMC} n'entraine pas de changements métaboliques et les souris répondent toujours à une forme à action prolongée de GLP-1 (Burmeister et al., 2017). Pour PYY, les souris invalidées pour son gène sont hyperphagiques et obèses tandis que la surexpression du peptide dans des souris transgéniques diminue la prise alimentaire et évite le développement d'une obésité induite par un régime riche (Boey et al., 2008). Du côté de CCK, un important polymorphisme du récepteur lié à une obésité est observé (Miller et al., 1995) suggérant que le peptide est également impliqué dans des processus à plus long terme.

Fortement activés, les neurones PVN^{MC4R} stimulent le noyau parabrachial latéral (LPBN) via des projections glutamatergiques. L'activation optogénétique de cet axe entraine la satiété (Garfield et al., 2015). Les neurones de LPBN représentent donc les neurones de 3ème ordre dans le contrôle de la prise alimentaire via le système hypothalamique des mélanocortines. Cependant, la population neuronale responsable de cette activité n'a pas encore été identifiée. Tout ce que l'on sait c'est qu'elle est sûrement glutamatergique comme les neurones $LBPN^{Vglut2}$ sont nécessaires pour le maintien de la satiété et qu'ils engagent sélectivement les neurones PVN^{MC4R} (Garfield et al., 2015; Shah et al., 2014). La stimulation optogénétique de l'axe PVN^{MC4R}-LPBN est vectrice d'une valence positive démontrée par le comportement d'auto-stimulation de ce circuit par des souris affamées (Garfield et al., 2015). Cette valence positive pourrait expliquer la sensation tardive de soulagement voire de plaisir après un repas. En outre, la population $LPBN^{Vglut2}$ est distincte des neurones exprimant le peptide lié au gène de la calcitonine. En conclusion, le LPBN apparait comme le centre intégrateur des signaux de la satiété et de la satiation avec un rôle intéroceptif évident via la réception d'informations sur les réserves énergétiques et la nourriture ingérée.

Pour résumer sur cette partie, la prise alimentaire comporte plusieurs phases : 1) recherche, 2) consommation de nourriture, 3) fin du repas/satiation, 4) satiété. La recherche et la consommation de nourriture sont intimement liées à la motivation à satisfaire un besoin (déficit énergétique) porteur d'une valence négative. Cette sensation désagréable associée à la faim dépend principalement de l'activité des neurones $ARC^{AgRP/NPY}$ de l'hypothalamus. Lors du repas, cette valence négative diminue au profit d'une valence positive renforcée par l'activité du système limbique, en particulier l'hypothalamus latéral permettant la poursuite de la prise alimentaire. Afin d'éviter une hyperphagie, d'autres systèmes dépendant de la quantité et de la nature de la nourriture ingérée se mettent en place. Ils impliquent des informations périphériques mécaniques et moléculaires (hormones, nutriments) qui seront transmis soit directement soit indirectement au niveau du noyau du tractus solitaire. Cette zone activera à son tour le noyau parabrachial latéral qui permettra la fin du repas autrement dit la satiation. Cependant d'autres signaux à long terme sont nécessaires pour éviter la multiplication des repas. Ces signaux de satiété traduisent les niveaux des réserves énergétiques via la sécrétion de la leptine et de l'insuline. Ces deux hormones activent le système des mélanocortines via la stimulation des neurones du noyau arqué de l'hypothalamus ARC^{POMC} qui activent à leur tour les neurones du noyau paraventriculaire PVN^{MC4R}. Ces derniers projettent enfin vers le noyau parabrachial latéral. Le système des mélanocortines joue un rôle majeur dans l'homéostasie énergétique en contrôlant sur le long terme la prise alimentaire. Néanmoins, ces systèmes doivent être équilibrés avec ceux impliqués dans les dépenses énergétiques.

3.5 Contrôle des dépenses énergétiques

Chez les animaux, l'oxydation des nutriments ingérés est couplée à la réduction de certaines coenzymes dans la chaine respiratoire mitochondriale. Cette dernière permet la formation d'un gradient de protons utilisé par la cellule pour la synthèse d'ATP. Cependant, de l'énergie est perdue au cours du processus et est dégagée sous forme de chaleur. L'hydrolyse d'ATP est également une source d'énergie et de déperdition de chaleur dans les processus de synthèse de protéines, le maintien du potentiel de membrane, la contraction musculaire… Ainsi les dépenses énergétiques totales (DET) sont thermogéniques et peuvent être divisées en trois composantes : les dépenses énergétiques au repos (DER) associées au métabolisme basal (fonction cardiorespiratoire, maintien du potentiel de membrane au repos,…), les dépenses adaptatives (DEA) liées à la digestion et aux changements environnementaux et enfin celles induites par l'activité physique (DIA). Sur une période de 24h, les DER réprésentent l'essentiel des dépenses énergétiques avec 60% des DET suivies des DIA (30-40% des DET) et les DIR avec 10% des DET (Rosenbaum and Leibel, 2010). Ces valeurs dépendent de la taille et de la composition du corps de l'individu, de l'environnement et des comportements. La taille et la composition du corps (% de graisses vs % de muscles) influent sur le métabolisme de base. Les personnes obèses ont par exemple un métabolisme de base plus important que des personnes minces. En outre, la corpulence influe bien sûr sur l'activité physique, les personnes obèses ayant souvent une mobilité réduite liée au coût énergétique associé au déplacement de leur propre masse. Les comportements alimentaires sont également un facteur important dans les dépenses énergétiques avec la nature et la quantité de nourriture consommée. Une augmentation de la part des protéines dans un régime résulte par exemple en une augmentation de la satiété et des dépenses énergétiques (Hochstenbach-Waelen et al., 2009). Concernant la quantité ingérée, la surconsommation répétée induit un gain de poids attribuable principalement à la masse graisseuse mais aussi à une augmentation de la masse musculaire non induite par une activité physique. Par contre, la surconsonsommation ponctuelle n'induit pas une augmentation de la thermogenèse (Westerterp, 2017). D'un point de vue évolutif, ces résultats sont compréhensibles. Les mécanismes économisant de l'énergie pour limiter la perte de poids et ceux favorisant le stockage d'énergie ont un avantage évident pour la survie de l'individu dans un environnement où la nourriture n'est pas toujours disponible. Cela explique la difficulté à perdre du poids et la facilité à en gagner. Par exemple, le maintien d'une réduction d'au moins 10% de la masse corporelle est accompagnée d'une diminution de 20-25% des DET que l'individu soit obèse ou non (Rosenbaum and Leibel, 2010). Par contre, un ancien obèse aura besoin de 300 à 400 calories de moins par jour pour maintenir son poids et son niveau d'activité physique par rapport à un individu qui n'a jamais été en surpoids mais présentant le même poids et composition corporelles (Wing and Hill, 2001). Il existe ainsi des systèmes de régulation qui permettent de limiter ou d'accroitre les dépenses énergétiques suivant les conditions environnementales.

Le premier système est médié par le système nerveux autonome. L'augmentation de l'activité du système nerveux parasympathique diminue par exemple la fréquence cardiaque et donc les DER. L'activité du système nerveux sympatique impliquant le nerf vague a des effets opposés. Le deuxième système est l'axe hypothalamo-hypophysaire. L'hormone thyroïde augmente les dépenses énergétiques en augmentant les DER via l'augmentation de la fréquence cardiaque, la presssion artérielle et la consommation musculaire d'ATP (Rosenbaum and Leibel, 2010). Le troisième système concerne le tissu adipeux brun. Au

contraire du tissu adipeux blanc, le tissu adipeux brun dissipe de l'énergie via un processus qui implique l'oxydation d'acides gras et la production de chaleur grâce à la proteine UCP1 (Klingenberg, 1999). Cette dernière est localisée dans la membrane interne de la mitochondrie des adipocytes du tissu adipeux brun. Elle permet la réentrée dans la matrice mitochondriale des protons issus de l'oxydation des acides gras. Ce processus dissipe ainsi le gradient électrochimique formé par les protons sous forme de chaleur (Cannon and Nedergaard, 2004). Le rôle du tissu adipeux brun est lié aux dépenses énergétiques de repos et aux dépenses adaptatives. Il est primordial pour le maintien de la température corporelle chez les enfants et les mammifères hibernants (Cannon and Nedergaard, 2004). Il est également présent chez l'Homme adulte mais de manière beaucoup moins importante que chez l'enfant. Il jouerait un rôle dans l'adaptation à des changements de température. En effet, un abaissement modéré de la température à 13-16°C pendant 1 à 2 h suffit à augmenter l'oxydation des acides gras et la production de chaleur (Chen et al., 2013). Cependant, ce rôle est encore discuté et les mécanismes de régulation ne sont pas bien compris. Le tissu adipeux brun est richement innervé par le système nerveux sympathique. Or des études génétiques ont identifié le récepteur MC4R au niveau de neurones sympathiques cholinergiques du noyau intermédiolatéral de la moelle épinière (IML). Ces neurones permettraient une réponse adaptative à un régime riche ou à un environnement froid via la stimulation du tissu adipeux brun (Krashes et al., 2016) (**Figure 9**). Cependant, certains résultats sont contradictoires. Par contre, les neurones du noyau arqué sont connus pour moduler l'activité du tissu adipeux brun. Les neurones $ARC^{AgRP/NPY}$ atténueraient l'activité du tissu adipeux brun via la production de NPY au niveau du noyau dorso-médian du nerf vague (DMV) (Shi et al., 2013) alors que les neurones ARC^{POMC} stimuleraient son activité par l'activation du MC4R au niveau de PVN et DMV (Enriori et al., 2011), de même que les récepteurs à la leptine de cette dernière zone. Les neurones du DMV projettent ensuite vers le raphe pallidus (Samuels et al., 2002) qui supprime la thermogenèse du tissu adipeux brun probablement via la libération de NPY (Chao et al., 2011). Néanmoins, il semble difficile de transposer à l'Homme les résultats obtenus chez la souris ou le rat. En effet, plus un animal est petit, plus son aire par rapport à son volume est grande donc associée à une plus grande déperdition de chaleur et une plus grande nécessité d'avoir des mécanismes de maintien de la température corporelle. Plus un animal est grand, plus ce sera le contraire.

En conclusion, l'homéostasie énergétique est un processus faisant intervenir des circuits neuronaux complexes intégrant de nombreux signaux périphériques traduisant non seulement l'état énergétique mais aussi l'état émotionnel de l'individu. La transduction du signal de la périphérie au SNC fait souvent intervenir des RCPGs. Parmi eux, le sous-type 4 des récepteurs aux mélanocortines (MC4R) apparait central tant au niveau du contrôle de la prise alimentaire qu'au niveau des dépenses énergétiques. De par son importance, le MC4R a été défini comme une cible thérapeutique de choix dans les troubles de l'homéostasie énergétique (Fani et al., 2014). La prochaine partie décrira différentes considérations à prendre en compte dans le cas du développement d'un futur médicament ciblant ce récepteur.

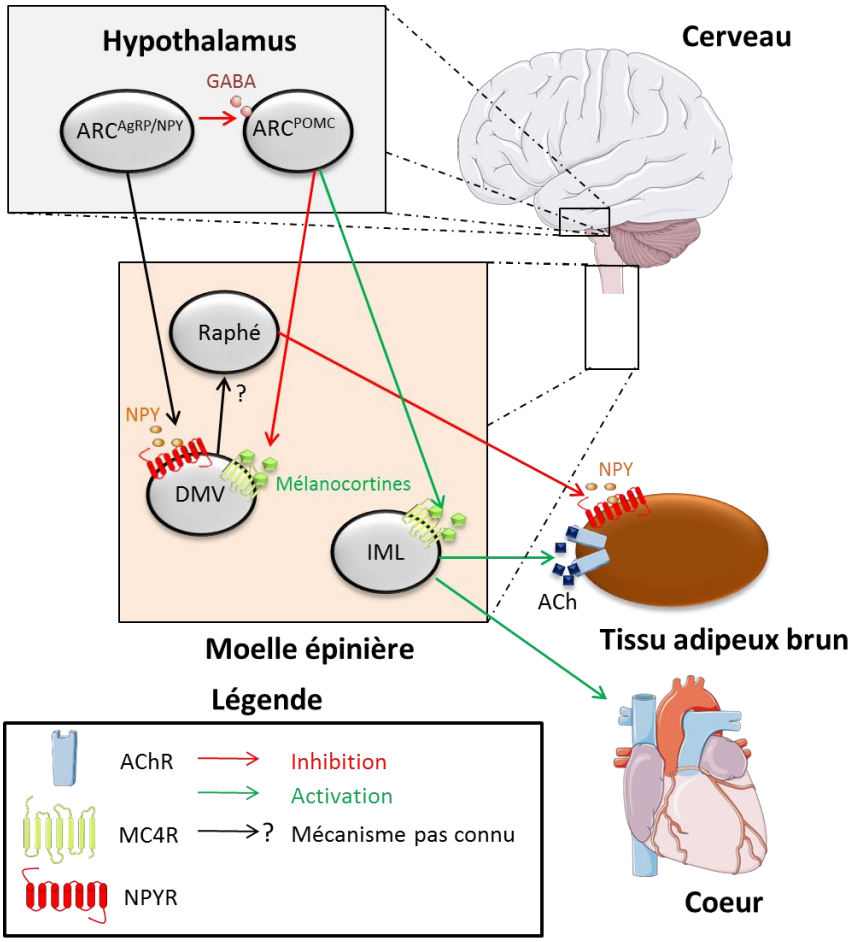


Figure 9. Rôle du sous-type 4 des récepteurs aux mélanocortines (MC4R) dans les dépenses énergétiques via son action sur le tissu adipeux brun. L'activité du tissu adipeux brun est stimulée via l'activation des neurones cholinergiques du noyau intermédiolatéral de la moelle épinière (IML) par les mélanocortines issues des neurones ARC^{POMC} du noyau arqué de l'hypothalamus. L'activation d'IML concourt à augmenter les dépenses énergétiques également via l'augmentation de la pression artérielle. Les neurones $ARC^{AgRP/NPY}$ ont l'effet opposé via la libération du neuropeptide Y (NPY) au niveau du noyau dorso-médian du nerf vague (DMV). Cependant le mécanisme exact n'est pas encore connu. ACh : acétylcholine ; AChR : récepteur à l'acétylcholine ; GABA : acide γ-aminobutyrique ; NPY2R : récepteur 2 au neuropeptide Y ; POMC : pro-opiomélanocortine.

4. MC4R comme cible thérapeutique : une signalisation complexe dans le système nerveux central

Le MC4R présente certaines caractéristiques qu'il est important de souligner dans un contexte de développement thérapeutique. Le rôle dans l'homéostasie énergtéique est porté majoritairement par le MC4R au niveau central. D'un point de vue thérapeutique, le futur médicament devra passer la barrière hématoencéphalique. Ensuite, la complexité de la signalisation doit être gardée à l'esprit pour identifier un ligand avec la fonction recherchée. De la compréhension de cette complexité peut dépendre la compréhension et la prévention des effets secondaires *in vivo*.

4.1 Passage de la barrière hématoencéphalique (BHE)

Classiquement, pour qu'une molécule atteigne sa cible au niveau central, elle doit traverser la barrière hématoencéphalique (BHE) qui isole le cerveau du sang. Le principal obstacle est constitué par les cellules endothéliales. Ces dernières présentent des jonctions serrées limitant le phénomène de transport passif et déterminant la perméabilité de la BHE. Cette propriété est également influencée par d'autres types cellulaires comme les astrocytes et les péricytes (Morita-Takemura and Wanaka, 2019). Cependant, le noyau arqué de l'hypothalamus est situé à proximité de l'éminence médiane, une zone présentant une BHE aux caractéristiques particulières. Cette zone présente par exemple des cellules endothéliales fenestrées (Broadwell et al., 1983) dépourvues de jonctions serrées (Ciofi et al., 2009). La perméabilité y est ainsi importante faisant de cette région un lieu d'échange important entre la circulation systémique et le cerveau. De plus, les membranes basales internes et externes forment un large espace périvasculaire nommé le parenchyme cérébral (**Figure 10**). L'absence de péricyte mature accroit également la perméabilité. Une autre composante est la présence de tanycytes, des cellules gliales polarisées situées autour de la base du troisième ventricule dans l'hypothalamus. Il en existe différents sous-types aux propriétés différentes. Les tanycytes β2 possèdent des jonctions serrées formant une barrière entre l'éminence médiane et le liquide cérébrospinal. Au contraire, les sous-types α2 au niveau du noyau arqué ne présentent pas ce type de jonction formant ainsi une couche perméable (Daneman et al., 2010; Mullier et al., 2010). En outre, l'environnement de l'éminence médiane est très proangiogénique via la production de facteur de croissance de l'endothélium vasculaire (VEGF) par les tanycytes et les astrocytes voisins et de NG2 par certaines cellules gliales.

La structuration particulière de l'éminence médiane n'est pas figée et l'accessibilité des molécules circulantes n'est pas constante. En effet, plusieurs études montrent une plasticité de la barrière de l'éminence médiane et du noyau arqué en fonction de la balance énergétique. Par exemple, des souris ayant subi une période de jeûne montrent une perméabilité accrue de la barrière de l'éminence médiane. D'une part, la fenestration des capillaires augmente et d'autre part, les tanycytes augmentent leurs jonctions serrées. Ces deux processus augmentent la surface d'échange possible entre le sang et l'éminence médiane permettant à certains neurones d'être nouvellement exposés aux signaux métaboliques en circulation, une situation inversée lors de la réalimentation des souris

(Langlet et al., 2013). Au niveau du noyau arqué, des phénomènes similaires ont lieu. Un régime riche en calories entraine une altération de la barrière du noyau arqué via une augmentation de l'angiogenèse, une hypervascularisation et la présence de cellules endothéliales dégénérées (Yi et al., 2012) ainsi qu'une atténuation de la neurogenèse (McNay et al., 2012).

Les mécanismes de passage des molécules de la circulation systémique au noyau arqué via l'éminence médiane restent cependant encore flous. En ce qui concerne la ghréline, des données *in vivo* (Schaeffer et al., 2013) et *in vitro* (Collden et al., 2015) suggèrent que la molécule est suffisamment petite (28 résidus) pour diffuser passivement à travers les vaisseaux de l'éminence médiane atteignant ainsi le parynchyme. Par la suite, elle serait prise en charge par les tanycytes β2 via un transport actif vers le liquide cérébrospinal. Le passage final vers le noyau arqué de l'hypothalamus (ARC) se ferait enfin par simple diffusion (Schaeffer et al., 2013). Mais ces résultats restent discutés. La leptine peut également subir un transport actif par les tanycytes de l'éminence médiane pour atteindre le noyau arqué (Balland et al., 2014). Des études ont également montré qu'elle pouvait être transportée à travers la BHE avec un rôle clé de son récepteur pour cibler les circuits de la récompense (Di Spiezio et al., 2018). La même chose a été observée pour l'insuline avec un transport actif à travers la BHE mais indépendamment de son récepteur (Rhea et al., 2018). Cependant, les mécanismes sous-jacents restent encore à être élucidés. Pour les autres peptides, aucune information n'est disponible. Pour GLP-1, le passage de la BHE est d'autant plus sujet à débat que son temps de demi-vie est très court (<2min) (Kim et al., 2018). Il est ainsi très difficile de prédire le passage de la BHE pour un peptide en développement pour cibler le MC4R hypothalamique.

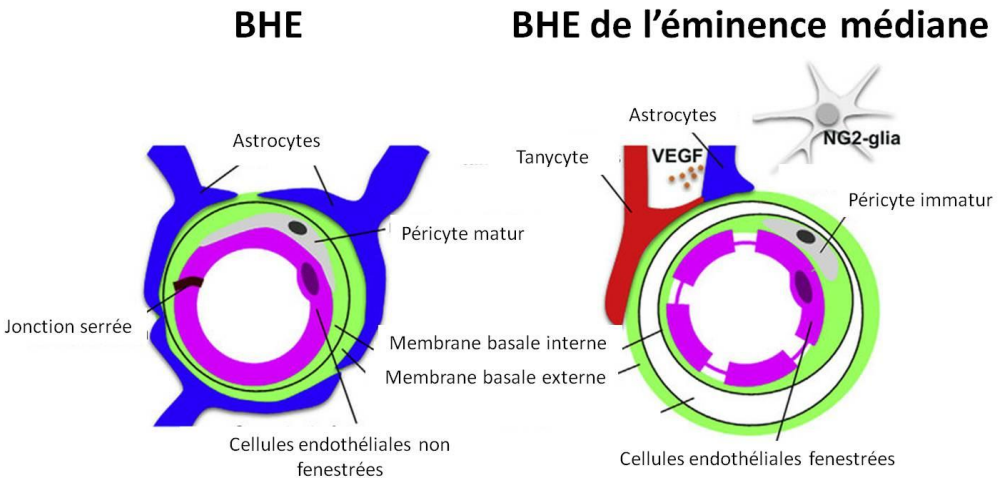


Figure 10. Représentation schématique d'une BHE peu perméable (à gauche) et d'une BHE perméable de l'éminence médiane (à droite). Une BHE est généralement peu perméable grâce à la présence notamment de jonctions serrées limitant la diffusion passive des molécules provenant de la circulation systémique. La BHE de l'éminence médiane est rendu perméable par la présence de cellules endothéliales fenestrées mais aussi par la production de facteurs pro-angiogéniques comme VEGF ou NG2. BHE : barrière hématoencéphalique. VEGF : facteur de croissance de l'endothélium vasculaire (*Vascular Endothelial Growth Factor*). (Adapté de Morita-Takemura et Wanaka, 2019)

4.2 Activité constitutive

L'activité constitutive d'un récepteur est définie comme une activité indépendante de la liaison d'un agoniste et résultant en la production de second messager. Plusieurs données *in vitro* démontrent une activité constitutive de MC4R de la voie G_s. Par exemple, Nijenhui et al. l'ont observée dans un clone de lignée cellulaire de mélanome murin MCB16/G4F transfectée pour surexprimer hMC4R. Dans ce cas, AgRP supprime de manière dose-dépendante jusqu'à 60% de cette activité basale (Nijenhuis et al., 2001). D'autres équipe ont obtenu par la suite des résultats similaires dans des cellules de rein d'embryon humain (HEK 293) et de cellules d'ovaires de hamster (CHO) (Asai et al., 2013; Chai et al., 2003; Haskell-Luevano and Monck, 2001). Le mécanisme sous-jacent à cette activité a été proposé par Ersoy et al en 2012. Dans leur étude, les auteurs ont montré qu'un récepteur MC4R avec une extrémité N-terminale tronquée présentait une activité constitutive réduite tout en répondant normalement à des agonistes comme l'α-MSH. Cette importance de l'extrémité N-terminale est confirmée par le rétablissement de l'activité basale avec la mise en présence d'un peptide mimant cette extrémité. Les auteurs ont également montré que l'effet agoniste inverse d'AgRP est expliqué en grande partie par l'inhibition de l'activation du récepteur par le domaine N-terminal de MC4R. Et cette auto-activation par le N-terminal est basée sur des résidus des TMs différents de ceux interagissant avec l'agoniste α-MSH (Ersoy et al., 2012). L'activation constitutive du MC4R par son extrémité N-terminale est ainsi complètement indépendante de celle engendrée par la liaison des agonistes endogènes. Cependant, ces résultats concernent uniquement des expressions hétérologues dans des cellules non hypothalamiques et ne représentent donc pas des systèmes physiologiques. On ne peut pas savoir quelles conséquences ces systèmes ont sur les conformations du récepteur. *In vivo*, certaines données sont en faveur d'une activité constitutive. En 2008, Tolle et Low observent par exemple chez des souris déficientes en POMC un effet retardé mais persistant d'AgRP sur la balance énergétque (Tolle and Low, 2008). Néanmoins, quelques années plus tard, Corander et al n'observent aucun effet de ce type chez des souris déficientes en POMC et en AgRP n'exprimant donc ni les agonistes ni l'antagoniste de MC4R (Corander et al., 2011). Les deux modèles sont hyperphagiques et il peut ainsi être difficile de montrer une stimulation de l'appétit induite par l'administration d'AgRP. En résumé, la signification physiologique d'une activité constitutive du MC4 *in vivo* reste encore largement incomprise.

4.3 Protéines accessoires

L'activité d'un récepteur dépend non seulement de sa régulation par les ligands endogènes mais aussi par l'interaction avec des protéines accessoires. Dans le cas de MC4R, la principale est MRAP2 de par sa forte expression hypothalamique. Dans une étude utilisant des neurones Sim1 du PVN de souris, MRAP2 diminue l'activité consitutive de la voie G_s mais augmente la réponse des cellules à l'agoniste α-MSH (Asai et al., 2013). Son absence dans les systèmes hétérologues pourrait expliquer l'activité constitutive observée de MC4R. En outre, les souris déficientes en MRAP2 développent une obésité tout en ayant un MC4R fonctionnel. Ce lien est d'ailleurs confirmé par des études génétiques chez un patients obèse poteur d'une mutation du gène mrap2. Néanmoins, MRAP2 est capable de stimuler la prise alimentaire indépendamment de son rôle sur MC4R via la régulation d'autres RCPGs comme en inhibant le récepteur de la prokinéticine 1 (Chaly et al., 2016) ou en potentialisant la

réponse du récepteur à la ghréline GHSR à son agoniste (Srisai et al., 2017). Cependant, toutes ces études se sont uniquement focalisées sur la voie G_s. Or le MC4R peut stimuler d'autres voies de signalisation pouvant être associées à différentes fonctions physiologiques.

4.4 Sélectivité fonctionnelle

Historiquement décrit pour être couplé à la protéine $G\alpha_s$ (Gantz and Fong, 2003), des études *in vitro* ont montré une plus grande diversité de voies de signalisation possible suite à la liaison des ligands endogènes sur le site orthostérique comme la voie G_i, la voie G_q, β-arrestine et MAPKinase (Breit et al., 2011). Bien que ces voies aient été observées dans des systèmes hétérologues, des études génétiques et *in vivo* apportent des réponses quant au rôle de chacune dans l'homéostasie énergétique. La première de 2015 utilise des modèles de souris n'exprimant pas les protéines $G\alpha_q$ et $G\alpha_{11}$ ou $G\alpha_s$ au niveau du PVN. Les auteurs montrent que les souris dépourvues de $G\alpha_q$ et $G\alpha_{11}$ au niveau du PVN sont hyperphagiques et développent une obésité sévère tout en ayant un métabolisme glucidique et des dépenses énergétques similaires aux contrôles. En outre, l'administration d'un agoniste de MC4R n'a pas d'effet orexigène mais entraine une augmentation de la pression artérielle. Chez les souris déficientes en $G\alpha_s$, les résultats sont opposés (Li et al., 2016b). Cela démontrerait un rôle physiologique différent suivant l'activation de la voie de signalisation : la voie de l'AMPc serait liée à des effets cardiovasculaires tandis que la voie de la phospholipase C serait impliquée dans la régulation de l'appétit. Récemment, une méta-étude sur des données génétiques montre que la variance de l'association entre l'indice de masse corporelle (IMC) de la population étudiée et les variants de MC4R observés s'expliquait à 88% par l'efficacité de recrutement de la β-arrestine au niveau du récepteur. De plus, des mutations gains de fonctions du MC4R entrainant une augmentation préférentielle de la voie β-arrestine par rapport à la voie G_s sont associées à une diminution jusqu'à 50% du risque de développer une obésité et un diabète de type 2. Ces mutations entrainent également une augmentation de l'expression des récepteurs à la membrane. Le mécanisme n'est pas encore connu mais les auteurs suggèrent une altération de l'internaliation du complexe récepteur-agoniste-β-arrestine ou un recyclage accru du récepteur (Lotta et al., 2019). Ainsi cette étude démontre une fois de plus que la fonction du MC4R dans l'homéostasie énergétique ne serait pas médiée par la voie G_s. Sachant qu'aucune des deux études n'étudie à la fois la voie $G_{q/11}$ et la voie β-arrestine, il est pour l'instant difficile de définir l'implication de chacune d'entre elles dans le maintien de l'homéostasie énergétique. Enfin, il a été montré que MC4R exprimé dans le PVN pouvait se coupler indépendamment de G_s à un canal potassique rectifiant entrant (Kir) possédant la sous-unité 7.1 (Ghamari-Langroudi et al., 2015). Dans cette étude, en plus d'une activité d'antagoniste compétiteur, d'agoniste biaisé G_i et d'agoniste inverse G_s, les auteurs décrivent l'AgRP comme un agoniste biaisé Kir7.1. En se fixant sur le récepteur, il entrainerait l'ouverture du canal potassique et donc l'hyperpolarisation des neurones du PVN de façon totalement indépendante de son activité d'antagoniste compétiteur vis-à-vis de α-MSH.

La signalisation du MC4R est ainsi complexe, chaque voie de signalisation pouvant moduler une fonction physiologique différente. Ces données sont extrêmement importantes dans un contexte thérapeutique afin d'éviter des effets secondaires pour le traitement des troubles de l'homéostasie énergétique. Ces troubles peuvent être liés à une balance énergétique positive comme l'obésité ou à une balance négative dans le cas de la cachexie.

5. État des lieux des principaux troubles de l'homéostasie énergétique

5.1 Obésité

- **Prévalence**

Selon l'Organisation Mondiale de la Santé (OMS), en 2016, plus d'un quart de la population mondiale était en surpoids (IMC>25kg/m^2), dont 650 millions d'obèses (IMC>30 kg/m^2). En 2016, un adulte sur 10 était donc obèse. Les cas d'obésité dans le Monde ont pratiquement triplé en l'espace de 40 ans. Et d'ici 2030, ce chiffre devrait atteindre 3.3 milliards de personnes. Or, l'obésité est souvent associée à d'autres problèmes de santé comme des maladies cardiovasculaires (notamment des cardiopathies et des accidents vasculaires cérébraux), le diabète de type 2, des troubles musculo-squelettiques comme l'arthrose et certains cancers. Le coût associé est donc très lourd. Par exemple, aux États-Unis, un des pays les plus touchés par l'obésité, ce coût est estimé entre 147 et 210 milliards de dollars par an (Cawley and Meyerhoefer, 2012). Cependant, ce chiffre date de 2012 et correspond à une proportion d'obèses de 33.8%. En 2016, cette proportion est passée à 39.6% suggérant un coût augmentant d'années en années.

- **Physiopathologie**

On a vu précédemment que les systèmes de maintien de l'homéostasie énergétique étaient très performants. Cependant, ils ont évolué principalement d'une part, pour faire face à des périodes de famine et d'autre part, pour induire une préférence envers une nourriture riche en sucres et en calories. Dès que la nourriture est accessible, la prise alimentaire est stimulée ainsi que la mise en réserves sous forme de graisse de l'énergie excédantaire. Or, dans nos sociétés occidentales, l'accès facile et rapide à une nourriture riche associé à un mode de vie sédentaire favorise l'hyperphagie. En outre, les publicités de nourriture rappellent contamment aux passants le plaisir que peuvent procurer des produits sucrés ou gras stimulant l'appétit même lorsque l'organisme n'a plus besoin d'apport énergétique. La première cause de l'épidémie mondiale d'obésité concerne donc essentiellement des facteurs environnementaux. Néanmoins, une cause génétique n'est pas à négliger. Responsable de 2% à 6% des cas d'obésité, la principale cause d'obésité génétique provient du MC4R avec près de 160 mutations majoritairement dominantes (Farooqi et al., 2003) (**Figure 11**). Elles entrainent une hyperphagie parfois à l'origine d'une obésité morbide (IMC>40kg/m^2). L'obésité peut également s'intégrer dans un syndrome comme celui de Prader-Willi en étant associée à d'autres signes cliniques comme des difficultés d'apprentissage et des troubles du comportement. Ce syndrome est lié à une anomalie du chromosome 15 causant un dysfonctionnement hypothalamohypohysaire à l'origine de l'hyperphagie dès l'enfance. Il représente la forme d'obésité syndromique la plus fréquente avec 1-9 enfants sur 100 000 (https://www.orpha.net).

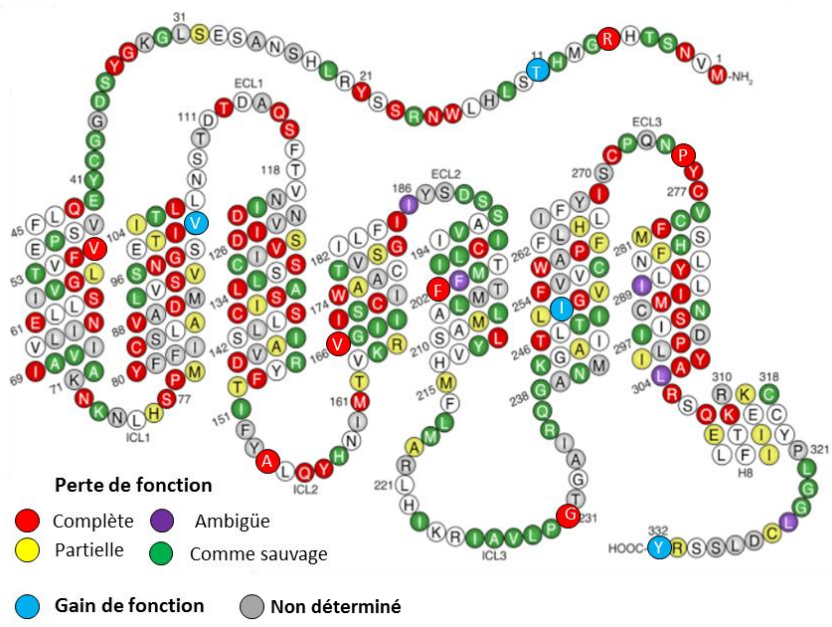


Figure 11. Variants perte de fonction ou gain de fonction du MC4R. La principale voie étudiée pour ces variants est celle de l'adénylate cyclase. (Adapté de Collet et al., 2017; Lotta et al., 2019)

- **Traitements actuels**

La principale recommendation de l'OMS est un mode de vie sain en limitant les aliments gras et sucrés et en ayant une activité physique régulière (60 minutes par jour pour un enfant et 150 minutes par semaine pour un adulte). Ce type de régime n'a des effets que sur le long terme. Pour les personnes souhaitant perdre rapidement du poids, le régime peut être plus restrictif. Néanmoins, s'il n'est pas associé à une activité physique, ce genre de régime est souvent peu efficace dû à la valence négative portée par la faim et la sensation de plaisir de manger. Un régime trop restrictif entrainera certes dans un premier temps une perte de poids rapide et importante. Cependant, la sensation désagréable de privation constante sera un facteur majeur d'arrêt du régime. L'individu reprendra alors ses anciennes habitudes. Or, son organisme se sera adapté pour favoriser une hyperphagie et le stockage des graisses entrainant une forte prise de poids par la suite. Pour des personnes en obésité morbide, l'association avec une activité physique peut être limitée à cause d'une mobilité réduite. Dans ce cas, la première solution est souvent la chirurgie bariatrique. Le but est de diminuer le calibre de l'estomac. L'opération la plus courante est le bypass. L'estomac est séparé en deux. Sa partie supérieure est directement reliée à l'intestin grêle évitant ainsi aux aliments de passer par l'estomac. Cela induit un ralentissement et une malabsorption du bol alimentaire ainsi qu'une diminution des taux de ghréline. La perte de poids est significative : plus de 50% de l'excès de poids à 12 mois chez les patients opérés (Puzziferri et al., 2014). Une autre méthode basée sur le même principe est la pose d'un anneau gastrique comprimant l'estomac. Contrairement au bypass, cette solution est ajustable et réversible d'où sa popularité depuis son autorisation en 2015 (Gautron et al., 2015).

Bien que très efficace, la chirurgie bariatrique reste une opération lourde et coûteuse nécessitant une anesthésie générale et plusieurs jours de rétablissement. Ces conditions peuvent ne pas être adaptées à des patients souffrant d'une hypertension sévère souvent liée à leur obésité. La solution résiderait alors dans un traitement pharmacologique. Plusieurs médicaments sont autorisés sur le marché aux États-Unis. Ils concernent principalement des petites molécules administrées par voie orale et ciblant le système sérotoninergique ou dopaminergique. En Europe, ces molécules font débat à cause de

risques cardiovasculaires et des problèmes de tolérance selon l'agence européenne du médicament. C'est le cas de Lorcaserin (Belviq®, Arena Pharmaceuticals) un agoniste du récepteur à la sérotonine 5HT2c ou de la combinaison entre le topiramate un anti-convulsif et la phentermine, un dérivé de l'amphétamine (Qsymia®, laboratoires Vivus). Des antidiabétiques peuvent également être prescrits pour traiter l'obésité. C'est le cas du liraglutide (Victoza®, Novo Nordisk) un agoniste de GLP-1R. Bien que cette indication soit autorisée aux États-Unis, elle ne l'est pas en France. Au final, un seul médicament contre l'obésité est réellement sur le marché et peut être pris de façon chronique. Il s'agit de l'orlistat (Xenical®, Roche ; Alli®, GlaxoSmithKline), une statine inhibitrice de la lipase gastrique et pancréatique limitant l'absorption des graisses. Ce tableau ne dépeint ainsi aucun agoniste du MC4R sur le marché. Néanmoins, une molécule est actuellement en phase clinique et son histoire révèle la complexité de développer un agoniste MC4R dans le traitement de l'obésité.

- **Traitements en développement**

Plusieurs molécules sont actuellement en essais cliniques comme des polyagonistes GLP-1, GIP, Glucagon. Mais une seule, le Setmélanotide®, un agoniste du MC4R, développée par Rythm Pharmaceuticals, est à un stade particulièrement avancé.

Le Setmélanotide® est issu comme tous les agonistes peptidiques artificiels du MC4R de l'ingénierie des ligands endogènes α- ou β-MSH (**Figure 12**) dans lesquels est toujours retrouvé le tétrapeptide HFRW essentiel pour l'activité agoniste. La première modification majeure a été la substitution de la phénylalanine du tétrapeptide par une D-Phe donnant la [NLe4,D-Phe7]α-MSH (NDP-α-MSH). Cette modification augmente l'effet agoniste sur la voie G_S et la stabilité du ligand. Ensuite pour améliorer la pénétration de la molécule dans le cerveau, le peptide a été cyclisé via les chaines latérales d'Asp et de Lys pour former le mélanotan II. Bien que ces deux molécules donnent des affinités sub-nanomolaires sur MC4R, elles restent peu sélectives du MC4R et entrainent des effets secondaires comme une aversion gustative (effet off-target), un brunissement de la peau et des cheveux (effet sur le MC1R), la stimulation de la fonction sexuelle (érection sponstanée du pénis) et une augmentation de la pression artérielle, toutes deux liées à une activité sur le système nerveux sympathique. Certains de ces effets sont intéressants d'un point de vue marché thérapeutique. Ainsi la NDP-α-MSH a été autorisée en 2014 en France et en 2019 aux États-Unis sous le nom de Scenesse® pour le traitement prophylactique de la protophyrie érythropoïétique, une intolérance sévère de la peau à la lumière. Des études sont en cours pour traiter le vitiligo. Quant au mélanotan II, il est commercialisé depuis 2019 aux États-Unis dans le traitement des troubles sexuels de la femme. À noter que ces deux peptides sont vendus sur Internet pour leur propriété auto-bronzante sans faire mention de risques cardiovasculaires. Les autres peptides issus de ces deux molécules, malgré une grande sélectivité pour le MC4R par rapport aux autres sous-types, présentent eux aussi les mêmes effets secondaires avec peu ou pas d'effet sur l'appétit (Fani et al., 2014). Du côté des molécules non peptidiques, il s'agit essentiellement de peptido-mimétiques tentant de représenter le pharmacophore avec une structure cyclique avec un noyau pyridine ou pyrrole. Malgré des formes actives par prise orale, aucune n'a franchi les phases cliniques à cause des effets cardiovasculaires. Pfizer a d'ailleurs complété deux phases II (NCT00862888, NCT00479570) avec sa molécule PF-00446687 pour le traitement des troubles de l'érection et des troubles de la sexualité chez les femmes ménopausées (données https://clinicaltrials.gov). Le Setmélanotide® se démarque ainsi de tous les ligands

MC4R disponibles. Il ne modifie pas la pression sanguine tout en ayant un effet significatif sur la diminution de l'appétit et du poids chez des patients obèses déficients en POMC, en récepteur à la leptine ou souffrant du syndrome de Prader-Willi (Clément et al., 2018; Kühnen et al., 2016). Il est actuellement en phase III (NCT03746522, NCT03287960, NCT03651765, NCT03013543, NCT02896192) pour évaluer l'efficacité et la toxicité de la molécule pour le traitement de ces pathologies (données https://clinicaltrials.gov). Au niveau structural, le Setmélanotide® se différencie du mélanotan II de par son cycle à 7 résidus formé par un pont disulfure contre un cycle à 6 (une D-Alanine en moins) dans le mélanotan II formé par la liaison entre les chaines latérales d'Asp et de la Lys. En dehors du cycle, Arg remplace une Nle dans le Setmélanotide. À part ces petites diifférences, le tétrapeptide His-D-Phe-Arg-Trp est toujours présent. Des études *in vitro* ont alors tenté d'expliquer l'absence d'effets cardiovasculaires par une sélectivité fonctionnelle de la molécule en faveur de la voie $G_{q/11}$ responsable des effets sur l'appétit par rapport à la voie G_s impliqué dans le tonus sympatique (Li et al., 2016b; Lotta et al., 2019). Le Setmélanotide® s'avère plus efficace que α-MSH pour activer les deux voies de signalisation et pour contrer l'effet antagoniste d'AgRP en particulier pour la voie $G_{q/11}$ (Clément et al., 2018). Bien que cela explique la grande efficacité dans la suppression de l'appétit, l'absence d'effets cardiovasculaires reste encore mal comprise.

Ainsi, le marché de l'obésité est actuellement très peu couvert avec des traitements peu efficaces et sugés à débat. L'approche multimodale associant régime alimentaire sain, activité physique et traitement pharmacologique semble être la stratégie la plus efficace. Mais d'un point de vue pharmacologique, peu de molécules passent avec succès les différentes phases cliniques par manque d'efficacité ou à cause d'effets indésirables. Bien qu'au début prometteur, le développement d'agonistes MC4R, s'est heurté à des effets cardiovasculaires importants potentiellement directement liés au rôle du MC4R dans le tonus sympathique. Sans ces effets, seul le Setmélanotide® franchit les différentes étapes avec succès et devrait arriver sur le marché dans les prochaines années. Il sera le « first in class » pour traiter les cas d'obésité génétiques mais il est certain que si la phase IV sur le marché est concluante, les indications seront élargies à des cas d'obésité plus fréquents.

L'obésité est ainsi le trouble de l'homéostasie énergétique le plus connu car le plus médiatisé. Mais un autre problème de santé publique moins connu touche également des millions de personnes dans le Monde, la cachexie, un bouleversement énergétique et métabolique lié au développement et l'aggravation d'une maladie chronique.

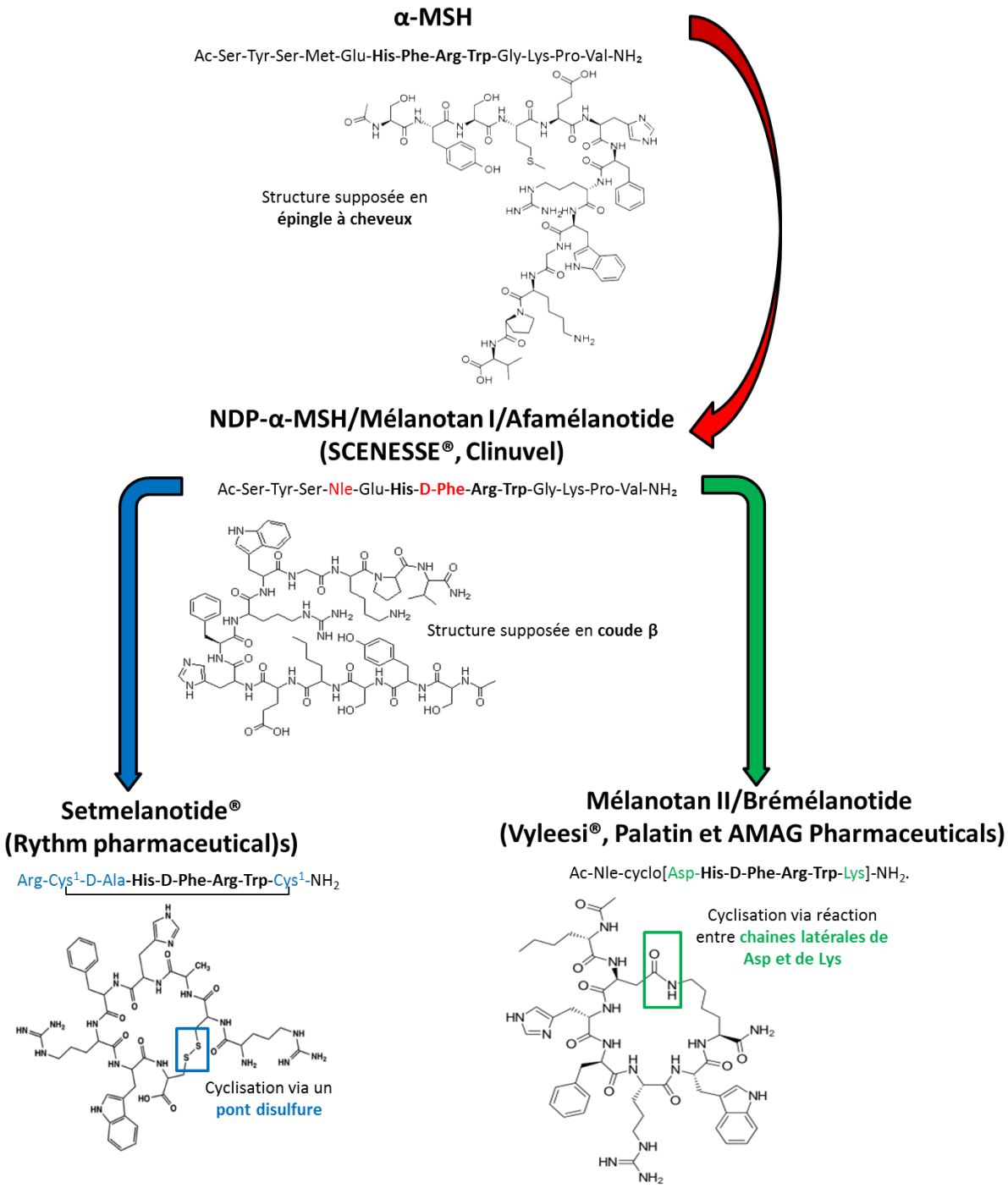


Figure 12. Développement à partir de l'hormone mélanotropes des principaux agonistes peptidiques MC4R sur le marché ou en phase clinique III. Seul le Setmélanotide® est en phase III pour le traitement de l'obésité. L'afamélanotide est indiqué pour le traitement de la protophyrie érythropoïétique. Le mélanotant II est autorisé aux États-Unis pour les troubles de la sexualité de la femme. Les positions modifiées par rapport à l'α-MSH sont indiqués en couleurs. Les noms commerciaux des différentes molécules ainsi que les laboratoires les développant sont spécifiés entre parenthèses.

5.2 Cachexie cancéreuse

- **Définition**

Plusieurs définitions ont été proposées au cours de la dernière décennie. La plus générale définit la cachexie comme un syndrome métabolique complexe associé à une maladie et caractérisé par une perte de la masse musculaire avec ou sans perte de masse graisseuse.(Evans et al., 2008). Elle représente une conséquence clinique sévère de presque toutes les maladies chroniques dans un stade avancé en particulier le cancer. On parle alors de cachexie cancéreuse.

- **Un problème de santé publique**

En 2018, l'incidence du cancer, tous types confondus, a été estimée à 18.1 millions de nouveaux cas (Ferlay et al., 2019) dans le Monde dont 382 000 en France (Defossez et al., 2019). La cachexie cancéreuse toucherait 50% à 80% des malades en fonction du type de tumeur (von Haehling et al., 2016) et serait responsable indirectement d'au moins 20% des décès (Argilés et al., 2014). Sa contribution dans la morbidité et la mortalité des cancers, l'altération de la qualité de vie des malades qu'elle induit et les importantes dépenses de santé associées font de la cachexie cancéreuse un problème de santé publique majeur (Farkas et al., 2013).

- **Un diagnostic difficile**

Le diagnostic est principalement basé sur une certaine perte de poids sur une période donnée : supérieure à 5% sur 6 mois sans volonté du malade de perdre du poids. Pour les personnes ayant un indice de masse corporelle déjà bas (inférieur à 20kg/m^2) ou déjà atteints de sarcopénie[4] comme les personnes âgées, le diagnostic se fait à partir d'une perte de poids de 2% (Fearon et al., 2011). Or, la perte de poids ne reflète pas à elle seule tous les changements pathophysiologiques liés à la cachexie cancéreuse. Plus que le poids, c'est la modification de la composition du corps qui compte avec notamment la perte de masse musculaire. Pour des malades recevant un traitement anti-cancéreux avec un poids stable ou augmenté, la sarcopénie peut être totalement occultée par l'accumulation de fluides (par exemple dans les ascites[5]). Le même problème est présent pour les patients obèses pour lesquels la sarcopénie est grandement négligée. Ainsi en 2011, un consensus a été pris sur les différents critères à évaluer. L'anorexie doit être évaluée fréquemment au minimum par le patient lui-même par rapport à sa prise alimentaire normale. Le métabolisme hypercatabolique lié à l'inflammation et à la tumeur peut être évalué par le dosage de la protéine C réactive. Les évaluations de la masse et de la force musculaire sont également majeures et peuvent être faites respectivement par imagerie à résonnance magnétique et par des tests d'efforts comme la dynamométrie de la poignée de main. Enfin, l'aspect psychologique ne doit pas être négligé.

- **Physiopathologie**

La présence de la tumeur et la prise d'un traitement anti-cancéreux (chimiothérapie et radiothérapie) altère le bon fonctionnement métabolique de l'organisme avec une perte d'appétit associée à un métabolisme inefficace aboutissant à une balance énergétique

[4] La sarcopénie est une perte progressive et élevée de la masse, de la force et de la fonction musculaires.
[5] Une ascite ou épanchement liquidien intra-abdominal est une accumulation de liquide dans l'abdomen.

négative et donc à une perte de la masse musculaire.

Plusieurs mécanismes rentrent en jeux. Le premier est un gaspillage énergétique avec une augmentation des dépenses énergétiques de repos. D'une part, la tumeur consomme de grandes quantités de glucose et d'acides aminés pour se développer. D'autre part, la fonction des mitochondries du muscle squelettique de patients cachectiques est totalement altérée. Par exemple, la production d'ATP via la chaine respiratoire est dégradée avec une production accrue de protéines découplantes UCPs (Argilés et al., 2014). Associé à une augmentation de la protéolyse musculaire et de l'apoptose, tout ceci favorise une atrophie musculaire. Le deuxième mécanisme important est une inflammation systémique avec la production de cytokines pro-inflammatoires par la tumeur ou par les cellules immunitaires de l'hôte. L'interleukine 1 joue un rôle clé en diminuant l'activité des neurones $ARC^{AgRP/NPY}$ et en stimulant celle des neurones ARC^{POMC}. Ces derniers libèrent alors le peptide anorexigène α-MSH et inhibe celle du peptide orexigène NPY (Scarlett et al., 2007). La conséquence est une diminution drastique de l'appétit conduisant à une anorexie. En outre l'activation du système des mélanocortines participerait également à l'atrophie musculaire (Cheung and Mak, 2012; Joppa et al., 2007). En périphérie, paradoxalement, le taux de ghréline est particulièrement élevé dans des cas de cachexie cancéreuse. Il représenterait peut-être une façon qu'a l'organsime de contrer l'anorexie mais il semble plus probable que ce soit le signe d'une résistance à la ghréline. Les mécanismes sous-jacents ne sont pas encore bien identifiés. On pourrait imaginer que la stimulation des voies anorexigènes soit trop forte pour que la ghréline ait un effet sur l'appétit. Par contre, des études ont montré le rôle que pouvait jouer la ghréline dans l'inhibition de la protéolyse et de l'apoptose de cellules musculaires (Reano et al., 2014). Cet effet serait lié à l'activation de son récepteur GSHR-1a au niveau de l'hypophyse entrainant la libération d'hormone de croissance qui active à son tour une cascade de signalisation (Reano et al., 2014). Bien que ce ne soit pas systématique, le tissu adipeux blanc peut également être touché avec notamment une augmentation de la lipolyse sans doute en réponse à la balance énergétique négative. En outre, l'étude de modèles animaux cachectiques montrerait une transformation d'une partie du tissu adipeux blanc en tissu adipeux brun augmentant ainsi les dépenses énergétiques et renforçant le gaspillage énergétique (Petruzzelli et al., 2014). Les autres mécanismes sont une résistance à l'insuline et une diminution des niveaux d'hormones anabolisantes comme la testostérone favorisant la perte musculaire.

Ce tableau clinique montre ainsi l'implication de différents organes dans une balance énergétique fortement négative. C'est pour cela que la cachexie cancéreuse peut être considérée comme un syndrome multi-organes et multi-factoriel (**Figure 13**).

- **Traitemenst actuels**

Les principales interventions ciblent une alimentation équilibrée associée à une activité physique. Cependant, ces approches doivent être adaptées à chaque patient et les bienfaits restent pour l'instant hypothétiques comme peu d'essais cliniques s'y intéressent vraiment. Concernant les traitements pharmacologiques, un seul est communément utilisé : l'acétate de mégestrol (MEGACE), un dérivé de la progestérone autorisé depuis 1993 par la FDA dans le traitement de l'anorexie, la cachexie et de la perte de poids inexpliquée chez des patients atteints du SIDA. Pris oralement, il augmente l'appétit. Le mécanisme précis d'action n'est pas connu mais il diminuerait la libération du neuropeptide Y (Aoyagi et al., 2015). Malgré une action orexigène, les effets sur le poids et la qualité de vie du malade ne sont pas significatifs (Roeland et al., 2016). En outre si une prise de poids est observée, elle

est souvent de courte durée et représente en réalité une augmentation du tissu adipeux associé à une perte de masse musculaire (Loprinzi et al., 1993). Enfin, la prise d'acétate de mégestrol augmente les risques de développer une thrombose veineuse, un rash et des problèmes gastro-intestinaux (Anderson et al., 2017). Ces effets sont particulièrement indésirables dans un contexte cancéreux où le malade reçoit déjà des traitements qui accroissent ces risques.

Les corticoïdes peuvent être également precrits mais leur utilisation est souvent limitée dans le temps pour prévenir les effets indésirables dermatologiques et gastro-intestinaux.

Au final, il n'existe pour l'instant aucun traitement efficace pour soigner la cachexie cancéreuse. Néanmoins deux molécules à prise orale présentent des résultats prometteurs en essai clinique.

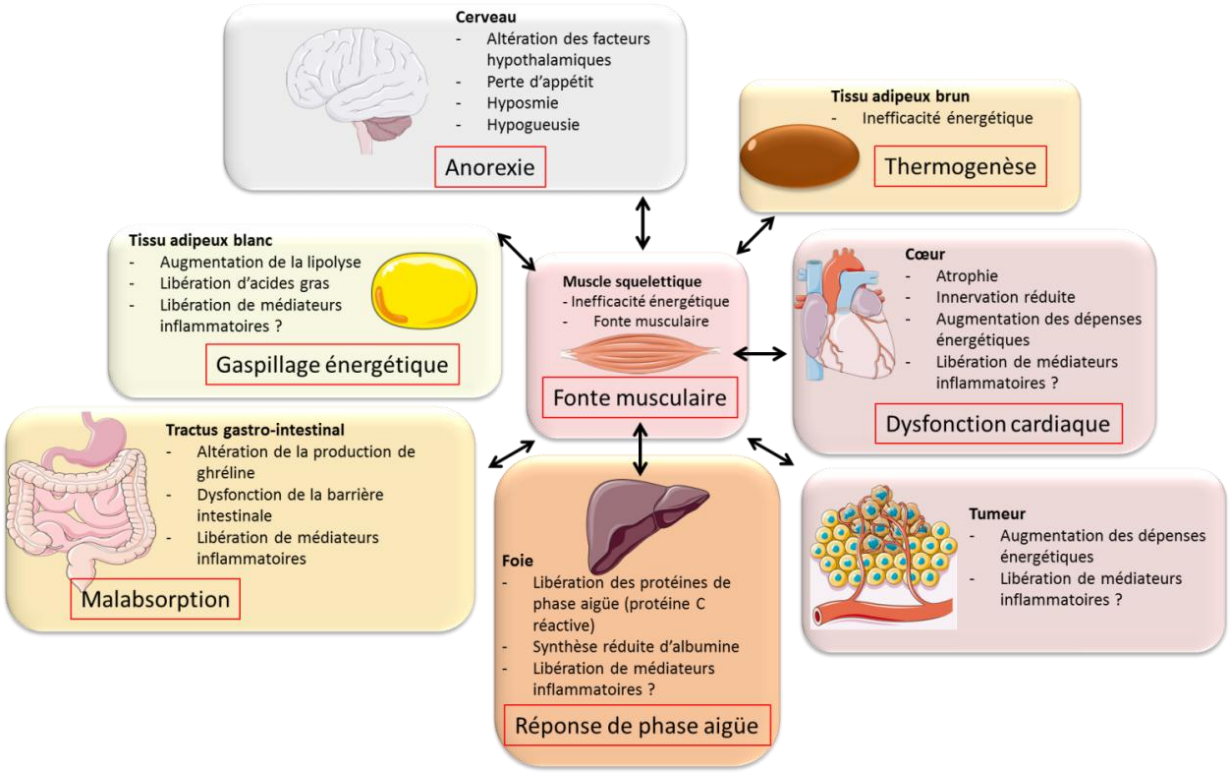


Figure 13. La cachexie cancéreuse, un syndrome multi-organes. Le principal signe clinique, la fonte musculaire, résulte en réalité de la conjonction de l'altération du bon fonctionnement de différents organes. Le premier facteur important est la modification du métabolisme vers un gaspillage énergétique au niveau du muscle squelettique et renforcé par le tissu adipeux blanc. Le deuxième, une inflammation systémique, cause la libération de médiateurs pro-inflammatoires altérant notamment les systèmes de l'homéostasie énergétique. Cette inflammation associée à la tumeur augmente fortement les dépenses énergétiques. Ainsi, l'association d'une'anorexie et d'un métabolisme inefficace dans un contexte de forts besoins énergétiques est à l'origine de la perte de masse musculaire. (Adapté de Argilés et al., 2014)

- **Traitements en développement : vers une approche multimodale**

Le premier traitement prometteur est l'Enobosarm (GTx-024; GTx, Memphis, TN, USA). C'est un modulateur non stéréoïdien sélectif du récepteur aux androgènes qui stimule l'activité anabolique et androgénique favorisant ainsi la fonction et la prise musculaire chez des patients atteints de cachexie cancéreuse tout en ayant des effets secondaires tolérables par les malades (Dobs et al., 2013). Néanmoins, des rapports préliminaires d'une phase III (NCT01355484) montrent qu'une dose de 3 mg n'a pas permis de montrer une amélioration significative de la masse musculaire (https://clinicaltrials.gov).

Le deuxième et le plus prometteur est l'anamoréline (Helsinn Healthcare SA), un analogue de la ghréline sélectif du récepteur GSHR et actif par voie orale. Une phase II (NCT01505764) a montré un effet anabolisant avec une augmentation de la masse musculaire mais aussi un effet positif significatif sur le poids, le score de qualité de vie mais pas sur le score d'appétit ni sur sur la force de la poignée de main (Garcia et al., 2015). Deux phases III ont confirmé ces résultats même si l'effet orexigène n'a pas été évalué (Temel et al., 2016). Ces résultats ne sont pas tout à fait inattendus. Les malades inclus dans la phase III présentaient une cachexie réfractaire (perte de poids supérieure à 5% à 6 mois) donc une perte importante de la masse musculaire. Or l'ampleur de la force générée par unité de masse musculaire chez ces personnes est diminuée par rapport à des sujets en bonne santé. Ainsi il leur faudra développer une masse musculaire plus importante pour faire le même exercice. À cause de cette absence d'amélioration de la force musculaire, l'agence européenne du médicament a refusé la mise sur le marché de l'anamoréline en 2018, ce qui pose plusieurs questions relatives à la définition des essais cliniques dans le cas de la cachexie cancéreuse. Dans un article de 2015, Fearon et al. demandent aux autorités de santé une meilleure concordance entre la significativité de l'effet demandé et ce qui est possible de démontrer en essai clinique pour un syndrome aussi multifactoriel que la cachexie cancéreuse (Fearon et al., 2015). Cela éviterait d'abandonner des traitements prometteurs pour pouvoir les inclure ensuite dans une approche multimodale. En effet, il semble peu probable qu'un seul médicament traite tous les aspects de la cachexie cancéreuse. La meilleure stratégie serait de combiner un médicament orexigène comme l'anamoréline avec un traitement anabolisant comme l'Enobosarm. Mais cela ne sera possible que si les deux molécules ne sont pas bloquées en phase clinique.

En conlusion, que ce soit l'obésité ou la cachexie cancéreuse, aucun traitement pharmacologique efficace n'est pour l'instant sur le marché. Bien que plusieurs molécules soient en développement, peu ciblent le MC4R. Pourtant, les connaissances actuelles montrent que ce récepteur pourrait être une cible clée dans le traitement de ces troubles métaboliques. Depuis une vingtaine d'années, la principale stratégie a été l'ingénierie des ligands endogènes afin de mimer leur pharmacophore. Bien que beaucoup de travail ait été fait dans ce sens, les mêmes problèmes qu'avec les ligands endogènes ont été rencontrés avec les ligands artificiels. Dans le cas des agonistes, des effets cardio-vasculaires ont limité par exemple le développement de plusieurs molécules dans le traitement de l'obésité. Concernant les antagonistes, ce serait plutôt des problèmes de rendements de synthèse et/ou de manque d'efficacité *in vivo*. Ainsi dans le cas du MC4R, mimer les ligands endogènes du MC4R s'avère être une stratégie peu efficace. Une solution serait d'explorer une source connue pour sa diversité structuralet et son potentiel thérapeutique, à savoir les venins.

6. Diversité et intérêt thérapeutique des toxines animales

6.1 Venin *versus* Poison

Les deux termes sont souvent utilisés comme synonymes. Pourtant ils correspondent à deux mélanges complexes totalement différents.

« Le venin est une sécrétion, produite par un animal dans une glande spécialisée et administrée à un autre organisme via la génération d'une blessure (aussi petite qu'elle soit), contenant des molécules qui perturbent les processus physiologiques et biochimiques normaux de la victime de telle sorte à faciliter l'alimentation, la défense ou la compétition de l'animal producteur » (Fry et al., 2009).

De cette définition, le venin diffère du poison principalement par son mode de production dans un organe dédié à sa synthèse. À l'opposé, les espèces animales sécrétant ou renfermant dans leur peau un poison accumulent des molécules produites par des mircoorganismes comme la tétrodotoxine retrouvée dans la chair du poisson-globe ou issues de leur alimentation comme l'épibatidine des petites grenouilles dendrobates. En outre, les espèces venimeuses injectent directement dans la circulation systémique leur cocktail de molécules via des systèmes d'inoculation variés, du cnidocyste microscopique des Cnidaires aux crochets de plusieurs centimètres de *Bitis gabonica* (**Figure 14**) (Lewis and Garcia, 2003). Récemment, un système d'inoculation surprenant a même été décrit chez deux espèces de grenouilles (*Corythomantis greeningi* et *Aparasphenodon brunoi*) (Jared et al., 2015). Ces amphibiens utilisent des épines osseuses de leur crâne pour administrer les sécrétions toxiques de leur peau chez un assaillant. Au contraire, un poison pénètre de façon passive dans l'organisme-cible par ingestion, inhalation ou absorption à travers la peau.

D'une façon générale, les molécules contenues dans les deux types de sécrétions sont appelées des toxines. C'est un terme générique qui ne prend en compte ni l'origine phylogénétique ni la nature chimique des molécules. Dans les venins, les toxines se réfèrent à la partie peptidique du mélange qui contient en outre des protéines, des sels et des composés organiques comme des acides aminés et des neurotransmetteurs (Casewell et al., 2013). Dans les poisons, elles sont représentées essentiellement par des alcaloïdes plus hydrophobes et plus à même de passer les barrières physiques de l'organisme (peau, muqueuses, …). Dans ce manuscrit, une toxine sera arbitrairement définie comme un peptide de moins de 120 résidus identifié dans le venin d'une espèce animale sans préjuger de son activité toxique.

Dans la dernière partie de la définition, la fonction est déterminante pour la survie de l'organisme. Le venin peut être défensif comme un poison mais peut également jouer un rôle dans la prédation ou plus rarement dans la compétition pour la reproduction, par exemple dans les combats entre deux ornithorynques mâles. Dans tous les cas, au cours de l'évolution les pressions de sélection ont façonné cette arme chimique. Comprendre l'histoire évolutive, c'est appréhender la richesse moléculaire des venins et leur potentiel en développement thérapeutique.

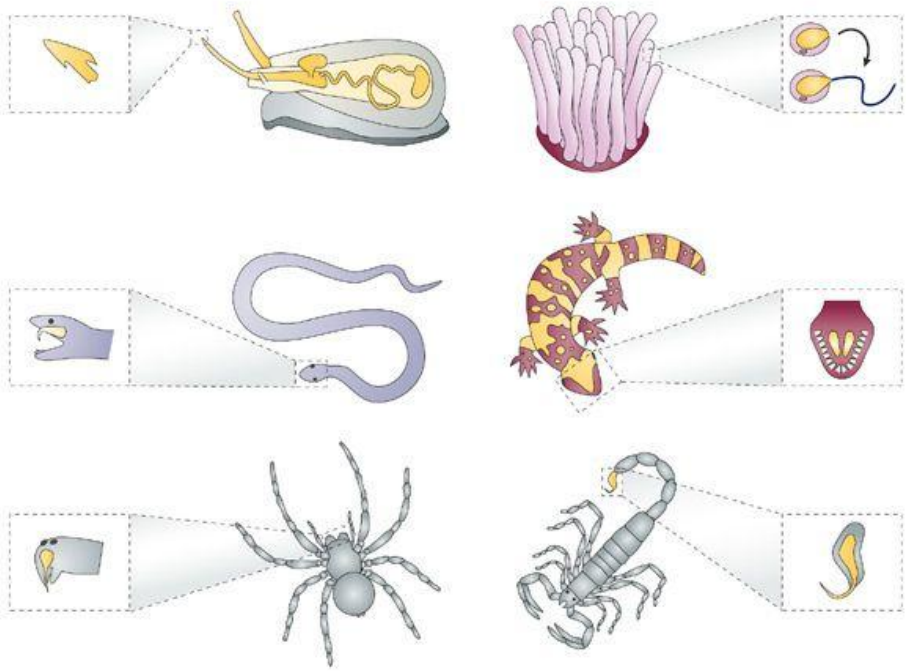


Figure 14. Diversité des principaux systèmes d'injection des animaux venimeux : harpon chez les cônes, cnidocystes chez les Cnidaires (anémones et méduses), dents chez les serpents et les lézards, chélicères chez les araignées, aiguillon chez les scorpions. (Adapté de Lewis and Garcia, 2003)

6.2 Histoire évolutive des venins

- **Une innovation apparue à plusieurs reprises**

L'innovation venimeuse est retrouvée dans un nombre impressionnant d'animaux parfois très éloignés d'un point de vue phylogénétique. L'essentiel des espèces venimeuses sur Terre est représenté par des arthropodes dont les Chélicérates (araignées, scorpions), les Chilopodes (myriapodes) et les Insectes. Ces derniers représentent à eux seuls 150 000 des 170 000 espèce venimeuses estimées (von Reumont et al., 2017). Outre ces espèces, d'autres taxons ou clades peuvent aussi être considérées comme venimeux : les sangsues, les tiques et les chauves-souris vampires qui injectent un cocktail anticoagulant pour se nourrir du sang de leur proie. De plus, grâce à l'observation de l'effet d'une morsure, à l'exploration de nouveaux biotopes et l'étude fine de l'anatomie des espèces, de nouveaux animaux continuent de compléter la liste des espèces venimeuses comme le Loris *Nycticebus sp.*, seul primate venimeux dont les glandes brachiales sécrètent des molécules allergisantes (Hagey et al., 2007) ou encore l'espèce de rémipède *Xibalbanus tulumensis,* seul crustacé venimeux découvert dans une doline du Yucatán au Mexique (von Reumont et al., 2017). Au total, la fonction venimeuse serait apparue de façon indépendante au moins à 20 reprises (**Figure 15**) (Casewell et al., 2013). Elle représente ainsi un parfait exemple d'évolution convergente donnant un avantage sélectif aux espèces venimeuses. Un tel succès évolutif est déterminé par des pressions de sélection fortes sur la composition du venin.

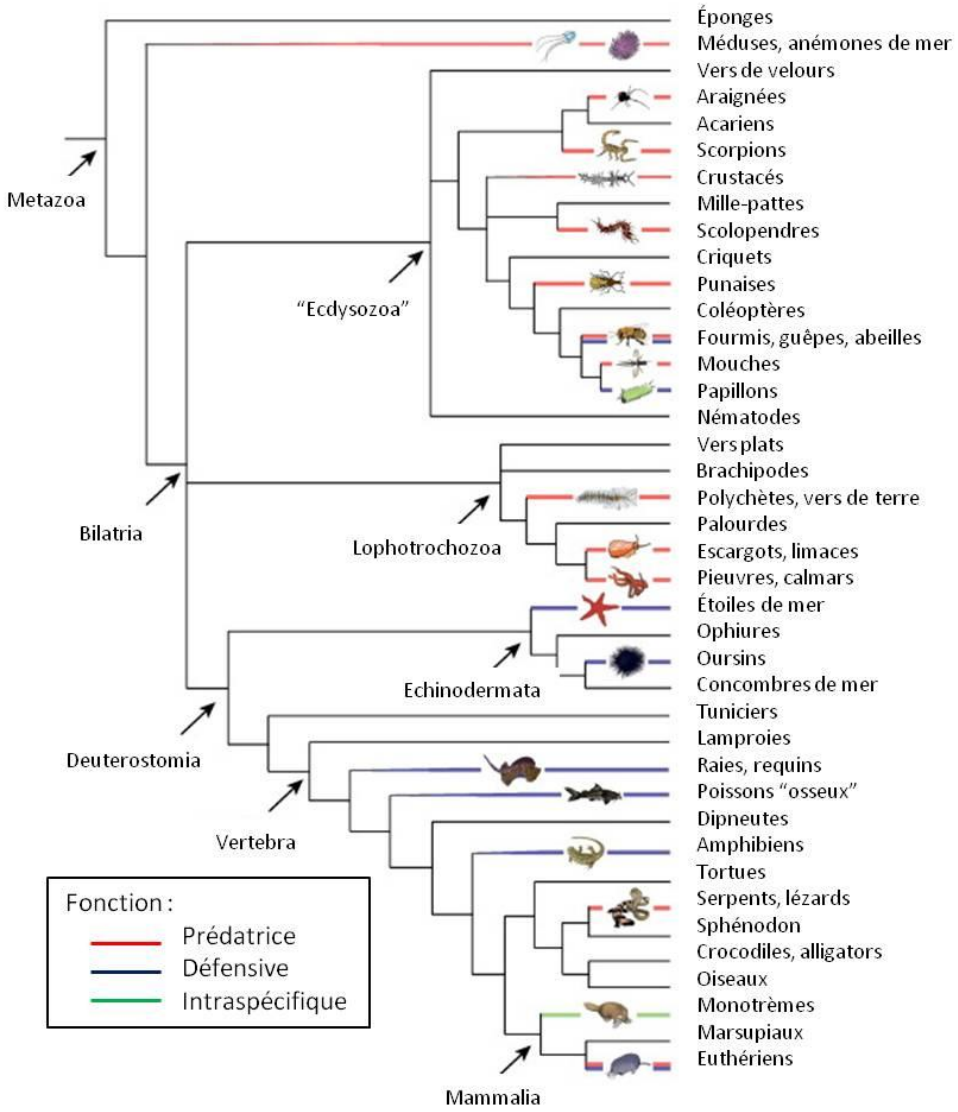


Figure 15. Arbre représentant les grands groupes d'animaux venimeux. Les branches colorées indiquent les lignées incluant des membres avec un système venimeux. (Adapté de Fry et al., 2015)

- **Pressions de sélection : moteur de la diversité moléculaire entre individus**

L'utilisation défensive ou prédatrice d'un venin détermine les pressions de sélection qui vont s'exercer sur l'espèce venimeuse. Un venin purement défensif comme chez les abeilles ou les poissons doit avoir une action immédiate causant une intense douleur localisée au point d'injection (Church and Hodgson, 2002; Peiren et al., 2005). Chez de nombreux groupes, la défense est une fonction secondaire du venin. Dans ce cas, aucune démonstration de l'effet de pressions de sélection sur la composition du venin n'a été faite à ce jour. Pour un venin à vocation prédatrice, c'est tout le contraire. Entre la proie et le prédateur venimeux s'exerce une réelle course à l'armement au cours de laquelle la résistance de la proie et l'arsenal de l'espèce venimeuse coévoluent par pression de sélection mutuelle selon l'hypothèse de la Reine Rouge (Van Valen, 1973). Cette hypothèse fait référence à un célèbre passage de « De l'autre côté du miroir », deuxième livre d'Alice au pays des merveilles de Lewis Carrol. Poursuivie par la Reine Rouge, Alice lui demande «Mais, Reine Rouge, c'est étrange, nous courons vite et le paysage autour de nous ne change

pas ?» La Reine répondit : «Ici il faut courir pour rester à la même place. Pour aller quelque part, il faudrait courir deux fois plus vite» (Carroll, 1871). La course à l'armement est continue jusqu'à l'extinction d'une des deux espèces : la proie (Alice) ou le prédateur venimeux (la Reine Rouge). La composition du venin est ainsi adaptative et peut dépendre des changements ontogéniques du régime alimentaire. Par exemple, le venin des juvéniles *Bothrops jararaca*, un crotale sud-américain, montre une toxicité plus accrue envers leurs proies de prédilection, les Amphibiens, que celui des adultes se nourrissant préférentiellement de Mammifères (Andrade and Abe, 1999). Des exemples similaires sont retrouvés chez les araignées mâles *Atrax robustus* qui produisent étant juvéniles un venin contre les Insectes mais qui une fois adultes, arrêtent de se nourrir pour se consacrer à la recherche d'une femelle et développent un venin défensif contre les vertébrés (Jelinek, 2001). Il était communément supposé qu'un prédateur venimeux injecte le même cocktail de toxines à une proie et à un assaillant. Une étude publiée en 2014 montre que la composition du venin injecté par un cône piscivore *Conus geographus* est en réalité différente suivant le stimulus extérieur (Dutertre et al., 2014). Dans une situation de défense, l'animal administre des toxines paralysantes bloquant les récepteurs neuromusculaires, expliquant l'effet mortel chez l'Homme. Au contraire, pour capturer une proie, le venin ne contient pas ces toxines et sa composition est spécifique de la proie, un poisson en l'occurrence. Un cas similaire est observé avec le prévenin du scorpion *Parabuthus* transvaalicus. Ce prévenin cause une douleur aigüe à des fins défensives alors que le venin plus riche en toxines est réservé à de grosses proies (Inceoglu et al., 2003).

Les pressions de sélection et notamment celles liées à la relation proie-prédateur conditionnent ainsi l'évolution du venin en maintenant ou pas dans les populations des évènements génétiques à l'origine de nouvelles toxines animales.

- **Quel est l'origine des gènes codant pour les toxines ?**

Dans ces évènements génétiques, la duplication génétique semble être un processus clé pour le recrutement de gènes dans la glande à venin. Cet évènement peut se produire à plusieurs reprises. Les mutations seront ensuite à l'origine de la néofonctionalistion du gène dans la glande à venin. Au cours de ce processus, certaines copies du gène peuvent dégénérer en pseudogènes ne conduisant plus à l'expression d'une protéine. Cependant la duplication est loin d'être le seul processus évolutif (**Figure 16**). Un gène physiologique peut accumuler suffisamment de mutations au cours du temps pour finir par coder pour une toxine, perdant ainsi sa fonction physiologique première. L'épissage alternatif est également un moteur clé d'évolution. C'est un mécanisme post-traductionnel impliquant la modification et le réarrangement d'exons. Cela a pour conséquence l'expression de différentes protéines (isoformes) à partir de la même séquence génétique réduisant ainsi le coût associé au maintien de familles multigéniques tout en augmentant la complexité du protéome. Ces isoformes peuvent agir sur différentes cibles et avoir des propriétés physicochimiques complètement différentes. En outre, des évènements plus complexes peuvent se produire comme la duplication ou la perte d'un ou plusieurs domaines d'un gène ancestral (Casewell et al., 2013). Tous ces évènements concourent à la diversité et la complexité moléculaire d'un venin. Si le gène code pour une toxine permettant une capture plus rapide de la proie, celui-ci sera sélectionné au cours des générations. On parle alors de pression de sélection positive.

Par ces différents mécanismes, les pressions évolutives ont façonné d'une part la biodiversité venimeuse et d'autre part la diversité moléculaire des venins. Un venin est efficace si les toxines atteignent rapidement leurs cibles pour permettre la capture de la proie ou la fuite de l'assaillant. Les venins ont ainsi évolué pour cibler des systèmes physiologiques atteignables par voie sanguine en particulier le système nerveux et le système cardiovasculaire.

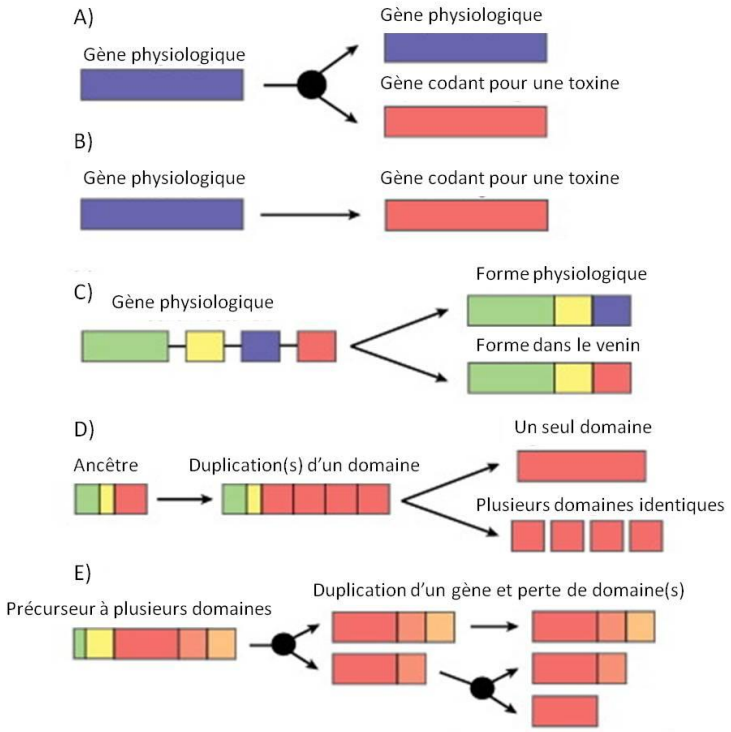


Figure 16. Mécanismes à l'origine de l'évolution des venins. A) Duplication d'un gène physiologique aboutissant à un gène codant pour une toxine. B) Modification d'un gène physiologique. C) Épissage alternatif d'un même gène donnant soit une toxine soit une autre protéine. D) Duplication d'un domaine ancestral issu d'un précurseur à plusieurs domaines. E) Perte de domaines d'un précurseur consécutive à la duplication d'un précurseur à plusieurs domaines donnant ainsi différentes possibilités de toxines. • : évènements de duplication de gène. (Adapté de Casewell et al., 2013)

- **Convergence des activités biologiques**

Comme le montre la **Figure 17**, il existe un incroyable degré de convergence évolutive entre les toxines de différentes espèces sur une même cible moléculaire. Par exemple, des toxines peuvent bloquer la transmission du signal nerveux et causer la paralysie ou la mort de la victime par asphyxie ou arrêt cardiaque. Cet effet peut être obtenu à différents niveaux de l'influx nerveux : avant les synapses en bloquant par exemple l'activité des canaux sodiques, au niveau de l'axone en agissant sur les canaux calciques ou en post-synaptique de la jonction neuromusculaire via les récepteurs à l'acétylcholine. Les effets cardiovasculaires sont quant à eux généralement liés à une perturbation à différentes étapes de l'hémostase conduisant à des hémorragies ou au contraire à la formation de caillots.

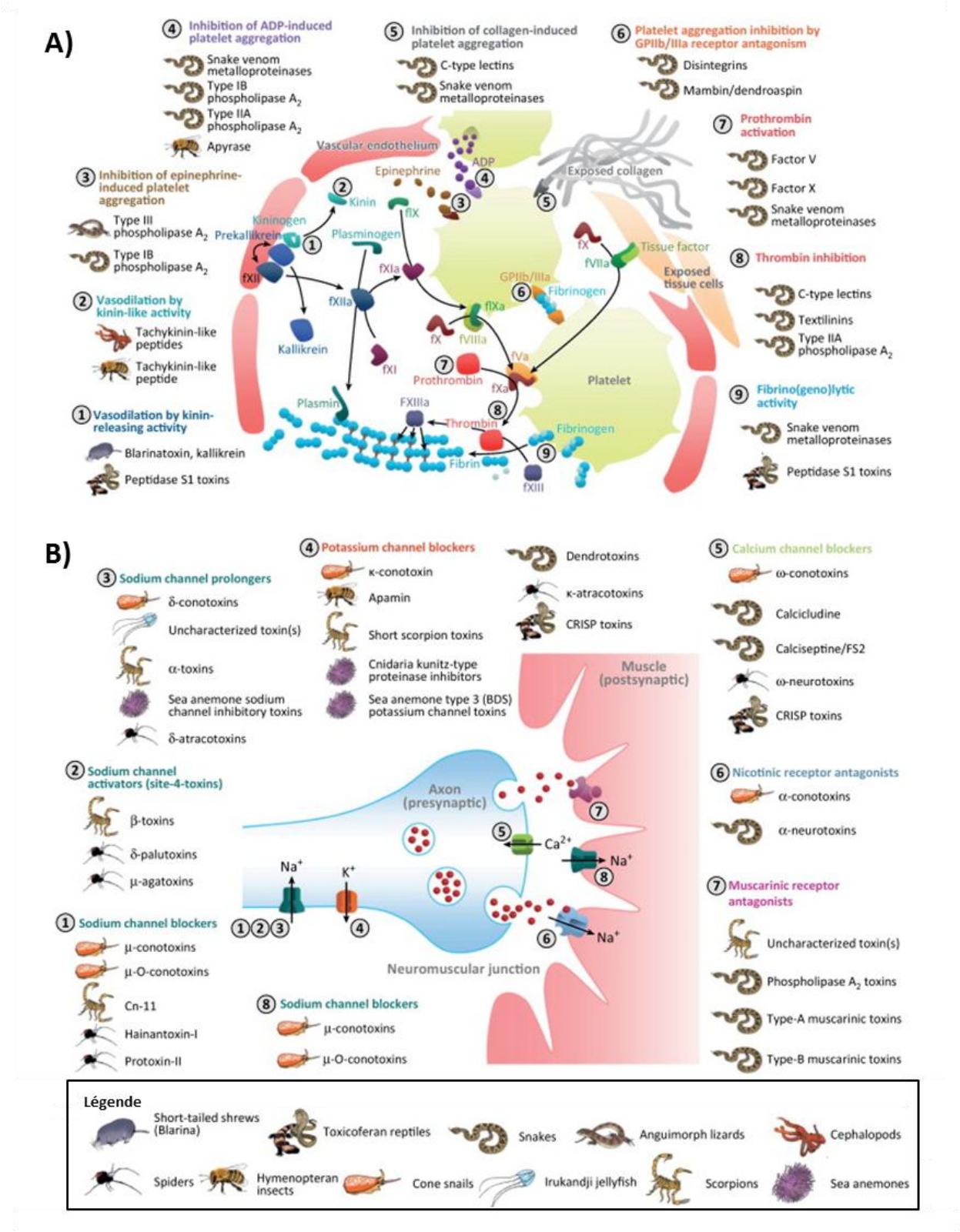


Figure 17. Convergence évolutive des activités des toxines animales. A) Sites de convergence des toxines hémotoxiques numérotés de 1 à 9. B) Sites de convergence des toxines neurotoxiques numérotés de 1 à 8. Les noms des toxines responsables de l'activité sur une cible sont listés sous chaque cible moléculaire et associés à une image de l'animal producteur. La signification des images est décrite dans l'encadré « légende ». (Adapté de Casewell et al., 2013)

Au cours de l'évolution, les toxines animales se sont également caractérisées par leur richesse en ponts disulfure, ces derniers étant formés par l'oxydation des fonctions thiols de deux cystéines. Nous allons voir dans la prochaine partie le rôle que jouent ces ponts et l'importance de leur arrangement dans l'activité d'une toxine.

6.3 Convergence et diversité des motifs de cystéines

- **Statistiques**

Les toxines animales sont connues pour leur forte réticulation. Le nombre de ponts disulfure peut grandement varier en fonction de la taille du peptide. Les contryphans (moins de 10 résidus) de cônes présentent par exemple un seul pont disulfure alors que les disintégrines (> 60 résidus) de serpents ou les Phospholipases A2 (PLA2) (>100 résidus) possèdent respectivement 12 et 14 cystéines. Les ponts disulfure impliquent le plus souvent deux cystéines d'une même séquence peptidique formant ainsi un pont intra-chaine structurant le peptide. Néanmoins, certaines toxines peuvent posséder un nombre impair de cystéines. Dans ce cas, la cystéine libre peut être impliquée dans un pont inter-chaine d'un peptide multimérique (Quinton et al., 2009) ou porter différentes modifications post-traductionnelles (PTMs) comme une cystéinylation (Robinson and Norton, 2014), des lipidations (par exemples : S-acylation or S-prénylation) (Tate et al., 2015), une ADP ribosylation (McDonald and Moss, 1994) ou encore une nitrosylation (Qin et al., 2013). Bien que ces modifications modulent potentiellement l'activité de la protéine, elles restent peu caractérisées chez les toxines et ne seront pas présentées dans cette partie. En s'affranchissant de ces cas particuliers, Lavergnes et al. ont identifié dans les bases de données UniProtKB/Swiss-Prot et TrEMBL (2015) 2 022 toxines animales. Au total, les auteurs ont comptabilisé 44 motifs différents, 86 en tenant compte de l'arrangement des ponts (Lavergne et al., 2015a). Plus de la moitié sont retrouvés dans au moins deux clades venimeux différents et correspondent à seulement 14 familles de toxines (Fry et al., 2009). On peut ainsi différencier les motifs largement répandus au sein de différentes lignées venimeuses de ceux qui sont propres à certains clades.

- **Motifs de cytéines les plus répandues**

Ces motifs témoignent d'un formidable recrutement convergent au sein de différentes lignées venimeuses. Parmi eux, deux sont particulièrement abondants dans les venins et couvrent deux cas de lien structure-fonction. Le premier, le motif Inhibitor Cystine Knot (ICK), montre une flexibilité dans la longueur des boucles associée à une diversité d'activités. Au contraire, le motif Kunitz est caractérisé par une conservation de la longueur des boucles avec une diversité d'activité beaucoup plus restreinte.

  - Motif Inhibitor Cystine Knot (ICK)
    - Origine animale

Historiquement, le motif ICK a été pour la première fois identifié dans l'inhibiteur PCI de la carboxypeptidase A de la pomme de terre (Rees and Lipscomb, 1982). Avec la découverte d'autres protéines possédant ce motif comme les inhibiteurs de la trypsine des Cucurbitaceae, le nom générique de « knottin » a fait son apparition (Le Nguyen et al., 1990). Par la suite, il a été retrouvé dans nombre de protéines et de clades différents : les cyclotides de plantes, les protéines Agouti-related peptide et Agouti-signaling peptide (ASIP) des mammifères et bien sûr dans les toxines animales (Postic et al., 2018). Ces dernières

représentent 15% des toxines identifiées dans UniProtKB/Swiss-Prot et TrEMBL en 2015 par Lavergnes et al. En 2019, la base de données KNOTTIN (Postic et al., 2018) comptabilise quant à elle 2 107 séquences ICK présentes dans les clades venimeux. Cependant, la base ne tient pas compte de la longueur des peptides et ne rassemble donc pas uniquement des toxines. Dans les deux cas, les toxines ICK sont retrouvées majoritairement dans les venins d'araignées mais aussi chez les cônes (Robinson and Norton, 2014) et les scorpions (Chen et al.; Gao et al., 2013).

- Structure

Les ICK sont définies par la séquence consensus suivante :
-C-x(3–7)-C-x(3–8)-C-x(0–7)-C-x(1–4)-C-x(4–13)-C- où x est n'importe quel acide aminé (Norton and Pallaghy, 1998). La structure qui en découle est caractérisée par deux ou trois feuillets Bêta antiparallèles stabilisés par trois ponts disulfure avec un arrangement en C1-C4, C2-C5 et C3-C6 (**Figure 18**). Les deux premiers ponts forment un anneau traversé par le troisième formant ainsi un nœud, d'où le nom de la famille. Cette structuration confère aux toxines ICK une forte stabilité chimique, thermique et biologique (Saez et al., 2010).

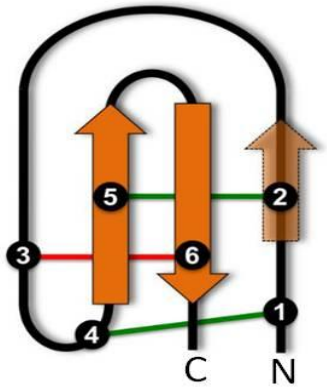


Figure 18. Représentation schématique de la structure canonique d'une toxine ICK. Les feuillets Bêta sont représentés en orange. Les chiffres correspondent à l'ordre de l'extrémité N-terminale (N) à l'extrémité C-terminale (C) des cystéines impliquées dans la formation des ponts disulfure. Le pont formant le nœud (knot) est représenté en rouge, les deux autres sont en vert. Le feuillet impliquant la seconde cystéine est souvent absent des toxines d'araignées. (Adapté de Saez et al., 2010)

- Activités pharmacologiques

Les toxines ICK ont été décrites exclusivement pour leur activité sur les canaux ioniques (Zhu et al., 2003). Celles issues d'araignées ou de cônes ciblent toute une diversité de canaux, incluant des canaux sodiques, potassiques, calciques, ... Celles de scorpions ne représentent qu'une faible part de la diversité moléculaire de leur venin avec une activité ciblée sur les canaux potassiques (Gao et al., 2013).

Canaux Na_v :

Parmi les canaux ioniques, les cibles préférentielles des ICKs d'araignées sont les canaux sodiques dépendant du potentiel (Na_v) suggérant une pression de sélection positive forte vers ces cibles pour capturer des insectes. Un pourcentage d'identité de séquences élevé entre les canaux sodiques Na_v d'insectes et de vertébrés expliquerait l'activité de ces toxines sur les canaux humains. Les canaux Na_v sont responsables de la dépolarisation rapide des cellules excitables nécessaire à l'initiation du potentiel d'action. Il existe différents sous-types de canaux Na_v, le plus étudié étant le $Na_v1.7$. Des données génétiques montrent son rôle crucial dans la nociception. En effet des mutations gain de fonction entrainent des neuropathies tandis que des pertes de fonctions sont responsables d'une insensibilité congénitale à toute forme de douleur (Cox et al., 2006; Fertleman et al., 2006). Ainsi $Na_v1.7$ est devenu depuis plusieurs années la cible thérapeutique de choix pour découvrir de nouveaux analgésiques, notamment à partir de venins. Sur les 12 familles de

toxines d'araignées actives sur les canaux Na_v (NaSpTx) (**Figure 19**), 8 possèdent un motif ICK (Klint et al., 2012). Toutes les araignées produisant ces toxines appartiennent à la famille des Theraphosidae du sous-ordre des Mygalomorphes (mygales).

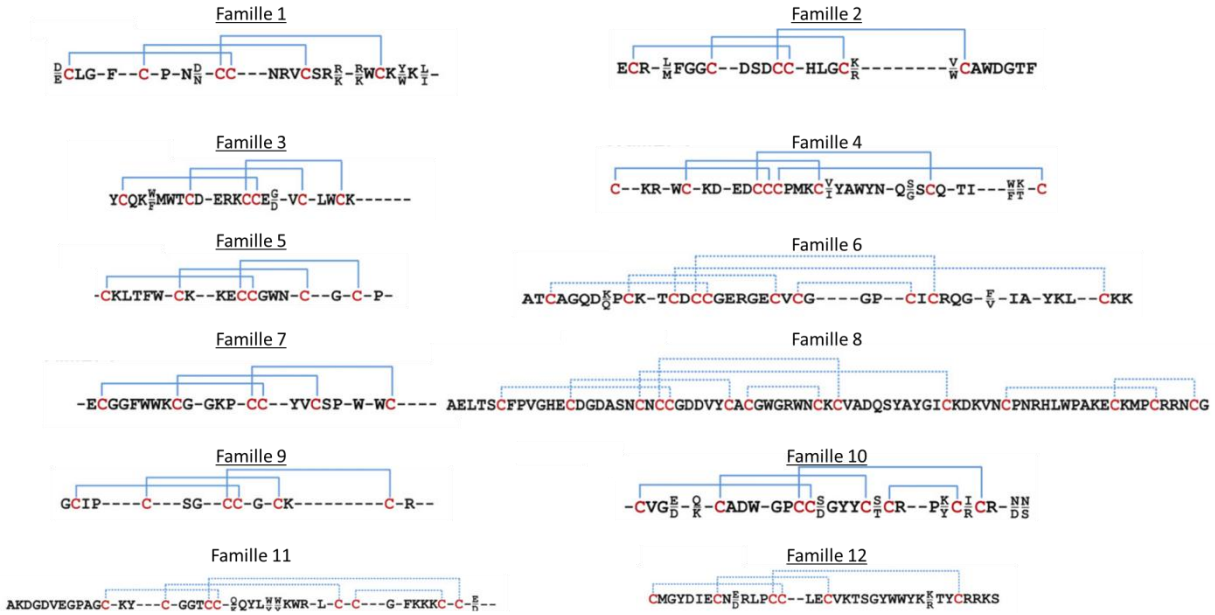


Figure 19. Séquences consensus des familles de NaSpTx. La connectivité des ponts est indiquée en bleu. Les familles présentant un motif ICK sont soulignées. Les pointillés représentent la connectivité théorique qui n'a pas été validée expérimentalement. (Adapté de Klint et al., 2012)

Canaux calciques Ca_v *:*

Tout comme les canaux Na_v, les canaux calciques dépendant du potentiel sont des cibles majeures des toxines d'araignées (King and Hardy, 2012). Sur le plan structural, elles sont très diversifiées. Comme exemple d'ICK, on peut citer l'ω-TRTX-Hh1a (Huwentoxine X) de *Cyriopagopus schmidti* (Liu et al., 2006). Cependant, l'essentiel des ICK caractérisées sur ces canaux proviennent des venins de cônes et plus particulièrement de la famille des ω-conotoxines retrouvées essentiellement chez les espèces piscivores. Cette famille a sans doute évolué pour faciliter la capture d'une proie via l'induction d'une paralysie flasque en agissant sur les canaux calciques (Olivera et al., 1985). Les plus connues sont MVIIA de *Conus magus* et GVIA de *Conus geographus*(Favreau et al., 2001; Wang et al., 2016). Bien que partageant un faible pourcentage d'identité (45% avec les cystéines, 18% sans les cystéines), elles ciblent toutes les deux le sous-type $Ca_v2.2$ au niveau nanomolaire pour MVIIA ou picomolaire pour GVIA. Le canal $Ca_v2.2$ est retrouvé majoritairement au niveau des terminaisons nerveuses présynaptiques du système nerveux central (SNC) et périphérique (SNP) où il initie la libération de neurotransmetteur dans la fente synaptique (Sousa et al., 2013). Il joue ainsi un rôle important dans la transmission ascendante du message douloureux. Le pharmacophore des ω-conotoxines sur $Ca_v2.2$ reposerait sur Tyr13 et Lys2, deux acides aminés hautement conservés. Néanmoins, ces résidus sont aussi présents dans les ω-conotoxines actives sur $Ca_v2.1$ (Sousa et al., 2013). Les déterminants de la sélectivité entre les deux sous-types n'ont pas encore été élucidés.

Canaux potassiques K_v:

Les canaux potassiques rassemblent une grande famille de canaux ioniques qui transportent sélectivement les ions K^+ à travers les membranes des cellules excitables (neurones ou cellules musculaires) et non excitables. Ce transport est crucial pour la génération du potentiel d'action et le maintien du potentiel membranaire de repos. L'inhibition des canaux potassiques dépendant du potentiel K_v par des toxines animales peut ainsi accroitre l'excitabilité des neurones, entrainant la mort rapide de la proie. L'avantage évolutif évident explique l'occurrence dans différents clades venimeux de toxines dirigées contre les canaux K_v comme les κ-conotoxines des cônes ou les κ-théraphotoxines d'araignées. Cependant, seule une partie d'entre elles possèdent un motif ICK à l'instar de PVIIA, une toxine de *Conus purpurascens* et de κ-theraphotoxin-Pg1a de *Chilobrachys guangxiensis* (Herrington et al., 2006; Shon et al., 1998). De façon plus anecdotique, des ICK de scorpions montrent une faible activité micromolaire sur ces canaux. ImKTx1 d'*Isometrus maculatus* a été le premier membre fondateur de la famille des toxines ICK de scorpion actives sur les canaux potassiques (λ-KTx) (Chen et al.). Un deuxième membre, MeuKTx, a par la suite été identifiée dans le venin de *Mesobuthus eupeus* (Gao et al., 2013).

Autres cibles

Outre les cibles décrites précédemment, les toxines ICK ont été caractérisées sur d'autres cibles liées à la nociception. Par exemple, π-TRTX-Pc1a (Psalmotoxin-1) cible le canal ASIC1a (Acid sensing ion channel 1a) avec un IC50 de 0.9 nM (Escoubas et al., 2000). Ce canal est sensible aux variations de pH liées à l'acidose tissulaire causée par une inflammation ou une blessure (Gründer and Chen, 2010). D'après la base de données Arachnoserver, cinq ICK d'araignées sont actives à des concentrations nanomolaires sur le récepteur 1 vanilloïde à potentiel transitoire (TRPV1), un canal ionotropique exprimé à l'extrémité périphérique des neurones sensitifs. Elles proviennent du venin des Theraphosidae *Grammmostola roseae et Psalmopeus cambridgei* (Siemens et al., 2006). Une autre toxine de tarentule, GsMTx4 de *Grammostola spatulata*, agit quant à elle sur les canaux mécanosenseurs des neurones sensitifs réduisant la douleur mécanique et neuropathique chez des rats (Park et al., 2008). Les toxines ICK présentent d'autres activités comme l'induction de l'érection chez des rats par la toxine d'araignée Tx2-6. Cet effet serait sans doute médié par la relaxation du muscle lisse du *corpus cavernosum* du pénis induit par la libération de monoxyde d'azote (Nunes et al., 2008). Enfin, les toxines Psalmopetoxines I et II de *Psalmopoeus cambridgei* ont un effet antiplasmodique en inhibant le développement intra-érytrocytaire du parasite *Plasmodium falciparum* (Choi et al., 2004).

- Motif Kunitz
  - Origine animale

L'inhibiteur pancréatique bovin de la trypsine (BPTI ou aprotinine), membre fondateur de la famille des protéases du type Kunitz, a été décrit en 1936 (Kunitz and John, 1936). Les peptides Kunitz découverts par la suite sont ubiquitaires et sont décrits dans de nombreux organismes que ce soient des animaux, des plantes ou des micro-organismes. En ce qui concerne les espèces venimeuses, la première toxine du type Kunitz (KTT) a été isolée du venin du mamba noir *Dendroaspis polylepis polylepis* en 1974. Après cette découverte, d'autres KTT ont été identifiées dans le venin de nombreux autres groupes incluant les serpents, les cônes, les araignées, les scorpions et les arthropodes hématophages comme les tiques (Mourão and Schwartz, 2013). À ce jour, 255 KTT sont répertoriées dans la base de

données UniProtKB/Swiss-Prot.

- Structure

Les KTT sont des peptides de 50-60 acides aminés stabilisés par trois ponts disulfure et contenant le motif -C-C-C-C-C-C- avec une connectivité en C1-C6, C2-C4 et C3-C5 (**Figure 20A**). Les longueurs des boucles sont hautement conservées : x2-C-x2-C-#-x3-C-x3-C-x6-C-x4-C-x2 où x est n'importe quel acide aminé et # le résidu du site actif responsable de l'inhibition des sérine protéases (généralement une lysine Lys) (Laskowski and Kato, 1980). Toutes les KTT adoptent une structure α/β appelé le domaine Kunitz. La structure est allongée avec une hélice alpha dans la région C-terminale connectée à un des deux/trois feuillets Bêta centraux ainsi qu'à l'extrémité N-terminale via les ponts disulfure C1-C6 et C3-C5. Le pont restant stabilise la région flexible située à l'opposé des extrémités N et C-terminales (**Figure 20B**). Ce dernier pont est parfois absent comme chez la Conkunitz de *Conus consor* sans affecter son activité (Bayrhuber et al., 2005).

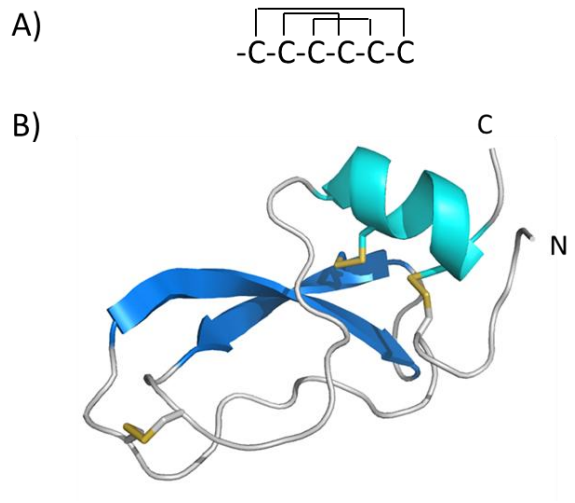


Figure 20. Caractéristiques structurales des toxines Kunitz. A) Motif de cystéines caractéristique avec 3 ponts disulfure. B) Structure RMN de la dendrotoxin I de *Dendroaspis polylepis polylepis*. Les extrémités N et C-terminales sont notées. Les ponts disulfure sont représentés en jaune. Numéro pdb : 1DEM (Lancelin et al., 1994)

- Activités pharmacologiques

Inhibiteur de sérines protéases

La boucle correspondant aux positions 13 à 18 du BPTI a été montrée comme importante dans l'activité inhibitrice de protéases. Selon la nomenclature des sites catalytiques peptidiques, ces positions sont notées P3-P2-P1-P'1-P'2-P'3 (Schechter and Berger, 1967), P1 étant la position la plus déterminante pour la reconnaissance de la protéase. Les inhibiteurs de la trypsine ont typiquement un résidu basique (arginine ou lysine) à cette position tandis que les inhibiteurs de la chymotrypsine ont une leucine ou une méthionine plus hydrophobe. Cette activité est souvent retrouvée dans les sécrétions salivaires des arthropodes hématophages comme les tiques, à l'exemple de l'espèce *Amblyomma cajennens* (Waxman et al., 1990). Dans son venin a été découvert par transcriptomique l'ambyomine X, un inhibiteur Kunitz du facteur Xa impliqué dans la coagulation mais qui montre également un effet cytotoxique sur plusieurs lignées tumorales (Chudzinski-Tavassi et al., 2010).

Canaux potassiques

Au cours de l'évolution, certains inhibiteurs de protéases ont pu acquérir une fonction neurotoxique en ciblant les canaux potassiques. C'est pourquoi on retrouve des KTT

avec une fonction duale dans les venins d'anémones de mer, de scorpions et d'araignées alors que celles ciblant spécifiquement les canaux potassiques, en particulier $K_v1.1$, sont présentes dans les venins de serpents. Découvertes d'abord dans les venins de mamba, elles portent le nom de dendrotoxines. La dendrotoxine K du mamba noir *Dendroaspis polylepis polylepis* bloquent par exemple le canal $K_v1.1$ à des concentrations picomolaires (Owen et al., 1997). Le pharmacophore est situé à l'extrémité N-terminale dont les résidus les plus importants sont la Lys5 et la Leu9 (Gasparini et al., 1998; Harvey et al., 1997).

Autres activités

Certaines KTT peuvent porter des activités autres que les principales présentées précédemment. Elles concernent principalement d'autres canaux ioniques. La calicludine de *Dendroaspis anguticeps* inhibe les canaux calciques Ca_v1 avec un IC50 nanomolaire sans montrer d'activité contre les protéases ou les canaux potassiques (Schweitz et al., 2006). Une activité visant les canaux TRPV1 a également été mise en évidence chez deux KTTs de l'anémone de mer *Heteractis crispa* possédant une thréonine en P1 d'où une faible activité inhibitrice de la trypsine (Gladkikh et al., 2015). Enfin une toxine du scorpion *Micrurus tener tener* active les canaux ASIC à l'origine d'une intense douleur en cas d'envenimation (Bohlen et al., 2011). Cependant, le motif Kunitz pourrait avoir un spectre d'activité plus large. En effet, une toxine de *Dendroaspis anguticeps,* la mambaquarétine*,* inhibe le récepteur à la vasopressine de type 2 sans présenter les activités des autres KTTs (Ciolek et al., 2017).

Contrairement aux toxines ICK et les KTTs, d'autres familles de toxines sont jusqu'à présent spécifiques d'un clade particulier. Certaines peuvent être associées à une grande diversité d'activité comme les toxines à trois doigts des serpents tandis que d'autres sont plus spécifiques d'une cible moléculaire comme les toxines Cysteine-stabilized-α-helix and β-sheet (CSαβ) de scorpions ou certains motifs de toxines de cônes.

- **Familles spécifiques à certains groupes d'espèces**
    - Toxines à trois doigts des serpents (3FTs) (clade : Serpentes)
        - *Structure*

Toutes les 3FTs décrites à ce jour ont été identifiées dans les venins de serpents et représentent l'une des familles de toxines les plus grandes de ce clade. Elles peuvent représenter jusqu'à 95% des toxines présentes dans le venin du serpent (Sanz et al., 2016). De 57 à 82 résidus, elles doivent leur nom à leur structure 3D particulière : trois boucles ou doigts issues d'un cœur globulaire hydrophobe, le tout stabilisé principalement par 4 ponts disulfure (**Figure 21**). Certaines ont un pont supplémentaire au niveau du doigt I ou II, la position affectant l'activité de la toxine. D'autres sont capables de former des homodimères liés de manière non covalente comme l'haditoxin ou covalente comme l'α-cobratoxine. Cette dernière est également capable de former un hétérodimère avec différentes cytotoxines modifiant son activité (Utkin, 2018). D'une manière générale, la dimérisation des 3FTs est requise pour la liaison sur le récepteur à l'acétycholine α3β2.

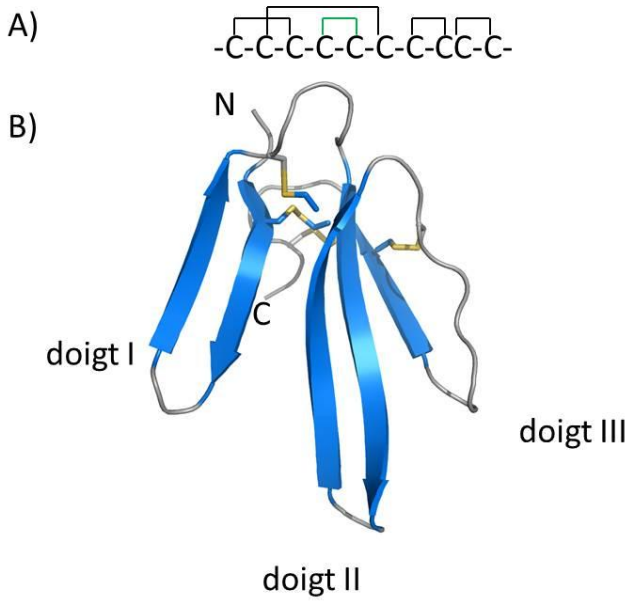


Figure 21. Caractéristiques structurales des toxines à 3 doigts de serpent. A) Motif de cystéines caractéristique avec 4 ponts disulfure. Un cinquième pont représenté en vert est parfois présent. L'arrangement des ponts est décrit. B) Structure RMN d'une toxine à trois doigts de serpent : ρDa1a de *Dendroaspis anguticeps.* Les trois doigts et les extrémités N et C-terminales sont indiquées. Les ponts disulfure sont représentés en jaune. Numéro pdb : 4IYE (Maïga et al., 2013)

- *Activités pharmacologiques*

Les toxines 3FTs représentent la famille de toxines avec le plus large spectre d'activité. Elles ciblent des récepteurs canaux, des canaux ioniques dépendant du potentiel, des RCPGs et des enzymes impliquées dans différentes fonctions physiologiques.

Signal cholinergique

Les 3FTs ont d'abord été décrites pour leur activité neurotoxique en ciblant la transmission cholinergique à différents sites post-synaptiques du SNC et SNP. L'α-bungarotoxine, première 3FT découverte en 1963 dans le venin de *Bungarus multicinctus*, montre une activité sur les récepteurs nicotiniques à l'acétylcholine. Ces neurotoxines peuvent être classées en trois groupes suivant la sélectivité pour un récepteur donné. Les α-neurotoxines se lient au récepteur nicotinique musculaire α1 nAChR, empêchant ainsi l'acétylcholine de se lier à son récepteur. Elles préviennent par conséquent la transmission neuromusculaire (Kini and Doley, 2010). Parmi les neurotoxines, on distingue les neurotoxines courtes (60-62 résidus et 4 ponts disulfure) des longues (66-74 résidus et 5 ponts disulfure) (Kini and Doley, 2010). Les deux lient le même site du récepteur mais le cinquième pont sur le doigt II et les différences au niveau de l'extrémité C-terminale pourraient expliquer des activités non partagées (Kini and Doley, 2010). Par exemple, les neurotoxines longues peuvent lier avec une haute affinité le récepteur neuronal α7 nAChR, ce qui n'est pas le cas pour les courtes. Le deuxième groupe, les κ-neurotoxines, reconnaissent les récepteurs neuronaux nAChR α3β4, α3β2 et α4β2 (Chiappinelli and Wolf, 1989; Grant and Chiappinelli, 1985). Elles sont proches structurellement des α-neurotoxines longues avec un cinquième pont disulfure dans le doigt II mais sont présentes sous forme de dimère. Le dernier groupe concerne les toxines muscariniques ciblant les récepteurs muscariniques à l'acétylcholine (mAChR). Certaines sont agonistes, d'autres antagonistes. MT7 issue de la même espèce de serpent est quant à elle active uniquement sur le récepteur M1 (Mourier et al., 2003). Bien que ne partageant pas le même profil pharmacologique, les toxines muscariniques sont proches structurellement des α-neurotoxines courtes.

Une autre classe de 3FT interfère également avec la transmission neuromusculaire en inhibant l'acétylcholine estérase entrainant une accumulation du neurotransmetteur dans la

synapse causant une fasciculation[6] du muscle (Bourne et al., 1995). Elles sont proches structurellement des α-neurotoxines courtes.

Les 3FTs peuvent aussi agir sur la transmission du signal nociceptif. La mambalgine, identifiée dans le venin du mamba noir, cible spécifiquement le canal ASIC1a. Elle montre un puissant effet analgésique chez les souris sans entrainer les effets secondaires observés avec des opioïdes comme une détresse respiratoire (Diochot et al., 2012).

D'autres activités liées à la transmission du signal nerveux ont été démontrées. Par exemple, l'α-bungaroxine ou l'α-cobratoxine agissent également sur les récepteurs ionotropiques à l'acide γ-aminobutyrique (GABA) (McCann et al., 2006). La calciseptine de *Dendroaspis polylepis*, proche des α-neurotoxines, bloquent les canaux calciques. Alors qu'on pensait les récepteurs muscariniques à l'acétylcholine, les seuls cibles RCPG des 3FTs, certaines toxines ciblent les récepteurs adrénergiques à l'instar de ρ-Da1a et ρ-Da1b respectivement sur les récepteur α1A et α2A (Maïga et al., 2012; Quinton et al., 2010; Rouget et al., 2010).

Effet cardiotoxique/cytotoxique

Les cardiotoxines sont retrouvées uniquement dans les venins de cobra et représentent le second plus grand groupe de 3FTs après les α-neurotoxines. Structurellement, elles ressemblent aux α-neurotoxines courtes (59-62 résidus, 4 ponts disulfure) (Kini and Doley, 2010). Injectées, elles tuent l'animal en induisant un arrêt cardiaque (Dufton and Hider, 1988). La plupart montrent également un effet cytolytique en formant des pores à travers la membrane (Dufton and Hider, 1988).

Coagulation du sang

Plusieurs 3FTs montrent un effet sur l'aggrégation plaquettaire (hémostase primaire) comme la Dendroaspine de *Dendroaspis jamesoni* (McDowell et al., 1992). Possédant le tripeptide RGD, elle inhibe l'interaction entre le fribirnogène et son récepteur.

L'hémostase secondaire peut aussi être ciblée par les 3FTs. Deux toxines du cobra *Hemachatus haemachatus* inhibent l'activation du facteur X en facteur Xa.

Inhibteurs de la Phospholipase A2 (PLA2)

En inhibant la PLA2, des 3FTs neutraliseraient l'effet toxique de cette enzyme présente dans le venin de l'animal (So et al., 2008) et constitueraient donc un moyen de défense en cas de morsures d'autres serpents.

De façon plus anecdotique, la cardiotoxine I du cobra *Naja kaouthia*, induit également la sécrétion d'insuline dans des cellules INS-1E sans affecter la viabilité et l'intégrité des cellules (Utkin, 2018). Un autre exemple sortant des activités toxiques est l'actiflagéline isolée du venin deu cobra *Walterinnesia aegyptia* et activant *in vitro* la motilité des spermatozoïdes de souris (Abd El-Aziz et al., 2018).

Contrairement à la famille précédente, d'autres groupes de toxines spécifiques d'un clade montrent une plus grande spécificité pour une cible moléculaire.

[6] Contraction anormale du muscle.

- Toxines Cysteine-stabilized-α-helix and β-sheet (CSαβ) de scorpions
  - *Structure*

Les toxines CSαβ de scorpions représentent la majorité des toxines actives sur les canaux potassiques de ce clade (KTx). Ce groupe est divisé en trois familles : les α-KTx de 20 à 40 résidus, les β-KTx de 45 à 75 résidus et les γ-KTx de 35 à 45 résidus ayant une sélectivité différente des α-KTx. La famille alpha est la plus connue avec 280 des 295 KTx décrites dans UniProtKB/SwissProt en 2019. Elle rassemble environ 27 sous-familles classées d'après leur structure primaire et notées αKTx1-27 (Tytgat et al., 1999). Le motif consensus est le suivant : C-x(4,5)-C-x(3)-C-x(6,9)-[GAS]-K-C-[IMQT]-x(3)-C-x-C (Prosite PS01138) avec trois ponts stabilisant une hélice alpha et 2 ou 3 feuillets Bêta selon une connectivité en C1-C4, C2-C5, C3-C6 (**Figure 22A**). Un quatrième pont disulfure est parfois observé comme c'est le cas pour la Butantoxine (α-KTx12.1) et la Maurotoxine (α-KTx2.3). La structure en CSαβ est également retrouvée dans les défensines d'autres arthropodes, de nématodes, de plantes et de champignons (Zhu et al., 2005). Les α-Ktx dériveraient de ces peptides aux propriétés antimicrobiennes. En effet, il a été montré qu'un petit nombre de mutations étaient nécessaires pour convertir une défensine CSαβ en neurotoxine (Zhu et al., 2014).

  - Activités pharmacologiques

Les α-KTx sont essentiellement connues pour leur activité sur les canaux K_v, notamment la famille K_v1, mais également sur les canaux potassiques activés par du calcium. Cette activité serait permise par une dyade fonctionnelle composée de la lysine K27 (numérotation de la Charybdotoxin α-KTx1.1) et d'un aromatique (F ou Y) ou un aliphatique (L) séparés de 6.6 Å (Dauplais et al., 1997; Gasparini et al., 1998). Dans ce modèle, la chaine latérale de la lysine entrerait dans le pore du canal et interagirait avec les chaines de résidus acides (D, E) empêchant ainsi l'activation du canal par l'ion K^+. L'importance de cette lysine a été confirmée à la fois par des études de Structure-Activity Relationship (SAR) dans lesquelles une inversion de charge entraine une baisse dramatique de l'affinité et par la résolution de la structure 3D du complexe entre la Charybdotoxine (α-KTx1.1) et un canal chimérique $K_v1.2$-$K_v2.1$ (Banerjee et al., 2013; Carlier et al., 2000). Dans la majorité des α-KTx, le rôle de l'aromatique serait joué par une tyrosine Y en extrémité C-terminale comme il a été démontré pour la Charybdotoxine (**Figure 22B**) ou par une phénylalanine comme c'est le cas pour la F25 de l'Agitoxine-2 (α-KTx3.2) (Dauplais et al., 1997; Mackinnon et al., 2011). Cependant, la découverte de nouvelles sous-familles de KTx montrent que la dyade n'est pas forcément requise pour une activité sur les canaux potassiques. Par exemple, la sous-famille α-KTx18 dont Tc32 est la représentante, possède une sérine à la place de la lysine et aucun aromatique tout en montrant une affinité nanomolaire sur $K_v1.3$ (Batista et al., 2002). Des expériences de docking montrent tout de même que K34 rentrerait dans le pore du canal (Stehling et al., 2012). Un autre exemple est donné par la toxine Pi1 (α-KTx6.1). Cette dernière inhibe avec un IC50 de 0.44 nM le canal $K_v1.2$. L'analogue de la toxine sans la dyade présente encore une haute affinité (30 nM) démontrant l'importance d'autres résidus qui formeraient un anneau basique (Mouhat et al., 2004). En effet, l'immense majorité des toxines de scorpions présentent un point isoélectrique basique (pI) de 8. L'importance de cet anneau basique est confirmée par l'étude des Parabutoxines 1,2 et 10 (respectivement α-KTx11.1, α-KTx11.2, α-KTx11.3) caractérisés par un pI acide de 3.82 et 3.88. Aucune ne présente la dyade et l'introduction de celle-ci dans la séquence de la Parabutoxine-2 à des positions similaires à celles de l'Agitoxine-2 n'améliore que d'un facteur 10 l'affinité micromolaire sur les canaux K_v1 (Huys et al., 2004).

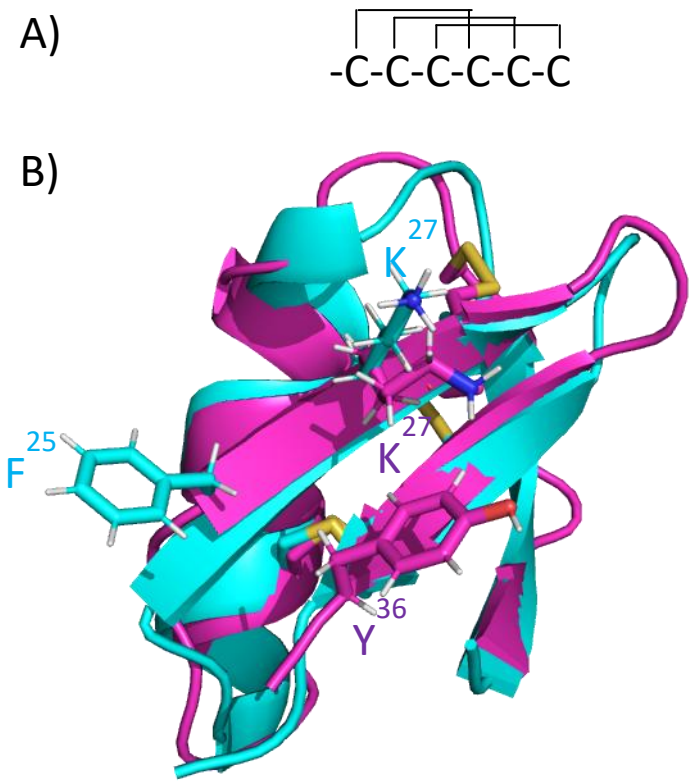


Figure 22. Caractéristiques structurales des α-Ktx de scorpion. A) Motif de cystéines caractéristique avec 3 ponts disulfure. L'arrangement des ponts est décrit. B) Superposition des structures RMN de la Charybdotoxine (α-KTx1.1, en magenta, pdb2CRD) et de l'Agitoxine-2 (α-KTx3.2, en cyan, pdb1AGT). Les ponts disulfure sont représentés en jaune. Les chaines latérales de la lysine et de l'aromatique formant la dyade fonctionnelles sont développées et leur position au sein de la séquence est notée. (Bontems et al., 1992; Krezel et al., 1995)

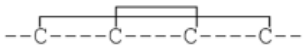
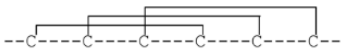
- Conotoxines

Les toxines de cônes ou conotoxines montrent plusieurs motifs de cystéines qui leur sont propres. Avec 293 peptides décrits en 2013, le premier est -CC-C-C- relatif aux α-conotoxines (Akondi et al., 2014). La majorité de ces peptides appartiennent à la superfamille A avec l'appariement globulaire en C1-C3, C2-C4. Un troisième pont peut également être présent à l'exemple de la toxine PIVA du cône piscivore *Conus purpurascens*. Comme pour la plupart des α-conotoxines, PIVA est un inhibiteur des récepteurs nicotiniques à l'acétylcholine post-synaptiques, en particulier le sous-type α1β1γδ (Teichert et al., 2006). Les ε-conotoxines représentent également une part importante avec 128 séquences décrites. Appartenant à la superfamille T, elles partagent le motif –CC-x(4-7)-CC- où x correspond à n'importe quel acide aminé. Contrairement aux α-conotoxines, les ε-conotoxines n'ont été que très peu étudiées (Walker et al., 1999). À l'heure actuelle, aucune cible ne leur a été attribuée. Seule une activité sur le récepteur à la somatostatine sst3 a été montrée pour la toxine de *Conus consor* τ-CnVA (Petrel et al., 2013).

D'après les données présentées précédemment, il existe clairement une convergence structurale au sein des toxines animales. Cependant, il apparait extrêmement difficile de prédire avec certitude l'activité d'une toxine en se basant uniquement sur son motif de cystéine et sa structure tridimensionnelle. La formation des ponts disulfure suivant un motif particulier permet de donner une certaine architecture au peptide lui conférant une stabilité et permettant d'exposer les résidus clés pour l'activité biologique. La longueur des boucles

et la présence de certains résidus au sein d'entre elles sont ainsi cruciales pour expliquer la convergence d'activité entre des motifs très différents. Par exemple, le motif RGD est largement répandu dans les toxines inhibant l'agrégation plaquettaire : les 3FTs et les toxines d'arthropodes hématophages (Fry et al., 2015). En outre, seule une infime partie des motifs actifs sur les canaux potassiques a été présenté précédemment. En réalité, il en existe une plus grande diversité (**Table 2**). À cette complexité structurale et fonctionnelle s'ajoutent les PTMs, en particulier au niveau des conotoxines. À ce jour, 15 PTMs différentes ont été identifiées dans ce groupe avec une majorité de 4-Hydroxyproline et d'amidation C-terminale et dans une moindre mesure la gamma-carboxylation du glutamate et le pyroglutamate (Conoserver database, 2019). Des PTMs plus rares incluent la sulfatation sur des tyrosines, la bromation de tryptophane, la glycosylation et la D-isomérisation d'acides aminés.

Table 2. Motifs de cystéines des toxines actives sur les canaux potassiques K_v. (Adapté de Mouhat et al., 2008)

	Toxine	Structure 3D	Clade
Toxine à deux ponts disulfure			
	Hefutoxine-1	αα	Scorpiones
Toxine à trois ponts disulfure			
	BgK	αα	Cnidaires
	ShK	3_{10}αα	Cnidaires
	Dendrotoxine-1	3_{10}ββα (3FT)	Serpentes
	PVIIA	βββ	Conus
	HaTX1	3_{10}ββ	Araneae
	Charybdotoxine	αββ	Scorpiones
	BDS-1	βββ	Cnidaires
Toxine à quatre ponts disulfure			
	Pi1	αββ	Scorpiones
	Maurautoxine	αββ	Scorpiones

En résumé, au cours de l'Évolution, des pressions de sélection fortes ont permis à certaines espèces d'élaborer un arsenal chimique constitué de peptides que l'on appelle les toxines animales. Ces dernières se caractérisent d'abord par une forte réticulation via des ponts disulfure qui jouent le rôle de charpente conférant aux toxines animales une grande stabilité et déterminant l'orientation des boucles intercystéines. Les motifs de cystéines à l'origine de ces ponts sont très diversifiés traduisant une grande diversité structurale. Ils peuvent être répandus dans un grand nombre de clades ou être spécifiques de certains groupes venimeux. Les pressions de sélection exercées sur les boucles inter-cystéines sont à l'origine d'une grande efficacité et sélectivité d'action des toxines animales sur leur cible moléculaire. La richesse structurale et fonctionnelle qui résulte fait ainsi des venins un vrai réservoir de candidats médicaments.

6.4 Applications thérapeutiques

- **Des toxines-médicaments sur le marché**

Avec l'avènement de la médecine moderne au XX^{ième} siècle, la recherche de médicaments à partir de venins a été marquée par différents succès avec la mise sur le marché de nouveaux médicaments dérivés d'une toxine animale. Le premier date des années 1970. Le Captopril, un inhibiteur de l'enzyme de conversion de l'angiotensine (ACE) a été développé à partir d'un pentapeptide du venin du crotale sud-américain *Bothrops jararaca*. Ce peptide est capable à la fois de potentialiser l'effet hypotenseur des bradykinines endogènes et d'inhiber l'ACE (FERREIRA, 1965). Après identification du pharmacophore formé du motif FAP, le peptide a donné un analogue portant une D-2-méthyl succinyl-L-proline, le Captopril pour le traitement de l'hypertension artérielle (Cushman et al., 1973). Depuis, il y a eu d'autres succès comme l'eptifibatide (Integrellin®) et le tirofiban (Aggrestat®). Ces deux molécules sont des dérivés de disintégrines des serpents, *Sistrusus miliarius* et *Echis carinatus* respectivement. Elles préviennent l'agrégation plaquettaire et sont ainsi utilisées pour diminuer le risque d'infarctus du myocarde chez des patients à risque (Koh and Kini, 2012). La bivalirudine (Hirulog®) a été développé à partir de l'hirudine de la sangsue médicinale *Hirudo medicinalis*. Il s'agit d'un inhibiteur de la thrombine approuvé en 2000, mais également d'un anticoagulant pour traiter des patients à risque élevé de développer une thrombocytopénie induite par l'héparine lors d'interventions chirurgicales (Warkentin et al., 2008). Le ziconotide ou Prialt® dérivé de la ω-conotoxine MVIIA a quant à lui été approuvé en 2004 par la FDA pour le traitement de la douleur. Sa fenêtre thérapeutique très étroite nécessite une injection intrathécale et une utilisation en dernière intention chez des patients résistants à la morphine (Miljanich, 2004). Enfin, l'étude du venin du monstre de Gila, *Heloderma horridum*, a permis la découverte des exendines (Eng et al., 1992). L'exendine-4 a donné le Byetta® approuvé en 2005 pour le traitement du diabète de type 2 (Furman, 2012). En outre, deux molécules issues de toxines sont actuellement en essais cliniques. La chlorotoxine, issue du scorpion *Leiurus quinquestriatus hebraeus*, a la particularité de bloquer les canaux chlore d'insecte mais surtout de se lier avec une forte affinité aux cellules gliales via la MMP-2, une métalloprotéase de la matrice extracellulaire surexprimée à la surface de cellules cancéreuses (Deshane et al., 2003). Cette particularité a été utilisée pour faire de la chlorotoxine, un outil d'imagerie des tumeurs. Le tozuleristide, une version fluorescente de la toxine, a complété une phase I (NCT02462629) pour les cancers du SNC chez l'adulte et l'enfant. Actuellement, une phase II/III (NCT03579602) est en cours de recrutement pour les mêmes indications (https://clinicaltrials.gov). La deuxième toxine est la ShK de l'anémone de mer *Stichodactyla helianthus* ciblant les canaux K_v. Une version améliorée, ShK-186 (dalazatide) est sélective du sous-type K_v1.3. Ce sous-type est impliqué dans l'activation des lymphocytes T mémoires connus pour leur rôle dans le développement de maladies auto-immunes (Chandy et al., 2004). Le dalazatide a passé avec succès une phase I (NCT02446340) et Ib (NCT02435342) chez des patients atteints de psoriasis et montre une bonne tolérance (Tarcha et al., 2017). Selon KPI Therapeutics, la société qui développe le médicament, une phase II serait bientôt prévue.

Au final, en près de 50 ans, cinq médicaments dérivés de toxines sont actuellement sur le marché. Bien que ce nombre paraisse prometteur, l'étude actuelle des toxines animales présente plusieurs limites ayant des conséquences sur l'exploration et l'exploitation de la richesse pharmacologique contenue dans les venins.

- **Limites de l'exploration actuelle des venins**
    - Diversité venimeuse

La principale limite concerne la diversité venimeuse étudiée. Depuis le début du XX^{ième} siècle, les efforts de recherche se sont concentrés principalement sur les grandes espèces venimeuses pour deux raisons principales. La première est liée au problème de santé publique qu'elles représentent. Selon l'Organisation Mondiale de la Santé (OMS), les serpents seraient à eux seuls responsables de 1.8 à 2.7 millions d'envenimations entrainant entre 81 410 et 137 880 décès par an. Ainsi les venins des espèces les plus dangereuses pour l'Homme ont été les plus étudiés. La deuxième raison concerne les méthodes d'analyse biochimique d'un venin. L'approche classique, le bioguidage, consiste à faire des cycles de fractionnement du venin et de tests d'activités, jusqu'à l'identification de la molécule responsable de l'activité étudiée. Cette stratégie nécessite cependant de grandes quantités de venins au départ (plusieurs mg) et est limitée aux espèces produisant une grande quantité de venins comme les serpents ou les araignées. Si la quantité de venin de départ n'est pas suffisante, l'activité recherchée peut être perdue car la concentration de la molécule testée sera devenue trop faible.

L'avènement des méthodes omiques dans les années 2000 a ouvert la voie à l'étude de petites quantités de venins (µg). Mais cette exploration reste encore focalisée sur les clades venimeux déjà les plus connus, notamment les Theraphosidae (mygales) et les Conidae. Cela s'explique sans doute par leur potentiel thérapeutique déjà démontré avec différentes toxines anti-douleurs. Ainsi, certains groupes restent largement négligés. C'est le cas des Terebridae, Mitridae (des mollusques apparentés aux cônes marins) mais également d'araignées avec la famille des *Salticidae*. Ces araignées « sauteuses » minuscules avec des tailles de 1 à 25 mm représentent la plus grande famille d'araignées avec 5 000 espèces décrites soit 13% des Araneae (Peng et al., 2002).

    - Cibles moléculaires étudiées

Les méthodes omiques ont également montré qu'un venin d'un seul individu pouvait renfermer de plusieurs centaines à plusieurs milliers de toxines différentes (Lavergne et al., 2015b). Aujourd'hui, on estime à 40 000 000 le nombre de toxines produites par les 170 000 espèces venimeuses sur Terre. La corne d'abondance de molécules actives que représentent les venins est cependant encore largement méconnue. Les méthodes omiques renseignent sur la diversité moléculaire mais pas sur la diversité d'activités biologiques. Ainsi seules 5 000 toxines ont été caractérisées pharmacologiquement et principalement sur les canaux ioniques, délaissant de potentielles cibles thérapeutiques importantes comme les RCPGs. Au total, Muratspahić et *al.* ont répertorié seulement 40 peptides issus de venins ciblant différents RCPGs, le dernier en date étant la mambaquarétine sur le V2R (Ciolek et al., 2017; Muratspahić et al., 2019). La moitié de ces peptides proviennent de serpents. Afin d'explorer et d'exploiter au mieux le potentiel des venins, il était donc nécessaire de totalement changer de paradigme dans l'étude des venins. C'est ce à quoi a répondu le projet européen Venomics.

PROJET VENOMICS

1. Stratégie

Le projet Venomics vise à reproduire *in vitro* la diversité des venins pour générer une banque de toxines synthétiques utilisable dans des programmes de découverte de médicaments. Pour ce faire, l'idée était de combiner la puissance des méthodes omiques pour l'analyse de la diversité moléculaire des venins avec une stratégie de production à haut débit des séquences les plus intéressantes. Pour chaque individu, le venin et la glande venimeuse ont été prélevés en respectant le bien-être de l'animal puis analysés respectivement en protéomique et en transcriptomique. Le recoupage des données permet de s'assurer de la présence de la séquence dans le venin de départ. Ensuite, les séquences les plus originales ont été sélectionnées pour être produites soit par voie chimique (synthèse peptidique sur support solide : SPPS) pour celles de moins de 35 résidus soit par voie recombinante pour les plus longues. L'ensemble des toxines produites forme ainsi la banque de peptides Venomics qui peut par la suite être criblée facilement sur des cibles thérapeutiques.

Cette stratégie est compatible avec la majorité des animaux venimeux, même de toute petite taille. Elle permet donc d'explorer et d'exploiter la grande majorité de la ressource naturelle venimeuse (**Figure 23**). Le criblage est réalisé sur des toxines synthétiques n'induisant qu'un faible pourcentage de faux positif. Lors de celui-ci, les concentrations en peptides testés sont connues et peuvent être adaptées en fonction des objectifs. Par exemple, si l'on veut identifier toute la diversité de peptides actifs sur une cible, les concentrations testées seront les plus fortes possibles. Au contraire, dans un contexte thérapeutique, la banque sera criblée à de plus faibles concentrations pour ne prendre en compte que les hits de forte affinité. La stratégie Venomics permet également d'analyser pour la première fois un grand nombre de venins.

La principale limite de ce genre d'étude est qu'elle ne se base que sur la structure primaire des peptides. Que ce soit pour des toxines proches de familles structurales bien connues comme les ICK ou possédant un motif totalement nouveau, on ne peut être certain de la correspondance de la connectivité des ponts entre le peptide naturel et le peptide synthétique d'autant plus que différents arrangements pourraient être présents dans le venin. C'est pour palier à cette limite que tous les peptides de la banque Venomics ont été oxydés de façon non dirigé. Ainsi après oxydation des ponts disulfure, le profil HPLC peut présenter différents pics associés à une seule masse. On peut en déduire que le produit final présentera différentes formes oxydées du même peptide, ces dernières pouvant être testées en même temps lors du criblage.

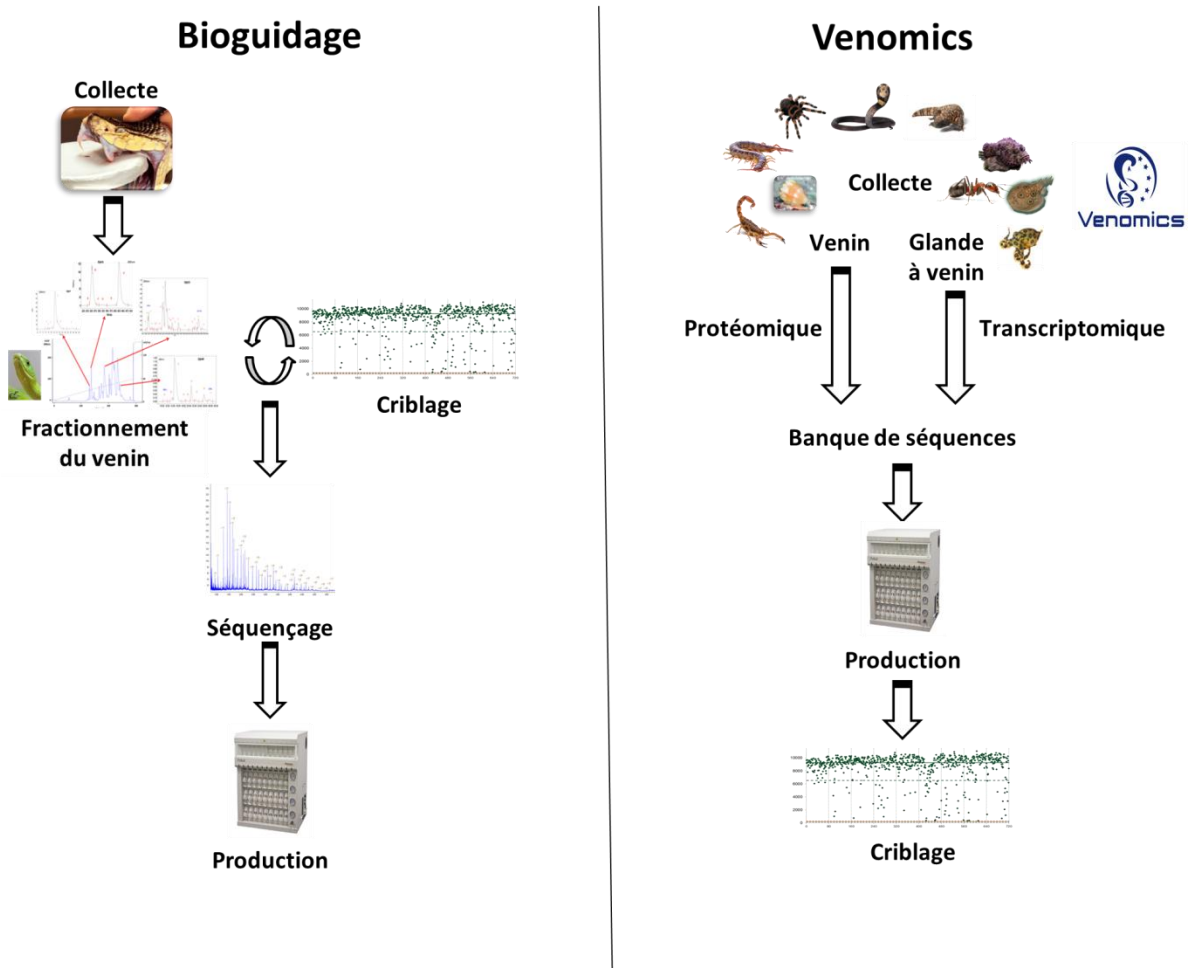


Figure 23. Comparaison de deux stratégies pour l'étude des venins : bioguidage (à gauche), Venomics (à droite). Concernant le bioguidage, le fractionnement du venin se fait par HPLC. Pour la stratégie Venomics, le venin et la glande venimeuse sont collectées puis analysés respectivement en protéomique et en transcriptomique. Pour les deux stratégies, la production des peptides peut se faire par voie chimique ou par expression recombinante. Le criblage est réalisé via un test de liaison.

2. Partenaires

Le projet Venomics s'inscrit dans le cadre du programme de biotechnologie "Recherche orientée PME pour le développement d'outils et de technologies pour la recherche à haut débit". Avec un budget total de 10 millions d'euros, Venomics est un consortium regroupant plusieurs partenaires, chacun ayant un rôle bien défini (**Figure 24**). Le CEA a coordonné l'ensemble du projet et a participé à la collecte des échantillons avec l'aide de la société AlphaBiotoxine en Belgique. Le laboratoire espagnol Sistemas Genomico s'est chargé ensuite des études transcriptomiques tandis que le laboratoire de spectrométrie de masse de l'université de Liège a conduit les études protéomiques. Le CEA a par la suite choisi les séquences à produire. La synthèse chimique a été prise en charge au CEA tandis que l'expression recombinante a été réalisée grâce à la collaboration de NZYTech au Portugal, fournissant les plasmides, et le laboratoire Architecture et Fonction des Macromolécules Biologiques de Marseille pour la partie production.

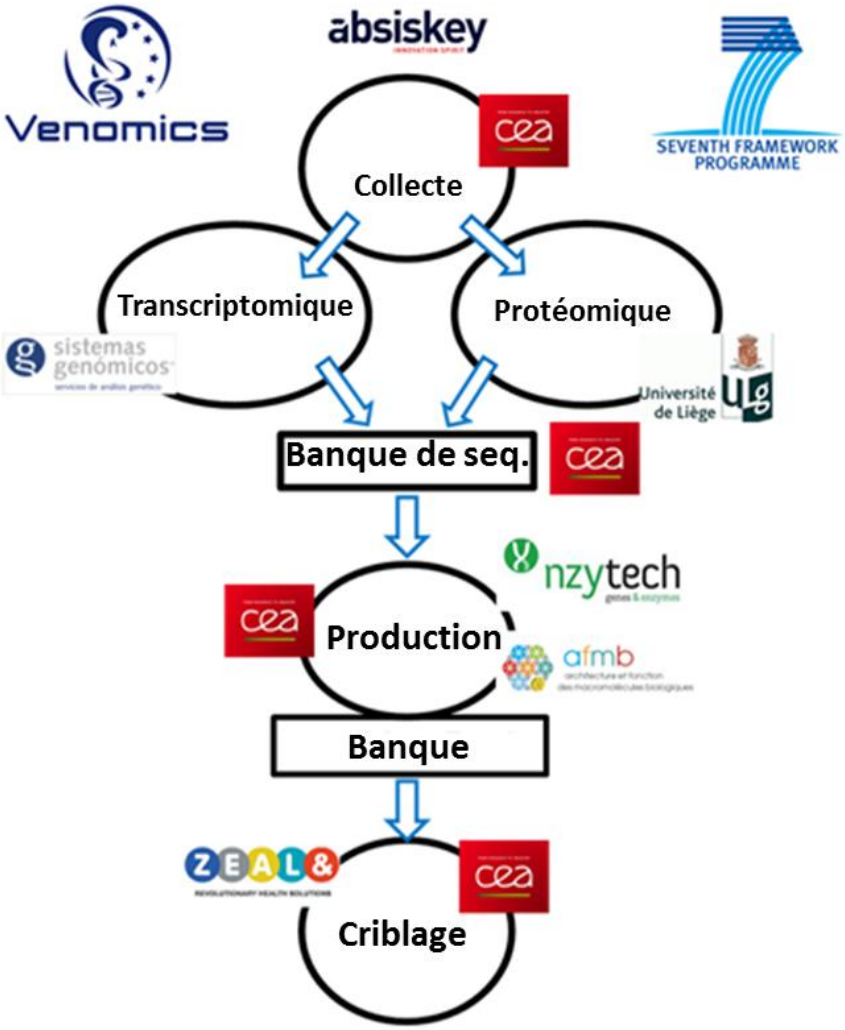


Figure 24. Consortium Venomics. L'implication de chaque partenaire est notée pour chaque étape du projet. seq. : séquence

3. Résultats

Le projet Venomics comportait au total trois grandes étapes qui ont abouti à trois principaux résultats. La première, correspondant aux analyses omiques des venins, a mené à une première banque de séquences représentant une exceptionnelle diversité structurale (parties 3.1 et 3.2). Puis, la sélection des séquences les plus originales a conduit en la production de la première banque de toxines synthétiques au Monde dont l'analyse est décrite dans la partie 3.3. Enfin, la dernière étape, le criblage de cette banque sur des cibles thérapeutiques, a permis de démontrer son potentiel dans un contexte de développement thérapeutique (partie 3.4) et plus particulièrement sur le sous-type 4 des récepteurs aux mélanocortines (hMC4R) impliqué dans les troubles de l'homéostasie énergétique (partie 3.5).

Ces résultats ont été obtenus avant le début de ma thèse. Néanmoins, leur analyse statistique a été réalisée dans le cadre de mon doctorat et fait partie de l'article soumis au *Journal of Medicinal Chemistry*. Pour cette raison, ils sont ici présentés en anglais. La totalité de l'article est disponible en **Annexe 1** et les informations supplémentaires en **Annexe 2**.

3.1 A large venomous animal collection

In total, 323 animal samples from 191 different species were gathered (**Table 3**). Almost every terrestrial and marine venomous clade was represented by at least one species such as ray (*Potamotrygon motoro*) or venomous lizard (*Heloderma horridum exasperatum*). Small terrestrial arthropods represented 95 species *i.e.* half of the sourcing. 55 species (29%) were studied for the first time, including 44 new species from the most infamous venomous clades (Araneae, Conus, Scorpiones and Serpentes). For instance, 12 previously-unexplored species of Neogastropods were sourced including *Conus canonicus* from the Conidae family, the most studied group of Neogastropoda. The Turridae and Terebridae groups also belong to the Neogastropoda order but have been largely neglected compared to their counterpart (Turner et al., 2018). For instance, among over 4,000 described species of Turridae, venom secretions of only twelve species have been studied so far. One of them, *Unedogemmula bisaya* was included in the sourcing. Concerning the Terebridae family, only seven of around 350 described species has been subject of venom analysis. In Venomics sourcing, two of them, *Terebra subulata* and *Terebra guttata,* were chosen. A third species, *Terebra maculata,* never explored before, completed the Terebridae species of the sourcing. Sourcing also included totally unexplored groups such as Mitridae. Two species considered non-venomous, the *Silurus glanis* fish and the *Trachycephalus resinifictrix* frog, but known to produce toxic skin secretions, were also part of the collection. The diverse collection of source animals encompassed various habitats, morphologies and animal sizes ranging from millimeters (*Manica rubida* ant) to several meters (*Ophiophagus hannah* snake).

Table 3. Selected taxa for sourcing.

Taxa	Total N° of species	Never studied species	Main groups
Neogastropoda			
Conus	38	10	*Conidae*
Other Neogastropoda	5	2	*Turridae, Terebridae & Mitridae*
Octopoda	3	0	*Octopodidae*
Arthropoda			
Chilopoda	6	2	*Scolopendridae*
Scorpiones	40	7	*Buthidae, Scorpionidae, Vaejovidae, Scorpionidae, Hemiscorpiidae & Euscorpiidae*
Araneae	41	23	*Theraphosidae, Ctenidae, Agelenidae, Araneidae, Diplurida,, Lycosidae, Nephilidae, Sparassidae & Theridiidae*
Hymenoptera	8	4	*Apidae, Vespida, Formicidae & Myrmicinae*
Chondrichthyes	1	0	*Potamotrygonidae*
Teleostei	8	2	*Synanceiidae, Plotosidae & Siluridae*
Squamata			
Serpentes	39	4	*Elapidae, Viperidae & Atractaspididae*
Anguimorpha	1	0	*Helodermatidae*
Anura	1	1	*Hylidae*

3.2 A huge sequence toxin diversity in animal venoms

De novo transcriptomics opens the door to the study of Nature without any previous knowledge of it and with a very high resolution, allowing the unbiased identification and quantification of transcripts in a broad spectrum, even for those of low expression. Total RNA was extracted from venom glands. In order to include species diversity, sample requirements were decreased from 2 µg to 100 ng of RNA, being even as low as 60 ng in specific samples. Illumina specific 2x100 pair-end libraries were constructed from each RNA following manufacturer´s protocols. Over 60 million reads were obtained per sample using Hiseq 2000 sequencing platform with an average length of 500 nucleotides.

Two proteomic strategies were used according to the size of the toxins (Morsa et al., 2019). The first one, for venoms presenting most of toxins with less than 25 amino acid residues (like Neogastropoda, Hymenoptera, Anura), analyzed the peptides post-disulfide-reduction by liquid chromatography coupled to tandem mass spectrometry. The second strategy, dedicated to the sequencing of the long venom toxins (> 25 amino acid residues (like Chilopoda, Scorpiones, Araneae, Squamata) required additional alkylation and trypsin digestion (short time and low enzyme concentration) steps after disulfide treatment leading to the generation of overlapping peptides (+/- 15 amino acids) (Morsa et al., 2019). These resulting peptides were then analyzed with the same mass spectrometry platform. The number of sequences varied between fifty to several hundred, depending on the venom. A high number of sequences are observed in Conidae venoms by proteomics. This diversity relies on the presence of post-translational modifications at different sites in a single toxin and has been discussed in previous publications (Degueldre et al., 2017; Dutertre et al., 2013). A total of 21,300,000 sequence tags were obtained.

Proteomic and transcriptomic data generated from the same animal were cross-validated to extract only sequences present in both the venom and venom gland transcripts, resulting in 20,206 validated toxin sequences (defined here as peptide up to 110 residues with at least one disulfide bridge). **Figure 25** describes these sequences in term of origins,

disulfide bridges (1 to 9), sizes (6 to 110) and PTMs. The quasi-totality of Conus sequences (98%) bears at least one PTM, in addition to disulfide bridges. Amidation is found in 2,889 sequences (14%) from all groups studied except Terebridae, Turridae, Mitridae and Teleostei. Hydroxylation on Pro, Met, Lys or Val residues was found in 2,630 sequences (13%) from various groups with a great majority from Conus (70%). In our study, 4-carboxyglutamic acid, pyroglutamic acid and bromo-tryptophan appeared to be Conus specific. Strangely, we did not find pyroglutamic acid in Arachnid venoms (Scorpiones and Araneae), although its presence has been described previously in many other studies (Lucchesi et al., 1989; Rong et al., 2013). The 20,206 toxin sequences were divided into 137 different cysteine patterns. Among the 97 patterns having an even number of cysteines, 38 do not correspond to any known peptide frameworks in any venomous species. Regarding the 59 known patterns, C-C-CC-C-C framework is the most represented comprising 19% of the bank (**Table 4**). It is related to Inhibitory Cystine Knot Toxins (ICK), which is known to be particularly abundant in spider, snail and scorpion venoms (Gao et al., 2013; Imperial et al., 2003; Robinson and Norton, 2014; Saez et al., 2010). For the first time, ICK peptides were identified at protein level from Mitridae, Turridae, *Manica rubida* ant, Chilopoda and the octopus *Hapalochlaena lunulata.*

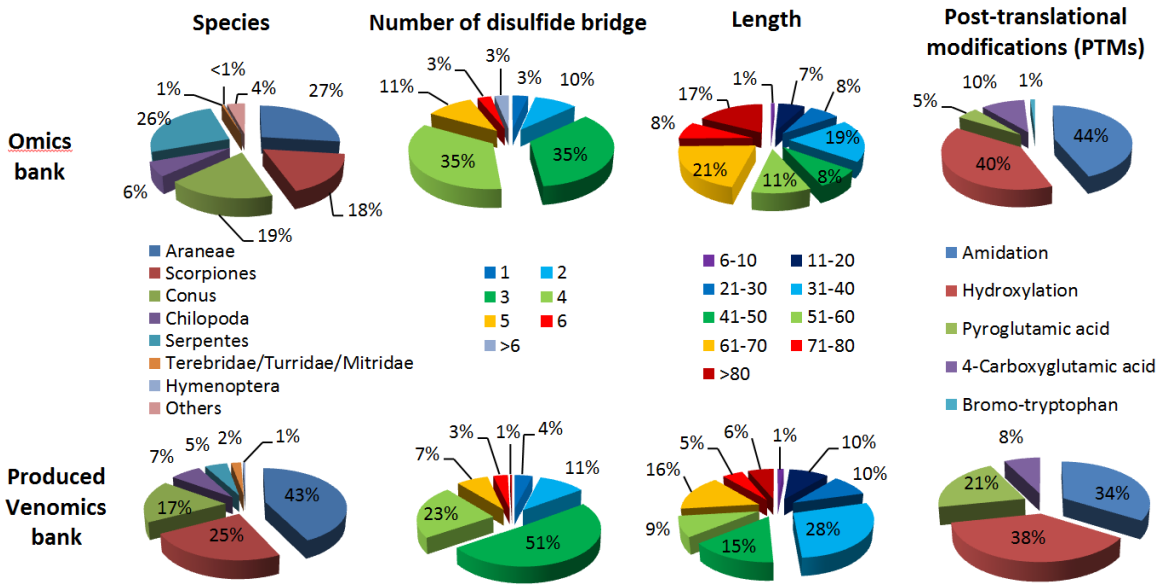


Figure 25. Statistical representation of the sequence banks built from omics analysis (upper panel) and produced Venomics bank (lower panel).

Table 4. Cysteine pattern diversity with an even number of cysteine identified in Omics bank. Newly described patterns are highlighted in grey.

Cystein framework	N° of Cys	Seq. number in : Data base	Peptide bank	Taxa in produced bank
C-C	2	637	142	Conus (108), Araneae (11), Serpentes (8), Chilopoda (6), Scorpiones (4), Undefined (4), Terebridae (1)
CC-CC	4	247	61	Conus (51), Araneae(5), Serpentes (4), Undefined (1)
CC-C-C	4	611	106	Conus (91), Araneae (11), Scorpiones (2),

				Serpentes (1), Undefined (1)
C-CCC	4	2	2	Conus (2)
C-CC-C	4	9	1	Conus (1)
C-C-CC	4	63	35	Conus (29), Araneae (5), Serpentes (1)
C-C-C-C	4	764	182	Conus (81), Chilopoda (49), Scorpiones (25), Araneae (21), Terebridae (3), Serpentes (2), Hymenoptera (1)
CCCC-CC	6	10	2	Conus (2)
CCC-C-C-C	6	1	0	
CC-CC-C-C	6	3	0	
CC-C-CC-C	6	2	1	Araneae (1)
CC-C-C-CC	6	450	99	Conus (68), Araneae (26), Serpentes (3), Undefined (2)
CC-C-C-C-C	6	10	1	Serpentes (1)
C-CC-CCC	6	1	0	
C-CC-C-C-C	6	20	5	Araneae (4), Conus (1)
C-C-CCC-C	6	1	0	
C-C-CC-CC	6	5	1	Serpentes (1)
C-C-CC-C-C	6	3,739	1,047	Araneae (887), Conus (77), Scorpiones (25), Serpentes (20), Chilopoda (14), Terebridae (12), Hymenoptera (7), Undefined (5)
C-C-C-CC-C	6	244	78	Chilopoda (67), Scorpiones (7), Araneae (3), Conus (1)
C-C-C-C-CC	6	83	20	Chilopoda (11), Conus (5), Scorpiones (4)
C-C-C-C-C-C	6	2,199	539	Scorpiones (367), Araneae (76), Conus (32), Serpentes (26), Chilopoda (20), Terebridae (15), Hymenoptera (2), Undefined (1)
CC-C-C-C-C-C-C	8	10	2	Conus (1), Scorpiones (1)
C-CCC-C-C-C-C	8	1	0	
C-CC-CC-C-C-C	8	3	0	
C-CC-C-CC-C-C	8	15	7	Araneae (7)
C-CC-C-C-C-C-C	8	11	3	Scorpiones (2), Araneae (1)
C-C-CCCC-C-C	8	1	0	
C-C-CCC-CC-C	8	1	0	
C-C-CC-CC-C-C	8	94	28	Conus (17), Araneae (8), Terebridae (2), Serpentes (1)
C-C-CC-C-CC-C	8	12	3	Araneae (2), Conus (1)
C-C-CC-C-C-CC	8	1	0	
C-C-CC-C-C-C-C	8	438	112	Araneae (68), Terebridae (13), Conus (11), Scorpiones (8), Serpentes (6), Hymenoptera (3), Chilopoda (3)
C-C-C-CCC-C-C	8	29	4	Conus (2), Serpentes (1), Araneae (1)
C-C-C-CC-C-CC	8	2	0	
C-C-C-CC-C-C-C	8	1,544	257	Araneae (190), Scorpiones (27), Chilopoda (17), Conus (10), Hymenoptera (6), Serpentes (5), Terebridae (2)
C-C-C-C-CC-CC	8	4	1	Chilopoda (1)
C-C-C-C-CC-C-C	8	101	43	Scorpiones (23), Serpentes (8), Conus (5), Araneae (4), Chilopoda (2), Terebridae (1)
C-C-C-C-C-CC-C	8	2,128	161	Serpentes (70), Chilopoda (44), Scorpiones (33), Araneae (12),Conus (1), Hymenoptera (1)
C-C-C-C-C-C-CC	8	1	0	
C-C-C-C-C-C-C-C	8	2,560	213	Scorpiones (171), Terebridae (23), Chilopoda (11),Conus (4), Serpentes (2), Araneae (2)
CC-C-C-CC-C-C-C-C	10	338	0	
CC-C-C-C-CC-C-C-C	10	2	0	
CC-C-C-C-C-CC-C-C	10	1	1	Scorpiones (1)
CC-C-C-C-C-C-C-C-C	10	1	1	Scorpiones (1)
C-CC-C-CC-C-C-C-C	10	54	7	Conus (4), Araneae (3)

C-CC-C-C-C-C-C-C-C	10	482	0	
C-C-CC-CC-CC-C-C	10	2	1	Scorpiones (1)
C-C-CC-CC-C-C-CC	10	1	0	
C-C-CC-CC-C-C-C-C	10	13	6	Conus (3), Araneae (1), Serpentes (1), Terebridae (1)
C-C-CC-C-CC-C-C-C	10	57	22	Araneae (20), Serpentes (2)
C-C-CC-C-C-C-CC-C	10	4	0	
C-C-CC-C-C-C-C-C-C	10	224	87	Araneae (46), Scorpiones (26), Chilopoda (8), Serpentes (4), Terebridae (2), Hymenoptera (1)
C-C-C-CCC-C-C-C-C	10	12	4	Araneae (2), Conus (1), Serpentes (1)
C-C-C-CC-C-C-CC-C	10	1	0	
C-C-C-CC-C-C-C-C-C	10	86	17	Araneae (14), Scorpiones (2), Terebridae (1)
C-C-C-C-CCCCC-C	10	1	0	
C-C-C-C-CCC-CC-C	10	1	0	
C-C-C-C-CC-C-C-C-C	10	3	1	Araneae (1)
C-C-C-C-C-CC-C-C-C	10	5	0	
C-C-C-C-C-C-CCC-C	10	1	0	
C-C-C-C-C-C-CC-C-C	10	1	1	Serpentes (1)
C-C-C-C-C-C-C-CC-C	10	605	9	Serpentes (7), Scorpiones (2)
C-C-C-C-C-C-C-C-C-C	10	253	107	Scorpiones (51), Araneae (41), Conus (7), Serpentes (4), Terebridae (4)
CC-C-C-C-CC-C-C-C-C	12	5	1	Conus (1)
CC-C-C-C-C-C-CC-C-C	12	2	1	Serpentes (1)
CC-C-C-C-C-C-C-C-C	12	1	1	Serpentes (1)
C-CC-C-C-C-C-CC-C-C	12	1	0	
C-CC-C-C-C-C-C-C-C	12	4	0	
C-C-CC-CC-CC-CC-C	12	2	1	Scorpiones (1)
C-C-CC-CC-C-C-C-C	12	2	0	
C-C-CC-C-C-CC-C-C	12	5	0	
C-C-CC-C-C-C-C-CC-C	12	15	4	Araneae (4)
C-C-CC-C-C-C-C-C-C	12	3	0	
C-C-C-CC-C-C-C-C-C	12	144	7	Araneae (7)
C-C-C-C-CCC-C-C-C	12	3	0	
C-C-C-C-CC-CC-CC-CC	12	1	1	Scorpiones (1)
C-C-C-C-CC-C-C-C-CC	12	1	1	Scorpiones (1)
C-C-C-C-CC-C-C-C-C	12	33	1	Hymenoptera (1)
C-C-C-C-C-CC-C-C-C	12	245	99	Scorpiones (90), Araneae (7), Chilopoda (2)
C-C-C-C-C-C-C-C-C-	12	1	0	

CC-C				
C-C-C-C-C-C-C-C-C-C-CC	12	1	0	
C-C-C-C-C-C-C-C-C-C-C-C	12	22	3	Conus (2), Scorpiones (1)
CC-C-C-C-C-C-CC-C-C-C-C-C	14	3	0	
C-CC-C-C-C-C-C-CC-C-C-C-C	14	1	0	
C-CC-C-C-C-C-C-CC-C-C-C-C	14	2	1	Conus (1)
C-C-CC-CC-C-C-C-C-C-C-C-C	14	11	0	
C-C-C-CC-C-C-C-C-CC-C-C-C	14	13	3	Conus (2), Araneae (1)
C-C-C-CC-C-C-C-C-C-C-C-C	14	574	7	Araneae (6), Terebridae (1)
C-C-C-C-CC-C-C-CC-C-C-C-C	14	1	1	Araneae (1)
C-C-C-C-C-CC-C-CC-C-C-C-C	14	3	0	
C-C-C-C-C-CC-C-C-C-C-C-C	14	2	0	
C-C-C-C-C-C-CC-C-C-C-C-C	14	4	0	
C-C-C-C-C-C-C-C-C-C-C-C-C	14	14	1	Scorpiones (1)
C-CC-C-C-C-C-C-CC-C-C-C-C-C-C	16	1	0	
C-C-C-C-CC-C-C-C-C-CC-C-C	16	2	0	
C-C-C-C-C-C-C-C-C-C-C-C-C-C	16	4	2	Scorpiones (1) , Terebridae (1)
C-C-C-C-C-C-C-C-C-C-C-C-C-C-C	18	1	1	Scorpiones (1)

Interestingly, 874 sequences covering 40 cysteine patterns display an odd number of cysteine residues (**Table 5**). Only thirteen were previously described. Two patterns (both lacking vicinal cysteine residues) are particularly abundant: C-C-C-C-C-C-C with 130 sequences (Scorpiones, Serpentes, Octopoda, Conus) and C-C-C-C-C specific to the Araneae taxon with 179 sequences. We did not identify any sequences from Conus sharing the latter pattern, despite it having been previously described (Lavergne et al., 2015b).

Table 5. Cysteine pattern diversity with an odd number of cysteine identified in Omics bank. Newly described patterns are highlighted in grey.

Cystein framework	**N° of Cys**	**Seq. Number in : Data base**	**Peptide bank**	**Taxa in produced bank**
CC-C	3	27	3	Araneae (3)
C-CC	3	3	0	
C-C-C	3	37	3	Scorpiones (1), Conus (1), Serpentes (1)
CC-CCC	5	2	0	
CC-C-CC	5	4	1	Conus (1)
CC-C-C-C	5	11	0	
C-CC-C-C	5	17	0	
C-C-CC-C	5	42	1	Araneae (1)

C-C-C-CC	5	3	0	
C-C-C-C-C	5	179	3	Araneae (3)
CC-C-C-CC-C	7	1	0	
C-CC-C-C-C-C	7	3	2	Araneae (2)
C-C-CC-CC-C	7	10	2	Araneae (2)
C-C-CC-C-C-C	7	97	20	Araneae (20)
C-C-C-CC-C-C	7	15	3	Araneae (3)
C-C-C-C-CC-C	7	55	0	
C-C-C-C-C-CC	7	3	0	
C-C-C-C-C-C-C	7	130	2	Chilopoda (1), Scorpiones (1)
CC-C-CC-C-C-C-C	9	2	0	
CC-C-C-CC-C-C-C	9	12	0	
CC-C-C-C-C-CC-C	9	2	0	
C-CC-C-C-C-CC-C	9	9	1	Serpentes (1)
C-CC-C-C-C-C-C-C	9	5	0	
C-C-CC-CCC-C-C	9	1	0	
C-C-CC-CC-C-C-C	9	1	0	
C-C-CC-C-C-C-C-C	9	23	0	
C-C-C-CC-C-C-C-C	9	3	0	
C-C-C-C-CC-C-C-C	9	12	0	
C-C-C-C-C-C-CC-C	9	5	0	
C-C-C-C-C-C-C-CC	9	1	0	
C-C-C-C-C-C-C-C-C	9	14	0	
C-CC-C-C-C-C-CC-C-C	11	1	0	
C-CC-C-C-C-C-C-C-C-C	11	8	0	
C-C-CC-C-C-C-C-C-C-C	11	1	0	
C-C-C-CC-C-C-C-C-C-C	11	1	0	
C-C-C-C-C-C-C-C-CC-C	11	3	0	
C-C-C-C-C-C-C-C-C-CC	11	1	0	
C-C-C-CC-C-C-C-C-C-C-C-C	13	121	0	
C-C-C-C-C-C-C-C-C-C-CC-C	13	6	0	
C-C-C-CC-C-C-C-CC-C-C-C-C-C	15	3	0	

Among these 20,206 sequences, 782 (4%) display a new cysteine pattern and 4,393 (22%) share less than 50% sequence identity with any described peptide or protein in the UniprotKB/SwissProt database using default settings (e-value max =1). It must be noted that the cysteine connectivity of the native toxins can not be explicitly determined using this pipeline.

3.3 From sequences to a high throughput screening compatible toxin bank

In order to allow screening and characterization of peptide activity, we developed two high-throughput production protocols. 5,923 sequences representing the broad diversity of the sequence database were selected and produced in the dedicated time of 18 months. Coming from 64 different species, they cover 91 different cysteine patterns, including 14 with an odd number of cysteine residues (Tables 4 and 5).

We selected 931 peptides shorter than 35 residues, with or without PTMs, to be produced by solid phase chemical synthesis (Upert et al., 2014). The crude synthetic product was cleaved from the resin, desalted and subjected to an optimized oxidation process, whatever the number or pattern of cysteines (Upert et al., 2014). After a final desalting step, liquid chromatography and mass analysis were performed to validate the presence of the expected oxidized mass (Figure S1). 893 toxins from 9 to 34 residues with 1 to 3 disulfide

bridges were successfully produced, representing a success rate of 96%.

The remaining 4,992 sequences containing more than 34 residues were selected for heterologous expression in *E. coli* (Duhoo et al., 2019; Sequeira et al., 2017; Turchetto et al., 2017). All synthetic genes encoding toxin sequences were subcloned into an optimized periplasmic expression plasmid containing an N-terminal *E. coli* Disulfide bond Isomerase C (DsbC) chaperone fusion tag, a 6xHIS tag for purification and a tobacco etch virus (TEV) protease cleavage site to release the native peptide from the fusion protein. Expressed fusion proteins are purified on a Nickel resin, subjected to overnight TEV cleavage and finally desalted. The correct mass was confirmed by mass analysis (Figure S2). 2,704 toxins, from 35 to 110 residues with 1 to 9 disulfide bridges were successfully produced representing a success rate of 54%.

Determining disulfide connectivity in native toxins is a challenge and the Venomics project did not have the resources to carry out this task. In the absence of this information, cysteine residues were randomly oxidized and the vast majority of toxins presented a unique peak in LC, suggesting that they adopted only one conformation, hopefully the natural one. In the case where LC profiles presented two or three peaks with the good expected oxidized mass, we assumed that it could be due to the generation of peptide isomers (Massonnet et al., 2018) (Figures S1 and S2).

Together, these two strategies resulted in a bank of 3,597 peptides composed of 71 different cysteine frameworks (**Tables 4 and 5**). In addition, 158 toxins (4%) display a novel cysteine pattern and 1,026 (29%) share less than 50% sequence identity with any previously described peptide or protein in the UniprotKB/SwissProt database. All together, we have successfully translated the high complexity of the omics bank into a physical toxin bank via high throughput methods.

3.4 A rich source of GPCR ligands

Animal toxins are mostly known for their bioactivity on ion channels. Few but very interesting toxins active on GPCRs have also been identified (Muratspahić et al., 2019). For example, the commercially available type 2 diabetes treatment, Byetta (derived from a venomous lizard) targets a class A GPCR. We selected four GPCRs linked to unmet medical needs in therapeutic areas such as diabetes, obesity, inflammation and allergy. A competitive binding strategy against a radio-labelled ligand was used for screening. For each receptor, binding conditions were optimized to obtain a Z'-factor above 0.5. The Venomics bank was tested once and 468 hits were obtained. The two first targets provided an amazing high hit rate of 4 and 9%. At the opposite, the two others gave very poor results (**Table 6**).

Table 6. Hit rate following the Venomics bank screening on different GPCRs.

	Hit rate Global (%)	Origin of hits (%)	
Targets	global	chemical synthesis	recombinant expression
GPCR1	9	57	43
GPCR2	4	11	89
GPCR3	0.3	64	36
GPCR4	0	0	0

3.5 First animal toxins on melanocortin receptors

We concentrated our study on the best-hit receptor, the human melanocortin 4 receptor (hMC4R). This receptor regulates food intake and its genetic invalidation is responsible of morbid obesity. Hit diversity is extraordinarily high in terms of toxin origin, number of disulfide bridges, length (**Figure 26**) and cysteine pattern (**Table 7**). Of the 286 hMC4R hits, 80 representative sequences were tested by dose-response characterization (data not shown), displaying affinities from nM to low µM range (Figure 2). Activity of eight hits could not be confirmed, giving a success rate of 90%. Half of the hMC4R-active peptides have an unknown cysteine pattern (**Table 7**) or share less than 50% sequence identity with toxins in other databases. The ICK-related and the CSαβ-related cysteine patterns are particularly well represented. Toxins with these patterns are renowned for their activities on ion channels. We selected one member of each of these structural families to perform a detailed structural and pharmacological characterization.

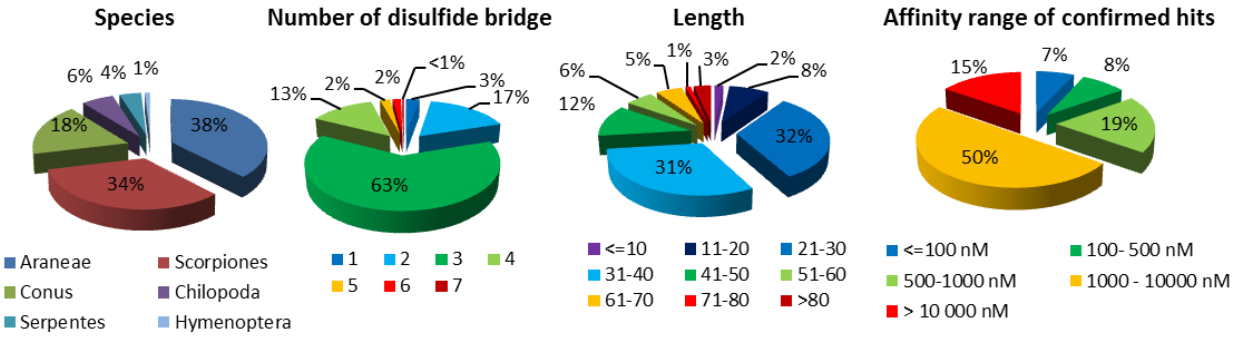


Figure 26. Statistical representation of the hits following Venomics bank screening on hMC4R. Affinities of 80 toxins representative of cysteine pattern diversity were measured by binding inhibition of ^{125}I-NDP-α-MSH on CHO cells stably expressing hMC4R.

Table 7. Cystein pattern diversity of MC4R hits. Newly described patterns are highlighted in grey.

Cystein frameworks	Nb of Cys	Nb of seq	% of total hit	Taxa (number of sequences)
C-C	2	8	3%	Conus (5), Araneae (1), Scorpiones (1), Serpentes (1)
CC-CC	4	16	6%	Conus (13), Araneae (1), Serpentes (1), Undefined (1)
CC-C-C	4	4	1%	Conus (3), Araneae (1)
C-C-CC	4	4	1%	Conus (4)
C-C-C-C	4	25	9%	Conus (14), Chilopoda (5), Scorpiones (5), Araneae (1)
CC-C-C-CC	6	7	2%	Araneae (4), Conus (3)
C-C-CC-C-C	6	97	34%	Araneae (86), Conus (6), Scorpiones (3), Hymenoptera (1), Serpentes (1)
C-C-C-C-CC	6	4	1%	Chilopoda (3), Conus (1)
C-C-C-C-C-C	6	71	25%	Scorpiones (62), Araneae (5), Chilopoda (2), Conus (1), Serpentes (1)
C-C-CC-CC-C	7	1	<1%	Araneae (1)
C-C-CC-C-C-C	7	1	<1%	Araneae (1)
C-C-C-CC-C-C	7	1	<1%	Araneae (1)
C-C-CC-C-C-C-C	8	5	2%	Scorpiones (4), Araneae (1)
C-C-C-CCC-C-C	8	1	<1%	Serpentes (1)
C-C-C-CC-C-C-C	8	9	3%	Scorpiones (5), Chilopda (3), Serpentes (1)
C-C-C-C-C-CC-C	8	10	3%	Serpentes (5), Chilopoda (3), Scorpiones (2)
C-C-C-C-C-C-C-C	8	10	3%	Scorpiones (9), Terebridae (1)
C-C-CC-C-C-C-C-C-C	10	2	1%	Araneae (2)
C-C-C-C-C-C-C-C-C-C	10	4	1%	Scorpiones (2), Araneae (1), Serpentes (1)
CC-C-C-C-C-C-C-C-C-C	12	1	<1%	Serpentes (1)
C-C-C-C-C-CC-C-C-C-C	12	4	1%	Scorpiones (3), Chilopoda (1)
C-C-C-CC-C-C-C-C-C-C-C	14	1	<1%	Araneae (1)

4. Discussion

To our knowledge, this is the first time that a thorough '-omic' analysis of venoms has been carried out on such a wide number and diversity of venomous species from various habitats (128 terrestrial, 55 aquatic and 8 aerial) including two European species (*Manica rubida* ant and *Buthus occitanus* scorpion). The *Manica rubida* ant was the smallest species of the project, studied for the first time and 35 sequences were validated. The numbers of validated sequences ranges from 0, due to experimental failures (some collecting conditions were not compatible with proper mRNA extraction), to 1,163, in the venom of the Mediterranean scorpion *Buthus occitanus.* Finally, no validated toxin sequences could be detected in the skin secretions of *Silurus glanis* fish and the *Trachycephalus resinifictrix* frog. An average of 106 toxin sequences per venom were validated, much less than what could be expected according to the assumed richness of venoms. This difference highlights the difficulties in obtaining full and unambiguous toxin sequences. The gotten sequence diversity also highlights our very limited state of knowledge of the toxin world. Indeed, almost one quarter of the toxins share less than 50% sequence identity with known toxin sequences, suggesting unpredictable and original biological activities. Furthermore, 65 new naturally occurring cysteine patterns have been found, suggesting 65 new structural organizations. This is an opportunity to discover peptides with attractive biophysical properties in terms of engineering and thermodynamic stability for optimal drugability. Another level of complexity, not investigated here, is the dimerization (homo- and/or hetero-) of toxins containing an odd number of cysteines. A specific proteomic analysis should be designed to address this point.

Converting a sequence database into physical peptides is challenging, in particular for the oxidation of disulfide bridges. Our recombinant expression strategy was able to directly produce oxidized peptide and has proven to be effective in terms of high throughput production. The two-step chemical synthesis strategy was also extremely efficient. This is mainly thanks to the use of a universal folding strategy applied on all toxins regardless of size or number of cysteines. These production strategies provided peptides with impurities, as no high-efficiency purification step was performed. While the presence of these contaminations could have caused numerous false positive results during the screening tests (Feng and Shoichet, 2006) we shown this was not the case as 90% of the tested hits were confirmed.

GPCRs constitute one of the largest and most versatile families of receptors. They are involved in a wide spectrum of physiological functions. Since the late 20^{th} century, pharmaceutical companies have significantly diminished their natural product programs, partly because of the advances in both HTS and combinatorial synthesis, thereby the creation of large synthetic libraries of small but artificial molecules. Natural product libraries consist of extracts and partially purified fractions and are perceived to be incompatible with HTS. Unfortunately, this loss of interest in natural products correlated with the overall reduction in new leads in the drug development pipeline and the substantial decline in new drug approval (Li and Vederas, 2009). The Venomics bank reconciles venoms, a particularly difficult exploitable resource of natural peptides, with HTS. The screening of the Venomics bank provided an extraordinary hit rate of 3%, much higher than the one reported in literature concerning artificial molecules and estimated at around 0.05% (Bender et al., 2008; Oprea, 2002). To our knowledge, this is the first large-scale demonstration of the power of natural peptide toxins to provide GPCR ligands. This strategy opens up an unexplored avenue for the discovery of new ligands on receptors linked to diseases with unmet medical needs.

5. Conclusion

Depuis quelques années, les analyses transcriptomiques et protéomiques sont devenues de puissants outils d'étude de la diversité des toxines animales. Elles requièrent de faibles quantités d'échantillon (respectivement, glande à venin et venin) et peuvent ainsi être adaptées aux plus petites espèces venimeuses. Dans le cadre du projet Venomics, cette puissance a été appliquée pour la première fois à près de 200 espèces différentes issues non seulement des principaux clades venimeux, de la fourmi au cobra mais également à des groupes jusqu'à présent largement négligés comme les Terebridae. Le recoupement des données des deux méthodes a abouti à l'identification d'une fabuleuse diversité structurale dont certains motifs de cystéines n'étaient jusqu'alors pas encore décrits. Cette diversité pourrait être en réalité encore plus importante au vu du faible nombre de séquences identifiées par espèce. Cela montre donc que la diversité moléculaire des venins est encore largement sous-estimée. Pour l'instant, aucune commercialisation de la banque des 20 206 séquences n'est prévue. Néanmoins, elle pourrait renfermer de précieuses informations quant à l'histoire évolutive des toxines et pourrait prochainement être exploitée dans ce sens via des accords de collaboration.

À partir des 20 206 séquences identifiées, une banque de 3 597 toxines synthétiques a été produite et respecte scrupuleusement l'immense diversité structurale observée dans l'analyse omique, y compris une certaine partie jamais décrite. Elle se présente sous forme de plaques de 96 puits, chaque puits contenant une toxine synthétique sous une ou plusieurs formes oxydées. Elle est ainsi parfaitement adaptée à un criblage à haut débit. Dans un contexte où les banques de petites molécules chimiques ont montré leurs limites, ces deux propriétés font de la banque Venomics une nouvelle source précieuse de nouvelles molécules pour des programmes de découvertes de nouveaux médicaments. Pouvant intéresser les industries pharmaceutiques, elle doit être commercialisée prochainement par la société grenobloise SmarTox avec un accord de licence. Néanmoins, elle reste la propriété exclusive des différents partenaires du projet et les séquences la consituant sont entièrement confidentielles.

Parmi le consortium Venomics, l'industrie pharmaceutique danoise ZealandPharma et le CEA, se sont réservés le droit de cribler la banque sur certaines cibles thérapeutiques. La protection d'une toxine de la banque ou d'une molécule dérivée pourra faire l'objet d'un brevet dont le seul inventeur sera celui ayant conduit le criblage primaire. Le taux de hits élevé obtenu par le CEA sur quatre RCPGs considérés comme des cibles thérapeutiques et le taux de succès supérieur à 90% sur l'un d'entre eux, le hMC4R, démontrent l'intérêt que représentent les toxines animales pour cibler d'autres cibles que les canaux ioniques. Cette stratégie permettrait non seulement de découvrir de nouvelles molécules thérapeutiques mais également de définir de nouvelles familles pharmacologiques de toxines animales.

RÉSULTATS et DISCUSSION

1. Choix des toxines à étudier

Le présent projet de thèse a débuté à partir des résultats décrits précédemment. Il concerne ainsi la sélection des toxines et les études pharmacologiques d'abord conduites dans le cadre d'un stage d'ingénieur de 6 mois puis poursuivies dans le cadre de la thèse. Au vu du grand nombre de hits identifiés lors du criblage de la banque Venomics sur hMC4R (286), nous avons sélectionné cinq toxines pour une caractérisation pharmacologique plus poussée selon deux principaux critères (**Table 8**).

Le premier critère est lié à la commodité de production. Seules des séquences de moins de 35 résidus ont été sélectionnées pour être produites en interne en synthèse chimique en phase solide. Contrairement à une expression recombinante dans *E.* coli, ce mode de production permet, si nécessaire, de contrôler facilement l'arrangement des ponts disulfure. Le second critère a été l'originalité de séquence et sera discuté plus en détail pour chacune des toxines étudiées. Par exemple, les séquences des toxines de cônes N-CTX-Can1a[7] et N-CTX-Ltg1a proviennent d'espèces jamais étudiées auparavant et ne se rapprochent d'aucun peptide connu dans la base de données UniprotKB/SwissProt en 2019. N-CTX-Mil1a, présente le motif CC-CC caractéristique de la superfamille T des conotoxines dont on connait peu d'activités pharmacologiques. Enfin, parmi les 286 hits identifiés sur hMC4R, deux motifs de cystéines sont particulièrement abondants : le motif Inhibitory Cystine Knot (ICK, C-C-CC-C-C) et le motif des toxines de scorpions en CSαβ (C-C-C-C-C-C), connus pour leur activité sur les canaux ioniques. Nous avons ainsi choisi un représentant de chacune de ces grandes familles de toxines pour les caractériser d'un pont de vue structural et pharmacologique.

Table 8. Toxines animales sélectionnées. Les cystéines sont représentées en gras et en orange dans la séquence. ? : famille structurale inconnue ; # : -$CONH_2$

Nom	Longueur	Nb de ponts disulfure	Famille structurale supposée	Séquence	Animal
N-TRTX-Preg1a	28	3	ICK	R**C**LHAGAA**C**SGPIQKIP**CC**GT**C**SRRK**C**T	Araignée
N-BUTX-Ptr1a	34	3	Short KCh	QMDMR**C**SASVE**C**KQK**C**LKAIGSIFGK**C**MNKK**C**K**C**	Scorpion
N-CTX-Can1a	16	3	**?**	G**C**R**C**PP**C**TGA**C**KYL**CC**#	Cône
N-CTX-Mil1a	19	2	Superfamille T	LQTHWNKR**CC**PRRPV**CC**FP	Cône
N-CTX-Ltg1a	25	2	**?**	LFHY**C**Q**C**Q**C**PPGFKGKF**C**QFKLRPP	Cône

Parmi les cinq toxines sélectionnées, N-TRTX-Preg1a et N-BUTX-Ptr1a ont fait l'objet d'un article scientifique soumis dans *Journal of Medicinal Chemistry* et seront d'abord présentées. Pour les toxines, N-CTX-Can1a et N-CTX-Mil1a, leur caractérisation a dû être stoppée. Enfin, nous verrons en quoi N-CTX-Ltg1a représente la toxine la plus prometteuse d'un point de vue thérapeutique.

[7] Les noms des toxines étudiées ont été donnés selon la nomenclature proposée par King et al., 2008. La première lettre correspond à l'activité biologique de la toxine suivie du nom générique indiquant de quelle toxine, la nouvelle molécule est dérivée, et enfin, de l'espèce représentée par une lettre majuscule pour le genre et une minuscule pour l'adjectif déterminant l'espèce. Dans le cas du présent projet, la lettre grecque nu majuscule (N) a été choisie comme descripteur de l'activité sur les récepteurs aux mélanocortines.

2. Les deux premières toxines actives sur les récepteurs aux mélanocortines

2.1 Contexte

Le travail de thèse devait être valorisé par une publication traitant de la découverte des premières toxines animales actives sur les récepteurs aux mélanocortines. Il devait ainsi intégrer une première partie sur le projet Venomics avec la description de la banque obtenue de toxines synthétiques et la démonstration de son potentiel via son criblage sur des cibles thérapeutiques. Une deuxième partie décrivant la caractérisation de plusieurs toxines sélectionnées sur une des cibles étudiées servirait à exemplifier l'originalité des ligands qui pouvaient être découverts. La totalité de l'article est mis en **Annexe 1** et les informations supplémentaires en **Annexe 2**.

Au départ, trois toxines issues des hits sur hMC4R avaient été sélectionnées. Elles présentaient toutes trois ponts disulfure avec des motifs de cystéines différents et provenaient de trois animaux différents : N-TRTX-Preg1a de l'araignée *Poecilotheria regalis*, N-BUTX-Ptr1a du scorpion *Parabuthus transvaalicus* et enfin la toxine N-CTX-Can1a du cône *Conus canonicus*. Les deux premières montrent des motifs de cystéines bien connus des toxines : motif ICK (-C-C-CC-C-C-) et motif des toxines courtes de scorpion en CSαβ (-C-C-C-C-C-x-C-) où x est n'importe quel acide aminé. N-CTX-Can1a devait être un exemple de toxine avec un motif beaucoup moins décrit dans la littérature (-C-C-C-C-CC-). Du fait de la complexité de sa synthèse et de sa faible affinité, cette dernière n'a finalement pas été incluse dans l'article. Les raisons seront développées en détail dans une prochaine partie. Concernant les deux toxines restantes, l'objectif était d'obtenir le plus d'informations possible à leur sujet sur le plan pharmacologique et structural. Le CEA ne pouvant pas couvrir l'ensemble de ces études, des collaborations ont été mises en place. L'étude de la fonction a été conduite par Karin Mörl de l'équipe d'Annette Beck-Sickinger de l'université de Leipzig. Steve Peigneur de l'université catholique de Louvain s'est occupé des tests de sélectivité sur les canaux ioniques. Ricardo Rodríguez de la Vega de l'université Paris-Sud a réalisé les études phylogénétiques des deux toxines. Enfin, les analyses structurales ont été faites par Hervé Meudal et Céline Landon du Centre de Biophysique Moléculaire d'Orléans. Au CEA, nous avons réalisé les alignements de séquences les plus proches des deux toxines et étudié leur affinité en test de liaison. La gestion des différentes collaborations a, en outre, fait partie intégrante du travail de thèse.

2.2 Résultats[8]

- **Caractérisation pharmacologique**

The ICK-related N-TRTX-Preg1a comes from the spider *Poecilotheria regalis* while the CSαβ-related N-BUTX-Ptr1a from the scorpion *Parabuthus transvaalicus*. N letter (uppercase nu greek letter) has been chosen as the activity descriptor for modulators of hMCRs. All the following work was done with synthetically produced toxins of > 95% purity (**Figure 27 and S3**). N-TRTX-Preg1a and N-BUTX-Ptr1a display low micromolar affinities on every studied

[8] Les résultats présentés ici sont la suite de l'article soumis dans *Journal of Medicinal Chemistry*. Ils seront ainsi présentés en anglais.

subtype of melanocortin receptors stably expressed in CHO-K1 cells for hMC1R, hMC4R, hMC5R and HEK293 cells for hMC3R (**Table 9**, **Figure 28**). Agonistic and antagonistic functions on the signaling pathways hMC1R/Gs and hMC4R/Gs were assessed on HEK293 cells. To control the experiments, it was demonstrated that neither toxin induced any effect on control hMCR-free HEK293. Both toxins tested at 100 µM induced cAMP production on hMC1R, similar to 1 nM of α-MSH, while no effect was observed on hMC4R (**Figure 28B**). In the presence of 1 nM of α-MSH, the toxins slightly enhance hMC1R related cAMP production but did not reduce the cAMP production related to hMC4R (**Figure 28B**). Both toxins were tested at 100 µM on several sodium channels ($Na_V1.1$, $Na_V1.2$, $Na_V1.3$, $Na_V1.4$, $Na_V1.5$, $Na_V1.6$, $Na_V1.7$ and $Na_V1.8$), potassium channels ($K_V1.1$, $K_V1.2$, $K_V1.3$, $K_V1.4$, $K_V1.5$, $K_V1.6$, $K_V2.1$, $K_V3.1$, $K_V4.3$, $K_V7.1$, $K_V7.3$, $K_V10.1$, $K_V11.1$(hERG), GIRK1), a voltage-gated calcium channel ($Ca_V3.3$) and ligand-gated ion channels nAChR $\alpha1_2\beta1\gamma\delta$ and nAChR α7). No activity could be detected (Figure S4 and S5).

Table 9. Inhibition constant values of N-TRTX-Preg1a, N-BUTX-Ptr1a on melanocortin receptors. mean (µM) ± SEM (n=3-10).

	hMC1R	hMC3R	hMC4R	hMC5R
N-TRTX-Preg1a	1.8 ± 0.3	19.8 ± 3.5	7.1 ± 1.5	10.0 ± 4.4
N-BUTX-Ptr1a	2.9 ± 0.6	3.9 ± 0.7	2.6 ± 0.4	2.2 ± 0.8

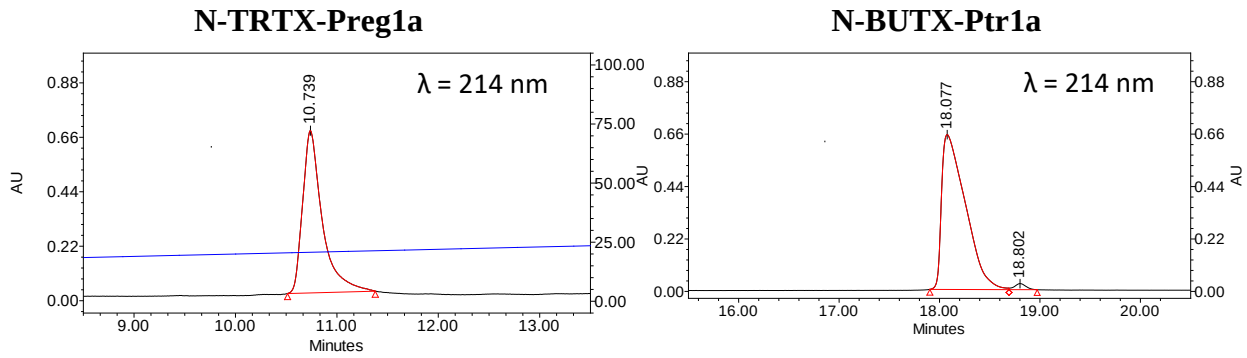


Figure 27. ***Analytical HPLC chromatogram of the final purified products of N-TRTX-Preg1a and N-BUTX-Ptr1a.*** Between 10 and 15 µg of each toxin were analyzed on a X Bridge BEH 300Å 5 µm 4.6 x 250 mm with a gradient of 10-50% CH_3CN/TFA 0.1% in 40 min at a flow rate of 1 mL/min. A spontaneous transformation of the N-terminal glutamine (Q form) residue of the peptide N-BUTX-Ptr1a into a pyroglutamic acid (Z form) can be detected after crude peptide deprotection. According to LC-MS analysis performed on Elute Sp Pump HPG 700 (Brucker) coupled with AmaZon SL LC/MS (Bruker), the peak at 18.077 min corresponds to the Q form (average mass of 3795.70 m/z) and represents 96.5% of the total area. The peak at 18.802 min with an average mass of 3778.70 m/z seems to correspond to the Z form of the peptide and represent 3.5% of the total area.

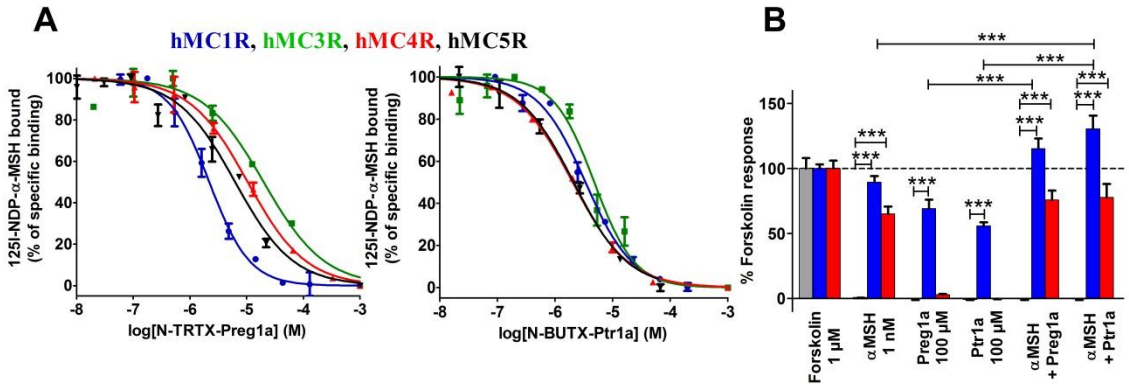


Figure 28. Pharmacological characterization of N-TRTX-Preg1a and N-BUTX-Ptr1a. (A) Representative curves of binding of ^{125}I-NDP-α-MSH on hMC1R, hMC4R and hMC5R (stably expressed in CHO cells) and on hMC3R (stably expressing in HEK293) in the presence of N-TRTX-Preg1a (left panel) or N-BUTX-Ptr1a (right panel). Mean (%) ± SD (n=2-3). (B) Functional assays on HEK293 (grey), HEK293 stably expressing hMC1R (blue) or HEK293 transitory expressing hMC4R (red). cAMP production expressed as % of the positive control (1 µM forskolin). Preg1a: N-TRTX-Preg1a. Ptr1a: N-BUTX-Ptr1a. Mean (%) ± SEM (n = 3), ***: p<0.001; one-way analysis of variance (ANOVA) followed by Dunnett's multiple comparison test.

- **N-TRTX-Preg1a et N-BUTX-Ptr1a présentent-elles réellement les structures ICK et CSαβ ?**

The two N-TRTX-Preg1a and N-BUTX-Ptr1a toxins have a cysteine pattern corresponding to the ICK-related and the CSαβ-related structure with the following pairing C1-C4, C2-C5, C3-C6. As both toxins display original pharmacological activities, we characterized both structures by NMR and determined experimentally their cysteine connectivity. Both peptides are well-folded, illustrated by their good dispersion of resonances. All the ^{1}H (except for the methyl groups of Met2 and Met28 of N-BUTX-Ptr1a), the ^{15}N chemical shifts and the ^{13}C resonances of Cα and Cβ were assigned. Depositions in the BioMagResBank were done with the entry code 34117 for N-TRTX-Preg1a and 34118 for N-BUTX-Ptr1a.

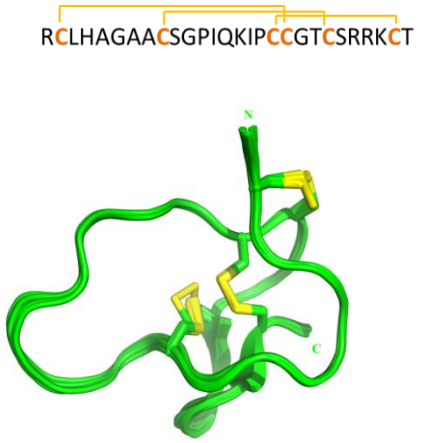
A set of 546 NOEs for N-TRTX-Preg1a were used in the final ARIA run, which includes 454 unambiguous restraints (Table S1). Additional constraints were introduced during the iterative ARIA procedure: 38 dihedral angle restraints, 10 H-bond constraints observed within the β-strands and 3 covalent bonds between sulphur atoms (C2-C19, C9-C22, C18-C27). Indeed, when enough experimental data had been introduced into the first runs of the iterative ARIA process, the C18-C27 disulfide bond is formed in two-thirds of the N-TRTX-Preg1a structures. The fine analysis of NOE constraints and the evaluation of distance between sulfur atoms allowed discrimination between the two remaining possibilities: C2-C19, C9-C22, C18-C27 (favorable) or C2-C22, C9-C19, C18-C27 (excluded). Therefore, the corresponding covalent bonds were added between sulfur atoms to refine the N-TRTX-Preg1a structure in the last runs of the calculations. The 10 models representative of the solution structure (**Figure 29A**) display a very low RMSD, small potential energy values, respect all the experimental data and are in perfect agreement with standard covalent geometry with 64.8% of the residues in the most favored regions of the Ramachandran plot, and 35.2% in additional allowed regions (Table S1). The N-TRTX-Preg1a structure is very well

defined, organized into the ICK motif without ambiguities regarding cysteine pairing. It is composed of a short antiparallel β-sheet made up of two short strands (22–23 and 26–27). The very short turn between these two strands, containing two arginine residues (R24 and R25), is highly constrained. Note the presence of an extended region (7–9), which faces the two-stranded β-sheet, resembling a third quasi-β-strand. Proline residue 12 is in a majority trans configuration (about 80%). N-TRTX-Preg1a sequence produced only a handful of significant hits in public databases and shares only 61% identity with the closest toxin with a known activity, Hs1a from *Cyriopagopus schmidti* (**Figure 29C**), a calcium channel blocker (Liu et al., 2006).

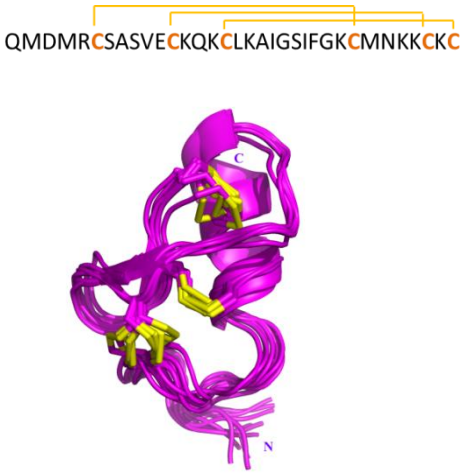
Regarding structural determination of N-BUTX-Ptr1a, 839 NOEs (including 558 unambiguous distance restraints) and 4 hydrogen bonds present within the β-sheet were used in the final ARIA run (Table S2). 92% of the 250 final structures display a C1-C4, C2-C5, C3-C6 disulfide bridge array, without adding any constraint on the bonds. The 10 models representative of the solution structure (**Figure 29B**) display low root mean square deviations (RMSD), small potential energy values, and respect all the experimental data, as attested by the absence of residual violations (Table S2). The Ramachandran plot exhibits 53% of the (ϕ,ψ) angles in the most favored regions, and 40% in the additional allowed regions. The N-BUTX-Ptr1a structure scaffold is a well-known motif within scorpion toxins (Kuzmenkov et al., 2016), plant defensins (Carvalho and Gomes, 2009) and insect defensins (Koehbach, 2017), and was designated as a cysteine-stabilized α-helix β-sheet motif (CSαβ motif) (Cornet et al., 1995). The structure is composed of an α-helix (11–19) and a 2-stranded short β-sheet (26–28; 31–33), stabilized by three disulfide bridges (C6-C27, C12-C32, C16-C34) unambiguously assigned. N-BUTX-Ptr1a shares the same sequence pattern as the majority of α-KTx toxins known to block potassium channels: C-x(4,5)-C-x(3)-C-x(6,9)-[GAS]-K-C-[IMQT]-x(3)-C-x-C (Prosite PS01138). Yet, N-BUTX-Ptr1a shares only 59% identity with the closest toxin with a known activity, Parabutoxin-3 from the same species *Parabuthus transvaalicus* (**Figure 29D**), a potassium channel blocker (Huys et al., 2002). The coordinates of structures have been deposited in the Protein Data Bank (http://www.pdb.org) with the entry codes 6SAA and 6SAB for N-TRTX-Preg1a and N-BUTX-Ptr1a, respectively.

While N-TRTX-Preg1a and N-BUTX-Ptr1a display the canonical ICK and CSαβ structure, respectively, their sequences diverged so much from most described toxins that homology searches or phylogenetic reconstruction failed to assign them into known toxin groups. On one hand, N-TRTX-Preg1a matched only a handful of spider toxins that could not be linked to the main group of spider ICK toxins. On the other hand, phylogenetic reconstruction with all known alpha-KTx (UniProt) placed N-BUTX-Ptr1a as sister to a clade consisting of several monophyletic alpha-KTx subfamilies (alpha-KTx 2, 4 and 12).

A



B



C

Name	Sequence	%Id	Targets
N-TRTX-Preg1a	RCLHAGAACSGPIQKIPCCG-TCSRRK--CT		MCRs
ω-TRTX-Hs1a	KCLPPGKPCYGATQKIPCCG-VCSHNK--CT	61%	Cav2.2 (Liu et al., 2006)
U2-TRTX-Pc1a	RCLPAGKTCVRGPMRVPCCG-SCSQNK--CT	54%	Unknown
U6-TRTX-Cg1a	CEPSGKPC-RPLMRIPCCG-SCVRGK--CA	46%	Unknown
ω-CTX-GVIA	CKSPGSSCSP--TSYNCCR-SCNPYTKRCY	29%	Cav2.2 (Favreau et al., 2001)
ω-CTX-MVIIA	CKGKGAKCSR--LMYDCCTGSCRSGK--C-	29%	Cav2.2 (Wang et al., 2016)

D

Name	Sequence	%Id	Targets
N-BUTX-Ptr1a	QMDMRCSASVECKQKCLKAIGSIFGKCMNKKCKC		MCRs
Parabutoxin-3	EVDMRCKSSKECLVKCKQATGRPNGKCMNRKCKCYPR	59%	Kv1.1, Kv1.2, Kv1.3 (Huys et al., 2002)
Hongotoxin-2	HFIDVKCTTSKECWPPCKAATGKAAGKCMNKKCKCQ	53%	Unknown
Rj1m1	IVIDVPCTVSKECWSACKKAVGTDRGKCMGKKCKCYP	53%	Unknown
Tst26	VFINAKCRGSPECLPKCKQAIGKAAGKCMNGKCKCYP	53%	Kv1.2, Kv1.3 (Papp et al., 2009)
Ts7	VFINAKCRGSPECLPKCKEAIGKAAGKCMNGKCKCYP	53%	Kv1.1, Kv1.2, Kv1.3, Kv1.6, Shaker Kv (Cerni et al., 2014; Rodrigues et al., 2003; Rogowski et al., 1994)
Tco30	WCSTCLDLECGASRECYDPCFKAFGRAHGKCMNNKCRCYT	50%	Shaker B KCh (Diego-García et al., 2005)
TtBut	WCSTCLDLACGASRECYDPCFKAFGRAHGKCMNNKCRCYT	50%	Unknown
Charybdotoxin	ZFTNVSCTTSKECWSVCQRLHNTSRGKCMNKKCRCYS	41%	Kv1.2, Kv1.3, Kv1.5, Kv1.6, KCa1.1, Shaker Kv (Garcia et al., 1994; Lucchesi et al., 1989)
Agitoxin-2	GVPINVSCTGSPQCIKPCKDA-GMRFGKCMNRKCHCTPK	38%	Kv1.1, Kv1.2, Kv1.3, Kv1.6 (Garcia et al., 1994; Takacs et al., 2009)

Figure 29. Structural characterization of N-TRTX-Preg1a and N-BUTX-Ptr1a. Primary and overlay of the 10 NMR models representative of the solution structure of N-TRTX-Preg1a (A) and N-BUTX-Ptr1a (B). Disulfide bridges are represented in yellow. N: N-terminal extremity; C: C-terminal extremity. Comparison of amino acid sequences between N-TRTX-Preg1a (C) or N-BUTX-Ptr1a (D) and the most similar toxins obtained with NCBI blast program (default settings). Cystein residues are in bold and orange, conserved residues from N-TRTX-Preg1a or N-BUTX-Ptr1a are in green. Sequence alignment was performed with Clustal Omega (version 1.2.4 from Emboss programs, EBlosum62 matrix for two pair alignment).

2.3 Discussion

MC4R is mainly expressed in the paraventricular nucleus of the hypothalamus. Its activity is finely regulated by both anorexigenic agonists, the melanocyte stimulating hormones (MSH), and the orexigenic antagonist, Agouti-related protein (AgRP) (Andermann and Lowell, 2017). Loss-of-function mutations on hMC4R are the primary cause of monogenic obesity (http://www.orpha.net/) making it a therapeutic target in eating disorders. Currently, only one MC4R ligand is in clinical trials (Setmelanotide, https://clinicaltrials.gov/). Subtype 1 (MC1R) is also of therapeutic interest with its protective role against melanoma. It elicits eumelanin production, a dark, UV radiation-absorbing, photoprotective pigment (Kobayashi et al., 1998) and promotes nucleotide excision repair of DNA damaged by UV radiation (Kadekaro et al., 2010). There is an urgent need for development of innovative ligands to better understand MCR functions and to generate new medicines. N-TRTX-Preg1a and N-BUTX-Ptr1a are the first MCR venom-peptide ligands, structurally unrelated to natural endogenous MCR agonists. Very interestingly, despite having two different structures, they both activate hMC1R without exhibiting the HFRW pharmacophore, previously described in endogenous agonists as the minimal active sequence (Hruby et al., 1987; Singh et al., 2010). In addition, they promote a synergistic effect with endogenous agonist α-MSH. This activity deserves to be finely studied as both ligands may activate hMC1R by a new molecular mechanism. N-TRTX-Preg1a and N-BUTX-Ptr1a inhibit ^{125}I-NDP-α-MSH binding on hMC4R but are not able to modulate activate or antagonize the hMC4R/Gs pathway, under our experimental conditions. This discrepancy can comes from our experimental procedure or from a biased function of the two toxins. A deeper exploration of the functions of the two toxins on various MC4R signaling pathways (Gq, β-arrestin, Kinases, etc.) will be performed to better understand their modes of action.

N-TRTX-Preg1a exhibits a common ICK fold only described so far for activity on ion channels. Its sequence is close to calcium channel toxins but does not possess the tyrosine residue in position 16 (N-TRTX-Preg1a numbering), described to be essential for blocking $Ca_V2.2$ channel conductivity (Sousa et al., 2013). N-BUTX-Ptr1a, close in sequence to potassium channel toxins, possesses a dyad composed of a basic and an aromatic residue (Dauplais et al., 1997; Gasparini et al., 1998) (F24 and K26, N-BUTX-Ptr1a numbering). This dyad, stated as essential for potassium activity (Dauplais et al., 1997; Gasparini et al., 1998) but we shown that is clearly not sufficient for N-BUTX-Ptr1a to bind potassium channels. N-TRTX-Preg1a and N-BUTX-Ptr1a sequences diverge so much within their own group that it was even impossible to phylogenetically anchor them into known toxin families. Thus, they constitute new clades for ICK and CSαβ families. Several other toxins of the produced bank sharing more than 80% of sequence identity with these two toxins are present in the Venomics database, suggesting that N-TRTX-Preg1a and N-BUTX-Ptr1a are not the only members of these new families. Production and pharmacological characterization of these natural variants should provide interesting insights with regard to the mode of action of these toxins on hMCRs. This may aid the generation of ligands with improved affinity and selectivity for development of new hMC1R agonists with applications in the treatment and/or prevention of melanoma.

2.4 Conclusion et précisions

Les résultats obtenus permettent de bien exemplifier dans l'article soumis dans *Journal of Medicinal Chemistry* l'originalité pharmacologique des toxines contenues dans la banque Venomics. Les premières toxines animales décrites sur les récepteurs aux mélanocortines sont ainsi deux toxines à trois ponts disulfure provenant de deux animaux et ayant des structures totalement différentes. N-TRTX-Preg1a produite par l'araignée *Poecilotheria regalis* possède un motif ICK. N-BUTX-Ptr1a du scorpion *Parabuthus transvaalicus* présente le motif des toxines courtes de scorpion en CSαβ. Chacune forme un nouveau groupe au sein de sa famille structurale respective. Malgré ces différences, elles partagent un même profil pharmacologique avec des affinités micromolaires sur les MCRs. Elles ne sont actives sur aucuns canaux ioniques testés, pourtant connus pour être des cibles des familles structurales des deux toxines. À noter que le canal $Ca_v2.2$, cible de la toxine connue Hs1a la plus proche de N-TRTX-Preg1a, n'a pas pu être testé car extrêmement difficile à exprimer dans les ovocytes de Xénope.

En outre, la grande majorité des toxines caractérisées à ce jour sont des inhibiteurs de l'activité de leur cible moléculaire. Dans le cas des deux présentes toxines, elles sont agonistes du récepteur MC1R. Plus intéressant encore, elles sont les premiers ligands peptidiques qui n'ont pas été développés à partir du pharmacophore HFRW de l'agoniste endogène α-MSH. Néanmoins, leur faible affinité et efficacité sur hMC1R a limité certaines études. Pour les tests de fonction, aucune courbe dose-réponse n'a été présentée car aucun effet n'a été observé à des concentrations inférieures ou égales à 10 µM. L'activation de la voie G_s du hMC1R avec 100 µM de toxine ne correspond sans doute pas à un effet maximal. Tester de plus fortes concentrations n'a pas été entrepris principalement à cause des quantités importantes de toxines nécessaires. Par contre, si les deux toxines avaient présenté de meilleures affinités sur MC1R, elles auraient fait l'objet d'un article à part entière. Cet article aurait intégré l'étude de stimulation de la pigmentation *in vitro* (Wen et al., 2013) ou *in vivo* chez les rongeurs (D'Orazio et al., 2006).

3. Deux conotoxines difficiles à caractériser

La caractérisation pharmacologique d'une toxine animale synthétique nécessite d'une part des arguments en faveur de l'équivalence avec la forme naturelle au niveau structural et pharmacologique et d'autre part une production permettant l'obtention de quantités suffisantes pour les différentes études. Ces deux exigences peuvent être particulièrement difficiles à respecter pour des toxines dont il n'existe aucune indication quant à l'arrangement des ponts disulfure de la toxine naturelle. C'est le cas de trois des cinq toxines sélectionnées, toutes des conotoxines. Parmis ces toxines, deux d'entre elles ont finalement été mises de côté principalement du fait de l'impossibilité de déterminer la ou les forme(s) succeptible(s) d'être dans le venin.

3.1 N-CTX-Can1a

- **Originalité de séquence**

La conotoxine N-CTX-Can1a (GCRCPPCTGACKYLCC#) a été identifiée dans le venin du cône *Conus canonicus*. Elle comporte 6 cystéines sur 16 résidus, ce qui en fait un peptide exceptionnellement réticulé même au sein des toxines animales. Ces cystéines forment le motif XXV (C-C-C-C-C-CC) des conotoxines partagé uniquement avec une autre toxine, as25a de *Conus cancellatus* (CKCPSCNFNDVTENCKCCIFRQP) (Aguilar et al., 2013) dont l'activité pharmacologique n'est pour l'instant pas connue. L'analyse du blast de N-CTX-Can1a ne donne aucun alignement satisfaisant avec les séquences connues de la base de données UniProtKB/SwissProt. Dans la banque des 20 206 séquences du projet Venomics, la toxine ne partage aucune identité de séquence avec les 82 autres séquences présentant le même motif de cystéine. Par contre, elle est proche de toxines au motif -CC-C-C-CC caractéristique de la superfamille M des conotoxines (Sato et al., 1983). La plus proche, P014846 (identifiant dans la banque Venomics), ne diffère de N-CTX-Can1a que par l'inversion de C et R en position 2 et 3 (**Table 10**).

Table 10. Comparaison de N-CTX-Can1a avec les séquences les plus proches de la banque Venomics. Les cystéines sont représentées en gras et en orange. Les résidus différents de N-CTX-Can1a dans les autres séquences sont en noir et en gras. Les séquences ont été obtenues avec le programme Blastp du site NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi) en interrogeant la base UniProtKB/SwissProt (paramètres par défaut, matrice Blosum62). # : -$CONH_2$

Nom	Séquence	%Id	Criblage primaire sur hMC4R	Espèce
N-CTX-Can1a	GCRCPPCTGACKYLCC#	100%	Active	*Conus canonicus*
P014846	GCC**R**PPCTGACKYLCC#	88%	Pas active	*C. canonicus*
P014850	GCC**R**PPCTGACKYLCC**G**	88%	Pas active	*C. canonicus*
P014842	GCC**RT**PCTGACKYLCC#	81%	Pas active	*C. canonicus*
P010959	GCC**R**PPC**SSV**CQYLCC#	63%	Pas active	*C. canonicus*

- **Production et caractérisation**

L'étape clé de la production d'un peptide réticulé est l'oxydation des ponts disulfure. En ce qui concerne N-CTX-Can1a, l'arrangement des ponts disulfure de la toxine naturelle n'est pas connu. La méthode d'oxydation aléatoire apparait ainsi la plus simple au départ en

permettant à la toxine d'atteindre une ou plusieurs formes énergétiquement stables et succeptibles de correspondre aux formes naturelles.

- Méthode d'oxydation aléatoire en présence d'un couple oxydo réducteur

Le peptide a d'abord été oxydé suivant les conditions du projet Venomics : HEPES 100 mM, ACN 20%, couple oxydo-réducteur cystéine/cystine (ratio molaire 1 :50 :5 peptide/ cystéine/cystine). Lorsque ces conditions sont utilisées, trois pics (T1, T2 et T3) sont observés en HPLC (**Figure 30A**) et présentent tous la masse monoisotopique du peptide oxydé ainsi que celle d'un dimère minoritaire (**Figure 30B**). Ce dernier correspondrait à deux toxines à deux ponts disulfure et un pont interchaine. Les trois pics correspondraient donc à trois arrangements de ponts disulfure différents sur les 15 possibles. En outre, l'oxydation est rapide puisqu'en moins de 4h, le profil HPLC n'évolue plus. Ces résultats compliquent grandement la caractérisation pharmacologique de la toxine. En effet, l'obtention de différentes formes diminue d'une part les quantités finales obtenues mais augmente d'autre part, le nombre de tests à faire pour leur caractérisation. La première solution à ce problème est de favoriser préférentiellement une forme par rapport aux autres en modifiant les conditions d'oxydation (**Annexe 3**). Plusieurs paramètres ont ainsi été modifiés : couple oxydo-réducteur, température, agents chaotropiques, pH. Toutes ces conditions aboutissent à un résultat similaire aux conditions Venomics avec trois formes oxydées. Ces derniers ont été purifiés mais seuls T2 et T3 ont été obtenus en quantité suffisante pour évaluer leur affinité sur hMC4R via un test de liaison.

Les deux pics T2 et T3 présentent de faibles affinités de l'ordre du micromolaire, respectivement 4.86 µM (n=1) et 1.12 µM (n=1) tandis que le peptide réduit alkylé n'est pas actif à 100 µM (n=2). La formation des ponts semble donc importante pour la liaison de la toxine sur le récepteur. Or l'arrangement des ponts disulfure n'est pas connu. En outre, il est probable que d'autres formes oxydées de la toxine avec des pontages différents présentent de meilleures affinités. Pour obtenir ces deux informations, une stratégie de pontage dirigé a été appliquée.

- 1 pont en pontage dirigé (entre la dernière cystéine C6 et C1/2/3/4)
  - *Stratégie*

Un peptide à trois ponts disulfure peut donner 15 arrangements différents. Or, la production en pontage dirigé d'un peptide à trois ponts disulfure requiert l'emploi de trois groupements protecteurs différents multipliant ainsi les étapes de déprotection, d'oxydation et de purification. Pour simplifier, nous avons appliqué une approche en deux temps grâce à l'utilisation de deux groupements protecteurs des fonctions thiols des cystéines. Le groupement trityl (Trt) est labile en milieu acide mais pas le groupement acétadométhyle (Acm). Après déprotection acide au TFA, seuls les groupements Acm seront donc toujours présents. Lors d'une première étape d'oxydation, les deux premiers ponts sont ainsi formés de manière aléatoire par les cystéines non protégées donnant trois possibilités d'arrangements de ponts disulfure (**Figure 31**). Dans un second temps, le dernier pont est formé en présence d'iode qui permet successivement le clivage des Acm et l'oxydation des fonctions thiols des deux cystéines restantes. On connait ainsi l'arrangement d'un pont sur les trois pour chacune des formes obtenues. Cette méthode requiert en premier lieu de choisir le pont qui sera dirigé. De façon arbitraire, c'est celui impliquant la dernière cystéine qui a été choisi. Au total, quatre formes de la toxine sont ainsi produites. Seule la forme pour laquelle le pont dirigé supposé le plus difficile à se former entre les deux dernières cystéines

adjacentes n'est pas produite.

- *Biochimie*

De façon intéressante, la première étape d'oxydation donne un seul pic majoritaire quelle que soit la position des cystéines protégées par des Acm. Sur les 12 formes oxydées possibles, seules 4 sont finalement obtenues.

Bien que les rendements globaux soient faibles (**Annexe 4**), les produits finaux présentent plus de 90% de pureté et des temps de rétention différents (**Figure 32**). Les spectres de masses sont présentés en **Annexe 5**.

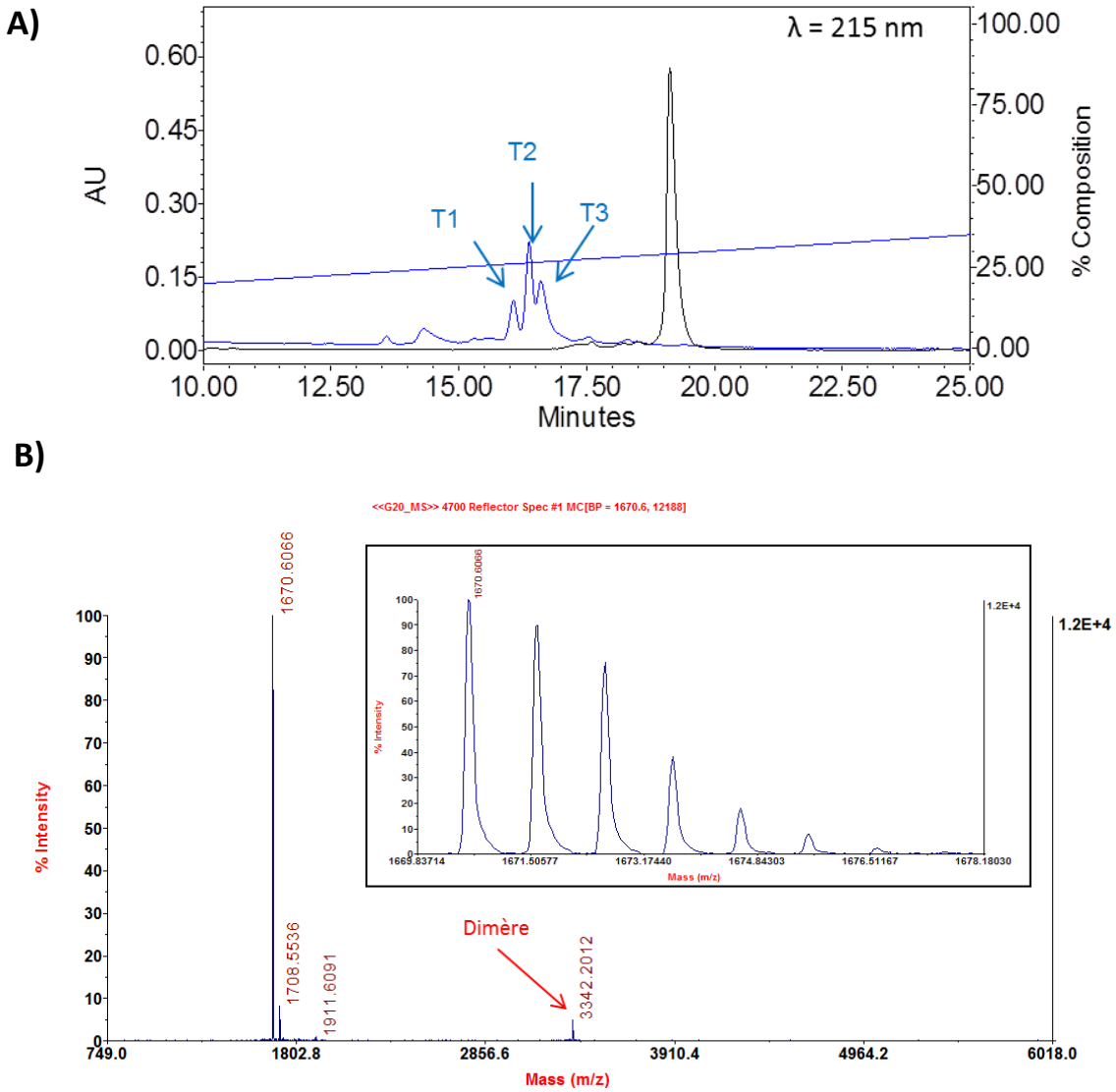


Figure 30. Oxydation aléatoire de N-CTX-Can1a. A) Profils HPLC de N-CTX-Can1a réduit (en noir) et après 4h d'oxydation (en bleu). Colonne : XBridge® BEH 300Å 5 µm, 4.6 x 250 mm ; Gradient : 10% d'ACN puis 1% ACN/min ; débit : 1 mL/min. B) Spectre de masse MALDI-TOF identique des pics purifiés T1, T2 et T3. La masse notée correspond à la masse monoisotopique du peptide. L'encadré est un agrandissement du massif isotopique de la masse majoritaire.

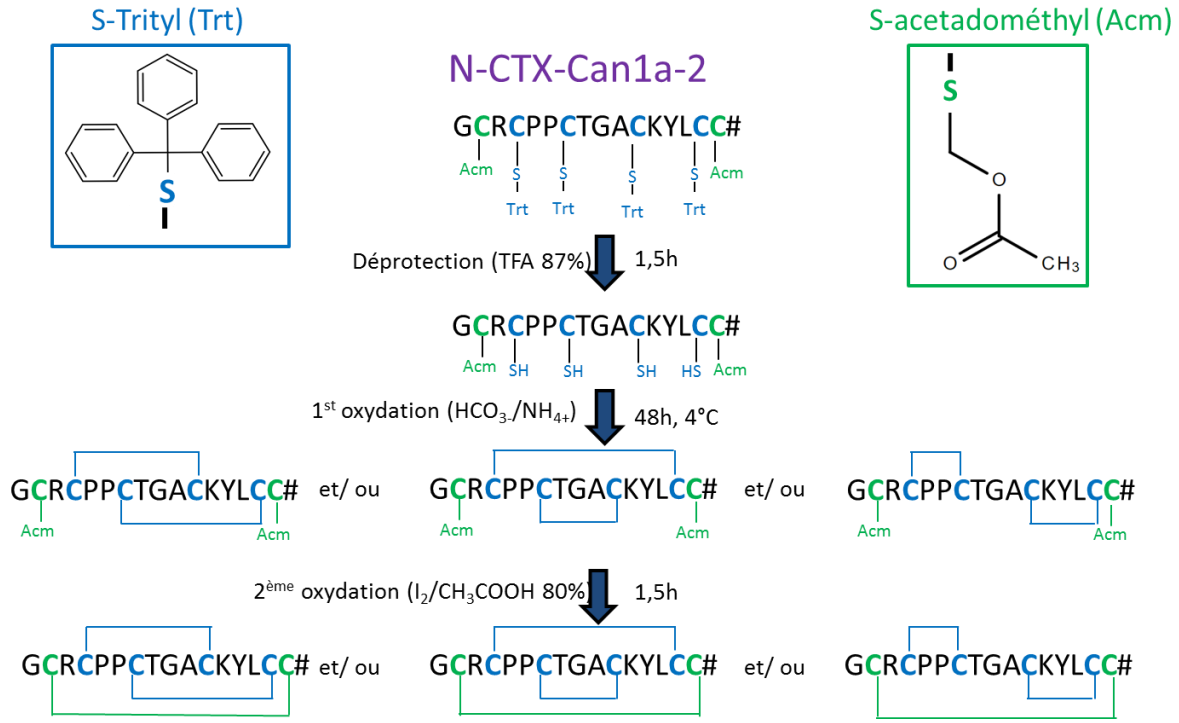


Figure 31. Stratégie de pontage dirigé partiel de N-CTX-Can1a (exemple de la forme Can1a-2). Dans une première étape, les deux premiers ponts sont formés de manière aléatoire par les cystéines non protégées (en bleu) par un groupement acétadométhyle (Acm) donnant trois possibilités d'arrangements des ponts disulfure. Une deuxième étape de pontage dirigé permet la fomation du dernier pont disulfure. # : -$CONH_2$

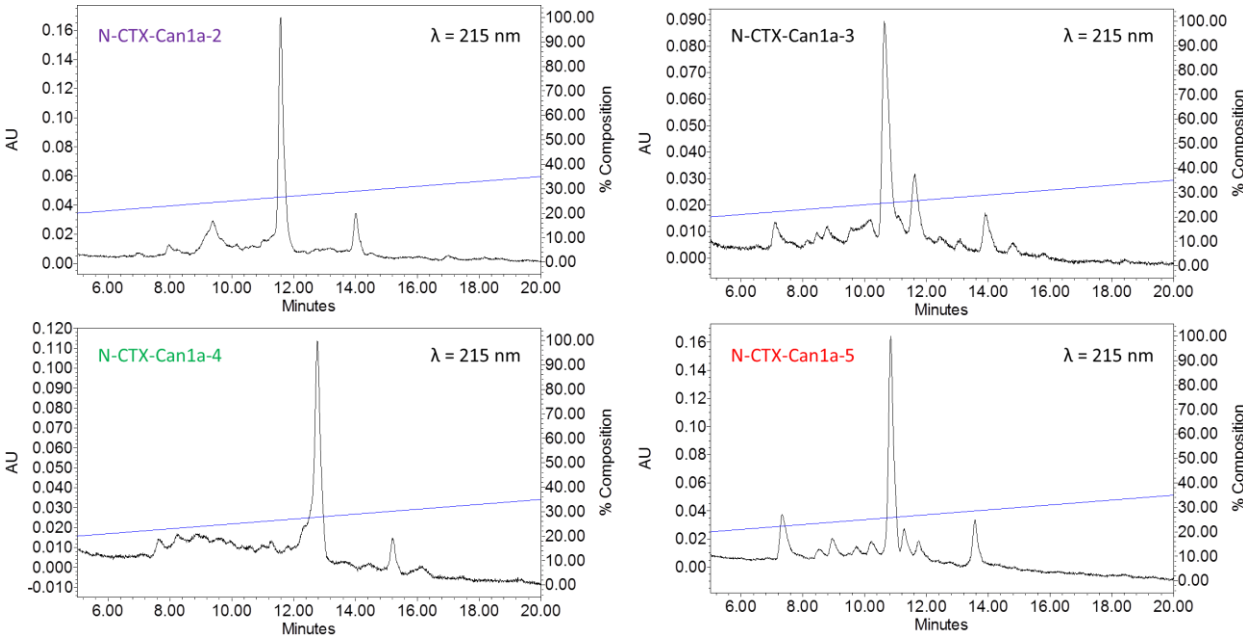


Figure 32. Profil en HPLC analytique des différents produits de N-CTX-Can1a à deux ponts disulfure obtenus après 3 jours d'oxydation aléatoire. Colonne : XBridge® BEH 300Å 5 µm, 4.6 x 250 mm ; Gradient : 20% d'ACN puis 0.75% ACN/min ; débit : 1 mL/min.

La différence de temps de rétention permet d'identifier à quelle forme en pontage dirigé correspond chacun des trois pics purifiés en oxydation aléatoire (**Figure 33**). Le premier pic T1 de l'oxydation aléatoire correspond à une forme oxydée avec un pont C2-C6, le deuxième pic T2 à une forme avec un pont C4-C6 et le dernier T3 à une forme avec un pont C1-C6. Seule la forme Can1a-4 (C3-C6) n'est pas observée avec la stratégie aléatoire d'oxydation.

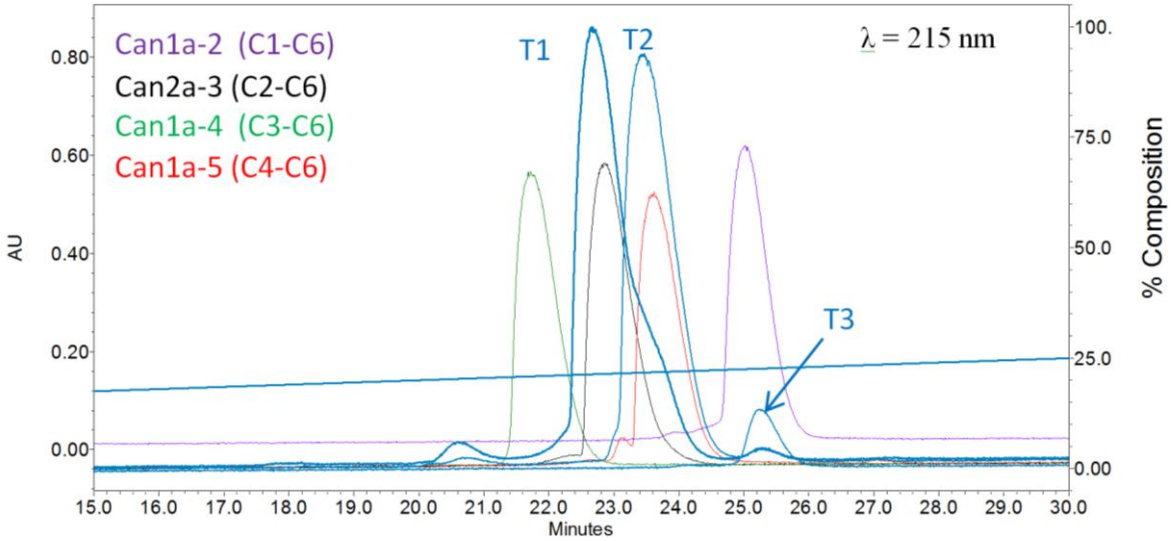


Figure 33. Attribution des pics obtenus en oxydation aléatoire à l'aide des formes oxydées après pontage dirigé de N-CTX-Can1a. Colonne : XBridge® BEH 300Å 5 µm, 4.6 x 250 mm ; Gradient : 10% d'ACN puis 1%/min ; débit : 1 mL/min.

- *Activité sur hMC4R*

Pour chacune des formes à produire, la première étape d'oxydation aboutit à un peptide à deux ponts disulfure dont l'arrangement n'est pas connu. Parmi ces peptides, seul celui de Can1a-2 (C1-C6 fixé) permet d'obtenir une affinité de 35.7 µM (n=1). Pour les autres, le Ki est supérieur à 70 µM (n=1).

Des peptides 3 ponts, tous ont des affinités de l'ordre de 10 µM sauf N-CTX-Can1a-5 dont le Ki est de 30 µM (**Table 11**). La formation du troisième pont est donc importante pour les formes -3/-4 et -5. Pour N-CTX-Can1a-2, la formation du 3ème pont permettrait d'améliorer l'affinité d'un facteur 3.

Table 11. Constantes d'inhibition (Ki) moyennes des différentes formes oxydées de N-CTX-Can1a sur hMC4R. Le peptide linéaire correspond au peptide ayant les cystéines alkylées à la suite d'un traitement à l'iodoacétamide. Moyenne (µM) ± SD (n=2-3) ; en italique : n=1.

	Pont dirigé	Ki sur hMC4R (µM)
N-CTX-Can 1a réduit	-	>100
N-CTX-Can1a T1	-	*4.86*
N-CTX-Can1a T2	-	*1.12*
N-CTX-Can1a-2	C1-C6	11.7 ± 0.5
N-CTX- Can1a -3	C2-C6	13.4 ± 0.2
N-CTX- Can1a -4	C3-C6	10.6 ± 3.4
N-CTX- Can1a -5	C4-C6	28.5 ± 4.6

- **Discussion**
  - Une toxine dérivée de la superfamille M des conotoxines ?

P014846 présente le motif de cystéine –CC-C-C-CC- de la superfamille M des conotoxines (Sato et al., 1983). Or, N-CTX-Can1a ne diffère de P014846 que par la simple inversion de Cys2 et Arg3. Une erreur lors de l'analyse du séquençage reste peu probable puisque la séquence a été retrouvée à la fois en transcriptomique et en protéomique. De plus, le changement de Cys en Arg peut être engendré par une simple mutation au niveau d'un nuclétotide. Dans l'ARNm, l'arginine est codée par CGN (N=A/C/G/U) et la cystéine par UGY (Y=C/U). Or dans l'ADN, la base méthylée 5-méthyl-cytosine peut subir une désamination donnant la thymine puis l'uracile (U) dans l'ARNm (**Figure 34A**) (Gehring and Henikoff, 2008). Il s'agit de la mutation d'un nucléotide simple la plus commune. La même réaction peut se faire directement au niveau de l'ARNm pour donner l'uracile (U) (**Figure 34A**). Ainsi le passage de Cys2-Arg3 en Arg2-Cys3 peut se faire en deux étapes selon deux possibilités (**Figure 34B**) en supposant que la séquence nucléotidique de départ soit TGYCGY. À noter que l'intermédiaire CGYCGY donnant Arg-Arg sans réaction de désamination n'a pas été détecté lors des analyses du projet Venomics. Comme les deux toxines ont été observées dans le même individu, cela suggère qu'il y ait également eu un évènement de duplication du gène. Bien que complexes, il est ainsi probable que certains processus évolutifs aient abouti à la toxine N-CTX-Can1a à partir du gène codant pour P014846.

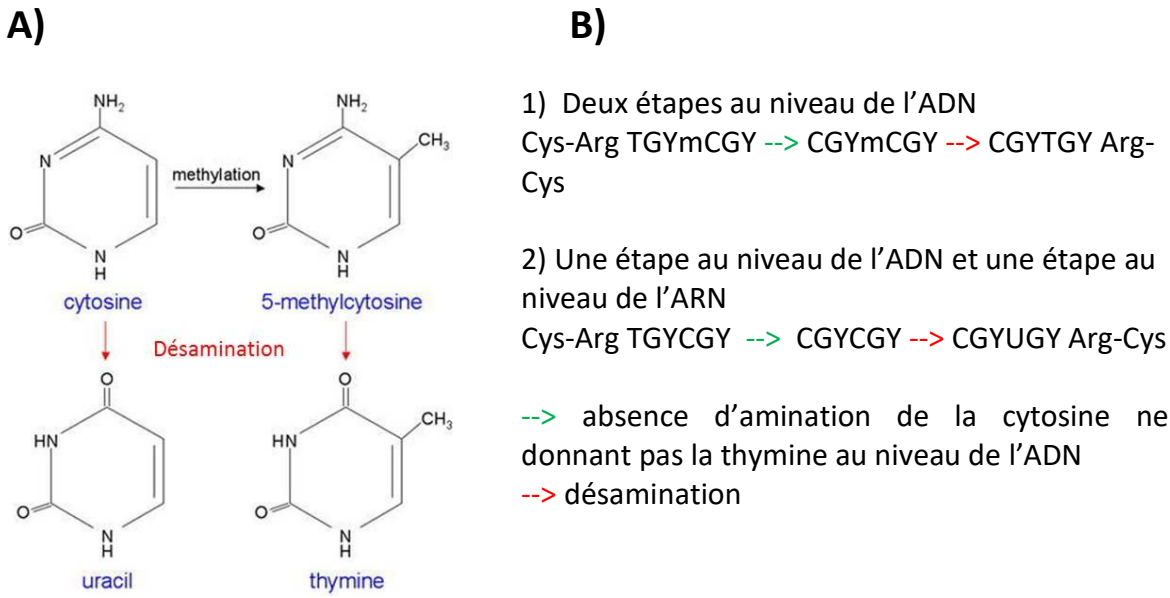


Figure 34. Mécanisme d'inversion de Cys et Arg dans la séquence de N-CTX-Can1a. A) Conversion de la cystosine en uracile ou de la méthyl-cytosine en thymine par désamination. B) Deux hypothèses expliquant l'inversion de Cys-Arg en Arg-Cys. S'il y a réaction de désamination, l'intermédiaire CGYCGY peut donner soit Arg-Cys ou redonner Cys-Arg. Sinon, il donnera Arg-Arg.

  - Une oxydation versatile

La formation d'un pont disulfure dépend directement de trois facteurs : la proximité ou l'accessibilité spatiale entre les cystéines formant le pont, la différence de pKa entre les fonctions thiol impliquées et le pH de l'environnement, et enfin l'environnement oxydo-réducteur (Mamathambika and Bardwell, 2008). En jouant sur ces facteurs en oxydation

aléatoire, on peut potentiellement modifier l'appariement des cystéines formant les ponts disulfure, ce qui ne semble pas être le cas de N-CTX-Can1a. Trois pics correspondant au peptide totalement oxydé sont toujours obtenues quelle que soit la condition testée. La faible longueur de la chaine peptidique associée à des éléments structuraux forts comme les deux prolines adjacentes orienteraient ainsi préférentiellement la formation de trois formes thermodynamiquement plus stables sur les 15 possibles. Cependant, chaque pic observé pourrait aussi être un mélange de différents arrangements comme on ne connait pas les temps de rétention de toutes les formes oxydées possibles.

- Différents appariements pour une même activité

Une absence d'activité MC4R du peptide réduit, montre que la liaison de la toxine sur hMC4R nécessite une structuration via les ponts disulfure. L'étude des formes issues du pontage dirigé montre quant à elle que la façon dont ces ponts sont arrangés n'est pas importante. Une forme n'est pas plus active que les autres.

- **Conclusion et perspectives**

Au final, N-CTX-Can1a cumule de nombreux défauts. Bien qu'intéressante, son histoire évolutive à partir de P014846 est complexe et difficile à prouver. Pour l'instant aucun protocole de production n'est au point. Or, au vue des faibles affinités sur hMC4R, plusieurs milligrammes seraient nécessaires pour réaliser toute la caractérisation pharmacologique et structurale. Mais la principale limite est qu'il n'y a aucun moyen à notre disposition pour nous assurer de la correspondance des formes produites avec la ou les formes naturelles. En effet, lors du projet Venomics, le venin a été analysé après réduction des ponts disulfure. Pour cette raison, la caractérisation de la toxine a été stoppée. Par contre, la toxine P014846 pourrait être une alternative pour une future publication. Dans la superfamille M, on distingue les mini-conotoxines M (moins de 20 résidus) des maxi-conotoxines M (plus de 20 résidus). P014846 appartiendrait donc aux mini-conotoxines M. Cependant, elle diffère des autres membres de par la faible longueur de ces boucles : -CC-x(3)-Cx(3)-Cx(3)-CC- contre -CC-x(4-6)-Cx(4-5)-Cx(1-5)-CC- (Jacob and McDougal, 2010). En outre, sa séquence ne se rapproche d'aucune autre conotoxine décrite dans la base de données UniProtKB/SwissProt faisant de P014846 potentiellement un membre fondateur d'une nouvelle branche des mini-contoxines M. De plus, P014846 testée à une concentration similaire à N-CTX-Can1a s'avère être un faux négatif du criblage primaire. En effet, le peptide purifié se révèle actif sur hMC4R (5.4 ± 1.6 µM). Ce résultat a un double intérêt. D'une part, elle ajoute un nouveau motif de cystéines actif sur ce récepteur. D'autre part, ce serait la première mini-conotoxine M avec une cible moléculaire connue. L'autre avantage de cette toxine est sa facilité de production. En oxydation aléatoire dans les conditions Venomics, P014846 ne donne qu'une seule forme oxydée (**Annexe 6**).

3.2 N-CTX-Mil1a

- **Originalité de séquence**

Provenant du cône *Conus* miliaris, la toxine N-CTX-Mil1a présente le motif de cystéine CC-CC de la superfamille T des conotoxines (Walker et al., 1999) : LQTHWNKRCCPRRPVCCFP. (**Table 12A**). Les conotoxines T sont couramment présentes dans les venins de cônes mais très peu d'informations fonctionnelles existent à leur sujet.

Table 12. Comparaison de N-CTX-Mil1a avec les séquences les plus proches de la base UniProtKB/SwissProt (A) et de la banque Venomics (B). Les cystéines sont représentées en gras et en orange. Les résidus différents de N-CTX-Can1a dans les autres séquences sont représentés en noir et en gras. La séquence du peptide maturé est soulignée en trait plein lorsque la présence du peptide mature a été prouvée au niveau protéique, en pointillé, par similarité. Les séquences ont été obtenues avec le programme Blastp du site NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi) en interrogeant la base UniProtKB/SwissProt ou la base de données Venomics (paramètres par défaut, matrice Blosum62). * : issue d'analyses transcriptionnelles ; [1](Liu et al., 2007)

A)

Nom	Séquence	%Id	Fonction	Affinité	Espèce
N-CTX-Mil1a	LQTHWNKR-CCPRRPVCCFP	100%	*MCRs*	1 µM	*Conus miliaris*
Ca5.3*	LQ**E**HWNKR-CCPRR**LA**CC**I**I**GRK**	64%			*C. caracteristicus*
Pu5.6*	LQ**Q**HWNKR-CCPRK**IW**CC**MIPR**	57%			*C. pulicarius*
Lt5d	LQ**MF**WNKRD**C**CP**AKLL**CC**N**P	55%	*Nav TTX-sensibles*[1]	0.2 µM	*C. litteratus*
Leo-T2*	LQ**EL**WNKR**P**CCP**LI**P**G**CC**R**	52%			*C. leopardus*
mr5.1a*	LQ**I**H**R**NKR-CCP**GWEL**CC**EWDEW**	50%			*C. marmoreus*

B)

Nom	Séquence	%Id	Criblage primaire sur hMC4R	Espèce
N-CTX-Mil1a	LQTHWNKRCCPRRPVCCFP	100%	Active	*C. miliaris*
P010344	**ST**LQTHWNKRCCPRRPVCCFP	100%	Active	*C. miliaris*
P010346	**DVPLASLLDNAKST**LQTHWNKRCCPRRPVCCFP	100%	Active	*C. miliaris*
P010345	**LLDNAKST**LQTHWNKRCCPRRPVCCFP	100%	Active	*C. miliaris*
P010342	CCPRRPVCCFP	58%	Pas active	*C. miliaris*
P010341	CCPRRPVCCF	53%	Active	*C. miliaris*

- **Production et caractérisation**

Comme pour la toxine précédente N-CTX-Can1a, la stratégie d'oxydation « aléatoire » des ponts disulfure a d'abord testée.

- Méthode d'oxydation aléatoire en présence d'un couple oxydo réducteur

Les conditions Venomics ont d'abord été utilisées : HEPES 100 mM, ACN 20%, couple oxydo-réducteur cystéine/cystine (ratio molaire 1 :50 :5 peptide/ cystéine/cystine). En HPLC, deux pics T1 et T2 correspondant au peptide totalement oxydé sont observées (**Figure 35**). Les deux pics purifiés ont été ensuite testés sur les récepteurs aux mélanocortines et

présentent des affinités similaires de l'ordre du micromolaire sur l'ensemble des sous-types (**Table 13**). Un peptide possédant 4 cystéines peut former trois arrangements de ponts disulfure différents nommés « Bead » (C1-C2, C3-C4), « Globular » (C1-C3, C2-C4) et « Ribbon » (C1-C4, C2-C3). Cela signifie soit que T1 ou T2 comporte en réalité deux de ces formes oxydées soit que la troisième possibilité de pontage n'est pas produite en oxydation aléatoire. Pour répondre à cette question et vérifier que la troisième forme partage des affinités similaires aux deux autres, les trois formes possibles du peptide oxydé sont produites en pontage dirigé.

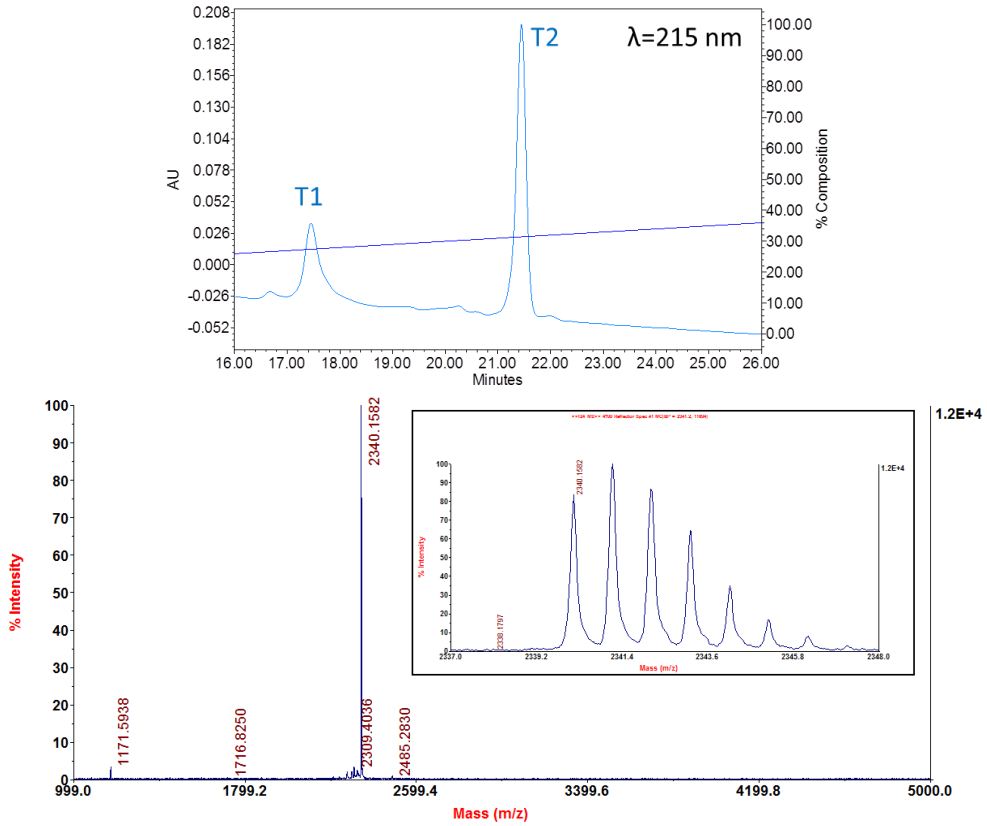


Figure 35. Oxydation aléatoire de N-CTX-Mil1a. A) Profils HPLC de N-CTX-Mil1a après 72h d'oxydation à 4°C. Colonne : Xbridge® peptide BEH C18 300 Å 3.5 µm, 4.6 x 150 mm ; Gradient : 10% d'ACN puis 1% ACN/min ; débit : 1 mL/min. B) Spectre de masse MALDI-TOF identique des pics purifiés T1 et T2. L'encadré est un agrandissement du massif isotopique de la masse majoritaire. Les masses notées correspondent aux masses monoisotopiques.

Table 13. Constantes d'inhibition (Ki) moyennes des différentes formes oxydées produites de N-CTX-Mil1a sur les récepteurs aux mélanocortines. Le peptide linéaire correspond au peptide ayant les cystéines alkylées à la suite d'un traitement à l'iodoacétamide. Moyenne (µM) ± SD (n=2-3) ; NT : non testé.

	Ki (µM)			
	hMC1R	**hMC3R**	**hMC4R**	**hMC5R**
N-CTX-Mil1a réduit	5.78 ± 1.10	3.36 ± 1.34	4.24 ± 0.58	9.07 ± 4.63
N-CTX-Mil1a T1	0.89 ± 0.08	2.39 ± 0.16	4.54 ± 4.35	0.90 ± 0.05
N-CTX-Mil1a T2	0.48 ± 0.02	2.96 ± 0.40	0.92 ± 0.39	1.00 ± 0.40
Bead	0.64 ± 0.21	1.52 ± 0.07	0.86 ± 0.19	0.63 ± 0.34
Globular	NT	NT	2.81	NT
Ribbon	0.87 ± 0.51	1.22 ± 0.43	0.74 ± 0.30	2.04 ± 1.36

- Pontage dirigé

Pour chacune des trois formes oxydées produites, le choix du premier pont, c'est-à-dire impliquant les deux cystéines ne portant pas le groupement protecteur Acm, a été fait de façon arbitraire : C3-C4 pour la forme « Bead », C1-C4 pour la forme « Ribbon » et C1-C3 pour la forme « Globular » (**Figure 36**). Les rendements des différentes étapes sont présentés en **Annexe 4**. Les formes « Ribbon » et « Bead » présentent, en HPLC analytique, un pic majoritaire correspondant à la masse du peptide oxydé et précédé d'un épaulement contenant la masse du peptide avec un incrément de +16 m/z (**Annexe 7**). Le profil HPLC de la forme « Globular » est un pic large dont la partie ascendante correspond à des incréments de +16, +32, +48 et +64 et la partie supérieure à un incrément de +32. Le peptide non modifié y est donc très minoritaire (**Annexe 7**). Dans tous les cas, la comparaison des temps de rétention des trois formes avec ceux de T1 et T2 dans les mêmes conditions ne permet de faire la correspondance qu'avec un seul pic en oxydation aléatoire (**Figure 37**). T1 correspondrait ainsi à la forme « Bead » tandis que T2 ne ferait référence à aucun des produits obtenus en pontage dirigé.

Testées sur les différents sous-types de récepteurs aux mélanocortines (à part la forme « Globular »), les trois formes issues du pontage dirigé présentent des affinités similaires entre elles sur l'ensemble des sous-types de récepteurs et restent toujours de l'ordre du micromolaire. En outre, la forme Bead et le pic T1 partagent bien les mêmes affinités sur les récepteurs au mélanocortines. La différence observée pour hMC4R est uniquement dû au faible nombre de réplicat de T1 sur ce sous-type. Par rapport au peptide réduit, le gain d'affinité engendré par l'oxydation des ponts disulfure des formes « Ribbon » et « Bead » est soit inférieur à un facteur 10 sur hMC5R soit inférieur à un facteur 5 sur les autres sous-types.

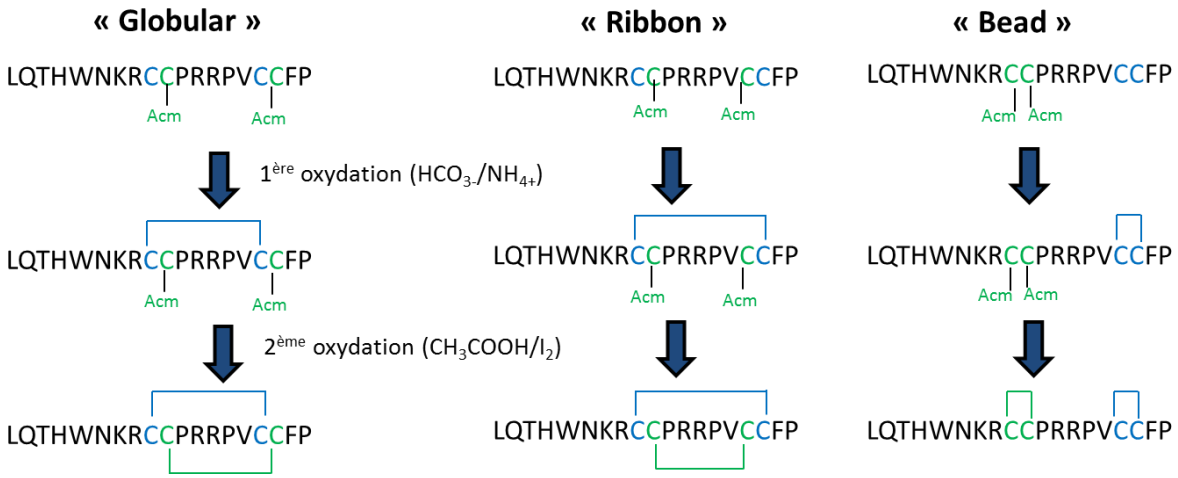


Figure 36. Stratégie de pontage dirigé de N-CTX-Mil1a. Dans une première étape, le premier pont est formé de manière aléatoire par les cystéines non protégées par un groupement acétadométhyle (Acm). Une deuxième étape de pontage dirigé permet la fomation du dernier pont disulfure.

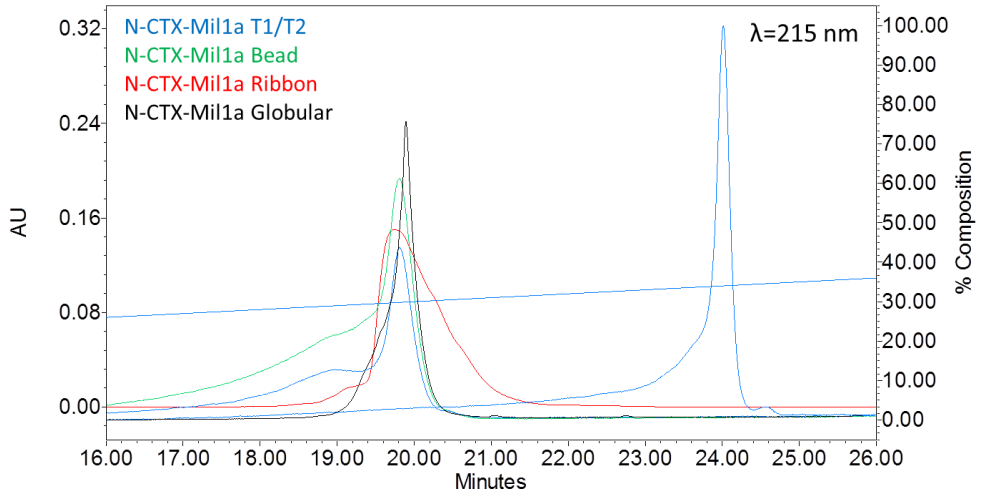


Figure 37. Attribution des pics obtenus en oxydation aléatoire à l'aide des formes oxydées après pontage dirigé de N-CTX-Mil1a. Colonne : XBridge® BEH 300Å 5 µm, 4.6 x 250 mm ; Gradient : 10% d'ACN puis 1%/min ; débit : 1 mL/min.

- **Discussion**
  - Une maturation incomplète du propeptide

Une conotoxine est produite sous forme de précurseur appelé prépropeptide composé de trois parties : une séquence signal « pré » en extrémité N-terminale suivie par une région « pro », et enfin le peptide mature (Terlau and Olivera, 2004). La séquence signal est extrêmement bien conservée et est généralement associée à un motif de cystéine particulier. Pour cette raison, elle est utilisée pour définir les grandes superfamilles de conotoxines dont la superfamille T à laquelle appartiendrait N-CTX-Mil1a (Biggs et al., 2010). La séquence « pro » jouerait quant à elle un rôle non seulement dans le passage de la membrane plasmique pour l'excrétion du peptide mature (Luo et al., 2006) mais également dans la modification post-traductionnelle (PTM) de celui-ci. Elle possèderait des signaux servant de point d'ancrage aux enzymes responsables des PTMs (Hooper et al., 2002). Le clivage du propeptide en peptide mature se fait, en règle générale, au niveau du motif KR. Or, ce dernier est retrouvé juste en amont de la première cystéine de N-CTX-Mil1a. En outre, la séquence située avant ce motif correspond à la région la mieux conservée parmi les séquences proches de N-CTX-Mil1a dans UniProtKB/SwissProt. Tout ceci est en faveur de l'hypothèse que N-CTX-Mil1a ne correpondrait pas au peptide mature. Cette hypothèse est renforcée par l'analyse des séquences similaires dans la banque Venomics (**Table 12B**). Elles proviennent toutes de la même espèce de cône *C. miliaris* et correspondent à un allongement en N-terminal ou un raccourcissement en C-terminal de la séquence de N-CTX-Mil1a. Théoriquement, l'analyse transcriptomique de la glande à venin fournit la séquence du prépropeptide tandis que l'analyse protéomique du venin devrait donner les séquences du propeptide et du peptide mature. Ainsi les différentes séquences proches de N-CTX-Mil1a seraient issues du même gène dont la région « pro » a été plus ou moins maturée et dégradée par des aminopeptidases, le peptide mature pouvant être P010341 ou P010342.

  - Production

Les incréments (+16, +32, +48 et +64) observées en pontage dirigé apparaissent uniquement après la formation du 2ème pont disulfure en présence d'I_2. Elles correspondraient à des hydroxylations. I_2 est connu pour son fort pouvoir oxydant. Des réactions secondaires peuvent ainsi se produire comme l'oxydation trop forte de la fonction

thiol engendrant un acide sulfonique ou des oxydations des résidus Tyr, Trp et Met (Laps et al., 2019). Or N-CTX-Mil1a présente un seul de ces résidus. L'oxydation de W5 en oxindole expliquerait la présence de l'incrément +16 pour les formes « Ribbon » et « Bead ». Cependant, cela n'explique toujours pas la présence des incréments +32, +48 et +64 retrouvées dans la forme « Globular ». D'autres réactions secondaires non décrites dans la littérature doivent donc se produire entrainant l'hydroxylation de plusieurs résidus.

Un peptide hydroxylé est moins hydrophobe que le même peptide non modifié. Son temps de rétention en HPLC en phase inverse est donc diminué. La forme « Globular » de N-CTX-Mil1a produite présenterait plusieurs hydoxylations. Sachant que le pic majoritaire T2 ne correspond ni à un dimère, ni aux formes « Ribbon » ou « Bead », il est donc fort probable qu'il représente en réalité la forme « Globular » non hydroxylée. L'arrangement des ponts en C1-C3, C2-C4 est d'ailleurs caractéristique des conotoxines T (Walker et al., 1999). Il est tout de même important de noter que pour beaucoup de ces toxines, l'attributation des ponts disulfure se fait par similarité par rapport aux premières conotoxines T décrites en 1999 et dont seule la forme « Globular » coélue avec la toxine naturelle. Pour confirmer la correspondance de T2 avec la forme « Globular », il est nécessaire d'utiliser une méthode sans I_2 afin d'éviter les réactions secondaires. Par exemple, Laps et al. ont développé récemment un protocole de pontage dirigé sans I_2 et plus rapide (20 min vs 1h30) en utilisant du $PdCl_2$ (10 eq.,10 min) pour la déprotection des groupements Acm puis du diethyldithiocarbamate (25 eq., 5 min) et enfin du disulfirame (2 eq., 5 min) pour favoriser la formation des ponts disulfure (Laps et al., 2019). Si la correspondance est confirmée, la forme « Globular » pourrait ensuite être facilement produite en oxydation aléatoire par purification du pic majoritaire.

- Des ponts disulfure dispensables

La perte d'affinité par réduction et alkylation du peptide est seulement d'un facteur 5-10. En outre, l'affinité du peptide oxydé est la même que ce soit pour la forme « Ribbon » ou « Bead ». Cela suggère donc que, quel que soit la structure tridimentionnelle du peptide, des interactions toxine-récepteur similaires sont toujours présentes.

- **Conclusion et perspectives**

Par rapport aux objectifs du projet de thèse, N-CTX-Mil 1a, avec des faibles affinités sur les récepteurs étudiés, n'est pas intéressante d'un point de vue thérapeutique. Cependant, elle pourrait être intéressante d'un point de vue fondamental. D'une part, sa caractérisation sur les récepteurs aux mélanocortines montre que la présence des ponts disulfure n'est pas requise pour lier une cible. D'autre part, elle ajouterait une activité à une famille de conotoxines décrite pour la première fois 20 ans auparavant mais très peu décrite au niveau pharmacologique depuis. Néanmoins, la présence de modifications des formes obtenues en pontage dirigé limitent l'exploitation des résultats dans un article scientifique. En outre, au vu des faibles affinités sur les récepteurs aux mélanocortines, l'exploitation dans un article de N-CTX-Mil1a nécessite plusieurs milligrammes de toxine oxydée. Pour cela, la solution la plus simple serait d'identifier la forme majoritaire en oxydation aléatoire. L'hypothèse la plus probable étant la forme « Globular », la production de celle-ci doit être entreprise en testant de nouvelles conditions de pontage dirigé limitant la modification du peptide. Mais, au vu des faibles affinités sur hMC4R, il a été décidé d'arrêter pour l'instant son étude pour se focaliser sur une autre toxine à deux ponts présentant de meilleures affinités.

4. D'une conotoxine à un potentiel lead sur hMC4R

4.1 Originalité de séquence

La toxine N-CTX-Ltg1a du cône *Conus lithoglyphus* présente le motif de cystéine -C-C-C-C-. Ce dernier a déjà été décrit chez les cônes pour la première fois en 2005 dans le venin de deux autres espèces vermivores *Conus floridanus floridensis* et *Conus villepinii* (Möller et al., 2005).

Néanmoins, la séquence de N-CTX-Ltg1a diffère fortement des toxines de ces cônes avec des motifs répétitifs « CQ », « PP », « GFK ». L'autre différence concerne la taille des boucles intercystéines selon le motif -C-x-C-x-C-x(8)-C- *versus* –C-x(3)-C-x(11)-C-x(3)-C- (où x est n'importe quel acide aminé) des toxines de *Conus floridanus floridensis* et *Conus villepinii*. Ce motif particulier est assimilé au domaine EGF-like (**Figure 38**) observé dans la région extracellulaire de protéines membranaires ou de protéines sécrétées dont les fonctions sont très diverses (immunité, apoptose, …). N-CTX-Ltg1a respecte bien celui-ci mais uniquement dans sa dernière partie. En effet, le domaine EGF-like contient normalement six cystéines dont l'arrangement des ponts disulfure formés est conservé parmi les protéines présentant ce domaine et ne peut être présent dans N-CTX-Ltg1a. Cette dernière ferait ainsi parti d'une nouvelle famille de toxines dont d'autres représentants sont présents dans la banque Venomics (**Table 14**). Ils représentent essentiellement une variation de la partie N-terminale en amont de la première cystéine et de la modification en position 8 de N-CTX-Ltg1a de Q en E. Dans l'ARNm, la glutamine (Q) est codée par CAN et l'acide glutamique par GAN avec N=A ou G. Une simple mutation pourrait ainsi expliquer cette variation. Néanmoins, il apparait peu probable qu'une base purique G mute en une base pyrimidique C ou inversement. Il s'agit donc sûrement d'une désamidation spontanée de la glutamine bien que le motif CQC ne soit pas connu pour favoriser cette réaction (Wright, 1991).

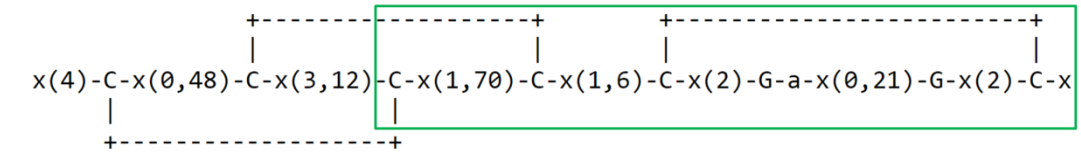


Figure 38. Domaine EGF-like (Prosite PDOC00021). L'encadrée en vert correspond à la partie du domaine respectée par la séquence de N-CTX-Ltg1a.

Table 14. Comparaison de N-CTX-Ltg1a avec les séquences les plus proches de la banque Venomics. Les cystéines sont représentées en gras et en orange. Les résidus différents de N-CTX-Ltg1a dans les autres séquences sont en noir et en gras.

Nom	Séquence	%Id	Criblage primaire sur hMC4R	Espèce
N-CTX-Ltg1a	LFHY**C**Q**C**Q**C**PPGFKGKF**C**QFKLRPP	100%	Active	*C. lithoglyphus*
P014791	**TH**LFHY**C**Q**C**Q**C**PPGFKGKF**C**QFKLRPP	100%	Active	*C. lithoglyphus*
P014788	**H**LFHY**C**Q**C**Q**C**PPGFKGKF**C**QFKLRPP	100%	Active	*C. lithoglyphus*
P014790	**TH**LFHY**C**Q**C****E****C**PPGFKGKF**C**QFKLRPP	96%	Active	*C. lithoglyphus*
P014787	**LH**-FHY**C**Q**C**Q**C**PPGFKGKF**C**QFKLRPP	96%	Active	*C. lithoglyphus*
P014780	FHY**C**Q**C**Q**C**PPGFKGKF**C**QFKLRPP	96%	Active	*C. lithoglyphus*
P014781	FHY**C**Q**C****E****C**PPGFKGKF**C**QFKLRPP	92%	Active	*C. lithoglyphus*
P014778	HY**C**Q**C**Q**C**PPGFKGKF**C**QFKLRPP	92%	Active	*C. lithoglyphus*
P014777	HY**C**Q**C****E****C**PPGFKGKF**C**QFKLRPP	88%	Active	*C. lithoglyphus*
P014773	Y**C**Q**C**Q**C**PPGFKGKF**C**QFKLRPP	88%	Active	*C. lithoglyphus*
P014774	**Q****C****Y****C**Q**C**PPGFKGKF**C**QFKLRPP	80%	Active	*C. lithoglyphus*
P014775	**C****Y****Q****C**Q**C**PPGFKGKF**C**QFKLRPP	80%	Active	*C. lithoglyphus*
P014710	**Q****C****C****Y**Q**C**PPGFKGKF**C**QFKLRPP	76%	Active	*C. lithoglyphus*

4.2 D'un peptide réticulé à un peptide linéaire

- **D'un peptide à 2 ponts disulfure …**

Ne sachant pas quelle(s) forme(s) oxydée(s) est/sont naturellement présente(s) dans le venin, nous avons réalisé une oxydation aléatoire (**Figure 39**). Seuls les deux premiers pics présentent la masse du monomère oxydé, le pic T3 correspondant à un dimère (**Figure 39**).

Les différentes formes ont été testées sur l'ensemble des sous-types de récepteurs aux mélanocortines (**Table 15**). Le nombre de réplicat étant trop faible (n=1 ou 2), il n'est pas possible de faire une analyse statistique rigoureuse. Néanmoins, que ce soit pour la forme réduite, les monomères oxydés T1 et T2 ou le dimère T3, les affinités sont similaires sur chacun des sous-types de récepteur étudiés sauf hMC5R. Ainsi, les différentes formes oxydées et réduite présentent toutes des affinités de l'ordre de 1 µM sur hMC1R et de 200 nM sur hMC3R et hMC4R. Par contre, l'oxydation des ponts disulfure semble plus importante pour la liaison sur hMC5R que sur les autres sous-types. En effet, les formes oxydées montrent une amélioration de l'affinité sur ce sous-type d'un facteur 5 par rapport à la forme réduite. Sachant que la toxine présente deux ponts disulfure, trois formes oxydées peuvent se former lors de leur oxydation : « Globular », « Bead » et « Ribbon ». Les trois formes ont été produites en pontage dirigé selon la même stratégie que la toxine N-CTX-Mil1a. Cette technique a permis notamment d'identifier les pics T1 et T2, respectivement comme étant les formes « Bead » et « Ribbon ». D'autre part, cela a permis également d'observer que la forme « Globular » est très peu produite lors de l'oxydation aléatoire (**Figure 40**). Toutes les formes oxydées présentent des affinités de l'ordre de 50 nM sur les sous-types hMC4R et hMC5R, 100 nM sur hMC3R et 500 nM sur hMC1R (**Table 16**). Des formes tronquées aux extrémités ont été testées pour la forme Ribbon. La perte de l'extrémité N-terminale fait perdre un facteur 6 d'affinité sur hMC1R et hMC3R, un facteur 28 sur hMC4R et un facteur 61 sur hMC5R. À l'opposé, lorsque l'extrémité C-terminale est tronquée, aucun effet sur l'affinité n'est observé quel que soit le récepteur considéré. Par contre, la perte simultanée des deux extrémités fait perdre un facteur 60 sur hMC1R et supérieur à 100 sur les autres sous-types. La partie N-terminale de la toxine est donc plus

importante que la partie C-terminale pour la liaison sur les MCRs. Par rapport au peptide réduit, le gain d'affinité du peptide à deux ponts disulfure est relativement faible : facteur 2 sur hMC1R, facteur 4-5 sur hMC3R et hMC4R et facteur 30 sur hMC5R. L'arrangement des ponts n'est donc pas primordial pour la liaison de la toxine sur les hMCRs. Partant de ce constat et des faibles rendements de production en pontage dirigé (**Annexe 4**), il a été décidé de ne garder que le pont formant la grande boucle pour chacune des trois formes oxydées possibles, c'est-à-dire celle supposée la plus structurante. Ceci a pour but de pouvoir produire facilement et rapidement un peptide tout en gardant des affinités similaires au peptide 2 ponts. Pour ce faire, les cystéines impliquées dans les petites boucles sont remplacées par des alanines (**Figure 41**).

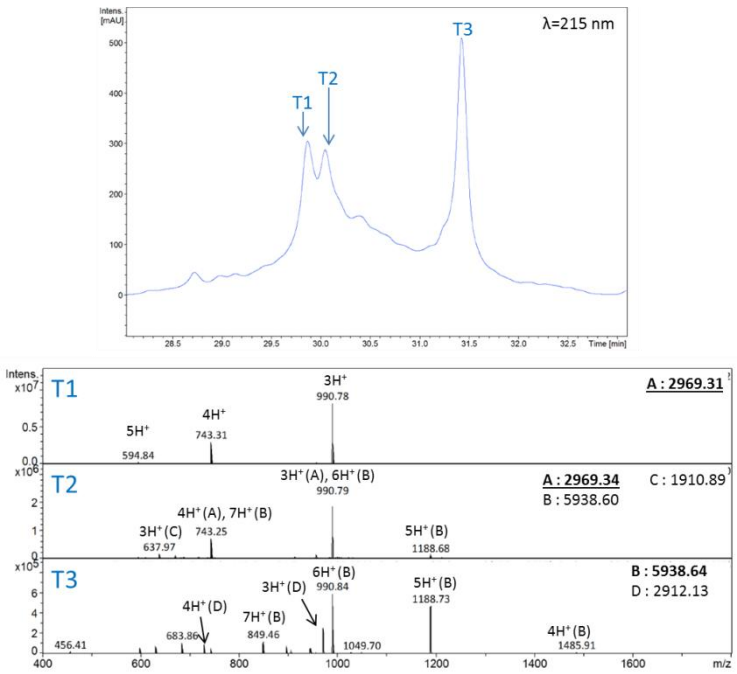


Figure 39. Analyse LC-MS (ESI-IT) du produit obtenu après oxydation aléatoire de la toxine N-CTX-Ltg1a. La masse majoritaire de chacun des spectres est signalée en gras. La masse de la toxine N-CTX-Ltg1a est soulignée lorsqu'elle est détectée. Colonne : XBridge® Peptide BEH C18 300Å 3.5 µm, 4.6 x 150 mm ; Gradient : 0% ACN/TFA0.1% puis 1% ACN/TFA0.1% /min ; débit : 1 mL/min.

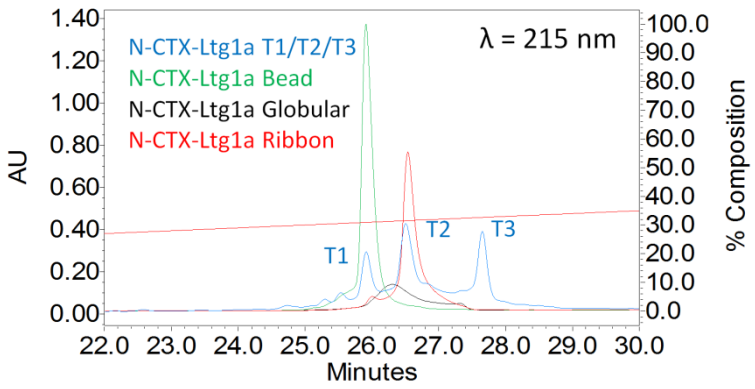


Figure 40. Attribution des pics obtenus en oxydation aléatoire (en bleu) de N-CTX-Ltg1a à l'aide des formes oxydées après pontage dirigé. Colonne : XBridge® BEH 300Å 5 µm, 4.6 x 250 mm ; Gradient : 5% ACN puis 1% d'ACN/min ; débit : 1 mL/min.

Table 15. Constantes d'inhibition (Ki) moyennes des différentes formes produites de N-CTX-Ltg1a et de ses variants sur les récepteurs aux mélanocortines aboutissant à la Hunger Toxin (HT1-0). Les variations de la séquence d'origine sont représentées en gras. 1 : alkylation des cystéines à la suite d'un traitement à l'iodoacétamide ; # : -$CONH_2$. N-CTX-Ltg1a T3 correspond à un dimère de la toxine. Moyenne (nM) ± SD (n=2-13) ; en italique : n=1 ; NT : non testé.

		Ki (nM)			
	Séquence	**hMC1R**	**hMC3R**	**hMC4R**	**hMC5R**
	1 5 10 15 20 25				
N-CTX-Ltg1a réduit	LFHY**1**Q**1**Q**1**PPGFKGKF**1**QFKLRPP	1020 ± 485	586 ± 146	202 ± 60	1083 ± 165
N-CTX-Ltg1a T1	LFHYCQCQCPPGFKGKFCQFKLRPP	*1096*	241 ± 50	187 ± 81	*178*
N-CTX-Ltg1a T2	LFHYCQCQCPPGFKGKFCQFKLRPP	*942*	187 ± 38	197 ± 14	*2151*
N-CTX-Ltg1a T3	LFHYCQCQCPPGFKGKFCQFKLRPP	*856*	134 ± 63	83 ± 9	*225*
« Bead »	LFHYCQCQCPPGFKGKFCQFKLRPP	530 ± 128	83 ± 49	45 ± 16	30 ± 8
« Globular »	LFHYCQCQCPPGFKGKFCQFKLRPP	649 ± 220	90 ± 15	49 ± 20	53 ± 60
« Ribbon »	LFHYCQCQCPPGFKGKFCQFKLRPP	554 ± 113	110 ± 56	48 ± 21	30 ± 15
« Ribbon » ΔNter	----CQCQCPPGFKGKFCQFKLRPP	*3377*	*724*	1332 ± 163	*1827*
« Ribbon » ΔCter	LFHYCQCQCPPGFKGKFC#------	663 ± 26	74 ± 17	46 ± 9	46 ± 21
« Ribbon » Δ(Nter+Cter)	----CQCQCPPGFKGKFC#------	32814 ± 5788	17217 ± 4148	24540 ± 8249	3008 ± 2317
Ltg1a-1SS-1	LFHYCQ**A**Q**A**PPGFKGKFCQFKLRPP	1060 ± 450	172 ± 40	130 ± 60	229 ± 169
Ltg1a-1SS-2	LFHY**A**QCQ**A**PPGFKGKFCQFKLRPP	1669 ± 597	NT	251 ± 57	NT
Ltg1a-1SS-3	LFHY**A**Q**A**QCPPGFKGKFCQFKLRPP	2856 ± 530	NT	259 ± 88	NT
Ltg1a-1SS-4	LFHYCQ**S**Q**S**PPGFKGKFCQFKLRPP	779 ± 79	229 ± 45	81 ± 47	235 ± 67
Ltg1a-1SS-5	LFHYCQ**V**Q**V**PPGFKGKFCQFKLRPP	732 ± 148	286 ± 63	68 ± 37	227 ± 51
Ltg1a-1SS-6	LFHYCQ**L**Q**L**PPGFKGKFCQFKLRPP	402 ± 87	281 ± 52	110 ± 60	122 ± 34
Ltg1a-1SS-7	LFHYCQ**K**Q**K**PPGFKGKFCQFKLRPP	89 ± 27	33 ± 12	9 ± 6	26 ± 8
Ltg1a-0SS-1	LFHY**1**Q**K**Q**K**PPGFKGKF**1**QFKLRPP	115 ± 38	38 ± 9	13 ± 5	139 ± 40
Ltg1a-0SS-2	LFHY**K**Q**K**Q**K**PPGFKGKF**K**QFKLRPP	13 ± 2	21 ± 10	5 ± 3	26 ± 4
Ltg1a-0SS-3 (HT1-0)	LFHY**Q**Q**K**Q**K**PPGFKGKF**Q**QFKLRPP	197 ± 104	41 ± 14	15 ± 9	195 ± 90

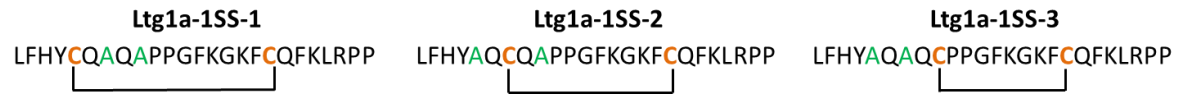


Figure 41. Séquences des trois formes à un pont disulfure issues de N-CTX-Ltg1a. Les acides aminés en vert représentent les anciennes positions des cystéines formant le deuxième pont disulfure dans la séquence de N-CTX-Ltg1a.

- **… à un peptide à 1 pont …**

Une perte d'affinité significative est observée entre le peptide à 1 pont et le peptide à 2 ponts disulfure correspondant sur hMC1R et hMC4R (**Table 15**). Néanmoins, la forme Ltg1a-1SS-1 avec un pont entre Cys5 et Cys18 est celle qui présente la perte d'affinité la plus faible sur hMC4R (facteur 2-3). Afin de revenir à des affinités inférieures à 100 nM, les alanines en positions 7 et 9 sont substituées par des acides aminés ayant des propriétés

physico-chimiques différentes : un résidu polaire non chargé, la sérine, un résidu hydrophobe, valine ou leucine ou un résidu basique chargé, la lysine.

Sur hMC1R et hMC4R, la substitution des alanines par des résidus hydrophobes ou polaires ne modifie pas l'affinité du peptide 1 pont. Par contre, l'affinité est significativement améliorée si des lysines sont ajoutées à la séquence (peptide Ltg1a-1SS-7) (**Table 15**).

- **… à un peptide linéaire nanomolaire**

La forme linéaire alkylée du peptide Ltg1a-1SS-7 (Ltg1a-0SS-1) a également été testée sur les hMCRs avec des affinités similaires à la forme 1 pont. En comparaison du peptide linéaire N-CTX-Ltg1a avec les quatre cystéines alkylées, la forme Ltg1a-0SS-1 présente des affinités 5 à 15 fois meilleures. Les lysines en position 7 et/ou 9 jouent donc un rôle majeur pour la liaison de ce peptide. Sachant cela, les autres cystéines en position 5 et 18 ont aussi été remplacées par des lysines (Ltg1a-0SS-2). Le profil du peptide est néanmoins moins intéressant que celui avec les cystéines alkylées. Bien que l'affinité sur hMC4R soit améliorée d'un facteur 2 sur hMC4R, les facteurs de sélectivité par rapport aux autres sous-types sont diminués. Cependant, Ltg1a-0SS-1 requiert une étape de traitement à l'iodoacétamide entrainant l'alkylation des fonctions thiols des cystéines. Afin d'éliminer cette étape, Cys5 et Cys18 sont remplacées par des glutamines dont la structure est proche de celle d'une cystéine alkylée (**Figure 42**). Le peptide qui en résulte, Ltg1a-0SS-3, conserve exactement le même profil que le peptide Ltg1a-0SS-1. Le remplacement des cystéines alkylées par des glutamines n'a donc aucun effet. Au niveau structural, l'analyse en dichroïsme circulaire révèle que l'alkylation des quatre cystéines entraine une augmentation de l'intensité des deux maximums (négatif à 200 nm et positif à 220 nm) (**Figure 43A**). Cette augmentation est encore plus importante lorsque les cystéines sont remplacées par deux glutamines et deux lysines (Ltg1a-0SS-1). Selon le logiciel BeStSel, cette variation serait associée à une disparition d'hélice et une diminution de turn tandis que les proportions des autres structures augmentent (**Figure 43B**). Néanmoins, l'analyse des spectres par un autre logiciel (Spectra Manager™ II (JASCO)) ne prédit aucune structure secondaire.

On obtient ainsi un peptide linéaire facile à produire (rendement de synthèse >60%) et présentant un meilleur profil d'affinité et de sélectivité que la toxine N-CTX-Ltg1a de départ. Ltg1a-0SS-3 a également été testé sur d'autres RCPGs. À 50 µM, le peptide inhibe 50% de la liaison du traceur sur le récepteur 1 à l'angiotensine mais n'est actif ni sur le récepteur 2 à l'angiotensine ni sur le récepteur 2 à la vasopressine.

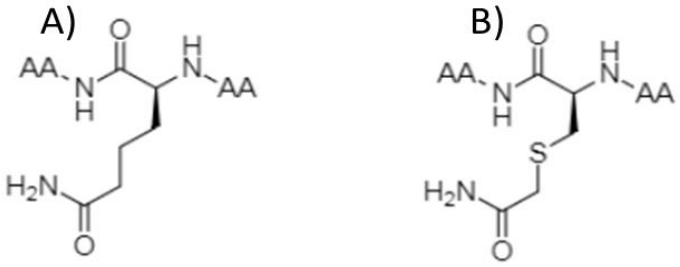


Figure 42. Comparaison des structures de la L-glutamine (A) et d'une cystéine alkylée à l'iodoacétamide (B) dans une chaine peptidique. AA : acide aminé.

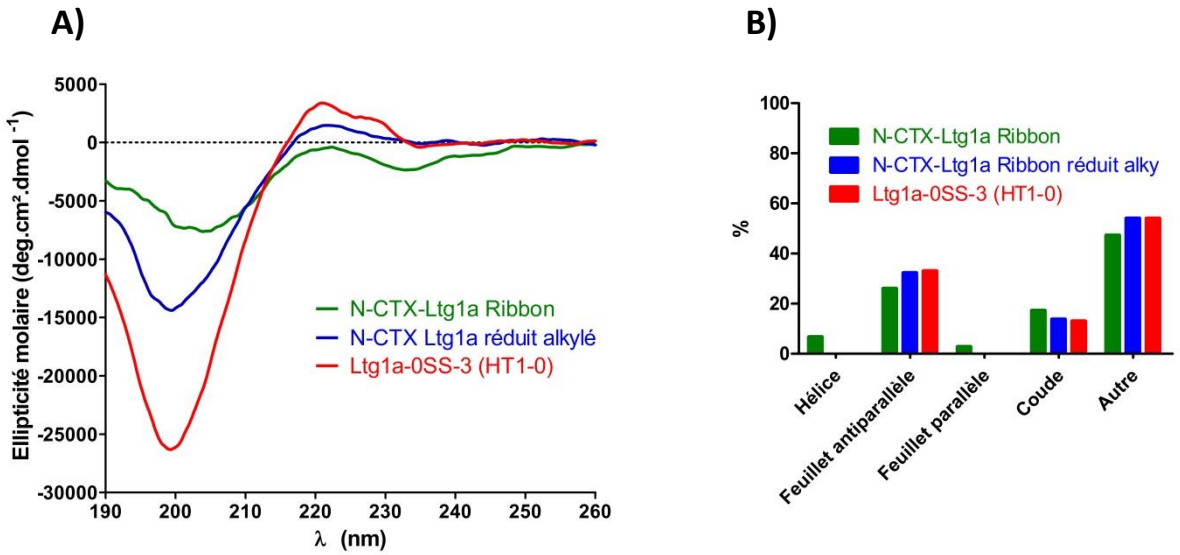


Figure 43. Analyse en dichroisme circulaire de la toxine N-CTX-Ltg1a et de son dérivé, la Hunger-Toxin (HT1-0). A) Spectre de dichroïsme circulaire. B) Prédiction des structures secondaires à l'aide du logiciel BeStSel.

Pour la suite de l'étude, Ltg1a-0SS-3 est renommé « Hunger-Toxin » (HT1-0).

La simplification de la production associée à l'amélioration de l'affinité par rapport à la toxine d'origine est un bon début dans le développement de la Hunger-Toxin. L'étape *in vitro* suivante est d'étudier plus dans le détail la liaison du peptide sur le récepteur hMC4R via notamment les différents paramètres cinétiques associés.

4.3 Étude la liaison de HT1-0 sur les récepteurs aux mélanocortines

- **Relation structure-activité (SAR)**

Le principal but des études de SAR est d'identifier les positions qui peuvent être modifiées sans affecter négativement l'affinité. Elles sont ainsi aussi indispensables dans le processus d'optimisation d'un lead que dans l'optique de marquer la molécule afin de connaitre ses paramètres cinétiques *in vitro* et sa biodistribution *in vivo*. La **Table 16** récapitule les différents variants testés. On estime ici qu'une position est importante si sa modification induit une perte significative ($p<0.001$) au moins d'un facteur 3. Les formes tronquées aux extrémités montrent une plus grande importance de l'extrémité N-terminale par rapport à celle en C-terminal. Cette observation est confirmée par les résultats de l'Alascan où chaque résidu est successivement remplacé par une alanine. Cette modification au niveau des cinq premiers résidus entraine une perte d'affinité supérieure à un facteur 4 par rapport à l'affinité du peptide d'origine. Au niveau de Tyr4, la substitution par une Phe entraine une perte d'un facteur 3. Bien que significative, cette perte reste relativement faible comparée à un facteur 7 avec une alanine. Par contre, une substitution par un résidu plus encombrant comme une diiodo-tyrosine ne fait pas perdre d'affinité. La position 4 doit donc être occupée par un résidu hydrophobe et aromatique et dans une moindre mesure porter une fonction polaire comme un hydroxyle. Q5A fait perdre un facteur 4 d'affinité.

Cette position nécessite donc un résidu polaire. La zone entre Q6 et F13 n'est au contraire pas cruciale pour la liaison. La charge portée par K7 n'est pas importante. Pour K9, les résultats sont similaires mais la variation des données pourrait masquer un effet significatif. En position 13, l'ajout d'une fonction hydroxyle en substituant Phe par Tyr n'a pas d'effet. Par contre, la présence d'un résidu plus encombrant comme une diiodo-tyrosine fait perdre significativement un facteur 30 d'affinité. Cela suggèrerait un encombrement stérique mais ce résultat reste à confirmer (n=1). Pour le reste de la séquence, seules les positions correspondant à des résidus basiques (Lys ou Arg) sont importantes. L'absence d'une longue chaine carbonnée chargée positivement a un effet négatif sur l'affinité (facteur 4-5). L'inversion de la charge via la substitution par un résidu acide entraine une perte plus importante (facteur 17 et 13) pour les positions 14 et 16. À ces positions, la présence de la charge est donc plus importante que la taille de la chaine latérale. Cela n'est pas le cas pour la position 21. En effet, la substitution de Lys en Ala ou Glu entraine une perte similaire d'affinité (facteur 5-6) alors que celle en azido-lysine ne fait perdre qu'un facteur 2. Dans une azido-lysine (**Figure 44**), la chaine carbonnée de la chaine latérale est identique à celle d'une lysine. Cependant, elle possède un groupement azido à la place d'une fonction amine primaire. Elle ne peut donc plus faire de liaison hydrogène via sa chaine latérale. Les résultats de substitution en position 21 suggèrent ainsi un rôle de l'hydrophobicité et/ou de la taille de la chaine latérale. Bien que l'inversion de charge n'ait pas été testée, les résultats en position 23 suggèrent un rôle similaire. Enfin, l'amidation en C-terminal n'aurait pas d'effet. Tous ces résultats permettent ainsi de construire une cartographie de la séquence délimitant les zones peu importantes dans la liaison du peptide sur hMC4R : LFHYQQKQKPPGFKGKFQQFKLRPP, avec les résidus importants en rouge, ceux peu impliqués en vert et ceux pas testés en noir.

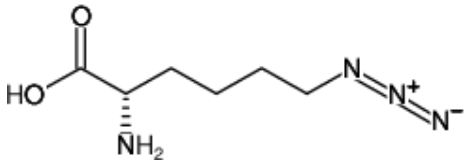


Figure 44. Structure de 6-Azido-L-Lysine

Ces résultats sont précieux pour la suite des expériences. En effet, l'étude des paramètres cinétiques de la liaison de HT1-0 sur hMC4R requiert le marquage du peptide. Ce marquage peut être soit radioactif (^{125}I, ^{3}H) soit réalisé avec un fluorophore. L'avantage d'un marquage radioactif est son coût relativement faible par rapport à un marquage avec un fluorophore. Cet avantage est cependant contre-balancé en ce qui concerne d'une part, l'exposition du manipulateur à des rayons ionisants, seule une personne formée à ce genre de risques peut effectuer le marquage, et d'autre part, la gestion des déchets générés. Les deux stratégies sont donc explorées en parallèle.

Table 16. Relation structure-affinité de la Hunger-Toxin HT1-0 sur hMC4R. Pour l'analyse statistique des données, les valeurs de Ki ont été transformées en pKi =-log(Ki) afin de respecter la loi normale et utiliser un test paramétrique : ANOVA à un facteur suivi du test post-hoc de Dunnett avec HT1-0 comme référence. ns : non significatif ; *: p<0.05 ; ** : p<0.01 ; ***: p<0.001. n=1-16 ; en italique : n=1 ; NT : non testé. Les variations de la séquence d'origine sont représentées en gras. Le code couleur correspond à l'effet significatif de la variation sur l'affinité : bleu, amélioration significative: (p<0.001) ; vert : pas d'effet (p>0.001) ; rouge : perte significative (p<0.001) d'au moins un facteur 3. 1 : diiodo-tyrosine ; 2 : 6-Azido-L-Lysine ; 3 : lysine-Fluorescéine ; 4 : lysine-Rhodamine ; # : -$CONH_2$.

	Séquence 1 5 10 15 20 25	Ki (nM)	± SD (nM)	Facteur de perte	p-value
HT1-0	LFHYQQKQKPPGFKGKFQQFKLRPP	15	± 9		
HT1-1	**-**FHYQQKQKPPGFKGKFQQFKLRPP	334	± 231	22	***
HT1-2	**--**HYQQKQKPPGFKGKFQQFKLRPP	2205	± 1035	143	***
HT1-3	**---**YQQKQKPPGFKGKFQQFKLRPP	1438	± 782	93	***
HT1-4	LFHYQQKQKPPGFKGKFQQFKLR**--**	14	± 7	1	ns
HT1-5	LFHYQQKQKPPGFKGKFQQFKL**---**	29	± 13	2	ns
HT1-6	LFHYQQKQKPPGFKGKFQQF**-----**	55	± 15	4	**
HT1-7	LFHYQQKQKPPGFKGKFQ**-------**	134	± 90	9	***
HT1-8	**Ac-**LFHYQQKQKPPGFKGKFQQFKLRPP	138	± 26	9	***
HT1-9	**A**FHYQQKQKPPGFKGKFQQFKLRPP	125	± 31	8	***
HT1-11	L**A**HYQQKQKPPGFKGKFQQFKLRPP	236	± 31	15	***
HT1-12	LF**A**YQQKQKPPGFKGKFQQFKLRPP	70	± 3	5	***
HT1-13	LFH**A**QQKQKPPGFKGKFQQFKLRPP	111	± 16	7	***
HT1-14	LFH**F**QQKQKPPGFKGKFQQFKLRPP	39	± 7	3	***
HT1-15	LFH**1**QQKQKPPGFKGKFQQFKLRPP	29	± 11	2	*
HT1-16	LFHY**A**QKQKPPGFKGKFQQFKLRPP	69	± 17	4	***
HT1-17	LFHYQ**A**KQKPPGFKGKFQQFKLRPP	24	± 3	2	ns
HT1-18	LFHYQQ**A**QKPPGFKGKFQQFKLRPP	33	± 13	2	ns
HT1-19	LFHYQQK**A**KPPGFKGKFQQFKLRPP	23	± 10	1	ns
HT1-20	LFHYQQKQ**A**PPGFKGKFQQFKLRPP	50	± 24	3	*
HT1-21	LFHYQQKQKP**A**GFKGKFQQFKLRPP	29	± 11	2	ns
HT1-22	LFHYQQKQKPP**A**FKGKFQQFKLRPP	36	± 4	2	**
HT1-23	LFHYQQKQKPPG**A**KGKFQQFKLRPP	14	± 8	1	ns
HT1-24	LFHYQQKQKPPG**Y**KGKFQQFKLRPP	11	± 2	1	ns
HT1-25	LFHYQQKQKPPG**1**KGKFQQFKLRPP	*485*	/	31	/
HT1-26	LFHYQQKQKPPGF**A**GKFQQFKLRPP	54	± 26	4	***
HT1-27	LFHYQQKQKPPGF**E**GKFQQFKLRPP	262	± 43	17	***
HT1-28	LFHYQQKQKPPGFK**A**KFQQFKLRPP	23	± 5	1	ns
HT1-29	LFHYQQKQKPPGFKG**A**FQQFKLRPP	75	± 23	5	***
HT1-30	LFHYQQKQKPPGFKG**E**FQQFKLRPP	203	± 36	13	***
HT1-31	LFHYQQKQKPPGFKGK**A**QQFKLRPP	16	± 8	1	ns
HT1-32	LFHYQQKQKPPGFKGK**1**QQFKLRPP	142	± 101	9	/
HT1-33	LFHYQQKQKPPGFKGKF**A**QFKLRPP	18	± 6	1	ns
HT1-34	LFHYQQKQKPPGFKGKFQ**A**FKLRPP	22	± 10	1	ns
HT1-35	LFHYQQKQKPPGFKGKFQQ**A**KLRPP	33	± 9	2	**
HT1-36	LFHYQQKQKPPGFKGKFQQ**Y**KLRPP	14	± 4	1	ns
HT1-37	LFHYQQKQKPPGFKGKFQQ**1**KLRPP	19	± 8	1	ns
HT1-38	LFH**F**QQKQKPPGFKGKFQQ**1**KLRPP	*122*	/	8	/
HT1-39	LFHYQQKQKPPGFKGKFQQF**A**LRPP	76	± 18	5	***

HT1-40	LFHYQQKQKPPGFKGKFQQF**E**LRPP	90	± 20	6	***
HT1-41	LFHYQQKQKPPGFKGKFQQF**2**LRPP	37	± 12	2	***
HT1-42	LFHYQQKQKPPGFKGKFQQF**3**LRPP	*153*	/	10	/
HT1-43	LFHYQQKQKPPGFKGKFQQF**4**LRPP	*230*	/	15	/
HT1-44	LFHYQQKQKPPGFKGKFQQFK**A**RPP	28	± 5	2	ns
HT1-45	LFHYQQKQKPPGFKGKFQQFKL**A**PP	61	± 19	4	***
HT1-46	LFHYQQKQKPPGFKGKFQQFKL**2**PP	36	± 14	2	**
HT1-48	LFHYQQKQKPPGFKGKFQQFKLR**A**P	30	± 3	2	*
HT1-49	LFHYQQKQKPPGFKGKFQQFKLRPP#	9	± 4	0,6	ns

- **Marquage**
  - Fluorescent
    - Choix de la position

Le marquage fluorescent en N-terminal de HT1-0 n'est pas possible au vu des résultats de l'acétylation de L1. La solution choisie est alors de marquer une lysine. Sa chaine latérale longue peut jouer le rôle d'espaceur, éloignant le fluorophore des autres résidus. En outre, l'amine terminale peut porter des groupements protecteurs permettant sa modification orthogonale ou être modifiée en azido afin de faire des réactions de Click. La modification de la lysine ne doit pas altérer fortement la liaison du peptide. D'après les résultats précédents de SAR, les positions K7 et K9 ne peuvent pas être modifiées. Ces positions sont proches de la principale région essentielle à la liaison du peptide. Une gêne stérique à cause de l'ajout d'un fluorophore est donc à craindre. Le marquage de la lysine modifiant la basicité de sa chaine latérale, K14 et K16 sont également exclues. Il reste alors K21. La perte d'affinité induite par une azido-lysine à cette position est faible (facteur 2). De plus, K21 est située dans une zone peu importante pour la liaison. Elle a donc été choisie pour être marquée.

- Méthode de marquage

La stratégie choisie pour marquer la lysine en position 21 fait intervenir une lysine dont la chaine latérale est protégée par un groupement ivDde. L'utilisation de ce groupement protecteur permet le marquage du peptide directement sur la résine (cf Matériel et Méthodes).

Le peptide HT1-0 a ainsi été marqué avec la fluorescéine (HT1-42). Le produit obtenu présente des épaulements en HPLC (**Figure 45**) et deux pics minoritaires mais uniquement la masse du peptide marqué. HT1-0 a aussi été marqué avec la rhodamine, un fluorophore rouge (HT1-43). Le produit final est propre avec un seul pic en HPLC (**Figure 46**) et présente bien la masse du peptide marqué. La masse du peptide non marqué n'est pas détectée.

- Influence sur l'affinité sur hMC4R

Avant de faire des cinétiques de liaison, l'affinité de HT1-42 non purifié est vérifiée en test de compétition avec le ^{125}I-NDP-α-MSH. L'affinité sur hMC4R est de 154 nM, soit une perte d'un facteur 10 d'affinité sur hMC4R. Il en est de même pour le peptide marqué à la rhodamine (facteur 15 de perte).

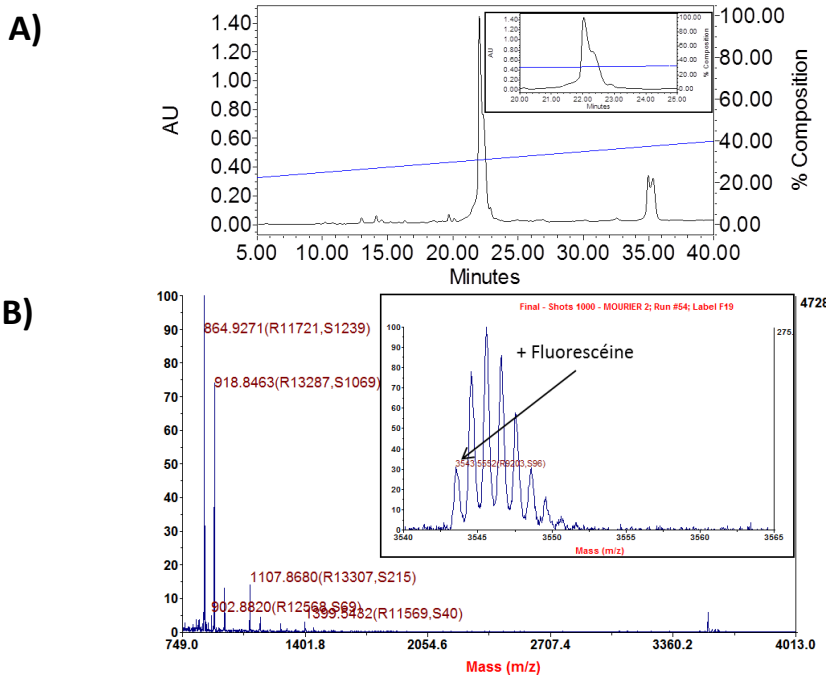


Figure 45. Analyse qualité de HT1-0 marqué à la fluorescéine (HT1-42) au niveau de K21. A) Profil HPLC. L'encadré noir est un zoom du pic majoritaire. Colonne : X Bridge BEH 300 5 µm 4.6 x 250 mm ; Gradient : 20% ACN puis 0.5% d'ACN/min ; débit : 1 mL/min. B) Spectre de masse MALDI-TOF du produit final non purifié. Les masses notées correspondent aux masses monoisotopiques.

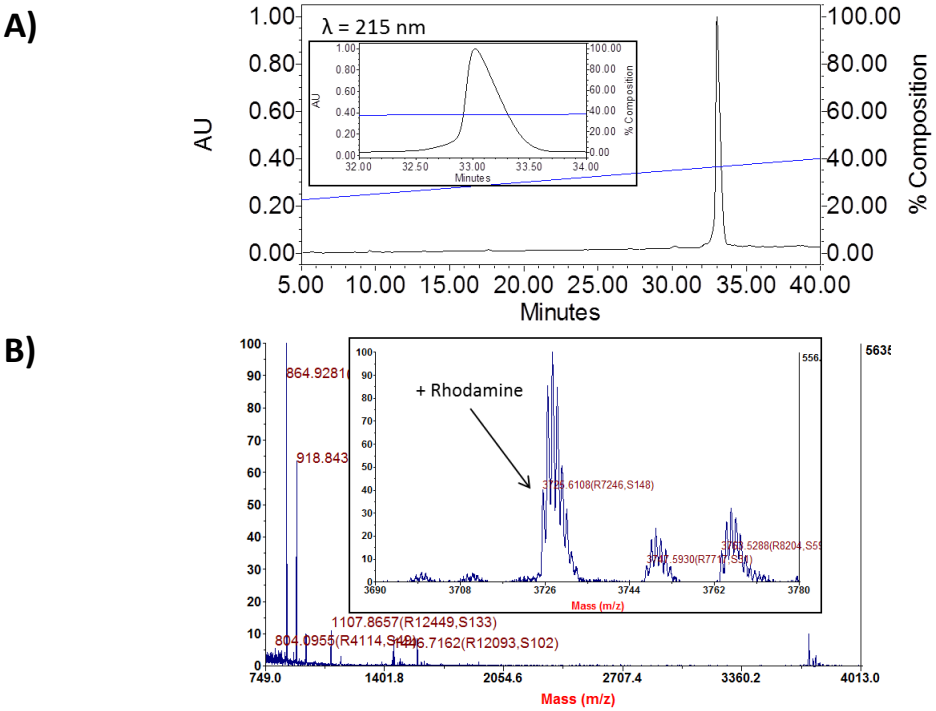


Figure 46. Analyse qualité de HT1-0 marqué à la rhodamine au niveau de K21 (HT1-43). A) Profil HPLC. L'encadré noir est un zoom du pic majoritaire. Colonne : X Bridge BEH 300 5 µm 4.6 x 250 mm ; Gradient : 20% ACN puis 0.5% d'ACN/min ; débit : 1 mL/min. B) Spectre de masse MALDI-TOF du produit final non purifié. Les masses notées correspondent aux masses monoisotopiques.

- Radioactif (^{125}I)
  - Choix de la position

Le marquage à l'iode se fait préférentiellement au niveau des tyrosines. L'iode est un élément d'encombrement stérique et d'hydrophobicité proches de ceux d'un groupement phényl. Le marquage d'un résidu près du site actif du peptide peut affecter l'affinité de celui-ci pour le récepteur. Afin d'estimer la conséquence d'une telle modification, certaines positions ont été remplacées par une diiodo-tyrosine. On se met ainsi dans le cas le plus défavorable stériquement. HT1-0 ne présente qu'une seule tyrosine en position 4. Néanmoins elle fait partie d'une zone importante dans la liaison du peptide sur hMC4R. Outre la position 4, d'autres ont donc été explorées en substituant chacune des phénylalanines par une une diiodo-tyrosine. Les positions 13 (HT1-25) et 17 (HT1-32) ont été abandonnées car leur modification induit un facteur de perte d'affinité supérieur ou égal à 9. Par contre, les substitutions en position 4 (HT1-15) ou 20 (HT1-37) altèrent peu l'affinité. Néanmoins, HT1-37 présente deux tyrosines qui peuvent être marquées radioactivement à l'iode 125. Or, la présence de deux iodes 125 sur la même molécule augmente les risques de dégradation de celle-ci par radiolyse. Sachant que l'activité biologique d'un seul atome d'iode 125 est suffisante (2 200 Ci/mmol), la solution est de remplacer Tyr4 par une phénylalanine (HT1-14). Cette modification n'entraine qu'une faible perte d'affinité (facteur 3). Par contre, la double modification Y4F F20Y (HT1-38) fait perdre un facteur 8 d'affinité. Au final, le marquage au niveau de la position 4 représente donc la meilleure solution.

  - Méthode de marquage radioactif

Le peptide radiomarqué est HT1-49 qui diffère de HT1-0 uniquement par l'amidation de son extrémité C-terminale. Cette différence ne modifie ni le temps de rétention ni l'affinité du peptide pour hMC4R.

Le marquage du peptide a été effectué en présence de la forme réduite d'iode 125 ($Na^{125}I$, 0.4 mCi) et d'un oxydant doux, l'iodogène. En HPLC, trois pics sont observés (**Figure 47A**). Les pics T1 et T2 présentent une radioactivité faible de l'ordre du bruit de fond de l'appareil de mesure. Pour T1, aucune masse n'est détectée en spectrométrie de masse. Par contre, le pic T2 correspond au peptide non marqué (**Figure 47B**). Enfin, le pic minoritaire T3 correspond au peptide monoiodé et présente une radioactivité forte. Il y a donc eu incorporation d'un seul atome d'iode 125.

- **Paramètres cinétiques kon et koff**
  - Avec le peptide marqué à la fluorescéine

La fluorescéine est un fluorophore vert compatible avec des études de transfert d'énergie par résonnance de Förster (FRET). Cette méthode est basée sur le transfert d'énergie d'un donneur préalablement excité, le cryptate de Terbium, à un accepteur, la fluorescéine. Ce transfert ne peut se faire que si les deux protagonistes sont suffisamment proches (1-10 nm) c'est-à-dire lorsque le peptide se lie au récepteur.

La forte perte d'affinité liée au marquage du peptide rend difficile les études de liaison en FRET. Classiquement, l'affinité d'un peptide marqué doit être de l'ordre du nanomolaire. Des études de FRET ont quand même été réalisées avec des concentrations de 10 nM, 100 nM et 1 µM sans succès. Aucun signal de FRET n'a été observé après 2 h d'incubation du peptide HT1-0 marqué ou de l'agoniste marqué fourni par Cisbio avec les cellules exprimant le récepteur hMC4R.

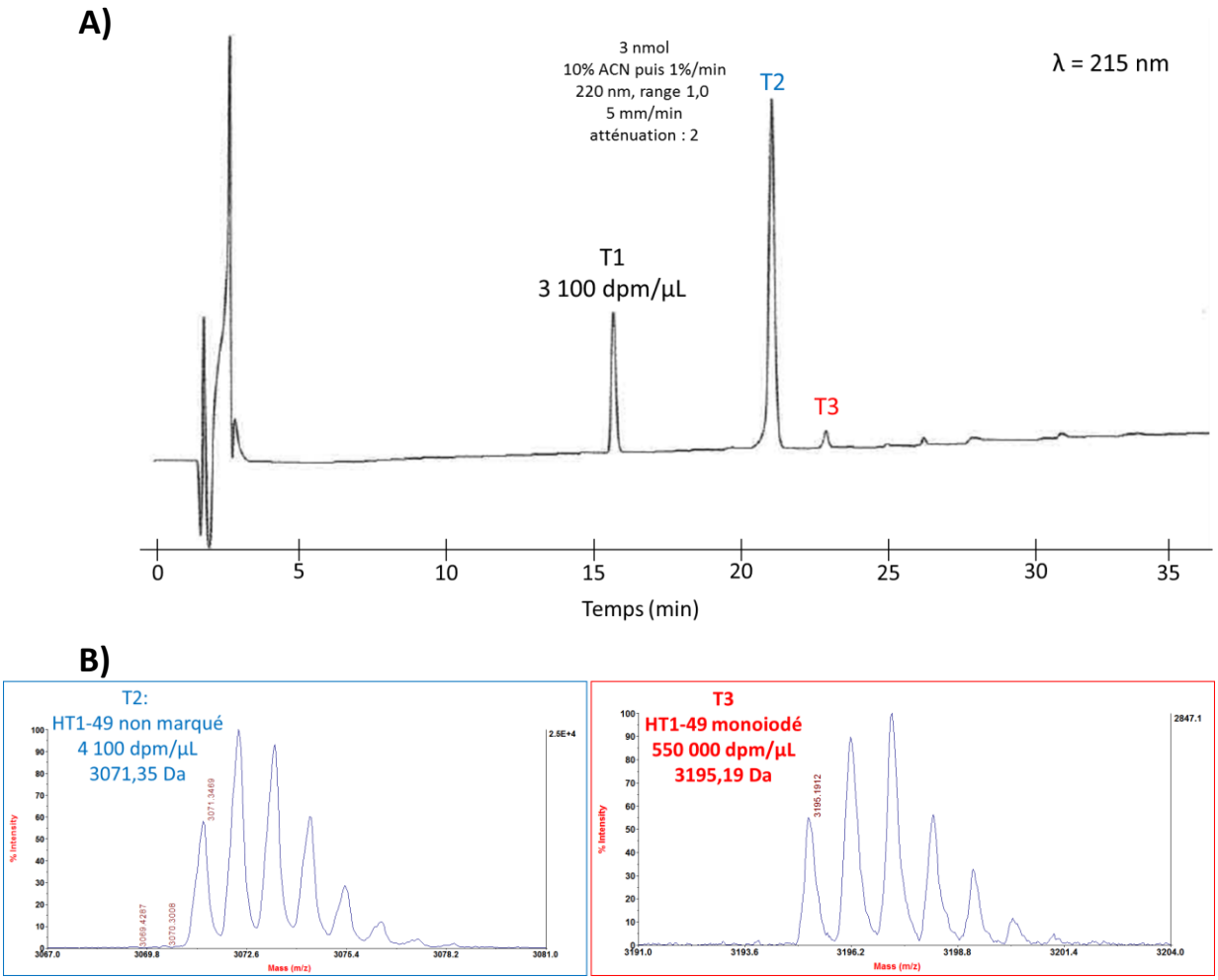


Figure 47. Marquage radioactif à l'iode 125 de HT1-49. A) Profil HPLC du mélange réactionnel. Colonne XBridge® Peptide BEH C18 300Å 3.5 µm, 4.6 x 150 mm; Gradient : 10% ACN puis 1% d'ACN/min ; débit : 1 mL/min. B) Spectre de masse MALDI-TOF des fractions T2 et T3 purifiées. La radioactivité et la masse monoisotopique détectées sont indiquées.

- Avec le peptide radiomarqué à l'iode 125
  - Par scintillation liquide

La première méthode employée est la scintillation liquide c'est-à-dire basée sur le même principe que celle utilisé dans les tests de compétitions. La liaison totale du peptide iodé est à un niveau similaire à celle de la liaison sur le filtre (BF) (**Figure 48**). Elle est maximale dès les premières minutes et n'évolue pas par la suite. Elle est cependant déplaçable par un excès de HT1-49 froid mais pas par NDP-α-MSH, un agoniste du MC4R. Ces résultats ont été confirmés deux fois de manière indépendante. Tout ceci est en faveur d'une liaison non spécifique de ^{125}I-HT1-49 sur le filtre et non sur le récepteur hMC4R. Une autre méthode sans filtration a donc été testée. Il s'agit du test de scintillation de proximité (SPA).

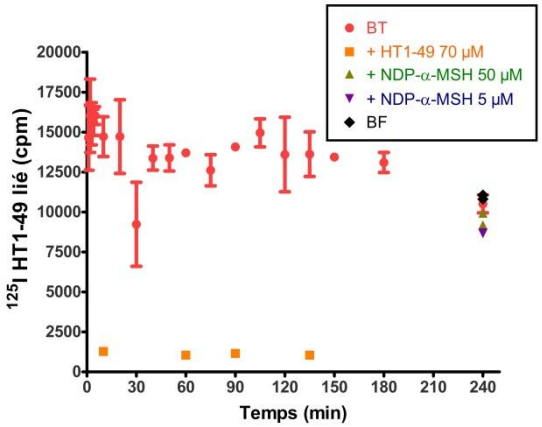


Figure 48. Cinétique d'association de ^{125}I-HT1-49 en scintillation liquide. Les membranes utilisées proviennent de cellules CHO exprimant de manière stable le récepteur hMC4R. BT : liaison totale ; BF : liaison sur le filtre en l'absence de membranes hMC4R. Moyenne ± SD (n=2).

- Par scintillation de proximité (SPA)

La scintillation de proximité est basée sur le principe suivant : des billes de scintillant convertissent en lumière l'énergie émise par les particules radioactives β^-. Cette conversion n'est possible que si la molécule radioactive émettrice est suffisamment proche des billes. Certaines billes peuvent lier des membranes contenant des récepteurs. Le transfert d'énergie peut alors se faire si la molécule radioactive reconnait son récepteur. Afin de vérifier que la conversion d'énergie est bien dûe à la liaison spécifique de ^{125}I-HT1-49 sur hMC4R, les mesures cinétiques de liaison ont été faites avec des billes présentant ou pas des membranes exprimant hMC4R (**Figure 49**). Les résultats sont identiques dans les deux cas. Le ligand radioactif se lie sur les billes et non sur le récepteur hMC4R. L'association est progressive et la liaison est déplaçable avec un excès de HT1-49 froid. Ces résultats ont été confirmés deux fois.

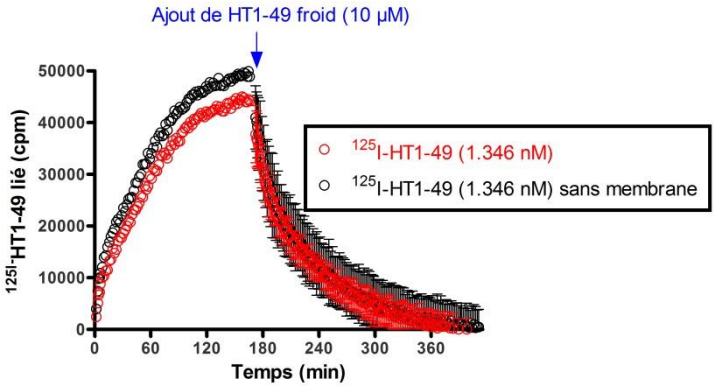


Figure 49. Exemple de cinétique d'association et de dissociation de ^{125}I-HT1-49 en scintillation de proximité (SPA). Le test a été effectué en présence ou non du récepteur hMC4R. Après trois heures, un excès de HT1-49 froid (10 µM) est ajouté. Les membranes utilisées proviennent de cellules CHO exprimant de manière stable le récepteur hMC4R. Moyenne ± SD (n=2).

- **Visualisation de la liaison de HT1-0 marqué à la rhodamine sur le récepteur hMC4R-GFP**

La fluorescence maximale de la rhodamine s'effectue à une longueur d'onde de 572 nm tandis que celle de la GFP se situe à 504 nm. Ce marquage pemet donc de visualiser au microscope à fluorescence la liaison du peptide sur le récepteur hMC4R-GFP exprimé à la surface de cellules. La perte d'affinité sur hMC4R est cependant importante (facteur 15). Pour les tests cellulaires, une forte concentration en peptide marqué (5 µM) a donc été utilisée. Les cellules utilisées expriment de manière transitoire le récepteur hMC4R marqué en C-terminal par la GFP (hMC4R-GFP). En présence du peptide marqué seul, la majorité des

cellules exprimant à la fois hMC4R-GFP (vert) et marquées par HT1-42 (rouge) sont rondes (**Figure 50**). Il pourrait s'agir de cellules en mort cellulaire dont la membrane plasmique est plus perméable. Les résultats sont similaires en présence d'un compétiteur froid (HT1-0 non marqué ou NDP-α-MSH). La concentration en peptide marqué devait être trop importante.

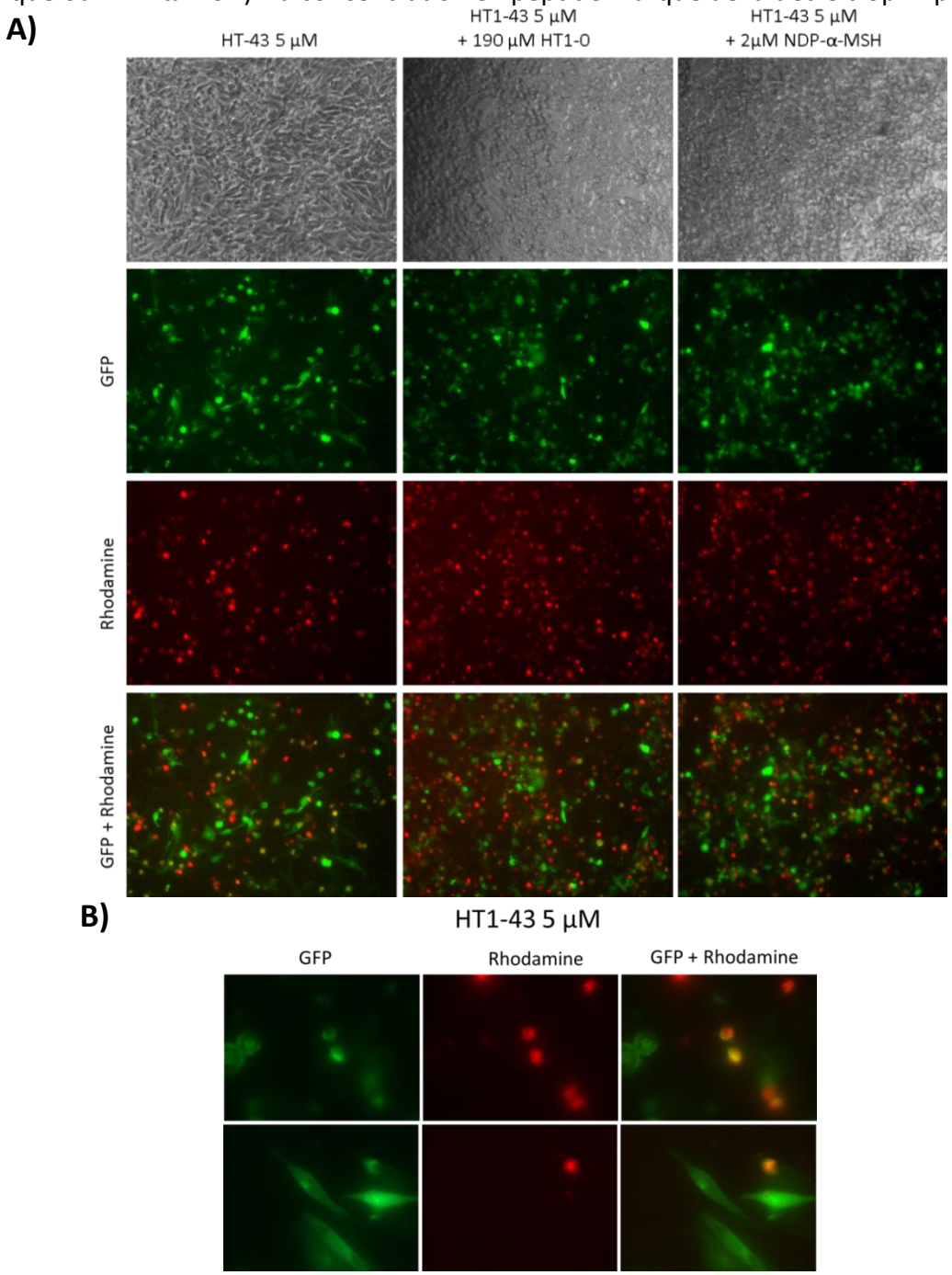


Figure 50. Visualisation de la liaison de HT1-0 marqué à la rhodamine sur le récepteur hMC4R-GFP. Les cellules étudiées sont des CHO exprimant de façon transitoire le récepteur hMC4R-GFP. A) Observation après 3h d'incubation avec le peptide marqué en présence ou en absence d'un compétiteur froid (HT1-0 non marqué ou NDP-α-MSH). B) Visualisation à plus fort grossissement des cellules incubées avec le peptide marqué seul.

Les résultats précédents ne permettent pas de prouver de façon claire la liaison du peptide HT1-0 sur le récepteur hMC4R. Une autre approche consiste à étudier sa fonction

agoniste ou antagoniste sur l'activité du récepteur. Si une activité est détectée, cela signifiera que le peptide se lie bien au récepteur. En outre, d'un point de vue thérapeutique, les principaux effets indésirables proviennent de l'activité même du récepteur MC4R et notamment de la voie de signalisation G_s. L'étude de la fonction *in vitro* de HT1-0 déterminera non seulement son application thérapeutique potentielle (obésité ou cachexie) mais sera également une indication précieuse des potentiels effets indésirables *in vivo*.

4.4 Fonction[9]

L'activité agoniste de HT1-0 a été étudiée sur différentes voies de signalisation de hMC4R *in vitro*. Ces tests sont basés sur le transfert d'énergie par résonance de bioluminescence (BRET). Pour les voies impliquant des protéines G, le biosenseur GRK2 a été utilisé (demande de brevet : US20170234870A1). Lorsque la Gα d'intérêt surexprimée dans les cellules couple avec le récepteur hMC4R, l'activation de celui-ci promeut la dissociation des complexes Gαβγ et éventuellement l'interaction entre la sous-unité Gγ (marquée avec le donneur RlucII) avec l'effecteur GRK2 (marqué avec l'accepteur GFP10), ce qui provoque une augmentation du signal de BRET. Pour la voie β-arrestine 2, le biosenseur β-arrestine-RlucII a été employé (Namkung et al., 2016). Si l'activation du récepteur promeut le recrutement de la β-arrestine d'intérêt, le tag RlucII marquant la β-arrestine peut interagir avec l'accepteur ancré à la membrane par le domaine CAAX de k-ras, provoquant de ce fait une augmentation du BRET. À 1 µM, HT1-0 n'active aucune des 15 voies de signalisation étudiées (**Figure 51A**). À 2 µM, α-MSH active les voies G_s, G_{15} (famille de G_q) et β-arrestine 2 (**Figure 51B**). L'étude de l'activité antagoniste de HT1-0 a donc été explorée sur ces trois voies. Pour ces études, le biosenseur GAPL a été préféré car il permet une meilleure fenêtre de lecture et une réponse faible pour l'activité endogène. Pour la voie G_s, la protéine G_s marquée avec RlucII quitte le senseur de localisation membranaire rGFP-CAAX lorsqu'elle est activée. L'activation du récepteur couplé à G_s entraine donc une baisse du signal de BRET. Pour les voies G_{15} et β-arrestine 2, l'activation du récepteur promeut le recrutement d'un effecteur spécifique marqué avec RlucII au senseur de localisation membranaire rGFP-CAAX. Dans le cas de la voie G_{15}, l'effecteur est le facteur d'échange nucléotidique p63. Pour la voie β-arrestine 2, il s'agit de la β-arrestine. Dans les deux cas, l'activation du récepteur conduit ainsi à une hausse du signal de BRET.

Pour les trois voies étudiées, HT1-0 inhibe l'activation du récepteur par son agoniste α-MSH ou son dérivé la NDP-α-MSH (**Figure 52**). La régression de Shild des courbes d'activation des trois voies de signalisation montrent une pente proche de 1, suggérant un antagonisme compétitif de HT1-0 vis-à-vis des agonistes α-MSH et NDP-α-MSH (**Table 17**). Quel que soit l'agoniste testé, HT1-0 montre une efficacité micromolaire : 16-27 µM pour la voie G_s, 8-9 µM pour la voie G_{15} et 9-14 µM pour la voie β-arrestine 2.

[9] Les études de fonction en BRET ont été réalisées par Suli-Anne Laurin, doctorante dans l'équipe du Professeur Michel Bouvier à l'Institut de recherche en immunologie et en cancérologie (IRIC), Université de Montréal, 2950 chemin de Polytechnique, Montréal (Québec) H3C 3J7, Canada. Les études en FRET ont été conduites par la société Eurofins Cerep, Le Bois L'Êveque, 86600 Celle-Lévescault France.

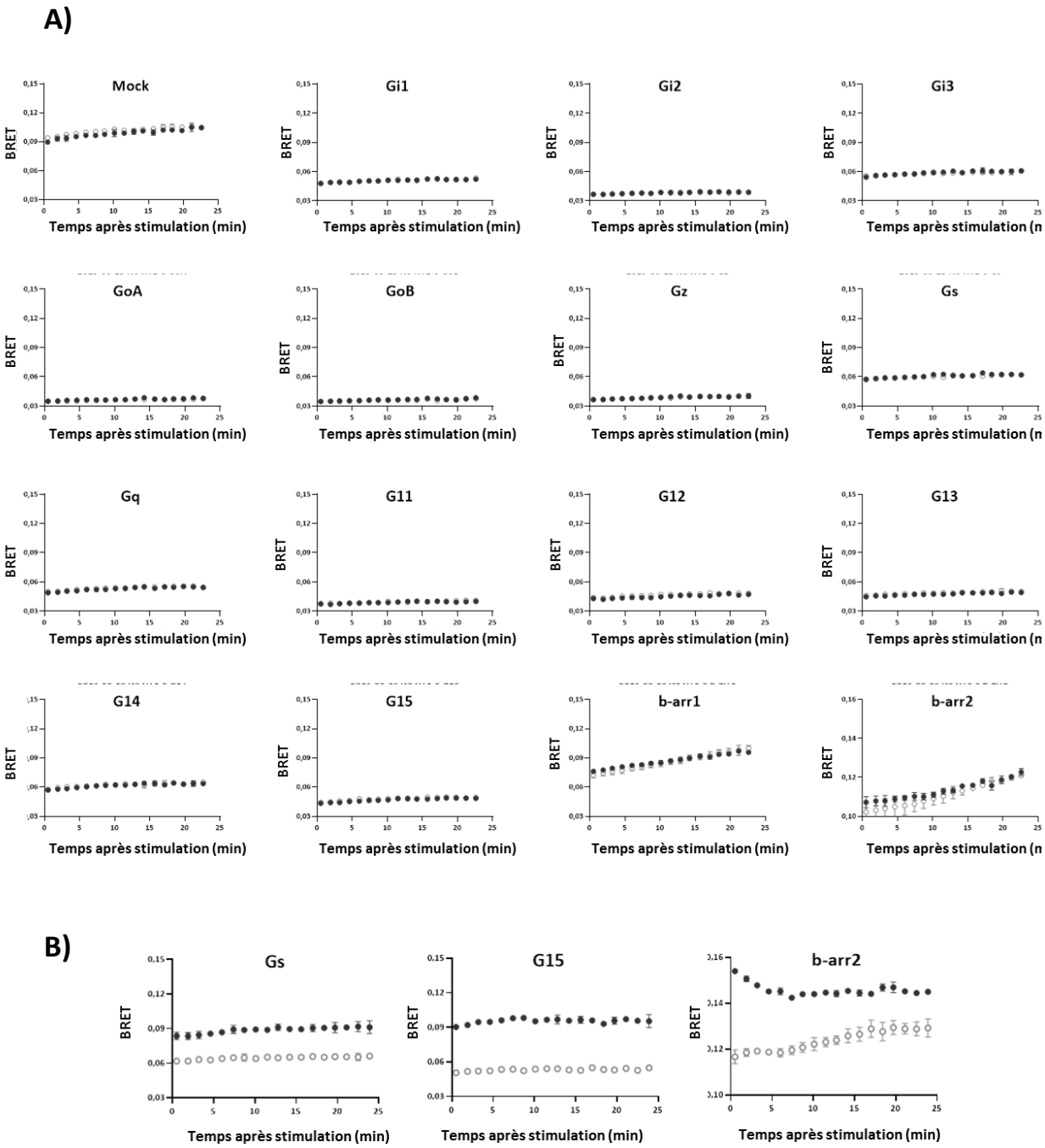


Figure 51. Cinétique d'activation de différentes voies de signalisation du hMC4R en présence (•) ou en absence (○) de HT1-0 (1 µM) (A) ou d'α-MSH (2 µM) (B). Les cellules utilisées sont des cellules HEK293 surexprimant la protéine Gα considérée ou la β-arrestine 2. Les résultats sont présentés en signaux de transfert d'énergie par résonnance de bioluminescence (BRET). Une activation de la voie considérée de hMC4R entraine une augmentation des signaux de BRET. Le Mock est sans surexpression de Gα et représente donc la contribution des Gα endogènes. Moyenne ± SD (n=2).

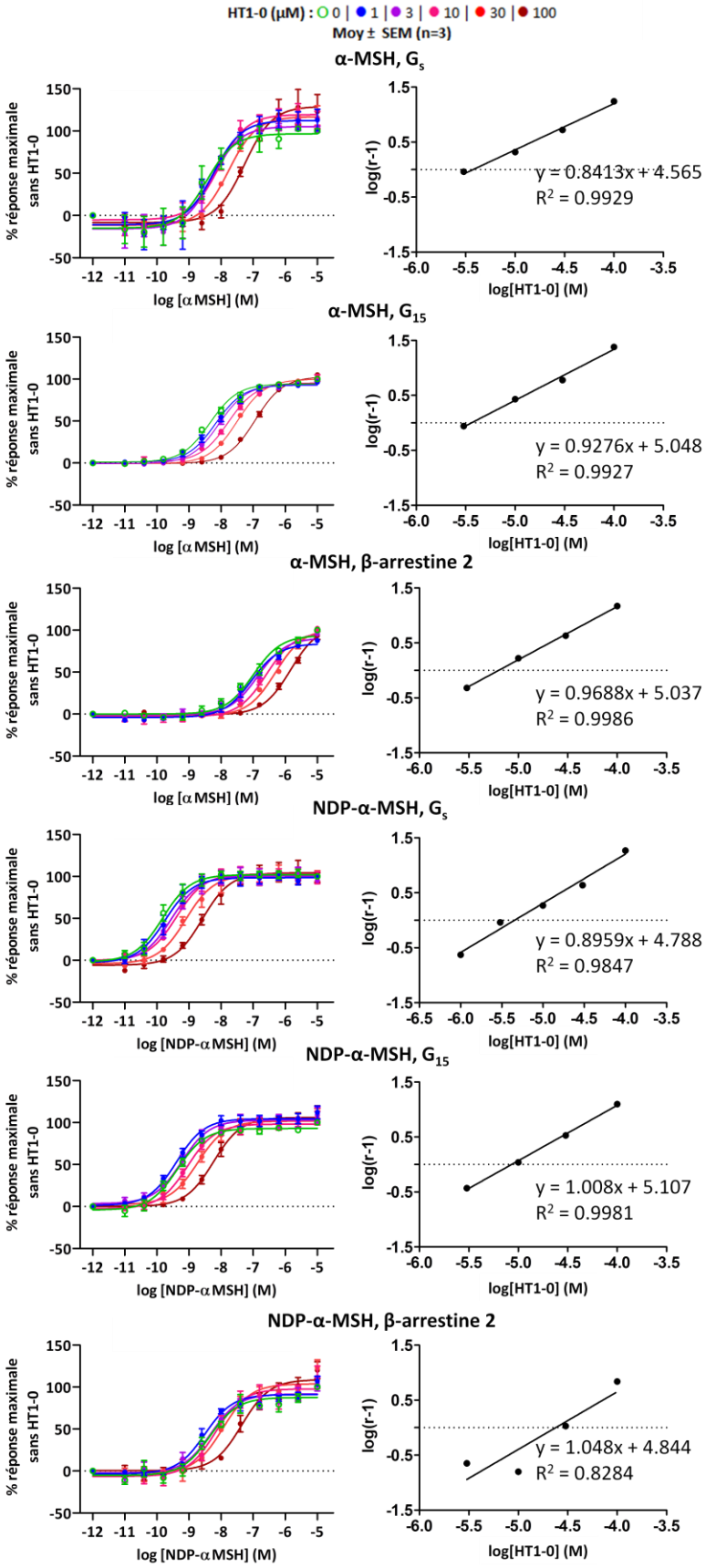


Figure 52. Étude de la fonction antagoniste de HT1-0 sur les voies G_s, G_{15} et β-arrestine 2 de hMC4R activées par α-MSH ou NDP-α-MSH. Les cellules utilisées sont des cellules HEK293 surexprimant la protéine Gα considérée ou la β-arrestine 2 et le récepteur hMC4R de manière transitoire. À gauche : Courbes d'activation de hMC4R par α-MSH ou NDP-α-MSH en présence de concentrations croissantes de HT1-0. Moyenne ± SEM (n=3). À droite : régression de Shild. L'équation de la courbe est indiquée.

Table 17. Statistiques issues de la régression de Shild des courbes d'activation de différentes voies de signalisation de hMC4R par α-MSH et NDP-α-MSH en présence de concentrations croissantes de HT1-0.

	α-MSH			NDP-α-MSH		
	Pente	K_B (µM)	R^2	Pente	K_B (µM)	R^2
Gs	0,8413	27,23	0,9929	0,8959	16,29	0,9847
G15	0,9276	8,95	0,9927	1,008	7,816	0,9981
β-arrestine 2	0,9688	9,18	0,9986	1,048	14,32	0,8284

En parallèle des études précédentes de BRET, l'activité agoniste et antagoniste du peptide sur la voie G_s a également été étudiée par une méthode de dosage direct du second messager, l'AMPc. Il s'agit d'un test de compétition basé sur le transfert d'énergie par résonnance de Förster (FRET) entre le fluorophore marquant un anticorps anti-AMPc (donneur) et D2 marquant l'AMPc (accepteur). Si la voie G_s est activée, les cellules vont produire de l'AMPc. Après lyse des cellules, l'accepteur et le donneur sont ajoutés. L'AMPc non marqué produit par les cellules et l'AMPc marqué entre alors en compétition pour la reconnaissance avec l'anticorps. Par conséquent, le signal de FRET mesuré est inversement proportionnel à la quantité d'AMPc produit par les cellules. À 10 µM, HT1-0 ne stimule pas la production d'AMPc dans des cellules CHO exprimant hMC4R. Par contre une activité antagoniste est bien observée pour HT1- 0 avec une amidation en C-terminal (HT1-49) avec un K_B de 2.3 µM (**Figure 53**). Sachant que l'amidation en Cter ne modifie pas significativement l'affinité du peptide sur hMC4R, on peut estimer que HT1-49 et HT1-0 présentent une fonction similaire.

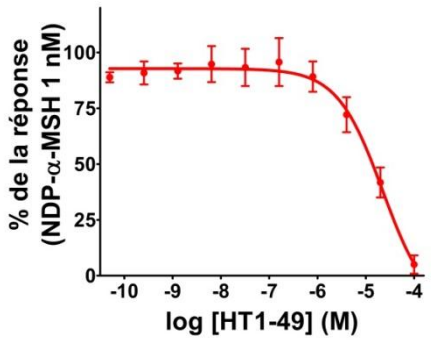


Figure 53. Inhibition de la production d'AMPc (voie G_s) de hMC4R stimulée par NDP-α-MSH en présence de concentrations croissantes de HT1-49. HT1-49 correspond à HT1-0 avec une extrémité C-terminale amidée. Les cellules utilisées sont des cellules CHO surexprimant le récepteur hMC4R de façon transitoire. Moyenne ± SD (n=2).

4.5 Discussion

- **D'un peptide réticulé à un peptide linéaire, conséquences structurales**

Les toxines animales sont riches en ponts disulfure. C'est cette richesse qui leur confère une stabilité structurale. Dans le cas des toxines à deux ponts disulfure, trois formes oxydées nommées « Bead », « Globular» et « Ribbon », sont possibles. Ces différents arrangements modifient la structure et généralement l'activité de la toxine sur sa cible moléculaire (Wu et al., 2014). Or, ce n'est pas le cas pour la conotoxine N-CTX-Ltg1a sur les récepteurs MCRs. Les trois formes oxydées présentent des affinités similaires sur l'ensemble des sous-types. Ce n'est pas très surprenant au vu de la séquence très particulière de la

toxine dans le sens où les trois premières cystéines sont situées à une glutamine les unes des autres. Les différents appariements de cystéines n'ont ainsi que peu d'impact sur la structure et donc sur les interactions entre la toxine et le récepteur. Plus intéressant encore, l'alkylation des quatre cystéines empêchant la formation des ponts ne supprime pas l'activité mais la diminue essentiellement sur le sous-type hMC5R. Les contraintes structurales exercées par les ponts disulfure modifieraient donc peu les interactions toxine-récepteurs MCRs. Les parties les moins sujettes à ces contraintes sont les extrémités N et C-terminales. Les études de troncation des extrémités ont ainsi révélé un rôle majeur de la partie N-terminale de la toxine dans la liaison aux récepteurs MCRs. En outre, les spectre en dichroïsme circulaire de la forme « Ribbon » et de la forme alkylée sont très similaires à ceux de protéines dénaturées (Venyaminov et al., 1993). L'analyse avec le logiciel Spectra Manager™ II (JASCO) confirme ces résultats en utilisant la référence de Reed (Reed and Reed, 1997). Par contre, le logiciel BeStSel prédit 7% d'hélice et 29% de feuillets β. La différence de prédiction provient sans doute de la base de données utilisée. Contrairement à la référence de Reed, la base utilisée par BeStSel ne contient aucun peptide de moins de 40 résidus (Lees et al., 2006) créant ainsi un biais dans l'analyse des résultats. Cette absence de structures secondaires rend ainsi difficile l'interprétation des conséquences structurales des modifications de la toxine aboutissant au peptide linéaire HT1-0. Néanmoins, les études de SAR sur ce dernier donnent plusieurs éclairages quant aux zones d'interaction du peptide avec hMC4R avec trois zones aux propriétés physico-chimiques particulières.

- **Définition des zones d'interaction de HT1-0**

Comme pour la toxine à deux ponts de départ, les résidus hydrophobes en N-terminal de HT1-0 sont cruciaux pour la liaison sur le récepteur. Néanmoins, les facteurs de perte sont plus importants pour le peptide linéaire. La troncature en N-terminal de la même zone entraine une perte d'un facteur 93 vs 28 pour la toxine. Celle en C-terminal modifie significativement l'affinité avec un facteur de perte de 9. Cela suggère donc que la substitution des cystéines par Q ou K renforcerait l'importance des extrémités. De nouvelles interactions pourraient alors se faire comme avec K21 ou R23 en C-terminal, expliquant le nouveau rôle de cette extrémité. D'après les résultats avec les 6-Azido-L-Lysine, l'interaction serait plus d'ordre hydrophobique qu'électrostatique au contraire de la troisième région d'interaction majeure avec K14 et K16. La présence d'une charge positive à ces positions est importante pour la liaison du peptide sur hMC4R. Au final, on peut dresser la carte du peptide suivante avec en couleur les principales propriétés physico-chimiques des différentes positions importantes pour la liaison de HT1-0 sur hMC4R :

LFHYQQKQKPPGFKGKFQQFKLRPP

avec en violet, l'hydrophobicité ; en orange, la polarité ; en rouge, la basicité.

Sachant qu'aucun mutant du récepteur hMC4R n'a été testé, il est difficile, d'après les seules informations précédentes de spéculer sur le site d'interaction du peptide sur le récepteur. Néanmoins, HT1-0 serait un antagoniste compétitif des agonistes α-MSH et NDP-α-MSH. De ce fait, HT1-0 partagerait un site d'interaction similaire à celui de ces deux ligands. En 2010, Chapman et al., via des expériences de ligation croisée et de mutagenèse (Chapman et al., 2010), confirment les résultats antérieurs de mutagenèse et de modélisation (Hogan et al., 2006; Pogozheva et al., 2005). Ce site fait intervenir le tétrapeptide HFRW présent dans les deux agonistes. H6 n'interagirait pas directement avec les résidus du TM3 mais stabiliserait la structure active de l'agoniste. F7 formerait une interaction hydrophobe avec $F^{6.52}$(262) du TM6 du récepteur. R8 interagirait avec au moins

un des résidus acides de hMC4R suivants : $D^{3.25}$(122), $D^{3.29}$(126) du TM3 et $E^{2.60}$(100) du TM2. Et W9 se logerait dans une poche hydrophobe semblant importante pour l'activation du récepteur, via notamment $I^{7.42}$(291) ou $M^{7.43}$(292) du TM7. Les interactions font ainsi intervenir des résidus acides et des résidus hydrophobes de hMC4R pouvant potentiellement interagir avec les résidus importants de HT1-0. Néanmoins, d'autres études montrent que le récepteur hMC4R peut être présent *in vitro* sous forme d'oligomères (dimères majoritairement), créant un deuxième site de fixation de l'agoniste NDP-α-MSH (Chapman and Findlay, 2013; Kopanchuk et al., 2006). Selon les auteurs, les deux sites sont dépendants l'un de l'autre. Les résultats des études cinétiques de l'agoniste NDP-α-MSH et de HT1-0 pourraient apporter une explication quant à la différence entre l'affinité nanomolaire de HT1-0 et sa faible efficacité à antagoniser les principales voies de signalisation de hMC4R.

- **^{125}I-NDP-α-MSH, un ligand pas adapté aux tests de liaison**

Il n'est pas rare d'observer une différence entre affinité obtenue en test de liaison et efficacité en test de fonction. En effet, il n'existe pas de relation linéaire entre la liaison du peptide sur le récepteur et la transduction du signal. Néanmoins, la différence observée ici entre une affinité nanomolaire et une efficacité micromolaire pourrait également s'expliquer par le mode de liaison de ^{125}I-NDP-α-MSH sur hMC4R. Selon les études cinétiques faites par Kopanchuk et al. en 2006, ^{125}I-NDP-α-MSH ne peut se dissocier que d'un seul site, le deuxième ayant une cinétique de dissociation très lente (k_{off} = 0.36 h^{-1}) (Kopanchuk et al., 2006). J'ai réalisé également des expériences de dissociation par dilution isotopique selon deux méthodes, en scintillation liquide et en scintillation de proximité. Les constantes cinétiques de dissociation (k_{off}) obtenues sont extrêmement faibles (**Figure 54**). Contrairement à ce qu'ont conclu Kopanchuk et al. en scintillation liquide, le ligand radioactif semble réversible à 92%. La différence d'interprétation vient sans doute de la durée de l'expérience. Dans l'article, elle est de 19h avec une réversibilité d'environ 30% (**Figure 54C**), ce qui correspond aux valeurs observées au bout de la même période dans notre expérience (**Figure 54A**). Dans cette dernière, une cinétique d'association en parallèle démontre la stabilité du complexe ligand radioactif-récepteur pendant les 52h de mesure. La dissociation observée est donc uniquement dû à la compétition avec le ligand froid en excès pour la liaison sur le récepteur. L'expérience a cependant été faite une seule fois avec des points très espacés dans le temps rendant la mesure du k_{off} peu précise. En scintillation de proximité (SPA), les mesures peuvent être effectuées régulièrement de façon automatisée. D'après les résultats de cette méthode, le ligand radioactif serait au contraire peu réversible avec seulement 10% de réversibilité. Néanmoins, la simple manipulation pour sortir et remettre la plaque dans le lecteur entraine une augmentation du signal mesuré faussant ainsi l'interprétation de la dissociation. Les paramètres cinétiques obtenus dans notre expérience de scintillation liquide montrent de toute façon que l'équilibre serait seulement atteint au bout de 200 d'heures ($5*t_{1/2}$). Au bout de 3h d'incubation en présence d'un compétiteur comme HT1-0, la mesure de l'affinité est donc faussée et dépend alors principalement de la différence de cinétique d'association (k_{on}) du compétiteur et du ligand radioactif. Si le compétiteur a une cinétique d'association plus rapide, plus le temps d'incubation sera long, plus il y aura de chance que ^{125}I-NDP-α-MSH se fixe au récepteur à la place du compétiteur. La dissociation du ligand radioactif étant lente, le compétiteur ne pourra plus se fixer au récepteur diminuant ainsi son affinité apparente au court du temps (**Figure 54D**).

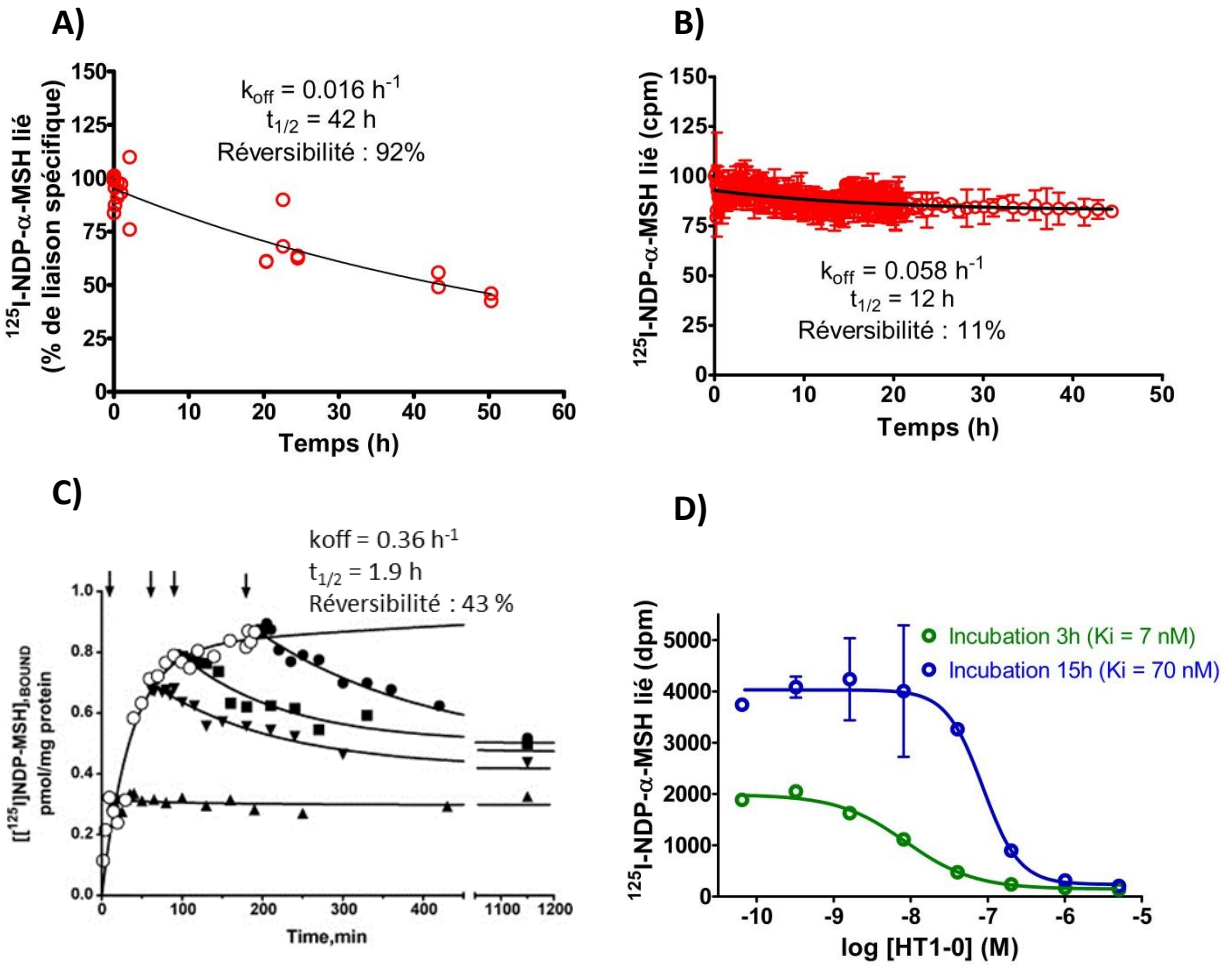


Figure 54. Dissociation lente de ^{125}I-NDP-α-MSH du complexe avec hMC4R et conséquence sur l'affinité de HT1-0 en compétition. Cinétique de dissociation de ^{125}I-NDP-α-MSH sur hMC4R par dilution isotopique en scintillation liquide (A) et en scintillation de proximité (B). Le complexe est formé en incubant pendant 3h le ligand radioactif avec des membranes de cellules CHO exprimant de manière transitoire le récepteur. La dissociation est ensuite initiée par ajout de NDP-α-MSH en large excès (1 µM). C) Figure extraite de Kopanchuk et al., 2006. Cinétique de dissociation en scintillation liquide de ^{125}I-NDP-α-MSH sur hMC4R exprimé à la surface de membranes de cellules d'insecte Sf9. La dissociation a été initiée à différents temps (flèches) avec 3 µM final d'I-NDP-α-MSH. Les valeurs de constante cinétique de dissociation k_{off}, de temps de demi-vie $t_{1/2}$ et de réversibilité sont indiquées sur chacun des graphes avec $t_{1/2} = \frac{ln2}{koff}$. D) Effet en scintillation liquide de la durée d'incubation en présence de différentes concentrations de HT1-0 sur la liaison de ^{125}I-NDP-α-MSH sur hMC4R. Les valeurs de constantes d'inhibition (Ki) sont indiquées pour chacune des conditions. Moyenne ± SD (n=2).

Une façon de s'affranchir de ce problème est d'obtenir l'affinité de HT1-0 en étudiant ses paramètres cinétiques d'association et de dissociation. Cela requiert le marquage du peptide puis l'étude de la liaison directe du peptide marqué sur le récepteur hMC4R. Malheureusement, ces paramètres n'ont pas pu être obtenus pour différentes raisons.

- **Une étude de la liaison directe de HT1-0 sur hMC4R compromise**

  Deux stratégies de marquage ont été explorées. La première visait à greffer un

fluorophore au peptide au niveau d'une lysine de la séquence, soit la fluorescéine pour des études cinétiques, soit la rhodamine pour visualiser la liaison du peptide à la surface de cellules. Voulant garder la séquence originale, le choix s'est porté sur la position 21 où l'hydrophobicité et l'encombrement stérique du fluorophore gênerait, en théorie, le moins la liaison. Ce choix s'est avéré mauvais puisque le marquage à la fluorescéine induit une perte d'affinité supérieure à un facteur 10 rendant difficile les expériences de cinétique d'association et de dissociation du peptide marqué. En effet, plus l'affinité est faible, plus la concentration en ligand marqué nécessaire sera grande. Or, une concentration trop élevée augmente le bruit de fond et peut masquer un signal de FRET. En outre, les expériences ont été faites sans contrôle positif. L'agoniste marqué fourni par Cisbio a subit une décongélation accidentelle de plusieurs jours à cause d'une panne de congélateur. Le ligand a pu se dégrader d'où une absence de signal. Il est alors difficile d'interpréter avec rigueur l'expérience et de conclure quant à l'absence ou non de liaison de HT1-0 marqué à la fluorescéine sur hMC4R. Concernant les tests en microscopie à fluorescence avec le peptide marqué à la rhodamine (rouge), ils ont été réalisés avec des cellules CHO exprimant le récepteur hMC4R marqué à la GFP (vert). Si le peptide marqué se fixe bien sur le récepteur, on devrait observer une colocalisation de la fluorescence rouge et verte. Or ce n'est le cas que pour des cellules rondes probablement en train de mourir. L'hypothèse la plus probable serait que le peptide marqué à la rhodamine ne reconnait pas le récepteur hMC4R-GFP mais se lie de manière non spécifique à des éléments intracellulaires. Le peptide ne franchissant pas la membrane de cellules vivantes, l'accès à ces éléments n'est possible que si la membrane est plus perméable comme c'est le cas lorsque les cellules sont en train de mourrir.

La seconde stratégie concernait le marquage radioactif à l'iode 125. Alors que les premières études d'Alascan montraient une importance de Y4, l'ajout de deux atomes d'iode en ortho de l'hydroxyle n'a pas induit de modification de l'affinité du peptide. Même si le peptide froid monoiodé n'a pas été testé, il est peu probable que l'ajout d'un seul atome d'iode soit plus délétère que la présence de deux iodes. Partant de ce postulat, nous avons réussi à marquer radioactivement avec un atome d'iode 125 le peptide HT1-0 amidé (HT1-49) sans modifier sa séquence. Néanmoins, la totalité de ^{125}I-HT1-49 lié se fixe de façon non spécifique au filtre en scintillation liquide et aux billes de scintillation de proximité. La qualité du peptide marqué a été vérifiée en spectrométrie de masse avant les expériences. Il ne s'agit donc pas d'une dégradation du peptide. Cette liaison non spécifique est uniquement diminuée en présence d'un excès de HT1-49 froid signifiant que HT1-49 se lie aussi fortement au filtre. Il sera donc impossible d'étudier ce peptide en scintillation liquide. HT1-49 a un point isoélectrique (pI) élevé (11.25) avec cinq résidus basiques et pourrait ainsi former des interactions électrostatiques avec les fibres de verre du filtre et le plastique des billes de scintillation. D'ailleurs une étude a démontré la forte adsorption de peptides cationiques à la surface du verre et du plastique (Kristensen et al., 2015). Les peptides étudiés sont similaires à HT1-0 en longueur et en pI. En scintillation liquide, ces interactions sur le filtre sont normalement minimisées par le pré-traitement du filtre au PEI 0.5% ou avec une solution de BSA 0.5%. Mais aucun de ces traitements ne diminue la liaison de ^{125}I-HT1-49 au filtre. Ayant lieu après l'incubation avec ^{125}I-NDP-α-MSH en test de compétition, cette fixation n'a donc aucune conséquence sur les valeurs de Ki obtenues. Cependant, on peut également imaginer une fixation sur le plastique dans les puits des plaques utilisées en tests de liaison et de fonction. La concentration en peptide libre rentrant véritablement en compétition avec NDP-α-MSH sur hMC4R serait alors surestimée et donc l'affinité calculée

sous-estimée. Elle est impossible à déterminer en scintillation liquide comme le peptide se fixe fortement au filtre. Une telle liaison rend ainsi extrêmement compliquées les études impliquant du matériel en plastique.

Les possibles biais dans l'estimation de l'affinité ne remettent cependant pas en cause les conclusions de la SAR quant aux positions de HT1-0 pouvant être modifiées. En effet, toutes les expériences ont été faites dans les mêmes conditions. Cela peut permettre notamment de proposer des modifications pour optimiser l'affinité et la sélectivité du peptide pour hMC4R.

- **Vers une optimisation de la séquence**

Avec un pI de 11.25, HT1-0 est très cationique. Certains résidus basiques peu importants pour la liaison sur hMC4R pourraient être remplacés afin de limiter d'une part, les interactions électrostatiques non spécifiques et d'autre part, pour augmenter potentiellement sa stabilité *in vivo*. En effet, certaines enzymes plasmatiques comme la trypsine peuvent cliver le peptide au niveau des résidus basiques (lysine, arginine) entrainant sa dégradation rapide. D'après les résultats de SAR, les positions 7 et 9 peuvent être modifiées en alanine sans perdre d'affinité contrairement aux lysines 14 et 16. K21 pourrait être remplacée par un résidu hydrophobe avec une chaine aliphatique de longueur similaire comme une leucine, une isoleucine ou une 6-azido-lysine. Enfin, le tripeptide RPP en C-terminal peut être enlevé et le dernier acide aminé amidé. Ces deux modifications n'altèrent pas l'affinité mais peuvent diminuer l'action d'exo-protéases (Craik and Adams, 2007). Pour améliorer l'affinité sur hMC4R, peu de modifications ont été explorées car l'objectif premier de la SAR visait ici le marquage du peptide. Aucune de ces modifications n'améliore significativement l'affinité sur hMC4R. Néanmoins, certaines données offrent des possibilités pour améliorer la sélectivité du peptide. La substitution de Q5 et Q18 par des lysines (Ltg1a-0SS-2) améliore surtout l'affinité pour les sous-types 1 et 5, respectivement d'un facteur 15 et 8. Sur les autres sous-types, l'amélioration est plus modérée avec un facteur 2-3. Cela signifie qu'au moins une de ces positions pourrait être modifiée pour améliorer la sélectivité de HT1-0 vis-à-vis des sous-types 1 et 5 via une inversion de charge. Or, d'après les résultats de SAR, la position 5 est importante. La substitution de la glutamine par un résidu acide devrait donc se faire en position 18. La séquence optimisée serait alors : LFHYQQAQAPPGFKGKFEQFLL#.

Néanmoins, cette question de sélectivité entre différents sous-types de MCRs est secondaire dans le cas d'un ligand hMC4R. En effet, les principaux effets pouvant limiter un futur développement thérapeutique proviendraient de l'implication du récepteur dans d'autres fonctions que l'homéostasie énergétique. Cette implication serait liée à l'activation de différentes voies de signalisation : G_s pour la fonction cardio-vasculaire, $G_{q/11}$ et/ou β-arrestine pour l'homéostasie énergétique (Li et al., 2016b; Lotta et al., 2019). Dans ce cas, la sélectivité fonctionnelle primerait sur la sélectivité moléculaire. L'étude de l'effet antagoniste de HT1-0 sur les voies de signalisation G_s, G_{15} et β-arrestine 2 ne montre cependant aucune sélectivité fonctionnelle du peptide pour une de ces voies. Mais il est très difficile de prédire des effets *in vivo* seulement à partir de ces résultats *in vitro* obtenus dans un système hétérologue. En effet, les résultats peuvent dépendre du modèle cellulaire utilisé et de la technique utilisée. Par exemple, dans notre système, aucune des voies G_q ou G_{11} n'est activée par NDP-α-MSH ou α-MSH alors que cela a été démontré dans des cellules hypothalamiques (Yang and Tao, 2017). Concernant la voie G_{15}, la protéine Gα associée

appartient également à la famille des protéines G_q. Jusqu'à présent, aucune publication n'a montré un couplage de $G\alpha_{15}$ avec le hMC4R. De plus, cette protéine G est surtout retrouvée au niveau de lignées hématopoïétiques et ne serait présente qu'à l'état de traces au niveau du cerveau (Wilkie et al., 1991).

CONCLUSION et PERSPECTIVES

Trois objectifs étaient fixés au début de la thèse. Le premier était de démontrer, via l'analyse statistique, l'intérêt des données générées par le projet Venomics. Le deuxième visait à caractériser d'un point de vue pharmacologique les premières toxines animales actives sur les récepteurs aux mélanocortines. Enfin, le dernier objectif était d'obtenir à partir d'une toxine animale, un potentiel lead dans le traitement de l'obésité ou de la cachexie cancéreuse.

Des toxines originales en tout point

Les analyses des données omiques issues du projet Venomics ont non seulement confirmé la richesse structurale que renferment les venins mais également élargi la connaissance que l'on en avait. Certains motifs de cystéines sont par exemple décrits pour la première fois. Et même au sein de grandes familles structurales comme les toxines ICK, une grande partie des séquences ne se rapprochent d'aucune déjà décrite. Cela suggère donc de nouvelles activités pharmacologiques. C'est ce qui a été démontré avec le criblage sur différentes cibles thérapeutiques de la 1ère banque de toxines synthétiques au Monde construite à partir des données omiques. Les toxines sélectionnées à partir du criblage sur le sous-type 4 des récepteurs aux mélanocortines (MC4R) montrent toutes une originalité structurale. Sur les cinq sélectionnées, trois ont pu être caractérisées de manière approfondie. Deux appartiennent à deux grandes familles structurales des toxines, la toxine d'araignée N-TRTX-Preg1a aux ICK et la toxine de scorpion N-BUTX-Ptr1a aux toxines courtes de scorpion en CSαβ. D'un point de vue phylogénétique, elles forment chacune un nouveau groupe au sein de leur famille structurale respective. Ce caractère à part est confirmé par un profil pharmacologique atypique. Alors que les ICK et les toxines courtes de scorpion sont connues pour leur activité sur les canaux ioniques, N-TRTX-Preg1a et N-BUTX-Ptr1a ne sont actives sur aucun des canaux testés. Elles représentent non seulement les premières toxines actives sur le sous-type 1 des récepteurs aux mélanocortines (MC1R) mais également les premiers agonistes de ce récepteur sans le pharmacophore HFRW. Le premier et le deuxième objectif de la thèse ont donc été remplis et ont fait l'objet d'un article soumis au *Journal of Medicinal Chemistry*. Néanmoins, l'affinité observée est faible et ne permet pas d'aller plus loin dans leur exploitation. En faire des outils d'imagerie du mélanome comme certains dérivés de l'agoniste endogène α-MSH serait trop compliqué (Yang et al., 2018; Zhang et al., 2018). De plus, un point resterait à éclaircir : l'absence d'activité sur hMC4R en test de fonction. Cela pourrait suggérer une sélectivité fonctionnelle pour une voie de signalisation autre que la voie G_s. Néanmoins, les faibles affinités restreignent l'intérêt de vérifier ce point d'autant plus que d'ici l'obtention de ces résultats, les autres données pharmacologiques devraient avoir été publiées.

HT1-0 : un potentiel lead contre la cachexie cancéreuse ?

L'étude de la toxine N-CTX-Ltg1a de *Conus lithoglyphus* a changé le paradigme sur l'importance des ponts disulfure pour l'activité des toxines animales. Même si cela pourrait n'être qu'un cas particulier, nous avons montré qu'une toxine animale sans pont peut toujours lier avec une grande efficacité une cible moléculaire. L'explication serait l'importance prépondérante de l'extrémité N-terminale, une zone moins affectée par la présence ou l'absence de ponts disulfure. Cette particularité a conduit à l'élaboration d'un peptide linéaire, la Hunger-Toxin (HT1-0). Ce peptide présente plusieurs caractéristiques remarquables. La principale est sa simplicité de production, critère crucial pour le développement thérapeutique d'une molécule. Ensuite, elle présente une affinité nanomolaire et un facteur 10 de sélectivité pour MC4R par rapport aux sous-types 1 et 5. Cependant, l'estimation de l'affinité est biaisée du fait de la dissociation lente du ligand radioactif utilisé en test de liaison, ^{125}I-NDP-α-MSH. Deux solutions peuvent alors être envisagées pour éviter ce problème. La première est d'utiliser un autre ligand radioactif avec une dissociation plus rapide comme ^{125}I-THIQ (Kopanchuk et al., 2006). La deuxième possibilité est d'utiliser le protocole décrit dans Antil-Delbeke et al., 2000 qui se base sur le dosage de sites non occupés par le peptide. Pour cela, le peptide à tester doit être préincubé avec le récepteur pour atteindre l'équilibre. Ensuite, le ligand radioactif ^{125}I-NDP-α-MSH est ajouté. Si le temps d'incubation est court (quelques minutes), l'équilibre entre la liaison du peptide et celle du ligand radioactif sur le récepteur n'est pas atteint et ^{125}I-NDP-α-MSH resterait alors lié uniquement sur les sites libres. Le dosage de ces sites en fonction de la concentration en peptide testée permet ainsi d'obtenir la constante de protection K_p du peptide équivalente à sa constante de dissociation K_d. Un autre point à souligner pouvant affecter les mesures d'affinité est le caractère cationique du peptide rendant son étude pharmacologique très complexe. De nouvelles études cinétiques de la liaison HT1-0-hMC4R ne sont ainsi pour l'instant pas prévues tant que ce problème de liaison au plastique ne sera pas résolu. Une solution est de produire la séquence optimisée d'après les études de SAR. Certains résidus basiques peuvent être supprimés ou remplacés. Le test futur de cette molécule déterminera si ces changements réduisent bien les interactions électrostatiques avec le plastique sans affecter la liaison sur hMC4R. Bien qu'affectant l'affinité, le caractère cationique de HT1-0 n'a normalement aucun effet sur sa nature antagoniste sur les différentes voies de signalisation étudiées de hMC4R.

HT1-0 est un antagoniste compétitif de l'agoniste endogène α-MSH pour les voies G_s, G_{15} et β-arrestine. La fonction antagoniste indique que l'application potentielle serait la cachexie cancéreuse avec un potentiel effet orexigène. Cependant, HT1-0 ne présente aucune sélectivité fonctionnelle pour la voie G_{15} ou β-arrestine par rapport à la voie G_s supposée être à l'origine d'effets cardiovasculaires. La faible efficacité du peptide sur ces trois voies n'est pas un problème en soi, seule la fenêtre thérapeutique compte. Nous savons par exemple qu'une dose de 50 mg/kg en injection intrapéritonéale chez des souris C57/BL/6J n'est pas létale. La priorité est maintenant de répondre à une question fondamentale : le peptide a-t-il un effet orexigène *in vivo* ?

Les premiers tests ont été conduits par l'équipe d'Alexandre Benani du Centre des Sciences du Goût et de l'Alimentation de Dijon. Le choix du mode d'administration a été fait au plus simple, à savoir en injection intrapéritonéale (ip). Au total, chacune des 20 souris

C57/BL/6J étudiées a reçu soit le véhicule (NaCl), soit HT1-49 (40-50 mg/kg). Néanmoins, aucune augmentation significative de la prise alimentaire n'a été détectée. Même si un effet significatif avait été observé, il aurait été impossible d'affirmer avec certitude qu'il soit causé par l'activité du peptide sur MC4R. Un effet orexigène peut être en effet causé par une activité sur le récepteur à la ghréline par exemple. Or, on ne sait pas si HT1-49 est actif ou pas sur ce récepteur. En s'affranchissant de ce cas, plusieurs hypothèses sont alors possibles pour expliquer l'absence d'effet : 1) le peptide n'a pas atteint le MC4R dans l'hypothalamus car trop rapidement dégradé et/ou incapable de franchir la barrière hématoencéphalique ; 2) le peptide a bien atteint sa cible mais n'induit aucun effet visible sur la prise alimentaire à cause d'une dose trop faible par exemple. Pour répondre à ces questions, des injections intracérébroventriculaires doivent être menées chez le rat au CEA et chez la souris à Dijon courant 2020. Restera la question du passage de la barrière hématoencéphalique (BHE). L'équipe d'Aloïse Mabondzo a développé un modèle *in vitro* de BHE (Lacombe et al., 2011). Cependant, aucun passage du peptide ^{125}I-HT1-49 n'a pu être détecté en utilisant ce modèle. Ce dernier représente une BHE peu perméable, donc peu représentative de l'éminence médiane au niveau de l'hypothalamus. L'expérience la plus précise programmée dans les prochains mois est la biodistribution du peptide marqué chez la souris. Pour des raisons administratives, le peptide iodé ne peut pas être utilisé. Il faut donc d'abord marquer le peptide au tritium. Une première stratégie a déjà été mise au point en deux étapes avec Héloïse Cahuzac, doctorante au CEA. La première consiste à faire réagir un DBCO-NH_2 avec du n-(propionyl oxy)succinimide pouvant être tritié. Puis le DBCO ainsi marqué peut être utilisé pour marquer à son tour par chimie click le peptide au niveau d'une azido-lysine en C-terminal. Le marquage en chaud devrait être entrepris prochainement. À noter que l'impact sur l'affinité n'est pas important dans ce contexte d'étude. En injectant ce peptide marqué chez la souris, on pourrait ainsi suivre sa bioditribution dans l'organisme et vérifier s'il atteint ou pas l'hypothalamus.

Au final, malgré quelques problèmes techniques à régler, HT1-0 pourrait bien représenter un potentiel lead dans le traitement de la cachexie cancéreuse. L'objectif 3 de la thèse est donc en cours de réalisation. Une demande de brevet est d'ailleurs en préparation et un article scientifique rassemblant les données *in vitro* obtenues pendant la thèse et les données *in vivo* futures est prévu. Néanmoins, le maintien de la demande de brevet et l'intérêt de l'article reposent entièrement sur les résultats *in vivo*, à savoir sur la démonstration de l'effet anorexigène du peptide. Si cet effet est confirmé, ce pourrait être le début du développement thérapeutique du premier peptide orexigène issu d'une toxine animale.

MATÉRIELS et MÉTHODES

1. Production des peptides

1.1 Matériel

Les Fmoc- acides aminés, c'est-à-dire possédant un groupement Fmoc en extrémité N-terminale (N-term), le 2-(6-chloro-1H-benzotriazole-1-yl)-1,1,3,3,-tetramethylaminium hexafluorophosphate (HCTU) proviennent d'Activotec (Cambridge, UK). La résine et les autres réactifs de synthèse peptidique (N-methylpyrrolidone (NMP), N-methylmorpholine (NMM), dichloromethane, piperidine, acide trifluoroactétique (TFA), anisole, thioanisole, and triisopropylsilane ont été commandés à la société Sigma (Saint-Quentin Fallavier, France).

Les résines Novabiochem® utilisées sont clivables en milieu acide à de fortes concentrations en TFA : résine Wang ou NovaPEG.

1.2 Synthèse chimique en phase solide Fluorénylméthoxycarbonyle (Fmoc)

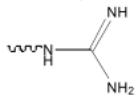
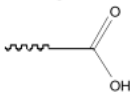
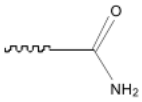
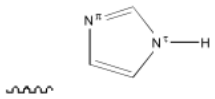
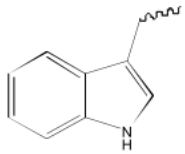
La synthèse chimique des différents peptides est réalisée sur un synthétiseur Prelude de Protein Technonologies, Inc à une échelle de 12.5 ou de 25 µmol en utilisant un excès d'un facteur 10 ou 20 en acides aminés-Fmoc par rapport à la résine utilisée. Le principe de la synthèse est décrit dans la **Figure 55A**. Le peptide est synthétisé de son extrémité C-terminale à son extrémité N-terminale. La fonction amine en alpha de l'acide aminé à incorporer est protégée par un groupement Fmoc labile en milieu basique. Les fonctions réactives des chaines latérales des acides aminés sont protégées par des groupements protecteurs stables en milieu basique mais labiles en milieu acide (**Table 18**). Les acides aminés sont couplés deux fois pendant 5 min en utilisant un mélange acide aminé/HCTU/NMM en proportion 1 :1 :2 dans du NMP. Cette étape permet l'incorporation d'un nouvel acide aminé via une réaction d'acylation entre l'ester activé par l'agent de couplage HCTU et l'amine libre du précédent résidu (**Figure 55B**).

Après incorporation de chaque résidu, la séquence est acétylée pendant 5 min en utilisant un excès d'un facteur 50 d'un mélange d'anhydride acétique et de NMM dans du NMP afin de bloquer définitivement les chaines aminées non réactives. La déprotection du Fmoc est réalisée deux fois pendant 2 min avec une solution de pipéridine 20% dans du NMP. Les lavages de 30s dans du NMP sont réalisés entre les étapes de déprotection et les étapes de couplage.

Une fois la séquence peptidique assemblée, la résine est traitée sous agitation pendant 2h avec un mélange TFA/anisole/thioanisole/triisopropylsilane/eau (82.5 :5 :5 :2.5 :5). Le TFA permet le clivage du peptide de la résine ainsi que la déprotection des chaines latérales. Les autres constituants sont des « scavengers » prévenant l'alkylation du peptide par les carbocations libérés lors de la déprotection. L'anisole est utilisé lorsque des groupements tert-butoxycarbonyl (Boc) sont présents. Le thioanisole est recommandé pour les peptides contenant des résidus Met, Cys et Arg. Il ne doit pas être utilisé lorsque la chaine peptidique contient des S-Acétamidométhyl-L-Cystéines (Acm) afin d'éviter le clivage précoce de ce groupement protecteur. Le triisopropylsilane sert à piéger les groupements

trityl (Trt), il remplace le malodorant éthane dithiol pour des peptides riches en résidus Arg et Trp. Enfin, l'eau prévient les réactions du peptide avec le groupement protecteur tert-Butyl. Après la déprotection, le brut de synthèse est obtenu après précipitation et lavages dans de l'éthyl éther froid puis dissous dans de l'eau/acide acétique 10%.

Table 18. Groupements protecteurs des chaines latérales des acides aminés en synthèse chimique en phase solide Fmoc. (Adapté de Mollica et al., 2014)

Amino-acid and side-chain functionality	Protecting groups	Removing
Arg 	Pmc	95% TFA
	Pbf	95% TFA
Asp/Glu 	OtBu	95% TFA
	OAll[a]	$Pd(Ph_3P)_4/PhSiH_3$
Asn/Gln 	Trt	95% TFA
	Trt	95% TFA
Cys -SH	Trt	95% TFA
	Trt	95% TFA
Histidine 	Trt(NH^{τ})	TFA
	Mtt[a]	1% TFA
Lysine $-NH_2$	Boc	TFA
	Boc Mtt[a] Alloc[a]	TFA TFA TFA
Ser/Thr/Tyr -OH	Trt[a]	1% TFA
	Boc	TFA
Tryptophan 	Boc	TFA

[a] Protection used for on-resin derivatization.

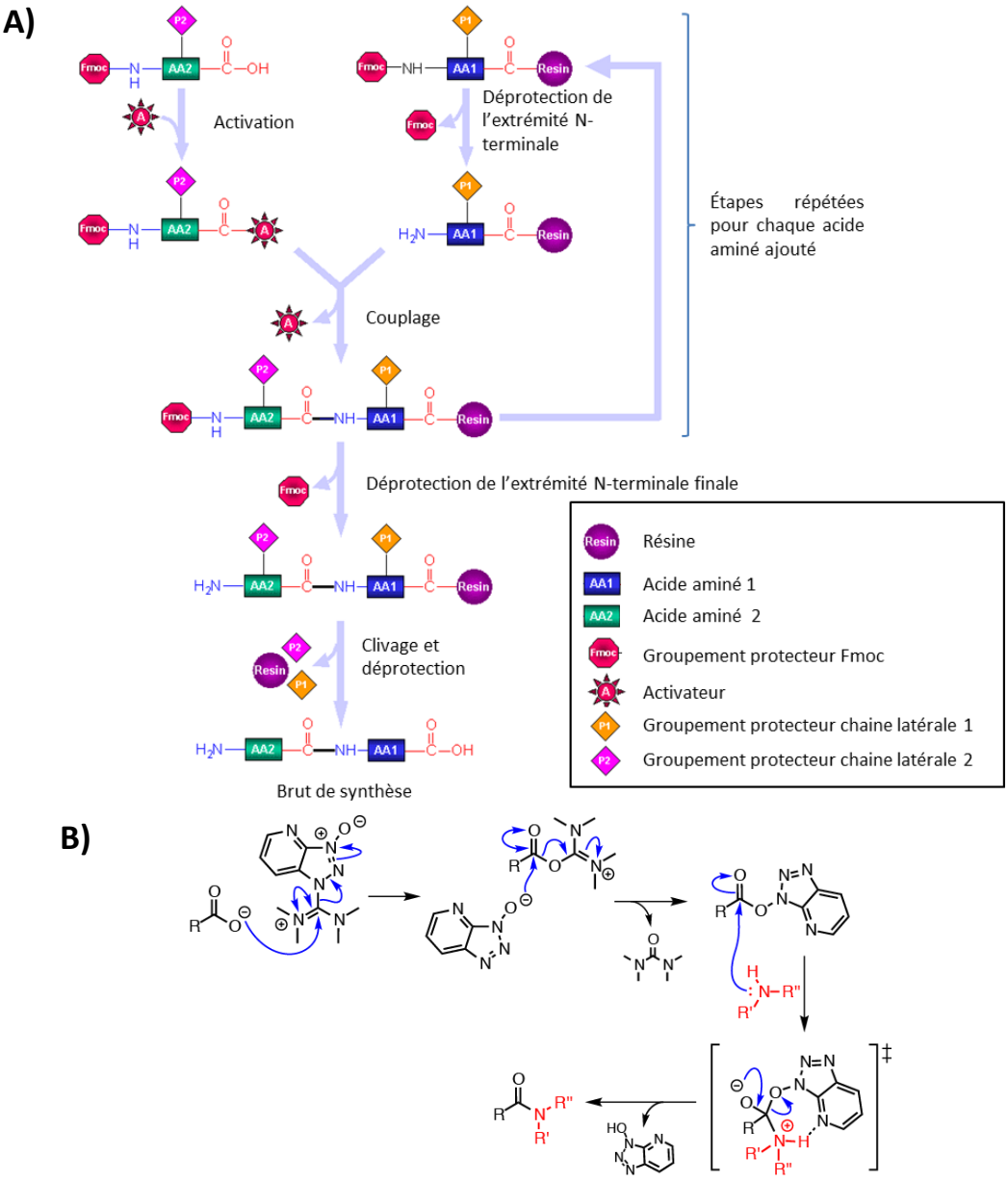


Figure 55. Principe de la synthèse chimique en phase solide. A) Pincipe général. Le premier acide aminé est fixé en extrémité C-terminale sur la résine. L'incorporation de l'acide aminé suivant se fait en deux étapes. La déprotection de la fonction NH_2 du premier acide aminé se fait en milieu basique (Pipéridine 20 % dans la N-méthylpyrrolidone (NMP)). La fonction COOH de l'acide aminé suivant est activée grâce à l'ajout d'un agent de couplage (HCTU ou HATU) en présence d'une base, la 4-méthylmorpholine (NMM). Cette étape permet la formation d'un ester actif augmentant ainsi les cinétiques de couplage des acides aminés. Entre chaque couplage, une étape d'acétylation des fonctions NH_2 des acides aminés libres n'ayant pas réagi à l'anhydride acétique minimise la formation de peptides porteurs de délétions. Ces étapes sont répétées jusqu'à l'obtention de la séquence désirée. Enfin, le traitement de la résine en milieu acide permet une déprotection des chaines latérales des acides aminés ainsi que le clivage du peptide de la résine aboutissant à l'obtention du brut de synthèse. B) Mécanisme réactionnel de N-acylation. L'agent de couplage HATU permet la formation d'un ester activé de l'acide aminé à incorporer. Cet ester réagit alors avec l'amine primaire du précédent résidu pour former une liaison peptidique.

1.3 Analyse LC/MS du brut de synthèse

Afin de s'assurer de la présence du peptide réduit souhaité dans le brut de synthèse, une analyse LC-MS est réalisée en chromatographie en phase inverse sur une HPLC Agilent couplée avec un spectromètre de masse Electrospray Ionization-Ion Trap (ESI-IT) (Brucker, Germany) avec une colonne analytique XBridge® peptide BEH C18 3.5 µm 4.6 x 150 mm (Waters) : 1mL/min ; solvant A, H_2O/TFA 0.1% ; solvant B, ACN/TFA 0.1% ; gradient : 5-50% solvant B en 25 min.

1.4 Purification

La purification du peptide réduit se fait en phase inverse sur une HPLC Préparative Waters 600 avec une colonne préparative XBridge® BEH 300Å Prep C18 19 x250 mm (Waters) à un débit de 10 mL/min. Suivant les quantités à purifier, la purification du peptide oxydé se fait comme précédemment ou sur HPLC Préparative Waters 600 avec une colonne semi-préparative BEH 300 Prep C18 5 µm 10 x 250 mm (Waters) à un débit de 5 mL/min.

1.5 Formation des ponts disulfure

- **Aléatoire en présence d'un couple oxydo réducteur**

Chaque peptide réduit dissous dans de l'eau est ajouté à une solution d'HEPES 100 mM ACN 20% contenant le couple oxydo-réducteur cystéine/cystine (ratio molaire 1 :50 :5 peptide/ cystéine/cystine avec une concentration finale en peptide réduit de 20 µM). Ce mélange est laissé à température ambiante pendant 48h. L'oxydation est arrêtée par acidification à un pH de 2-3 par ajout d'une solution de TFA 20%.

- **par pontage dirigé**

Dans le cas où l'on souhaite contrôler la formation successive des ponts disulfure, les paires de cystéines impliquées dans un des ponts du peptide sont protégées respectivement par un groupement acétamidométhyle (Acm) et un groupement trityl (Trt) (**Figure 56**). Contrairement au groupement Trt, le groupement protecteur Acm n'est pas clivé lors de la déprotection acide au TFA. Lors d'une première oxydation, seules les cystéines qui ne sont plus protégées formeront un pont disulfure. Dans un second temps, les deux cystéines restantes protégées par des Acm formeront le dernier pont disulfure (**Figure 56**).

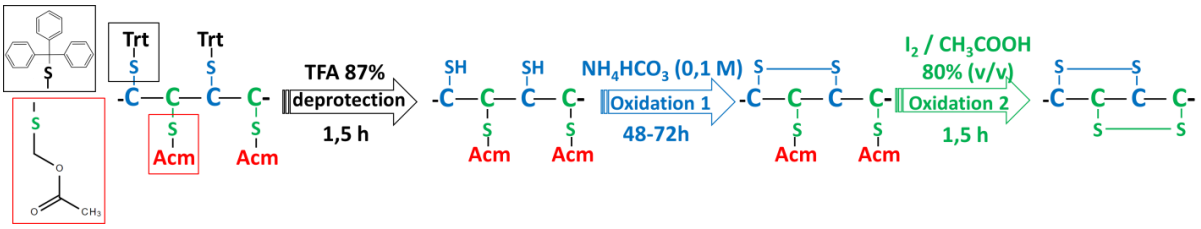


Figure 56. Exemple de la production en pontage dirigé d'un peptide à 2 ponts disulfure. Trt : groupement trityl labile en milieu acide ; Acm : groupement acétamidométhyle non labile en milieu acide**.**

Dans un premier temps, le peptide linéaire est dissous sous agitation à l'air dans une solution d'hydrogénocarbonate d'ammonium à 0.1 M pH 8.3, de telle sorte à avoir une concentration finale en peptide de 0.1 mg/mL. Lors d'une première oxydation, seules les cystéines non protégées forment un pont disulfure. Le mélange réactionnel est laissé plusieurs jours à température ambiante. L'oxydation est ensuite arrêtée par acidification au TFA.

Dans un second temps, après purification du peptide, la formation du dernier pont avec les cystéines Acm peut être réalisée. Le peptide issu de la 1ère oxydation est ajouté à une solution d'iode (25 fois en excès par rapport au peptide) dissoute dans de l'acide acétique 80%. Le mélange est mis sous agitation pendant 1h30. L'iode a un fort pouvoir oxydant permettant non seulement la déprotection des cystéines mais également l'oxydation de leurs fonctions thiol. La réaction est arrêtée par dilution du mélange dans de l'eau. L'iode est extrait avec du dichlorométhane dans une ampoule à décanter ou neutralisé avec une solution de thiosulfate de sodium 1M. La solution contenant le peptide entièrement oxydé peut ensuite être purifiée en HPLC.

1.6 Évaluation de la qualité du peptide produit

La qualité de chaque toxine purifiée est vérifiée par :

- HPLC analytique en phase inverse sur un appareil Alliance (Waters) avec une colonne analytique XBridge® peptide BEH C18 5 µm 4.6x250mm (Waters) : 1 mL/min ; solvant A, H_2O/TFA 0.1% ; solvant B, ACN/TFA0.1% ; gradient : 1% solvant B /min.
- Spectrométrie de masse MALDI-TOF (Matrix-assisted laser desorption/ionization – Time of flight) sur l'appareil 4800 MALDI TOF-TOF Analyser (Applied Biosystems, USA). L'acide α-Cyano-4-hyroxy-cinnamique (CHCA) est utilisé comme matrice (Quinton et al., 2007). Les fractions issues de l'HPLC en phase inverse ou le produit purifié final est mélangé en proportions 1 :1 (v/v) avec la matrice (7.5 mg/ml dans 50/50 ACN/H_2O, 0.1% TFA) sur une plaque MALDI. Les spectres MALDI-TOF sont obtenus en mode réflectron et les masses enregistrées correspondent aux masses monoisotopiques $M+H^+$.

1.7 Détermination de la concentration en peptide produit

- **par analyse de la composition en acides aminés**

La composition en acides aminés est réalisée dans les conditions suivantes : l'échantillon est lyophilisé et mis dans des tubes en verre utilisant le système PicoTag (Waters Associates, Milford, MA) puis hydrolysé en phase gazeuse de 6N HCl en présence de phénol cristallisé pendant 17h à 110°C. L'échantillon hydrolysé est ensuite dissous dans 100 µL d'eau milliQ (contenant au minimum 200 pmol de chaque acide aminé) puis analysé et quantifié via dérivatisation à la ninhydrine à l'aide d'un aminoTac JLC-500/V amino acids analyzer (JEOL, Japan). Une calibration avec des solutions standards d'acides aminés est réalisée au début de chaque analyse et permet de quantifier le peptide analysé.

- **par spectrophotométrie**

L'absorption de la solution à 280 nm est mesurée à l'aide d'un spectrophotomètre. La concentration C de la solution est ensuite calculée selon la loi de Beer-Lambert :

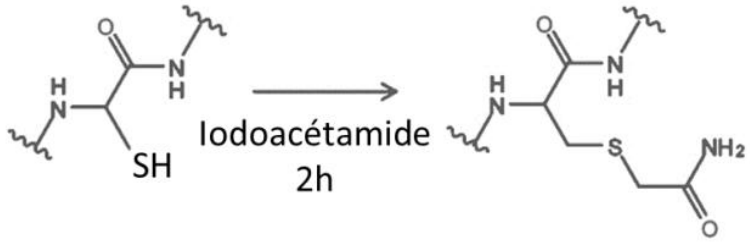
$$C = \frac{A}{l * \varepsilon}$$

avec C la concentration de l'échantillon en mol/l, l la longueur du trajet optique correspondant à l'épaisseur de la cuve en cm et ε le coefficient d'extinction molaire à 280 nm en l/mol/cm.

2. Modifications chimiques du peptide

2.1 Alkylation des cystéines du peptide réduit

Le peptide linéaire purifié est traité à l'iodoacétamide (10 fois en excés par rapport au peptide) dissous dans du chlorure de Guanidinium 6M pendant 2h sous agitation à l'obscurité. À la suite de cette réaction, les fonctions thiols des cystéines sont protégées et ne peuvent plus former de ponts disulfure **(Figure 57)**.



Cys Cys alkylée

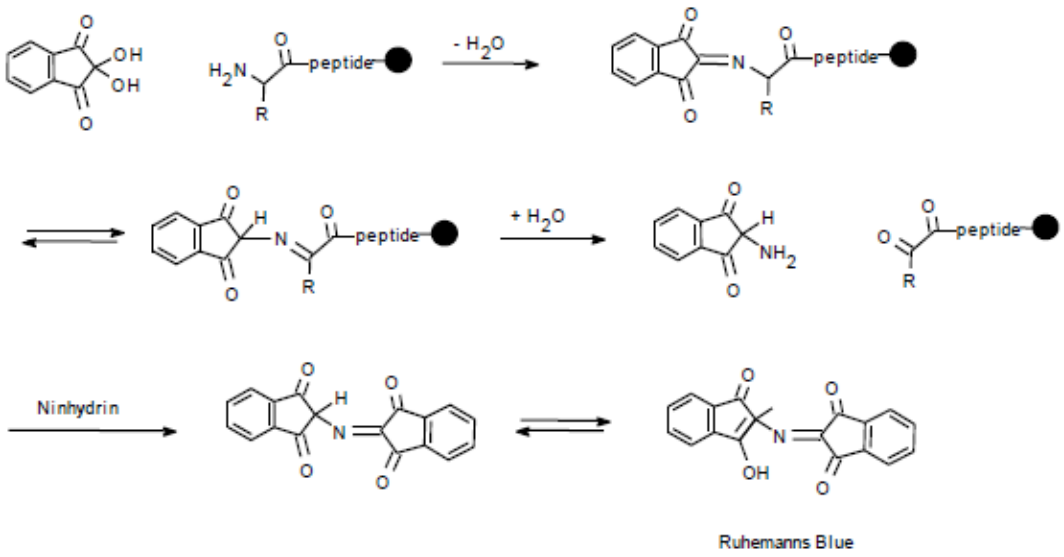
Figure 57. Protection des cystéines du peptide réduit par traitement à l'iodoacétamide. Cys : cystéine.

2.2 Marquage

- **Fluorescent (Fluorescéine, Rhodamine)**

Après synthèse de la séquence peptidique, le marquage se fait en trois étapes directement sur la résine. La première étape consiste en la déprotection sélective en milieu basique de l'amine primaire de K21 portant un groupement ivDde. La séquence synthétisée doit donc présenter en N-terminal un groupement Boc à la place d'un Fmoc évitant de ce fait une amine N-terminale libre. La déprotection de K21-ivDde se fait en présence d'une solution basique d'hydrazine 5% dans du NMP. La résine est ensuite lavée 10 fois avec 5 mL de NMP pur, puis 10 fois avec 5 mL de Dichlorométhane (DCM). Un test de Kaiser permet de vérifier la présence de l'amine libre (**Figure 58**)**.**

A)



B)

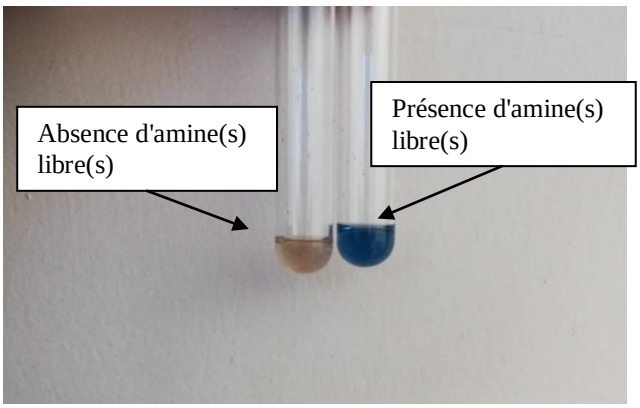


Figure 58. Principe du test de Kaiser. A) Mécanisme réactionnel. Une partie de la résine sur lequel est fixé le peptide est traitée avec une solution de ninhydrine. La ninhydrine réagit alors avec l'amine libre du peptide pour donner une solution colorée (B).

La seconde étape est celle du couplage du fluorophore sur la chaîne latérale de la lysine. Celui-ci se fera par une réaction d'acylation entre le carboxylate activé du fluorophore et l'amine primaire de la chaîne latérale de la lysine. Dans le cas de la 5-Carboxyfluorescéine (**Figure 59A**), la fonction carboxylate doit être activée en présence de l'agent de couplage HCTU. Le marquage se fait dans un mélange peptide/HCTU/NMM en proportion 1 :5 :10 dans du NMP. Pour la Rhodamine Red™-X, Succinimidyl Ester (**Figure 59B**), le carboxylate est déjà activé. Le marquage se fait alors dans un mélange peptide/NMM en proportion 7 :1 dans du NMP. Un second test de Kaiser peut être réalisé pour s'assurer de l'absence d'amines libres en solution.

A) **B)**

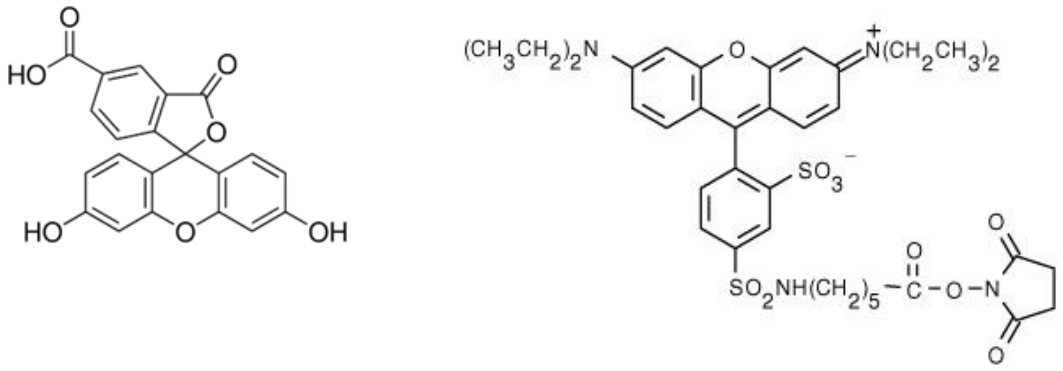


Figure 59. Structure des fluorophores utilisés. A) 5-Carboxyfluorescéine. B) Rhodamine Red™-X, Succinimidyl Ester.

Enfin, la résine est traitée en milieu acide comme décrit précédemment permettant le clivage du peptide de la résine et la déprotection des chaines latérales et de l'extrémité N-terminale.

- **Radioactif**

Le marquage se fait via l'utilisation de $Na^{125}I$. Cette forme réduite de l'iode vient réagir avec le groupement phénol d'une tyrosine. Ces groupements doivent être préalablement oxydés à l'aide d'un agent oxydant, ici l'iodogène (agent oxydant doux). Le marquage se fait en ortho de la fonction hydroxyle de la tyrosine selon une substitution nucléophile.

De facon pratique, une solution de peptide froid à 30 µM dans du TRIS 100 mM pH 7.8 est mise en présence de 0.4 mCi de $Na^{125}I$ pendant 10 min. Après purification, la radioactivité et les masses présentes dans chaque fraction sont analysées. Après identification, la fraction représentant le peptide monoiodé est ensuite diluée d'un facteur 4 dans de l'HEPES 10 mM BSA1% pour prévenir la radiolyse.

3. Dichroïsme circulaire

3.1 Principe

L'intérêt du dichroïsme circulaire (DC) est d'obtenir des informations sur les structures secondaires et tertiaires d'un peptide. Cette méthode est basée sur l'absorption inégale d'une lumière polarisée circulairement droite (E_{droite}) par rapport à une lumière polarisée circulairement gauche (E_{gauche}). Ces deux ondes sont émises par un faisceau monochromatique rectiligne de direction connue. Si la molécule étudiée est chirale, celle-ci absorbera et réfractera différemment E_{droite} et E_{gauche}. La résultante des deux ondes est alors une onde polarisée elliptique dont l'ellipticité dépend de la longeur d'onde et de la structure du peptide analysé (Greenfield, 2006).

3.2 Mode opératoire

Les analyses sont faites sur un JASCO J-815 CD Spectometer couplé à un Peltier JASCO CDF-426S/15 permettant le réglage de la température d'analyse. Le peptide est dilué dans l'eau à une concentration de 20 µM. Le pH n'est pas contrôlé. Chaque solution est ensuite analysée à 20°C dans une cuve de quartz de 1 mm de trajet optique. L'analyse est répétée trois fois entre 190 et 250 nm. Les mesures de dichroïsme circulaire (mdeg) sont ensuite transformées en ellipticité molaire et traitées avec le logiciel Prism, v5.0.0 (GraphPad, La Jolla, CA, USA). La prédiction des structures secondraires est réalisée grâce aux logiciels BeStSel (http://bestsel.elte.hu/index.php) ou Spectra Manager™ II (JASCO). Pour ce dernier, la référence de Reed est utilisée (Reed and Reed, 1997).

L'ellipticité molaire est calculée comme suit :

$$\textit{Ellipticité molaire} = \frac{DC}{L*C*l}$$

avec DC : dichroïsme circulaire (mdeg) ; L : Longueur du peptide ; C : concentration de la solution analysée ; l : largeur de la cuve

4. Caractérisation pharmacologique

4.1 Tests de liaisons

- **Compétition**
  - Principe

Pour mesurer l'affinité et la spécificité du ligand pour le récepteur, on réalise des tests de liaison avec un ligand radiomarqué. Lors des tests, on met en présence le ligand et son récepteur. La liaison spécifique entre les deux est mesurée grâce à l'élimination par filtration des ligands libres en excès. La radioactivité mesurée (BT) correspond à des liaisons spécifiques (BS) et non spécifiques (NS) du traceur radioactif:

BS = BT – NS

Avec :

BT : Liaison totale. C'est la somme de la liaison spécifique du ligand à son récepteur (forte affinité) et à d'autres sites de liaison non spécifique (faible affinité).

NS : Liaison non spécifique. Elle est mesurée en présence d'une quantité de ligand non radioactif (ligand froid : L) suffisante (en excès) pour empêcher la fixation du ligand radioactif sur ses sites spécifiques.

BS = BT- NS : Liaison spécifique.

En utilisant une gamme de quantité croissante de toxine, on détermine ainsi à l'équilibre la capacité de la molécule testée à rentrer en compétition avec le ligand radiomarqué pour la liaison sur son récepteur.

  - Mode opératoire

Les expériences de compétition sont réalisées en utilisant des membranes de cellules CHO-K1 exprimant les récepteur humains hMC1R, hMC4R ou hMC5R ou de cellules HEK293 exprimant le MC3R humain. Le ligand radioactif utilisé est le [^{125}I]-[Nle4, D-Phe7]-α-MSH. Tous les produits utilisés proviennent de PerkinElmer (Courtaboeuf, France). Le seul agoniste connu du MC2R est l'ACTH. Le test de liaison employé ici ne permet donc pas d'étudier ce sous-type de MCR.

Dans chaque puits d'une plaque 96 puits (Perkin Elmer, Courtaboeuf, France), des concentrations croissantes en toxine sont ajoutées à 18 000 désintégrations par minute (dpm) de ligand radioactif dissous dans le tampon de liaison (HEPES 25mM pH 7.4, 1mM $MgCl_2$, 1mM $CaCl_2$, 0.2% Bovine Serum Albumin (BSA) et à une quantité définie de membrane (hMC1R : 6 µg/puits ; hMC3R : 0.192 µg/puits ; hMC4R : 0.015 µg/puits ; hMC5R : 0.04 µg/puits). Du tampon de liaison est ajouté pour un volume final de 100 µL.

Contrôles :

- Contrôle positif : la toxine est remplacée par des concentrations croissantes en NDP-α-MSH (Tocris bioscience, UK).
- NS : la toxine est remplacée par NDP-α-MSH (Tocris bioscience, UK) à une concentration finale de 10 µM.
- BT : aucune toxine n'est ajoutée.
- BF : les membranes ne sont pas ajoutées.

Après 3h d'incubation, la plaque est filtrée à travers un filtre GF/C. Le filtre est préalablement saturé avec une solution de polyéthylèneimine (PEI) 0.5% afin de limiter la fixation du ligand radioactif libre sur le filtre. Le filtre est ensuite lavé trois fois avec 300 µL

de tampon de liaison puis séché avant d'ajouter 25 µL d'Ultimagold F (PerkinElmer, Courtaboeuf, France) dans chaque puits. L'enregistrement de la radioactivité est réalisé sur un compteur Packard TopCount NXT (rendement de comptage : 65%).

Les courbes d'inhibition de la liaison sont ensuite obtenues grâce au logiciel Prism, v5.0.0 (GraphPad, La Jolla, CA, USA) en utilisant un modèle à un site de liaison. Les valeurs d'IC50 sont converties en Ki suivant l'équation de Cheng-Prussof (Cheng and Prusoff, 1973) : $\boldsymbol{Ki} = \frac{\boldsymbol{IC50}}{\mathbf{1}+\frac{(\boldsymbol{L})}{\boldsymbol{Kd}}}$ avec IC50, la concentration de ligand inhibant 50 % de la liaison spécifique de ^{125}I-NDP-α-MSH ; L, la concentration en ^{125}I-NDP-α-MSH utilisée dans l'expérience ; Kd, la constante de dissociation de ^{125}I-NDP-α-MSH ; Ki, la constante d'inhibition du compétiteur, considérée ici comme une estimation de l'affinité du compétiteur pour un récepteur donné.

- **Cinétique d'association/dissociation**
  - Radioactif
    - Par scintillation liquide

Le principe est similaire à celui décrit précédemment. Seul le temps d'incubation diffère.

^{125}I-HT1-49 :

Dans une plaque 96 puits, des membranes présentant le récepteur hMC4R sont ajoutés à intervalle de temps régulier à une solution de ligand radioactif seul ou en présence d'un excès de ligand froid (NDP-α-MSH ou HT1-49). Le premier puits correspondra donc au temps d'incubation le plus long tandis que le dernier correspondra au temps le plus court. La filtration de la plaque à travers un filtre préalablement saturé au PEI permet ainsi de retracer la cinétique de liaison du ligand radioactif. Le filtre est ensuite lavé trois fois avec 300 µL de tampon de liaison puis séché avant d'ajouter 25 µL d'Ultimagold F (PerkinElmer, Courtaboeuf, France) dans chaque puits. L'enregistrement de la radioactivité est réalisé sur un compteur Packard TopCount NXT (rendement de comptage : 65%).

Le ligand radioactif utilisé est issu de l'iodation du peptide HT1-49.

^{125}I-NDP-α-MSH:

Contrairement à la méthode précédente, la réaction n'est pas faite dans plusieurs puits en parallèle. Dans un tube contenant 4 mL de ligand radioactif avec ou sans NDP-α-MSH en excès (1 µM final), 4 mL d'une solution de membranes hMC4R est ajoutée. À intervalle de temps régulier, 100 µL de la solution est prélevé puis filtré à travers un filtre GF/C préalablement saturé au PEI 0.5%. La radioactivité du filtre est ensuite mesurée à l'aide d'un compteur gamma LKB Wallac 1261 Multigamma. Après 3h d'incubation, la dissociation du ligand radioactif est initiée par ajout de NDP-α-MSH en excès (1 µM final). Les mesures sont réalisées de la même façon que lors de l'étude de l'association.

Le ligand radioactif utilisé provient de Perkin Elmer.

- Par scintillation de proximité

La scintillation de proximité est basée sur le principe suivant : des billes de scintillant convertissent en lumière l'énergie émise par les particules radioactives β^-. Cette conversion n'est possible que si la molécule radioactive émettrice est suffisamment proche des billes (17 µm dans le cas d'une molécule marquée à l'iode 125). Elle ne nécessite donc pas la séparation du ligand radioactif lié du libre via une étape de filtration. En outre, les mesures

peuvent facilement être automatisées.

Les billes utilisées sont des billes WGA PVT. Le scintillant, le diphénylanthracine, est copolymérisé dans une matrice de polyvinyltoluène (PVT). Les billes sont recouvertes d'aglutinine de germe de blé permettant la liaison de membranes à leur surface. Elles sont incubées 30 min avec les membranes hMC4R dans le tampon de liaison (identique à la scintillation liquide) puis distribuées dans les puits d'une plaque OptiPlate 96 (Perkin Elmer). Le ligand radioactif est ensuite ajouté à différentes concentrations pour lancer l'association. La dissociation est initiée par ajout de ligand froid en excès. Les mesures sont réalisées de manière automatisées à l'aide d'un compteur Packard TopCount NXT.

Les données sont ensuite traitées avec le logiciel Prism, v5.0.0 (GraphPad, La Jolla, CA, USA). La courbe d'association est tracée selon un modèle à un site de liaison. La courbe de dissociation est obtenue selon une décroissance exponentielle monophasique.

- Fluorescent
  - Principe du transfert d'énergie par résonance de Förster (FRET)

Le FRET est basé sur le transfert d'énergie d'un fluorophore donneur préalablement excité à un fluorophore accepteur. Ce transfert ne peut avoir lieu que si les deux protagonistes sont suffisamment proches (1-10 nm). Le choix du couple de fluorophores doit être fait de telle sorte que la longueur d'onde d'émission du donneur corresponde à la longueur d'onde d'excitation de l'accepteur (**Figure 60**). En marquant un récepteur par un flurorophore donneur et un peptide avec un fluorophore accepteur, on peut ainsi étudier l'interaction de ce peptide sur le récepteur d'intérêt. La limite de cette méthode est le bruit de fond causé par des émissions non spécifiques à court terme. La société Cisbio a ainsi développé une technologie de mesure de fluorescence résolue en temps homogène (HTRF). Elle est basée sur l'introduction d'un délai de 50 à 150 µs entre l'excitation et la mesure de fluorescence et l'emploi d'accepteurs émettant une fluorescence prolongée comme le cryptate de terbium (Tb^{2+}) (**Figure 61**).

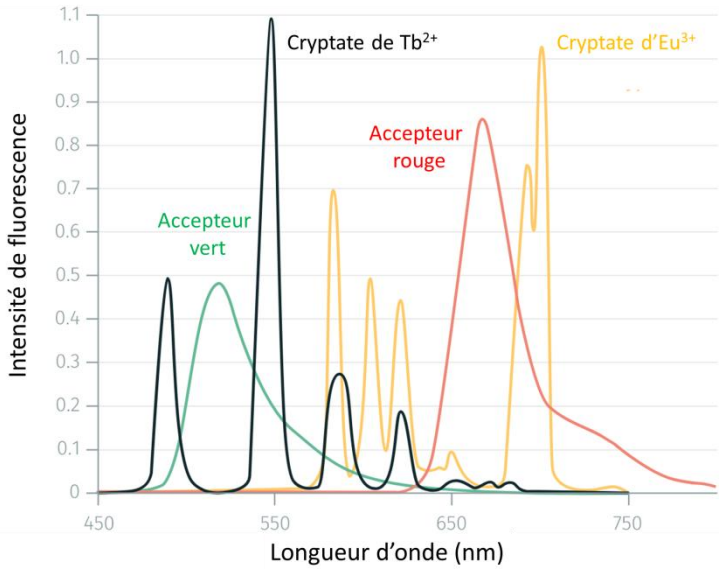


Figure 60. Spectres d'émission d'accepteurs vert ou rouge et de donneur HTFR (cryptate de Tb^{2+} ou d'Eu^{3+}). D'après Cisbio.

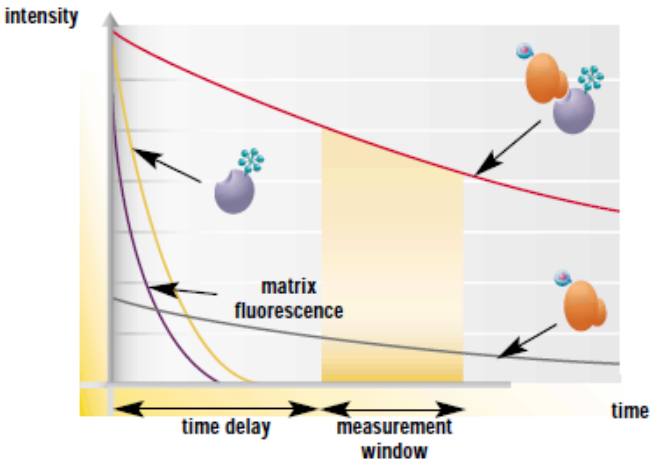


Figure 61. Principe de la fluorescence résolue en temps homogène (HTRF). Suite à l'excitation du donneur (energy pulse), l'enregistrement de la fluorescence émise (measurement window) du donneur seul (gris) et de l'accepteur (rouge) est retardé de 50-150 µs (time delay), évitant ainsi la mesure de la fluorescence non spécifique (milieu, protéines, ...) (violet) et de l'accepteur seul (jaune). D'après Cisbio.

- Mode opératoire

Les cellules exprimant le récepteur hMC4R marqué au cryptate de Tb^{2+} (Cisbio) sont mises en présence de plusieurs concentrations de peptide marqué à la fluorescéine. La fluorescence à 520 nm (émission de la fluorescéine) et à 620 nm (émission du cryptate de Tb^{2+}) est mesurée à intervalle régulier et de façon automatisée grâce à un CLARIOstar (BMG LABTECH). Les mesures sont faites dans une plaque GREINER 384 SMALL VOLUME**.**

4.2 Tests de fonction

- **Dosage direct (FRET) (réalisé par la société Eurofins Cerep, France)**
  - Principe

La quantification de l'AMPc par le kit Cisbio (cAMP – GS dynamic kit), repose sur un test de compétition permettant de doser la concentration intracellulaire d'AMPc. La première étape consiste en l'incubation pendant 30 min de cellules CHO exprimant le récepteur hMC4R avec une concentration fixe d'agoniste connu, ici NDP-α-MSH à 1 nM, et différentes concentrations de peptide HT1-0. Suit l'étape de dosage. Les cellules sont lysées afin de stopper la réaction et libérer dans le milieu l'AMPc produit. Un mélange d'AMPc marqué avec le fluorophore accepteur D2 et d'anticorps anti-AMPc marqué avec le fluorophore donneur est ajouté au milieu. L'AMPc produit non marqué et l'AMPc marqué entrent alors en compétition pour la reconnaissance par l'anticorps. Par conséquent, le signal de FRET mesuré est inversement proportionnel à la quantité d'AMPc produit par les cellules au cours de la réaction avec les ligands (**Figure 62**).

  - Protocole

Pour le test d'agoniste, les cellules de l'ovaire de hamster chinois (CHO) exprimant de façon transitoire hMC4R sont mises en suspension dans un tampon HBSS (Invitrogen) complété avec 20 mM HEPES (pH 7.4) et IBMX 500 µM, puis réparties en plaques à une densité de 5 103 cellules par puits et incubées pendant 30 minutes à 37°C en présence du HBSS (témoin basal), du composé à analyser ou de l'agoniste de référence. Pour la mesure du contrôle stimulé, des puits séparés contiennent 30 nM NDP-α-MSH. Après incubation, les

cellules sont lysées et l'accepteur de fluorescence (AMPc marqué D2) et le donneur de fluorescence (anticorps anti-cAMP marqué au cryptate d'europium) sont ajoutés. Après 60 min à température ambiante, le transfert de fluorescence est mesuré à λex=337 nm et λem=620 et 665 nm en utilisant un lecteur de microplaques (Envision, Perkin Elmer). La concentration de l'AMPc est déterminée en divisant le signal mesuré à 665 nm par celui mesuré à 620 nm (rapport). Les résultats sont exprimés en pourcentage de la réponse du témoin à 30 nM NDP-α-MSH. Pour le test de l'effet antagoniste, les cellules sont préincubées pendant 5 minutes à température ambiante en présence de HBSS (témoins basaux et stimulés), du composé à tester ou de l'antagoniste de référence. Par la suite, l'agoniste de référence NDP-α-MSH est ajouté à une concentration finale de 1 nM. Pour les mesures de contrôle basal, des puits séparés ne contiennent pas NDP-α-MSH. Enuite, le protocole est le même que pour l'étude des agonistes. L'antagoniste de référence, SHU 9119, est testé dans chaque expérience à plusieurs concentrations pour générer une courbe concentration-réponse à partir de laquelle sa valeur d'IC50 est calculée. L'agoniste de référence standard est NDP-α-MSH testé dans chaque expérience à plusieurs concentrations pour générer une courbe concentration-réponse à partir de laquelle sa valeur d'EC50 est calculée. L'EC5O correspond à la concentration d'agoniste nécessaire pour obtenir 50 % de la réponse maximale

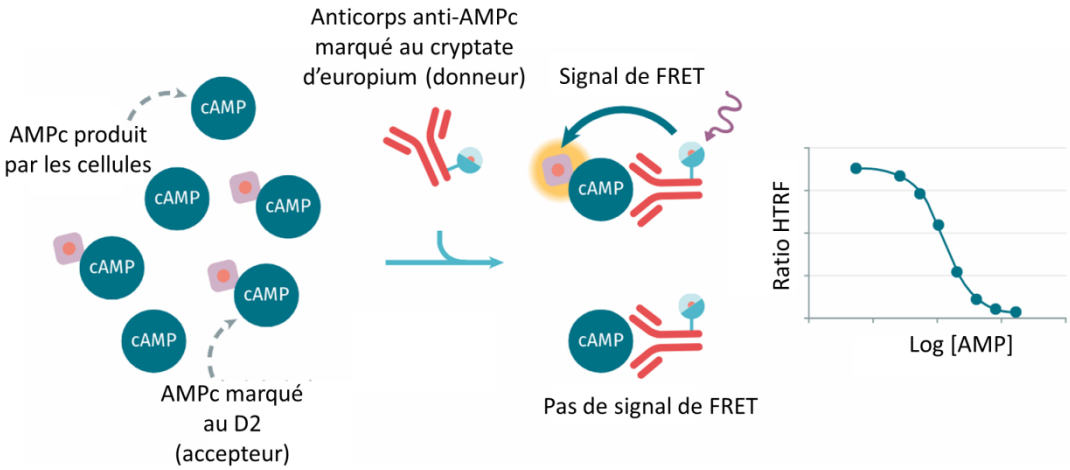


Figure 62. Principe de la quantification de l'AMPc produit par les cellules par compétition de liaison à l'anticorps anti-AMPc. L'AMPc produit par la cellule au cours de la réaction rentre en compétition avec l'AMPc marqué au D2 (accepteur) pour la liaison à l'anticorps anti-AMPc marqué par un cryptate (donneur). La liaison de l'anticorps à l'AMPc produit par les cellules ne donne lieu à aucun signal. Le ratio HTRF est inversement proportionnel à la quantité d'AMPc produite par les cellules. Schéma d'après Cisbio.

- **Dosage indirect par transfert d'énergie par résonance de bioluminescence (BRET) (réalisé par Suli-Anne Laurin, doctorante à l'université de Montréal dans l'équipe du Prof Michel Bouvier)**

  - Principe

Le principe du BRET est identique à celui du FRET sauf qu'il nécessite une réaction enzymatique. Le BRET utilise comme donneur une enzyme. En présence de son substrat, l'enzyme catalyse sa transformation avec émission de lumière à une certaine longueur

d'onde. Si un accepteur est suffisamment proche, la lumière émise y est transférée et on peut ainsi mesurer le signal de BRET à la longueur d'onde d'émission de l'accepteur. Dans le cas présent, l'enzyme est une luciférase RlucII catalysant la transformation du DeepBlueC (DBC) ou du Prolume Purple avec émission maximale de lumière vers 400 nm. L'accepteur est la GFP10 avec une longueur d'onde d'émisson maximale à 515 nm. Contrairement au FRET décrit précédemment, les cellules HEK293 surexpriment de manière transitoire non seulement le récepteur hMC4R mais aussi la protéine Gα d'intérêt ou la β-arrestine 2 et le senseur de localisation membranaire CAAX.

L'activité agoniste de HT1-0 a été testée avec le biosenseur GRK2 (demande de brevet : US20170234870A1). Lorsque la Gα d'intérêt surexprimée dans les cellules se couple avec le récepteur hMC4R, l'activation de celui-ci promeut la dissociation du complexe Gαβγ et éventuellement l'interaction entre la sous-unité Gγ surexprimée (marquée avec le donneur RlucII) et l'effecteur GRK2 surexprimé (marqué avec l'accepteur GFP10), ce qui provoque une augmentation du signal de BRET. Pour la voie β-arrestine 2, le biosenseur β-arrestine-RlucII a été employé (Namkung et al., 2016). Si l'activation du récepteur promeut le recrutement de la β-arrestine d'intérêt, le tag RlucII marquant la β-arrestine surexprimée peut interagir avec l'accepteur GFP10 surexprimé et ancré à la membrane par le domaine CAAX de k-ras, provoquant de ce fait une augmentation du BRET. Les enregistrements sont faits sur 25 min dès l'ajout du peptide HT1-0 aux cellules. L'étude de l'activité antagoniste de HT1-0 a été explorée à l'aide du biosenseur GAPL (non publié, confidentiel). Pour la voie de signalisation G_s, la protéine $G\alpha_s$ marquée avec RlucII quitte le senseur de localisation membranaire rGFP-CAAX lorsqu'elle est activée. L'activation du récepteur couplé à $G\alpha_s$ entraine donc une baisse du signal de BRET. Pour les voies G_{15} et β-arrestine 2, l'activation du récepteur promeut le recrutement d'un effecteur spécifique marqué avec RlucII au senseur de localisation membranaire rGFP-CAAX. Pour la voie G_{15}, l'effecteur est p63. Pour la voie β-arrestine 2, il s'agit de la β-arrestine. Dans les deux cas, l'activation du récepteur conduit à une hausse du signal de BRET.

- Protocole

Concrètement, pour tester l'activité antagoniste de HT1-0, les cellules sont préincubées avec HT1-0 pendant 30 min puis sont stimulées pendant 45 min avec l'agoniste α-MSH ou NDP-α-MSH. Pour les voies G_s et G_{15}, le substrat DBC est ensuite incubé avec les cellules pendant 5min. Pour la voie β-arrestine 2, le substrat est le Prolume Purple incubé pendant 6 min. Enfin, le signal de BRET est mesuré à t = 3 min pour α-MSH et 10 min pour NDP-α-MSH pour les voies G_s et G_{15} et à tmax pour la voie β-arrestine afin de compenser l'internalisation du récepteur. Les lectures se font grâce au lecteur Tristar² de Berthold technology (BRET2 : 37°, integration 0,1 sec, channels 400 nm and 515 nm).

5. Obtention de cellules exprimant le récepteur hMC4R marqué à la GFP

5.1 Amplification du plasmide codant pour hMC4R-GFP

Le plasmide utilisé est un plasmide pCMV6-AC-GFP codant pour le récepteur hMC4R (plMC4RGFP) marqué en C-terminal par la GFP. Il est amplifié dans des bactéries *E. coli* XL1 Blue comme suit. Les bactéries sont transformées après choc thermique avec 0.5 µg de plasmide et sont ensuite mises en culture à 37°C sous agitation (200 rpm) en présence d'ampicilline (100 µg/mL). Après 24h, la préculture sert à ensemencer 2 x 200 mL de milieu LB. Les cultures sont ensuite mises en culture à 37°C sous agitation (200 rpm) en présence d'ampicilline (100 µg/mL) pendant 24h. Le lendemain, les cultures sont centrifugées à 4 000 rpm pendant 30 min à 4°C. Après élimination du surnageant, le plasmide est extrait selon le protocole « Midi Prep High copy » du kit NucleoBond®. Après évaporation à l'éthanol, le culot obtenu est repris dans du tampon TRIS 5 mM pH 8.5. La concentration en plasmide de la solution est déterminée par mesure d'absorbance au CLARIOstar (BMG LABTECH). Le rapport d'absorbance 260/280 est vérifié et doit être proche de 1.8. La présence du plasmide est vérfié par migration sur gel d'agarose 1% pendant 1h dans du tampon TAE.

5.2 Transfection de cellules CHO K1

Des cellules CHO K1 sont transfectées avec le plasmide plMC4RGFP suivant le protocole décrit dans le kit Lipofectamine 3000®. La transfection s'effectue dans une plaque 6 puits adaptée pour la culture de cellules adhérentes. Les cellules sont laissées en culture à 37°C pendant 24h-48h dans du milieu DMEM/F12 complémenté avec du sérum de veau fœtal (10%) et du glutamax.

6. Visualisation de la liaison de HT1-0 marqué à la rhodamine sur le récepteur hMC4R-GFP

Le milieu des puits contenant les cellules CHO transfectées est enlevé et remplacé par du milieu contenant le peptide marqué avec un fluorophore. Les cellules sont laissées à 37°C pendant 3h. Les puits sont ensuite vidés, puis rincés avec du milieu DMEM non complémenté. 50 µL de DMEM sont finalement ajoutés à chaque puits avant l'observation au microscope à fluorescence.

BIBLIOGRAPHIE

Abd El-Aziz, T.M., Al Khoury, S., Jaquillard, L., Triquigneaux, M., Martinez, G., Bourgoin-Voillard, S., Sève, M., Arnoult, C., Beroud, R., and De Waard, M. (2018). Actiflagelin, a new sperm activator isolated from Walterinnesia aegyptia venom using phenotypic screening. J. Venom. Anim. Toxins Incl. Trop. Dis. *24*, 2.

Adamantidis, A.R., Zhang, F., Aravanis, A.M., Deisseroth, K., and De Lecea, L. (2007). Neural substrates of awakening probed with optogenetic control of hypocretin neurons. Nature *450*, 420–424.

Aguilar, M.B., Zugasti-Cruz, A., Falcón, A., Batista, C.V.F., Olivera, B.M., and Heimer De La Cotera, E.P. (2013). A novel arrangement of Cys residues in a paralytic peptide of Conus cancellatus (jr. syn.: Conus austini), a worm-hunting snail from the Gulf of Mexico. Peptides *41*, 38–44.

Akondi, K.B., Muttenthaler, M., Dutertre, S., Kaas, Q., Craik, D.J., Lewis, R.J., and Alewood, P.F. (2014). Discovery, Synthesis, and Structure–Activity Relationships of Conotoxins. Chem. Rev. *114*, 5815–5847.

Alon, T., and Friedman, J.M. (2006). Late-onset leanness in mice with targeted ablation of melanin concentrating hormone neurons. J. Neurosci. *26*, 389–397.

An, J.J., Rhee, Y., Kim, S.H., Kim, D.M., Han, D.-H., Hwang, J.H., Jin, Y.-J., Cha, B.S., Baik, J.-H., Lee, W.T., et al. (2007). Peripheral effect of alpha-melanocyte-stimulating hormone on fatty acid oxidation in skeletal muscle. J. Biol. Chem. *282*, 2862–2870.

Anand, B.K., and Brobeck, J.R. (1951). Localization of a “Feeding Center” in the Hypothalamus of the Rat. Exp. Biol. Med. *77*, 323–325.

Andermann, M.L., and Lowell, B.B. (2017). Toward a Wiring Diagram Understanding of Appetite Control. Neuron *95*, 757–778.

Anderson, L.J., Albrecht, E.D., and Garcia, J.M. (2017). Update on Management of Cancer-Related Cachexia. Curr. Oncol. Rep. *19*, 3.

Andrade, D. V, and Abe, A.S. (1999). Relationship of venom ontogeny and diet in Bothrops. Herpetologica *55*, 200–204.

Antil-Delbeke, S., Gaillard, C., Tamiya, T., Corringer, P.J., Changeux, J.P., Servent, D., and Ménez, A. (2000). Molecular determinants by which a long chain toxin from snake venom interacts with the neuronal α7-nicotinic acetylcholine receptor. J. Biol. Chem. *275*, 29594–29601.

Aoyagi, T., Terracina, K.P., Raza, A., Matsubara, H., and Takabe, K. (2015). Cancer cachexia, mechanism and treatment. World J. Gastrointest. Oncol. *7*, 17–29.

Aponte, Y., Atasoy, D., and Sternson, S.M. (2011). AGRP neurons are sufficient to orchestrate feeding behavior rapidly and without training. Nat. Neurosci. *14*, 351–355.

Argilés, J.M., Busquets, S., Stemmler, B., and López-Soriano, F.J. (2014). Cancer cachexia: understanding the molecular basis. Nat. Rev. Cancer *14*, 754–762.

Asai, M., Ramachandrappa, S., Joachim, M., Shen, Y., Zhang, R., Nuthalapati, N., Ramanathan, V., Strochlic, D.E., Ferket, P., Linhart, K., et al. (2013). Loss of function of the melanocortin 2 receptor accessory protein 2 is associated with mammalian obesity. Science *341*, 275–278.

Atasoy, D., Betley, J.N., Su, H.H., and Sternson, S.M. (2012). Deconstruction of a neural circuit for hunger. Nature *488*, 172–177.

Atasoy, D., Betley, J.N., Li, W.-P., Su, H.H., Sertel, S.M., Scheffer, L.K., Simpson, J.H., Fetter, R.D., and Sternson, S.M. (2014). A genetically specified connectomics approach applied to long-range feeding regulatory circuits. Nat. Neurosci. *17*, 1830–1839.

Attwood, T.K., and Findlay, J.B. (1994). Fingerprinting G-protein-coupled receptors. Protein Eng. *7*, 195–203.

Baillie, G.S., Sood, A., McPhee, I., Gall, I., Perry, S.J., Lefkowitz, R.J., and Houslay, M.D. (2003). beta-Arrestin-mediated PDE4 cAMP phosphodiesterase recruitment regulates beta-adrenoceptor switching from Gs to Gi. Proc. Natl. Acad. Sci. U. S. A. *100*, 940–945.

Balland, E., Dam, J., Langlet, F., Caron, E., Steculorum, S., Messina, A., Rasika, S., Falluel-Morel, A., Anouar, Y., Dehouck, B., et al. (2014). Hypothalamic tanycytes are an ERK-gated conduit for leptin into the brain. Cell Metab. *19*, 293–301.

Ballesteros, J.A., and Weinstein, H. (1995). [19] Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. In Methods in Neurosciences, pp. 366–428.

Balthasar, N., Dalgaard, L.T., Lee, C.E., Yu, J., Funahashi, H., Williams, T., Ferreira, M., Tang, V., McGovern, R.A., Kenny, C.D., et al. (2005). Divergence of melanocortin pathways in the control of food intake and energy expenditure. Cell *123*, 493–505.

Baltoumas, F.A., Theodoropoulou, M.C., and Hamodrakas, S.J. (2013). Interactions of the α-subunits of heterotrimeric G-proteins with GPCRs, effectors and RGS proteins: A critical review and analysis of interacting surfaces, conformational shifts, structural diversity and electrostatic potentials. J. Struct. Biol. *182*, 209–218.

Banerjee, A., Lee, A., Campbell, E., and MacKinnon, R. (2013). Structure of a pore-blocking toxin in complex with a eukaryotic voltage-dependent K+ channel. Elife *2*, 1–22.

Batista, C.V.F., Gómez-Lagunas, F., Rodríguez de la Vega, R.C., Hajdu, P., Panyi, G., Gáspár, R., and Possani, L.D. (2002). Two novel toxins from the Amazonian scorpion Tityus cambridgei that block Kv1.3 and Shaker B K(+)-channels with distinctly different affinities. Biochim. Biophys. Acta *1601*, 123–131.

Bayrhuber, M., Vijayan, V., Ferber, M., Graf, R., Korukottu, J., Imperial, J., Garrett, J.E., Olivera, B.M., Terlau, H., Zweckstetter, M., et al. (2005). Conkunitzin-S1 is the first member of a new Kunitz-type neurotoxin family: Structural and functional characterization. J. Biol. Chem. *280*, 23766–23770.

Bender, A., Bojanic, D., Davies, J.W., Crisman, T.J., Mikhailov, D., Scheiber, J., Jenkins, J.L., Deng, Z., Hill, W.A.G., Popov, M., et al. (2008). Which aspects of HTS are empirically correlated with downstream success? Curr. Opin. Drug Discov. Devel. *11*, 327–337.

Berthoud, H.-R. (2004). Anatomy and function of sensory hepatic nerves. Anat. Rec. A. Discov. Mol. Cell. Evol. Biol. *280*, 827–835.

Berthoud, H.-R., and Münzberg, H. (2011). The lateral hypothalamus as integrator of metabolic and environmental needs: from electrical self-stimulation to opto-genetics. Physiol. Behav. *104*, 29–39.

Berthoud, H.R., Blackshaw, L.A., Brookes, S.J.H., and Grundy, D. (2004). Neuroanatomy of extrinsic afferents supplying the gastrointestinal tract. Neurogastroenterol. Motil. *16 Suppl 1*, 28–33.

Betley, J.N., Xu, S., Cao, Z.F.H., Gong, R., Magnus, C.J., Yu, Y., and Sternson, S.M. (2015). Neurons for hunger and thirst transmit a negative-valence teaching signal. Nature *521*, 180–185.

Biggs, J.S., Watkins, M., Puillandre, N., Ownby, J.-P., Lopez-Vera, E., Christensen, S., Moreno, K.J., Bernaldez, J., Licea-Navarro, A., Corneli, P.S., et al. (2010). Evolution of Conus peptide toxins: analysis of Conus californicus Reeve, 1844. Mol. Phylogenet. Evol. *56*, 1–12.

Boey, D., Lin, S., Enriquez, R.F., Lee, N.J., Slack, K., Couzens, M., Baldock, P.A., Herzog, H., and Sainsbury, A. (2008). PYY transgenic mice are protected against diet-induced and genetic obesity. Neuropeptides *42*, 19–30.

Bohlen, C.J., Chesler, A.T., Sharif-Naeini, R., Medzihradszky, K.F., Zhou, S., King, D., Sánchez, E.E., Burlingame, A.L., Basbaum, A.I., and Julius, D. (2011). A heteromeric Texas coral snake toxin targets acid-sensing ion channels to produce pain. Nature *479*, 410–414.

Bontems, F., Gilquin, B., Roumestand, C., Ménez, A., and Toma, F. (1992). Analysis of side-chain organization on a refined model of charybdotoxin: structural and functional implications. Biochemistry *31*, 7756–7764.

Bourne, Y., Taylor, P., and Marchot, P. (1995). Acetylcholinesterase inhibition by fasciculin: Crystal structure of the complex. Cell *83*, 503–512.

Braun, T.P., and Marks, D.L. (2010). Pathophysiology and treatment of inflammatory anorexia in chronic disease. J. Cachexia. Sarcopenia Muscle *1*, 135–145.

Breit, A., Wolff, K., Kalwa, H., Jarry, H., Büch, T., and Gudermann, T. (2006). The natural inverse agonist agouti-related protein induces arrestin-mediated endocytosis of melanocortin-3 and -4 receptors. J. Biol. Chem. *281*, 37447–37456.

Breit, A., Büch, T.R.H., Boekhoff, I., Solinski, H.J., Damm, E., and Gudermann, T. (2011). Alternative G protein coupling and biased agonism: new insights into melanocortin-4 receptor signalling. Mol. Cell. Endocrinol. *331*, 232–240.

Broadwell, R.D., Balin, B.J., Salcman, M., and Kaplan, R.S. (1983). Brain-blood barrier? Yes and no. Proc. Natl. Acad. Sci. U. S. A. *80*, 7352–7356.

Büch, T.R.H., Heling, D., Damm, E., Gudermann, T., and Breit, A. (2009). Pertussis toxin-sensitive signaling of melanocortin-4 receptors in hypothalamic GT1-7 cells defines agouti-related protein as a biased agonist. J. Biol. Chem. *284*, 26411–26420.

Burmeister, M.A., Ayala, J.E., Smouse, H., Landivar-Rocha, A., Brown, J.D., Drucker, D.J., Stoffers, D.A., Sandoval, D.A., Seeley, R.J., and Ayala, J.E. (2017). The Hypothalamic Glucagon-Like Peptide 1 Receptor Is Sufficient but Not Necessary for the Regulation of Energy Balance and Glucose Homeostasis in Mice. Diabetes *66*, 372–384.

Cannon, B., and Nedergaard, J. (2004). Brown adipose tissue: function and physiological significance. Physiol. Rev. *84*, 277–359.

Carlier, E., Avdonin, V., Geib, S., Fajloun, Z., Kharrat, R., Rochat, H., Sabatier, J.M., Hoshi, T., and De Waard, M. (2000). Effect of maurotoxin, a four disulfide-bridged toxin from the chactoid scorpion Scorpio maurus, on Shaker K+ channels. J. Pept. Res. *55*, 419–427.

Carroll, L. (1871). Chapter Two – The Garden of Live Flowers. In Through the Looking-Glass, and What Alice Found There, (Macmillan), pp. 22–36.

Carter, M.E., Soden, M.E., Zweifel, L.S., and Palmiter, R.D. (2013). Genetic identification of a neural circuit that suppresses appetite. Nature *503*, 111–114.

Carter, M.E., Han, S., and Palmiter, R.D. (2015). Parabrachial calcitonin gene-related peptide neurons mediate conditioned taste aversion. J. Neurosci. *35*, 4582–4586.

Carvalho, A. de O., and Gomes, V.M. (2009). Plant defensins--prospects for the biological functions and biotechnological properties. Peptides *30*, 1007–1020.

Casewell, N.R., Wüster, W., Vonk, F.J., Harrison, R.A., and Fry, B.G. (2013). Complex cocktails: the evolutionary novelty of venoms. Trends Ecol. Evol. *28*, 219–229.

Catania, A., Gatti, S., Colombo, G., and Lipton, J.M. (2004). Targeting melanocortin receptors as a novel strategy to control inflammation. Pharmacol. Rev. *56*, 1–29.

Cawley, J., and Meyerhoefer, C. (2012). The medical care costs of obesity: an instrumental variables approach. J. Health Econ. *31*, 219–230.

Cerni, F.A., Pucca, M.B., Peigneur, S., Cremonez, C.M., Bordon, K.C.F., Tytgat, J., and Arantes, E.C. (2014). Electrophysiological characterization of Ts6 and Ts7, K+ channel toxins isolated through an improved Tityus serrulatus venom purification procedure. Toxins (Basel). *6*, 892–913.

Chai, B., Li, J.-Y., Zhang, W., Ammori, J.B., and Mulholland, M.W. (2007). Melanocortin-3 receptor activates MAP kinase via PI3 kinase. Regul. Pept. *139*, 115–121.

Chai, B.-X., Neubig, R.R., Millhauser, G.L., Thompson, D.A., Jackson, P.J., Barsh, G.S., Dickinson, C.J., Li, J.-Y., Lai, Y.-M., and Gantz, I. (2003). Inverse agonist activity of agouti and agouti-related protein. Peptides *24*, 603–609.

Chakraborty, A.K., Funasaka, Y., Slominski, A., Ermak, G., Hwang, J., Pawelek, J.M., and Ichihashi, M. (1996). Production and release of proopiomelanocortin (POMC) derived peptides by human melanocytes and keratinocytes in culture: regulation by ultraviolet B. Biochim. Biophys. Acta *1313*, 130–138.

Chaly, A.L., Srisai, D., Gardner, E.E., and Sebag, J.A. (2016). The Melanocortin Receptor Accessory Protein 2 promotes food intake through inhibition of the Prokineticin Receptor-1. Elife *5*.

Chan, L.F., Webb, T.R., Chung, T.-T., Meimaridou, E., Cooray, S.N., Guasti, L., Chapple, J.P., Egertová, M., Elphick, M.R., Cheetham, M.E., et al. (2009). MRAP and MRAP2 are bidirectional regulators of the melanocortin receptor family. Proc. Natl. Acad. Sci. U. S. A. *106*, 6146–6151.

Chandy, K.G., Wulff, H., Beeton, C., Pennington, M., Gutman, G.A., and Cahalan, M.D. (2004). K+ channels as targets for specific immunomodulation. Trends Pharmacol. Sci. *25*, 280–289.

Chao, P.-T., Yang, L., Aja, S., Moran, T.H., and Bi, S. (2011). Knockdown of NPY expression in the dorsomedial hypothalamus promotes development of brown adipocytes and prevents diet-induced obesity. Cell Metab. *13*, 573–583.

Chapman, K.L., and Findlay, J.B.C. (2013). The melanocortin 4 receptor: Oligomer formation, interaction sites and functional significance. Biochim. Biophys. Acta - Biomembr. *1828*, 535–542.

Chapman, K.L., Kinsella, G.K., Cox, A., Donnelly, D., Findlay, J.B.C., and Chapman, K.L. (2010). Interactions of the Melanocortin-4 Receptor with the Peptide Agonist NDP-MSH. J. Mol. Biol *401*, 433–450.

Chemelli, R.M., Willie, J.T., Sinton, C.M., Elmquist, J.K., Scammell, T., Lee, C., Richardson, J.A., Williams, S.C., Xiong, Y., Kisanuki, Y., et al. (1999). Narcolepsy in orexin knockout mice: molecular genetics of sleep regulation. Cell *98*, 437–451.

Chen, W., Kelly, M.A., Opitz-Araya, X., Thomas, R.E., Low, M.J., and Cone, R.D. (1997). Exocrine gland dysfunction in MC5-R-deficient mice: evidence for coordinated regulation of exocrine gland function by melanocortin peptides. Cell *91*, 789–798.

Chen, Y., Lin, Y.-C., Kuo, T.-W., and Knight, Z.A. (2015). Sensory detection of food rapidly modulates arcuate feeding circuits. Cell *160*, 829–841.

Chen, Y., Lin, Y.-C., Zimmerman, C.A., Essner, R.A., and Knight, Z.A. (2016). Hunger neurons drive feeding through a sustained, positive reinforcement signal. Elife *5*, 1–21.

Chen, Y.-C.I., Cypess, A.M., Chen, Y.-C., Palmer, M., Kolodny, G., Kahn, C.R., and Kwong, K.K. (2013). Measurement of human brown adipose tissue volume and activity using anatomic MR imaging and functional MR imaging. J. Nucl. Med. *54*, 1584–1587.

Chen, Z., Hu, Y., Han, S., Yin, S., He, Y., Wu, Y., Cao, Z., and Li, W. ImKTx1, a new Kv1.3 channel blocker with a unique primary structure. J. Biochem. Mol. Toxicol. *25*, 244–251.

Cheng, Y., and Prusoff, W.H. (1973). Relationship between the inhibition constant (K1) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. Biochem. Pharmacol. *22*, 3099–3108.

Cheung, W.W., and Mak, R.H. (2012). Melanocortin antagonism ameliorates muscle wasting and inflammation in chronic kidney disease. Am. J. Physiol. Renal Physiol. *303*, F1315-24.

Chiappinelli, V.A., and Wolf, K.M. (1989). Kappa-neurotoxins: heterodimer formation between different neuronal nicotinic receptor antagonists. Biochemistry *28*, 8543–8547.

Chida, D., Nakagawa, S., Nagai, S., Sagara, H., Katsumata, H., Imaki, T., Suzuki, H., Mitani, F., Ogishima, T., Shimizu, C., et al. (2007). Melanocortin 2 receptor is required for adrenal gland development, steroidogenesis, and neonatal gluconeogenesis. Proc. Natl. Acad. Sci. U. S. A. *104*, 18205–18210.

Choi, S.-J., Parent, R., Guillaume, C., Deregnaucourt, C., Delarbre, C., Ojcius, D.M., Montagne, J.-J., Célérier, M.-L., Phelipot, A., Amiche, M., et al. (2004). Isolation and characterization of Psalmopeotoxin I and II: two novel antimalarial peptides from the venom of the tarantula Psalmopoeus cambridgei. FEBS Lett. *572*, 109–117.

Chudzinski-Tavassi, A.M., De-Sá-Júnior, P.L., Simons, S.M., Maria, D.A., de Souza Ventura, J., Batista, I. de F.C., Faria, F., Durães, E., Reis, E.M., and Demasi, M. (2010). A new tick Kunitz type inhibitor, Amblyomin-X, induces tumor cell death by modulating genes related to the cell cycle and targeting the ubiquitin-proteasome system. Toxicon *56*, 1145–1154.

Church, J.E., and Hodgson, W.C. (2002). The pharmacological activity of fish venoms. Toxicon *40*, 1083–1093.

Ciofi, P., Garret, M., Lapirot, O., Lafon, P., Loyens, A., Prévot, V., and Levine, J.E. (2009). Brain-endocrine interactions: a microvascular route in the mediobasal hypothalamus. Endocrinology *150*, 5509–5519.

Ciolek, J., Reinfrank, H., Quinton, L., Viengchareun, S., Stura, E.A., Vera, L., Sigismeau, S., Mouillac, B., Orcel, H., Peigneur, S., et al. (2017). Green mamba peptide targets type-2 vasopressin receptor against polycystic kidney disease. Proc. Natl. Acad. Sci. *114*, 7154–7159.

Clément, K., Biebermann, H., Farooqi, I.S., Van Der Ploeg, L., Wolters, B., Poitou, C., Puder, L., Fiedorek, F., Gottesdiener, K., Kleinau, G., et al. (2018). MC4R agonism promotes durable weight loss in patients with leptin receptor deficiency. Nat. Med. *24*, 551–555.

Collden, G., Balland, E., Parkash, J., Caron, E., Langlet, F., Prevot, V., and Bouret, S.G. (2015). Neonatal overnutrition causes early alterations in the central response to peripheral ghrelin. Mol. Metab. *4*, 15–24.

Collet, T.-H., Dubern, B., Mokrosinski, J., Connors, H., Keogh, J.M., Mendes De Oliveira, E., Henning, E., Poitou-Bernert, C., Oppert, J.-M., Tounian, P., et al. (2017). Evaluation of a melanocortin-4 receptor (MC4R) agonist (Setmelanotide) in MC4R deficiency. Mol. Metab. *6*, 1321–1329.

Comeras, L.B., Herzog, H., and Tasan, R.O. (2019). Neuropeptides at the crossroad of fear and hunger: a special focus on neuropeptide Y. Ann. N. Y. Acad. Sci. nyas.14179.

Corander, M.P., Rimmington, D., Challis, B.G., O'Rahilly, S., and Coll, A.P. (2011). Loss of agouti-related peptide does not significantly impact the phenotype of murine POMC deficiency. Endocrinology *152*, 1819–1828.

Cornet, B., Bonmatin, J.M., Hetru, C., Hoffmann, J.A., Ptak, M., and Vovelle, F. (1995). Refined three-dimensional solution structure of insect defensin A. Structure *3*, 435–448.

Corp, E.S., Woods, S.C., Porte, D., Dorsa, D.M., Figlewicz, D.P., and Baskin, D.G. (1986). Localization of 125I-insulin binding sites in the rat hypothalamus by quantitative autoradiography. Neurosci. Lett. *70*, 17–22.

Cowley, M.A., Pronchuk, N., Fan, W., Dinulescu, D.M., Colmers, W.F., and Cone, R.D. (1999). Integration of npy, agrp, and melanocortin signals in the hypothalamic paraventricular nucleus: Evidence of a cellular basis for the adipostat. Neuron *24*, 155–163.

Cowley, M.A., Smart, J.L., Rubinstein, M., Cerdán, M.G., Diano, S., Horvath, T.L., Cone, R.D., and Low, M.J. (2001). Leptin activates anorexigenic POMC neurons through a neural network in the arcuate nucleus. Nature *411*, 480–484.

Cox, J.E. (1998). Cholecystokinin satiety involves CCKA receptors perfused by the superior pancreaticoduodenal artery. Am. J. Physiol. *274*, R1390-6.

Cox, J.J., Reimann, F., Nicholas, A.K., Thornton, G., Roberts, E., Springell, K., Karbani, G., Jafri, H., Mannan, J., Raashid, Y., et al. (2006). An SCN9A channelopathy causes congenital inability to experience pain. Nature *444*, 894–898.

Craik, D.J., and Adams, D.J. (2007). Chemical modification of conotoxins to improve stability and activity. ACS Chem. Biol. *2*, 457–468.

Creemers, J.W.M., Pritchard, L.E., Gyte, A., Le Rouzic, P., Meulemans, S., Wardlaw, S.L., Zhu, X., Steiner, D.F., Davies, N., Armstrong, D., et al. (2006). Agouti-related protein is posttranslationally cleaved by proprotein convertase 1 to generate agouti-related protein (AGRP)83-132: interaction between AGRP83-132 and melanocortin receptors cannot be influenced by syndecan-3. Endocrinology *147*, 1621–1631.

Cushman, D.W., Pluscec, J., Williams, N.J., Weaver, E.R., Sabo, E.F., Kocy, O., Cheung, H.S., and Ondetti, M.A. (1973). Inhibition of angiotensin-coverting enzyme by analogs of peptides from Bothrops jararaca venom. Experientia *29*, 1032–1035.

D'Orazio, J.A., Nobuhisa, T., Cui, R., Arya, M., Spry, M., Wakamatsu, K., Igras, V., Kunisada, T., Granter, S.R., Nishimura, E.K., et al. (2006). Topical drug rescue strategy and skin protection based on the role of Mc1r in UV-induced tanning. Nature *443*, 340–344.

Daneman, R., Zhou, L., Kebede, A.A., and Barres, B.A. (2010). Pericytes are required for blood-brain barrier integrity during embryogenesis. Nature *468*, 562–566.

Darling, R.A., Zhao, H., Kinch, D., Li, A.-J., Simasko, S.M., and Ritter, S. (2014). Mercaptoacetate and fatty acids exert direct and antagonistic effects on nodose neurons via GPR40 fatty acid receptors. Am. J. Physiol. Regul. Integr. Comp. Physiol. *307*, R35-43.

Dauplais, M., Lecoq, A., Song, J., Cotton, J., Jamin, N., Gilquin, B., Roumestand, C., Vita, C., de Medeiros, C.L., Rowan, E.G., et al. (1997). On the convergent evolution of animal toxins. Conservation of a diad of functional residues in potassium channel-blocking toxins with unrelated structures. J. Biol. Chem. *272*, 4302–4309.

Defossez, G., Le Guyader-Peyrou, S., Uhry, Z., Grosclaude, P., Remontet, L., and Colonna, M. (2019). Estimations nationales de l'incidence et de la mortalité par cancer en France métropolitaine entre 1990 et 2018 (Saint-Maurice : Santé publique France).

Degueldre, M., Verdenaud, M., Legarda, G., Minambres, R., Zuniga, S., Leblanc, M., Gilles, N., Ducancel, F., De Pauw, E., and Quinton, L. (2017). 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 *130*, 116–125.

Delgado, J.M.R., and Anand, B.K. (1953). Increase of food intake induced by electrical stimulation of the lateral hypothalamus. Am. J. Physiol. *172*, 162–168.

Denis, R.G.P., Joly-Amado, A., Webber, E., Langlet, F., Schaeffer, M., Padilla, S.L., Cansell, C., Dehouck, B., Castel, J., Delbès, A.-S., et al. (2015). Palatability Can Drive Feeding Independent of AgRP Neurons. Cell Metab. *22*, 646–657.

Deshane, J., Garner, C.C., and Sontheimer, H. (2003). Chlorotoxin inhibits glioma cell invasion via matrix metalloproteinase-2. J. Biol. Chem. *278*, 4135–4144.

Diego-García, E., Batista, C.V.F., García-Gómez, B.I., Lucas, S., Candido, D.M., Gómez-Lagunas, F., and Possani, L.D. (2005). The Brazilian scorpion Tityus costatus Karsch: genes, peptides and function. Toxicon *45*, 273–283.

Diochot, S., Baron, A., Salinas, M., Douguet, D., Scarzello, S., Dabert-Gay, A.-S., Debayle, D., Friend, V., Alloui, A., Lazdunski, M., et al. (2012). Black mamba venom peptides target acid-sensing ion channels to abolish pain. Nature *490*, 552–555.

Dobs, A.S., Boccia, R. V, Croot, C.C., Gabrail, N.Y., Dalton, J.T., Hancock, M.L., Johnston, M.A., and Steiner, M.S. (2013). Effects of enobosarm on muscle wasting and physical function in patients with cancer: a double-blind, randomised controlled phase 2 trial. Lancet. Oncol. *14*, 335–345.

Duc, N.M., Kim, H.R., and Chung, K.Y. (2015). Structural mechanism of G protein activation by G protein-coupled receptor. Eur. J. Pharmacol. *763*, 214–222.

Dufton, M.J., and Hider, R.C. (1988). Structure and pharmacology of elapid cytotoxins. Pharmacol. Ther. *36*, 1–40.

Duhoo, Y., Sequeira, A.F., Saez, N.J., Turchetto, J., Ramond, L., Peysson, F., Brás, J.L.A., Gilles, N., Darbon, H., Fontes, C.M.G.A., et al. (2019). High-Throughput Production of Oxidized Animal Toxins in Escherichia coli. Methods Mol. Biol. *2025*, 165–190.

Dupré, D.J., Robitaille, M., Rebois, R.V., and Hébert, T.E. (2009). The role of Gbetagamma subunits in the organization, assembly, and function of GPCR signaling complexes. Annu. Rev. Pharmacol. Toxicol. *49*, 31–56.

Dutertre, S., Jin, A., Kaas, Q., Jones, A., Alewood, P.F., and Lewis, R.J. (2013). Deep venomics reveals the mechanism for expanded peptide diversity in cone snail venom. Mol. Cell. Proteomics *12*, 312–329.

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. (2014). Evolution of separate predation- and defence-evoked venoms in carnivorous cone snails. Nat. Commun. *5*, 3521.

Eichel, K., Jullié, D., and von Zastrow, M. (2016). β-Arrestin drives MAP kinase signalling from clathrin-coated structures after GPCR dissociation. Nat. Cell Biol. *18*, 303–310.

Elias, C.F., Saper, C.B., Maratos-Flier, E., Tritos, N.A., Lee, C., Kelly, J., Tatro, J.B., Hoffman, G.E., Ollmann, M.M., Barsh, G.S., et al. (1998). Chemically defined projections linking the mediobasal hypothalamus and the lateral hypothalamic area. J. Comp. Neurol. *402*, 442–459.

Eng, J., Kleinman, W.A., Singh, L., Singh, G., and Raufman, J.P. (1992). 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. *267*, 7402–7405.

Enriori, P.J., Sinnayah, P., Simonds, S.E., Garcia Rudaz, C., and Cowley, M.A. (2011). Leptin Action in the Dorsomedial Hypothalamus Increases Sympathetic Tone to Brown Adipose Tissue in Spite of Systemic Leptin Resistance. J. Neurosci. *31*, 12189–12197.

Enriori, P.J., Chen, W., Garcia-Rudaz, M.C., Grayson, B.E., Evans, A.E., Comstock, S.M., Gebhardt, U., Müller, H.L., Reinehr, T., Henry, B.A., et al. (2016). α-Melanocyte stimulating hormone promotes muscle glucose uptake via melanocortin 5 receptors. Mol. Metab. *5*, 807–822.

Ersoy, B.A., Pardo, L., Zhang, S., Thompson, D.A., Millhauser, G., Govaerts, C., and Vaisse, C. (2012). Mechanism of N-terminal modulation of activity at the melanocortin-4 receptor GPCR Suppl data. Nat. Chem. Biol.

Escoubas, P., De Weille, J.R., Lecoq, A., Diochot, S., Waldmann, R., Champigny, G., Moinier, D., Ménez, A., and Lazdunski, M. (2000). Isolation of a tarantula toxin specific for a class of proton-gated Na+ channels. J. Biol. Chem. *275*, 25116–25121.

Evans, W.J., Morley, J.E., Argilés, J., Bales, C., Baracos, V., Guttridge, D., Jatoi, A., Kalantar-Zadeh, K., Lochs, H., Mantovani, G., et al. (2008). Cachexia: a new definition. Clin. Nutr. *27*, 793–799.

Eysselein, V.E., Eberlein, G.A., Schaeffer, M., Grandt, D., Goebell, H., Niebel, W., Rosenquist, G.L., Meyer, H.E., and Reeve, J.R. (1990). Characterization of the major form of cholecystokinin in human intestine: CCK-58. Am. J. Physiol. *258*, G253-60.

Fani, L., Bak, S., Delhanty, P., van Rossum, E.F.C., and van den Akker, E.L.T. (2014). The melanocortin-4 receptor as target for obesity treatment: a systematic review of emerging pharmacological therapeutic options. Int. J. Obes. (Lond). *38*, 163–169.

Farkas, J., von Haehling, S., Kalantar-Zadeh, K., Morley, J.E., Anker, S.D., and Lainscak, M. (2013). Cachexia as a major public health problem: frequent, costly, and deadly. J. Cachexia. Sarcopenia Muscle *4*, 173–178.

Farooqi, I.S., Keogh, J.M., Yeo, G.S.H., Lank, E.J., Cheetham, T., and O'Rahilly, S. (2003). Clinical Spectrum of Obesity and Mutations in the Melanocortin 4 Receptor Gene. N. Engl. J. Med. *348*, 1085–1095.

Favreau, P., Gilles, N., Lamthanh, H., Bournaud, R., Shimahara, T., Bouet, F., Laboute, P., Letourneux, Y., Ménez, A., Molgó, J., et al. (2001). A new omega-conotoxin that targets N-type voltage-sensitive calcium channels with unusual specificity. Biochemistry *40*, 14567–14575.

Fearon, K., Strasser, F., Anker, S.D., Bosaeus, I., Bruera, E., Fainsinger, R.L., Jatoi, A., Loprinzi, C., MacDonald, N., Mantovani, G., et al. (2011). Definition and classification of cancer cachexia: an international consensus. Lancet. Oncol. *12*, 489–495.

Fearon, K., Argiles, J., Baracos, V., Bernabei, R., Coats, A., Crawford, J., Deutz, N., Doehner, W., Evans, W., Ferrucci, L., et al. (2015). Request for regulatory guidance for cancer cachexia intervention trials. J. Cachexia. Sarcopenia Muscle *6*, 272–274.

Feng, B.Y., and Shoichet, B.K. (2006). Synergy and antagonism of promiscuous inhibition in multiple-compound mixtures. J. Med. Chem. *49*, 2151–2154.

Ferlay, J., Colombet, M., Soerjomataram, I., Mathers, C., Parkin, D.M., Piñeros, M., Znaor, A., and Bray, F. (2019). Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. Int. J. Cancer *144*, 1941–1953.

FERREIRA, S.H. (1965). A BRADYKININ-POTENTIATING FACTOR (BPF) PRESENT IN THE VENOM OF BOTHROPS JARARCA. Br. J. Pharmacol. Chemother. *24*, 163–169.

Fertleman, C.R., Baker, M.D., Parker, K.A., Moffatt, S., Elmslie, F. V., Abrahamsen, B., Ostman, J., Klugbauer, N., Wood, J.N., Gardiner, R.M., et al. (2006). SCN9A Mutations in Paroxysmal Extreme Pain Disorder: Allelic Variants Underlie Distinct Channel Defects and Phenotypes. Neuron *52*, 767–774.

Fredriksson, R. (2005). The Repertoire of G-Protein-Coupled Receptors in Fully Sequenced Genomes. Mol. Pharmacol. *67*, 1414–1425.

Fredriksson, R., Lagerström, M.C., Lundin, L.-G., and Schiöth, H.B. (2003). The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Mol. Pharmacol. *63*, 1256–1272.

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. (2009). The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. Annu. Rev. Genomics Hum. Genet. *10*, 483–511.

Fry, B.G., Koludarov, I., Jackson, T.N.W., Holford, M., Terrat, Y., Casewell, N.R., Undheim, E.A.B., Vetter, I., Ali, S.A., Low, D.H.W., et al. (2015). CHAPTER 1. Seeing the Woods for the Trees: Understanding Venom Evolution as a Guide for Biodiscovery. In Venoms to Drugs : Venom as Source for the Development of Human Therapeutics, Glenn F. King, ed. (RSC Drug Discovery), pp. 1–36.

Furman, B.L. (2012). The development of Byetta (exenatide) from the venom of the Gila monster as an anti-diabetic agent. Toxicon *59*, 464–471.

Gantz, I., and Fong, T.M. (2003). The melanocortin system. Am. J. Physiol. Endocrinol. Metab. *284*, E468-74.

Gao, B., Harvey, P.J., Craik, D.J., Ronjat, M., De Waard, M., and Zhu, S. (2013). Functional evolution of scorpion venom peptides with an inhibitor cystine knot fold. Biosci. Rep. *33*, 513–527.

Garcia, J.M., Boccia, R. V., Graham, C.D., Yan, Y., Duus, E.M., Allen, S., and Friend, J. (2015). Anamorelin for patients with cancer cachexia: an integrated analysis of two phase 2, randomised, placebo-controlled, double-blind trials. Lancet. Oncol. *16*, 108–116.

Garcia, M.L., Garcia-Calvo, M., Hidalgo, P., Lee, A., and MacKinnon, R. (1994). Purification and characterization of three inhibitors of voltage-dependent K+ channels from Leiurus quinquestriatus var. hebraeus venom. Biochemistry *33*, 6834–6839.

Garfield, A.S., Li, C., Madara, J.C., Shah, B.P., Webber, E., Steger, J.S., Campbell, J.N., Gavrilova, O., Lee, C.E., Olson, D.P., et al. (2015). A neural basis for melanocortin-4 receptor-regulated appetite. Nat. Neurosci. *18*, 863–871.

Gasparini, S., Danse, J.M., Lecoq, A., Pinkasfeld, S., Zinn-Justin, S., Young, L.C., de Medeiros, C.C., Rowan, E.G., Harvey, A.L., and Ménez, A. (1998). Delineation of the functional site of alpha-dendrotoxin. The functional topographies of dendrotoxins are different but share a conserved core with those of other Kv1 potassium channel-blocking toxins. J. Biol. Chem. *273*, 25393–25403.

Gautron, L., Elmquist, J.K., and Williams, K.W. (2015). Neural control of energy balance: translating circuits to therapies. Cell *161*, 133–145.

Gehring, M., and Henikoff, S. (2008). DNA Methylation and Demethylation in Arabidopsis. Arab. B. *6*, e0102.

Ghamari-Langroudi, M., Digby, G.J., Sebag, J.A., Millhauser, G.L., Palomino, R., Matthews, R., Gillyard, T., Panaro, B.L., Tough, I.R., Cox, H.M., et al. (2015). G-protein-independent coupling of MC4R to Kir7.1 in hypothalamic neurons. Nature *520*, 94–98.

Ghamari-Langroudi, M., Cakir, I., Lippert, R.N., Sweeney, P., Litt, M.J., Ellacott, K.L.J., and Cone, R.D. (2018). Regulation of energy rheostasis by the melanocortin-3 receptor. Sci. Adv. *4*, eaat0866.

Gladkikh, I., Monastyrnaya, M., Zelepuga, E., Sintsova, O., Tabakmakher, V., Gnedenko, O., Ivanov, A., Hua, K.F., Kozlovskaya, E., and Jacobson, P.B. (2015). New kunitz-type HCRG polypeptides from the sea anemone heteractis crispa. Mar. Drugs *13*, 6038–6063.

González, J.A., Iordanidou, P., Strom, M., Adamantidis, A., and Burdakov, D. (2016). Awake dynamics and brain-wide direct inputs of hypothalamic MCH and orexin networks. Nat. Commun. *7*, 11395.

Grant, G.A., and Chiappinelli, V.A. (1985). kappa-Bungarotoxin: complete amino acid sequence of a neuronal nicotinic receptor probe. Biochemistry *24*, 1532–1537.

Greenfield, N.J. (2006). Using circular dichroism spectra to estimate protein secondary structure. Nat. Protoc. *1*, 2876–2890.

Grill, H.J., Schwartz, M.W., Kaplan, J.M., Foxhall, J.S., Breininger, J., and Baskin, D.G. (2002). Evidence that the caudal brainstem is a target for the inhibitory effect of leptin on food intake. Endocrinology *143*, 239–246.

Gründer, S., and Chen, X. (2010). Structure, function, and pharmacology of acid-sensing ion channels (ASICs): Focus on ASIC1a. Int. J. Physiol. Pathophysiol. Pharmacol. *2*, 73–94.

Gutierrez, J.A., Solenberg, P.J., Perkins, D.R., Willency, J.A., Knierman, M.D., Jin, Z., Witcher, D.R., Luo, S., Onyia, J.E., and Hale, J.E. (2008). Ghrelin octanoylation mediated by an orphan lipid transferase. Proc. Natl. Acad. Sci. U. S. A. *105*, 6320–6325.

von Haehling, S., Anker, M.S., and Anker, S.D. (2016). Prevalence and clinical impact of cachexia in chronic illness in Europe, USA, and Japan: facts and numbers update 2016. J. Cachexia. Sarcopenia Muscle *7*, 507–509.

Hagey, L.R., Fry, B.G., and Fitch-Snyder, H. (2007). Talking Defensively, a Dual Use for the Brachial Gland Exudate of Slow and Pygmy Lorises. In Primate Anti-Predator Strategies, (Boston, MA: Springer US), pp. 253–272.

Halatchev, I.G., and Cone, R.D. (2005). Peripheral administration of PYY(3-36) produces conditioned taste aversion in mice. Cell Metab. *1*, 159–168.

Hanlon, C.D., and Andrew, D.J. (2015). Outside-in signaling - a brief review of GPCR signaling with a focus on the Drosophila GPCR family. J. Cell Sci. *128*, 3533–3542.

Harvey, A.L., Rowan, E.G., Vatanpour, H., Engström, A., Westerlund, B., and Karlsson, E. (1997). Changes to biological activity following acetylation of dendrotoxin I from Dendroaspis polylepis (black mamba). Toxicon *35*, 1263–1273.

Haskell-Luevano, C., and Monck, E.K. (2001). Agouti-related protein functions as an inverse agonist at a constitutively active brain melanocortin-4 receptor. Regul. Pept.

Haslach, E.M., Schaub, J.W., and Haskell-Luevano, C. (2009). Beta-turn secondary structure and melanocortin ligands. Bioorg. Med. Chem. *17*, 952–958.

Hassani, O.K., Lee, M.G., and Jones, B.E. (2009). Melanin-concentrating hormone neurons discharge in a reciprocal manner to orexin neurons across the sleep-wake cycle. Proc. Natl. Acad. Sci. U. S. A. *106*, 2418–2422.

Hauser, A.S., Attwood, M.M., Rask-Andersen, M., Schiöth, H.B., and Gloriam, D.E. (2017). Trends in GPCR drug discovery: new agents, targets and indications. Nat. Rev. Drug Discov. *16*, 829–842.

Havrankova, J., Roth, J., and Brownstein, M. (1978). Insulin receptors are widely distributed in the central nervous system of the rat. Nature *272*, 827–829.

Hawes, B.E., Kil, E., Green, B., O'Neill, K., Fried, S., and Graziano, M.P. (2000). The melanin-concentrating hormone receptor couples to multiple G proteins to activate diverse intracellular signaling pathways. Endocrinology *141*, 4524–4532.

Hayes, M.R., Leichner, T.M., Zhao, S., Lee, G.S., Chowansky, A., Zimmer, D., De Jonghe, B.C., Kanoski, S.E., Grill, H.J., and Bence, K.K. (2016). Intracellular Signals Mediating the Food Intake-Suppressive Effects of Hindbrain Glucagon-like Peptide-1 Receptor Activation. Cell Metab. *23*, 745.

Herisson, F.M., Waas, J.R., Fredriksson, R., Schiöth, H.B., Levine, A.S., and Olszewski, P.K. (2016). Oxytocin Acting in the Nucleus Accumbens Core Decreases Food Intake. J. Neuroendocrinol. *28*.

Herrington, J., Zhou, Y.-P., Bugianesi, R.M., Dulski, P.M., Feng, Y., Warren, V.A., Smith, M.M., Kohler, M.G., Garsky, V.M., Sanchez, M., et al. (2006). Blockers of the delayed-rectifier potassium current in pancreatic beta-cells enhance glucose-dependent insulin secretion. Diabetes *55*, 1034–1042.

Hochstenbach-Waelen, A., Veldhorst, M.A.B., Nieuwenhuizen, A.G., Westerterp-Plantenga, M.S., and Westerterp, K.R. (2009). Comparison of 2 diets with either 25% or 10% of energy as casein on energy expenditure, substrate balance, and appetite profile. Am. J. Clin. Nutr. *89*, 831–838.

Hogan, K., Peluso, S., Gould, S., Parsons, I., Ryan, D., Wu, L., and Visiers, I. (2006). Mapping the binding site of melanocortin 4 receptor agonists: A hydrophobic pocket formed by I3.28(125), I3.32(129), and I7.42(291) is critical for receptor activation. J. Med. Chem. *49*, 911–922.

Hooper, D., Lirazan, M.B., Schoenfeld, R., Cook, B., Cruz, L.J., Olivera, B.M., and Bandyopadhyay, P. (2002). Post-translational modification: A two-dimensional strategy for molecular diversity of Conus peptides. In Peptides for the New Millennium, (Dordrecht: Kluwer Academic Publishers), pp. 727–729.

Hruby, V.J., Wilkes, B.C., Hadley, M.E., Al-Obeidi, F., Sawyer, T.K., Staples, D.J., de Vaux, A.E., Dym, O., Castrucci, A.M., and Hintz, M.F. (1987). alpha-Melanotropin: the minimal active sequence in the frog skin bioassay. J. Med. Chem. *30*, 2126–2130.

Huys, I., Dyason, K., Waelkens, E., Verdonck, F., van Zyl, J., du Plessis, J., Müller, G.J., van der Walt, J., Clynen, E., Schoofs, L., et al. (2002). Purification, characterization and biosynthesis of parabutoxin 3, a component of Parabuthus transvaalicus venom. Eur. J. Biochem. *269*, 1854–1865.

Huys, I., Olamendi-Portugal, T., Garcia-Gómez, B.I., Vandenberghe, I., Van Beeumen, J., Dyason, K., Clynen, E., Zhu, S., van der Walt, J., Possani, L.D., et al. (2004). A subfamily of acidic alpha-K(+) toxins. J. Biol. Chem. *279*, 2781–2789.

Imperial, J.S., Watkins, M., Chen, P., Hillyard, D.R., Cruz, L.J., and Olivera, B.M. (2003). The augertoxins: biochemical characterization of venom components from the toxoglossate gastropod Terebra subulata. Toxicon *42*, 391–398.

Inceoglu, B., Lango, J., Jing, J., Chen, L., Doymaz, F., Pessah, I.N., and Hammock, B.D. (2003). One scorpion, two venoms: prevenom of Parabuthus transvaalicus acts as an alternative type of venom with distinct mechanism of action. Proc. Natl. Acad. Sci. U. S. A. *100*, 922–927.

Jackson, D.S., Ramachandrappa, S., Clark, A.J., and Chan, L.F. (2015). Melanocortin receptor accessory proteins in adrenal disease and obesity. Front. Neurosci. *10*.

Jacob, R.B., and McDougal, O.M. (2010). The M-superfamily of conotoxins: A review. Cell. Mol. Life Sci. *67*, 17–27.

Jacobson, K.A. (2015). New paradigms in GPCR drug delivery. Biochem Pharmacol. *98*, 541–555.

Jared, C., Mailho-Fontana, P.L., Antoniazzi, M.M., Mendes, V.A., Barbaro, K.C., Rodrigues, M.T., and Brodie, E.D. (2015). Venomous Frogs Use Heads as Weapons. Curr. Biol. *25*, 2166–2170.

Jean-Charles, P.Y., Kaur, S., and Shenoy, S.K. (2017). G Protein-Coupled Receptor Signaling Through β-Arrestin-Dependent Mechanisms.

Jego, S., Reed, S.J., Glasgow, S.D., Friedman, J., Adamantidis, A.R., Herrera, C.G., Burdakov, D., Boyce, R., and Ekstrand, M. (2013). Optogenetic identification of a rapid eye movement sleep modulatory circuit in the hypothalamus. Nat. Neurosci. *16*, 1637–1643.

Jelinek, G.A. (2001). Australian Animal Toxins: The Creatures, Their Toxins and Care of the Poisoned Patient. Emerg. Med. Australas. *13*, 473–474.

Jennings, J.H., Rizzi, G., Stamatakis, A.M., Ung, R.L., and Stuber, G.D. (2013). The inhibitory circuit architecture of the lateral hypothalamus orchestrates feeding. Science *341*, 1517–1521.

Jennings, J.H., Ung, R.L., Resendez, S.L., Stamatakis, A.M., Taylor, J.G., Huang, J., Veleta, K., Kantak, P.A., Aita, M., Shilling-Scrivo, K., et al. (2015). Visualizing hypothalamic network dynamics for appetitive and consummatory behaviors. Cell *160*, 516–527.

Joppa, M.A., Gogas, K.R., Foster, A.C., and Markison, S. (2007). Central infusion of the melanocortin receptor antagonist agouti-related peptide (AgRP(83-132)) prevents cachexia-related symptoms induced by radiation and colon-26 tumors in mice. Peptides *28*, 636–642.

Kadekaro, A.L., Leachman, S., Kavanagh, R.J., Swope, V., Cassidy, P., Supp, D., Sartor, M., Schwemberger, S., Babcock, G., Wakamatsu, K., et al. (2010). Melanocortin 1 receptor genotype: an important determinant of the damage response of melanocytes to ultraviolet radiation. FASEB J. *24*, 3850–3860.

Kamegai, J., Tamura, H., Shimizu, T., Ishii, S., Sugihara, H., and Wakabayashi, I. (2000). Central effect of ghrelin, an endogenous growth hormone secretagogue, on hypothalamic peptide gene expression. Endocrinology *141*, 4797–4800.

Katritch, V., Cherezov, V., and Stevens, R.C. (2013). Structure-function of the G protein-coupled receptor superfamily. Annu. Rev. Pharmacol. Toxicol. *53*, 531–556.

Khan, S.M., Sleno, R., Gora, S., Zylbergold, P., Laverdure, J.-P., Labbé, J.-C., Miller, G.J., and Hébert, T.E. (2013). The expanding roles of Gβγ subunits in G protein-coupled receptor signaling and drug action. Pharmacol. Rev. *65*, 545–577.

Kim, K.-S., Seeley, R.J., and Sandoval, D.A. (2018). Signalling from the periphery to the brain that regulates energy homeostasis. Nat. Rev. Neurosci. *19*, 185–196.

King, G.F., and Hardy, M.C. (2012). Spider-Venom Peptides: Structure, Pharmacology, and Potential for Control of Insect Pests. Annu. Rev. Entomol. *58*, 475–496.

King, G.F., Gentz, M.C., Escoubas, P., and Nicholson, G.M. (2008). A rational nomenclature for naming peptide toxins from spiders and other venomous animals. Toxicon *52*, 264–276.

Kini, R.M., and Doley, R. (2010). Structure, function and evolution of three-finger toxins: mini proteins with multiple targets. Toxicon *56*, 855–867.

Klein Herenbrink, C., Sykes, D.A., Donthamsetti, P., Canals, M., Coudrat, T., Shonberg, J., Scammells, P.J., Capuano, B., Sexton, P.M., Charlton, S.J., et al. (2016). The role of kinetic context in apparent biased agonism at GPCRs. Nat. Commun. *7*, 10842.

Klingenberg, M. (1999). Uncoupling protein--a useful energy dissipator. J. Bioenerg. Biomembr. *31*, 419–430.

Klint, J.K., Senff, S., Rupasinghe, D.B., Er, S.Y., Herzig, V., Nicholson, G.M., and King, G.F. (2012). Spider-venom peptides that target voltage-gated sodium channels: pharmacological tools and potential therapeutic leads. Toxicon *60*, 478–491.

Kobayashi, N., Nakagawa, A., Muramatsu, T., Yamashina, Y., Shirai, T., Hashimoto, M.W., Ishigaki, Y., Ohnishi, T., and Mori, T. (1998). Supranuclear melanin caps reduce ultraviolet induced DNA photoproducts in human epidermis. J. Invest. Dermatol. *110*, 806–810.

Koch, M., Varela, L., Kim, J.G., Kim, J.D., Hernández-Nuño, F., Simonds, S.E., Castorena, C.M., Vianna, C.R., Elmquist, J.K., Morozov, Y.M., et al. (2015). Hypothalamic POMC neurons promote cannabinoid-induced feeding. Nature *519*, 45–50.

Koehbach, J. (2017). Structure-Activity Relationships of Insect Defensins. Front. Chem. *5*, 45.

Koh, C.Y., and Kini, R.M. (2012). From snake venom toxins to therapeutics--cardiovascular examples. Toxicon *59*, 497–506.

Kopanchuk, S., Veiksina, S., Mutulis, F., Mutule, I., Yahorava, S., Mandrika, I., Petrovska, R., Rinken, A., and Wikberg, J.E.S. (2006). Kinetic evidence for tandemly arranged ligand binding sites in melanocortin 4 receptor complexes. Neurochem. Int. *49*, 533–542.

Krashes, M.J., Koda, S., Ye, C., Rogan, S.C., Adams, A.C., Cusher, D.S., Maratos-Flier, E., Roth, B.L., and Lowell, B.B. (2011). Rapid, reversible activation of AgRP neurons drives feeding behavior in mice. J. Clin. Invest. *121*, 1424–1428.

Krashes, M.J., Shah, B.P., Koda, S., and Lowell, B.B. (2013). Rapid versus delayed stimulation of feeding by the endogenously released AgRP neuron mediators GABA, NPY, and AgRP. Cell Metab. *18*, 588–595.

Krashes, M.J., Shah, B.P., Madara, J.C., Olson, D.P., Strochlic, D.E., Garfield, A.S., Vong, L., Pei, H., Watabe-Uchida, M., Uchida, N., et al. (2014). An excitatory paraventricular nucleus to AgRP neuron circuit that drives hunger. Nature *507*, 238–242.

Krashes, M.J., Lowell, B.B., and Garfield, A.S. (2016). Melanocortin-4 receptor–regulated energy homeostasis. Nat. Neurosci. *19*, 206–219.

Krezel, A.M., Kasibhatla, C., Hidalgo, P., MacKinnon, R., and Wagner, G. (1995). Solution structure of the potassium channel inhibitor agitoxin 2: caliper for probing channel geometry. Protein Sci. *4*, 1478–1489.

Kristensen, K., Henriksen, J.R., and Andresen, T.L. (2015). Adsorption of cationic peptides to solid surfaces of glass and plastic. PLoS One *10*, e0122419.

Kublaoui, B.M., Gemelli, T., Tolson, K.P., Wang, Y., and Zinn, A.R. (2008). Oxytocin deficiency mediates hyperphagic obesity of Sim1 haploinsufficient mice. Mol. Endocrinol. *22*, 1723–1734.

Kühnen, P., Clément, K., Wiegand, S., Blankenstein, O., Gottesdiener, K., Martini, L.L., Mai, K., Blume-Peytavi, U., Grüters, A., and Krude, H. (2016). Proopiomelanocortin Deficiency Treated with a Melanocortin-4 Receptor Agonist. N. Engl. J. Med. *375*, 240–246.

Kumar, K.G., Sutton, G.M., Dong, J.Z., Roubert, P., Plas, P., Halem, H.A., Culler, M.D., Yang, H., Dixit, V.D., and Butler, A.A. (2009). Analysis of the therapeutic functions of novel melanocortin receptor agonists in MC3R- and MC4R-deficient C57BL/6J mice. Peptides *30*, 1892–1900.

Kunitz, M., and John, H.N. (1936). Isolation from beef pancreas of crystalline trypsinogen, trypsin, a trypsin inhibitor, and an inhibitor-trypsin compound. J. Gen. Physiol. *19*, 991–1007.

Kuzmenkov, A.I., Grishin, E. V., and Vassilevski, A.A. (2016). Diversity of Potassium Channel Ligands: Focus on Scorpion Toxins. Biochem. *80*, 1764–1799.

Lacombe, O., Videau, O., Chevillon, D., Guyot, A.-C., Contreras, C., Blondel, S., Nicolas, L., Ghettas, A., Bénech, H., Thevenot, E., et al. (2011). In vitro primary human and animal cell-based blood-brain barrier models as a screening tool in drug discovery. Mol. Pharm. *8*, 651–663.

Lancelin, J.M., Foray, M.F., Poncin, M., Hollecker, M., and Marion, D. (1994). Proteinase inhibitor homologues as potassium channel blockers. Nat. Struct. Biol. *1*, 246–250.

Langlet, F., Levin, B.E., Luquet, S., Mazzone, M., Messina, A., Dunn-Meynell, A.A., Balland, E., Lacombe, A., Mazur, D., Carmeliet, P., et al. (2013). Tanycytic VEGF-A boosts blood-hypothalamus barrier plasticity and access of metabolic signals to the arcuate nucleus in response to fasting. Cell Metab. *17*, 607–617.

Laps, S., Sun, H., Kamnesky, G., and Brik, A. (2019). Palladium-Mediated Direct Disulfide Bond Formation in Proteins Containing S-Acetamidomethyl-cysteine under Aqueous Conditions. Angew. Chem. Int. Ed. Engl. *58*, 5729–5733.

Laskowski, M., and Kato, I. (1980). Protein inhibitors of proteinases. Annu. Rev. Biochem. *49*, 593–626.

Latorraca, N.R., Venkatakrishnan, A.J., and Dror, R.O. (2017). GPCR Dynamics: Structures in Motion. Chem. Rev. *117*, 139–155.

Lavergne, V., Alewood, P.F., Mobli, M., and King, G.F. (2015a). CHAPTER 2. The Structural Universe of Disulfide-Rich Venom Peptides. In Venoms to Drugs : Venom as Source for the Development of Human Therapeutics, G.F. King, ed. (RSC Drug Discovery), pp. 37–79.

Lavergne, V., Harliwong, I., Jones, A., Miller, D., Taft, R.J., and Alewood, P.F. (2015b). Optimized deep-targeted proteotranscriptomic profiling reveals unexplored Conus toxin diversity and novel cysteine frameworks. Proc. Natl. Acad. Sci. U. S. A. *112*, E3782-91.

Lawson, E.A. (2017). The effects of oxytocin on eating behaviour and metabolism in humans. Nat. Rev. Endocrinol. *13*, 700–709.

Lawson, E.A., Marengi, D.A., DeSanti, R.L., Holmes, T.M., Schoenfeld, D.A., and Tolley, C.J. (2015). Oxytocin reduces caloric intake in men. Obesity (Silver Spring). *23*, 950–956.

Lees, J.G., Miles, A.J., Wien, F., and Wallace, B.A. (2006). A reference database for circular dichroism spectroscopy covering fold and secondary structure space. Bioinformatics *22*, 1955–1962.

Lewis, R.J., and Garcia, M.L. (2003). Therapeutic potential of venom peptides. Nat. Rev. Drug Discov. *2*, 790–802.

Li, J.W.-H., and Vederas, J.C. (2009). Drug discovery and natural products: end of an era or an endless frontier? Science *325*, 161–165.

Li, A.-J., Wang, Q., Dinh, T.T., Simasko, S.M., and Ritter, S. (2016a). Mercaptoacetate blocks fatty acid-induced GLP-1 secretion in male rats by directly antagonizing GPR40 fatty acid receptors. Am. J. Physiol. Regul. Integr. Comp. Physiol. *310*, R724-32.

Li, Y.-Q., Shrestha, Y., Pandey, M., Chen, M., Kablan, A., Gavrilova, O., Offermanns, S., and Weinstein, L.S. (2016b). G(q/11)α and G(s)α mediate distinct physiological responses to central melanocortins. J. Clin. Invest. *126*, 40–49.

Liu, J., Wu, Q., Pi, C., Zhao, Y., Zhou, M., Wang, L., Chen, S., and Xu, A. (2007). Isolation and characterization of a T-superfamily conotoxin from Conus litteratus with targeting tetrodotoxin-sensitive sodium channels. Peptides *28*, 2313–2319.

Liu, Z., Dai, J., Dai, L., Deng, M., Hu, Z., Hu, W., and Liang, S. (2006). 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.

Livneh, Y., Ramesh, R.N., Burgess, C.R., Levandowski, K.M., Madara, J.C., Fenselau, H., Goldey, G.J., Diaz, V.E., Jikomes, N., Resch, J.M., et al. (2017). Homeostatic circuits selectively gate food cue responses in insular cortex. Nature *546*, 611–616.

Loprinzi, C.L., Schaid, D.J., Dose, A.M., Burnham, N.L., and Jensen, M.D. (1993). Body-composition changes in patients who gain weight while receiving megestrol acetate. J. Clin. Oncol. *11*, 152–154.

Lotta, L.A., Mokrosiński, J., Mendes de Oliveira, E., Li, C., Sharp, S.J., Luan, J., Brouwers, B., Ayinampudi, V., Bowker, N., Kerrison, N., et al. (2019). Human Gain-of-Function MC4R Variants Show Signaling Bias and Protect against Obesity. Cell *177*, 597–607.e9.

Lucchesi, K., Ravindran, A., Young, H., and Moczydlowski, E. (1989). Analysis of the blocking activity of charybdotoxin homologs and iodinated derivatives against Ca2+-activated K+ channels. J. Membr. Biol. *109*, 269–281.

Ludwig, D.S., Tritos, N.A., Mastaitis, J.W., Kulkarni, R., Kokkotou, E., Elmquist, J., Lowell, B., Flier, J.S., and Maratos-Flier, E. (2001). Melanin-concentrating hormone overexpression in transgenic mice leads to obesity and insulin resistance. J. Clin. Invest. *107*, 379–386.

Luo, R., Jeong, S.-J., Jin, Z., Strokes, N., Li, S., and Piao, X. (2011). G protein-coupled receptor 56 and collagen III, a receptor-ligand pair, regulates cortical development and lamination. Proc. Natl. Acad. Sci. U. S. A. *108*, 12925–12930.

Luo, S., Zhangsun, D., Wu, Y., Zhu, X., Xie, L., Hu, Y., Zhang, J., and Zhao, X. (2006). Identification and molecular diversity of T-superfamily conotoxins from Conus lividus and Conus litteratus. Chem. Biol. Drug Des. *68*, 97–106.

Lutz, T.A. (2013). The interaction of amylin with other hormones in the control of eating. Diabetes. Obes. Metab. *15*, 99–111.

Mackinnon, R., Guy, C.L., Rev, A., Physiol, P., Mol, P., Hughes, M.A., Dunn, M.A., Cohen, S.L., Kuo, A., Lee, A., et al. (2011). Channels Structural Conservation in Prokaryotic and Eukaryotic Potassium Channels. Spring *106*, 106–110.

Maïga, A., Mourier, G., Quinton, L., Rouget, C., Gales, C., Denis, C., Lluel, P., Sénard, J.-M., Palea, S., Servent, D., et al. (2012). G protein-coupled receptors, an unexploited animal toxin targets: Exploration of green mamba venom for novel drug candidates active against adrenoceptors. Toxicon *59*, 487–496.

Maïga, A., Vera, L., Marchetti, C., Lorphelin, A., Bellanger, L., Mourier, G., Servent, D., Gilles, N., and Stura, E.A. (2013). Crystallization of recombinant green mamba ρ-Da1a toxin during a lyophilization procedure and its structure determination. Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun. *69*, 704–709.

Mamathambika, B.S., and Bardwell, J.C. (2008). Disulfide-linked protein folding pathways. Annu. Rev. Cell Dev. Biol. *24*, 211–235.

Mandelblat-Cerf, Y., Ramesh, R.N., Burgess, C.R., Patella, P., Yang, Z., Lowell, B.B., and Andermann, M.L. (2015). Arcuate hypothalamic AgRP and putative POMC neurons show opposite changes in spiking across multiple timescales. Elife *4*, 1–25.

Manglik, A., and Kruse, A.C. (2017). Structural Basis for G Protein-Coupled Receptor Activation. Biochemistry *56*, 5628–5634.

Margues, D.L., and Olds, J. (1962). Identical “feeding” and “rewarding” systems in the lateral hypothalamus of rats. Science *135*, 374–375.

Massonnet, P., Haler, J.R.N., Upert, G., Smargiasso, N., Mourier, G., Gilles, N., Quinton, L., and De Pauw, E. (2018). Disulfide Connectivity Analysis of Peptides Bearing Two Intramolecular Disulfide Bonds Using MALDI In-Source Decay. J. Am. Soc. Mass Spectrom. *29*, 1995–2002.

Masuda, Y., Tanaka, T., Inomata, N., Ohnuma, N., Tanaka, S., Itoh, Z., Hosoda, H., Kojima, M., and Kangawa, K. (2000). Ghrelin stimulates gastric acid secretion and motility in rats. Biochem. Biophys. Res. Commun. *276*, 905–908.

McCann, C.M., Bracamontes, J., Steinbach, J.H., and Sanes, J.R. (2006). The cholinergic antagonist alpha-bungarotoxin also binds and blocks a subset of GABA receptors. Proc. Natl. Acad. Sci. U. S. A. *103*, 5149–5154.

McDonald, L.J., and Moss, J. (1994). Enzymatic and nonenzymatic ADP-ribosylation of cysteine. Mol. Cell. Biochem. *138*, 221–226.

McDowell, R.S., Dennis, M.S., Louie, A., Shuster, M., Mulkerrin, M.G., and Lazarus, R.A. (1992). Mambin, a potent glycoprotein IIb-IIIa antagonist and platelet aggregation inhibitor structurally related to the short neurotoxins. Biochemistry *31*, 4766–4772.

McNay, D.E.G., Briançon, N., Kokoeva, M. V., Maratos-Flier, E., and Flier, J.S. (2012). Remodeling of the arcuate nucleus energy-balance circuit is inhibited in obese mice. J. Clin. Invest. *122*, 142–152.

McNulty, J.C., Jackson, P.J., Thompson, D.A., Chai, B., Gantz, I., Barsh, G.S., Dawson, P.E., and Millhauser, G.L. (2005). Structures of the agouti signaling protein. J. Mol. Biol. *346*, 1059–1070.

Metherell, L.A., Chapple, J.P., Cooray, S., David, A., Becker, C., Rüschendorf, F., Naville, D., Begeot, M., Khoo, B., Nürnberg, P., et al. (2005). Mutations in MRAP, encoding a new interacting partner of the ACTH receptor, cause familial glucocorticoid deficiency type 2. Nat. Genet. *37*, 166–170.

Miljanich, G.P. (2004). Ziconotide: neuronal calcium channel blocker for treating severe chronic pain. Curr. Med. Chem. *11*, 3029–3040.

Miller, L.J., Holicky, E.L., Ulrich, C.D., and Wieben, E.D. (1995). Abnormal processing of the human cholecystokinin receptor gene in association with gallstones and obesity. Gastroenterology *109*, 1375–1380.

Milligan, G., and Kostenis, E. (2006). Heterotrimeric G-proteins: a short history. Br. J. Pharmacol. *147 Suppl*, S46-55.

Möller, C., Rahmankhah, S., Lauer-Fields, J., Bubis, J., Fields, G.B., and Marí, F. (2005). A novel conotoxin framework with a helix-loop-helix (Cs aa/aa) fold. Biochemistry *44*, 15986–15996.

Mollica, A., Pinnen, F., Azzurra, S., and Costante, R. (2014). The Evolution of Peptide Synthesis: From Early Days to Small Molecular Machines. Curr. Bioact. Compd. *9*, 184–202.

Moran, T.H., Shnayder, L., Hostetler, A.M., and McHugh, P.R. (1988). Pylorectomy reduces the satiety action of cholecystokinin. Am. J. Physiol. *255*, R1059-63.

Morita-Takemura, S., and Wanaka, A. (2019). Blood-to-brain communication in the hypothalamus for energy intake regulation. Neurochem. Int. *128*, 135–142.

Morsa, D., Baiwir, D., La Rocca, R., Zimmerman, T.A., Hanozin, E., Grifnée, E., Longuespée, R., Meuwis, M.-A., Smargiasso, N., Pauw, E. De, et al. (2019). Multi-Enzymatic Limited Digestion: The Next-Generation Sequencing for Proteomics? J. Proteome Res. *18*, 2501–2513.

Mouhat, S., Mosbah, A., Visan, V., Wulff, H., Delepierre, M., Darbon, H., Grissmer, S., De Waard, M., and Sabatier, J.-M. (2004). The "functional" dyad of scorpion toxin Pi1 is not itself a prerequisite for toxin binding to the voltage-gated Kv1.2 potassium channels. Biochem. J. *377*, 25–36.

Mouhat, S., Andreotti, N., Jouirou, B., and Sabatier, J.-M. (2008). Animal toxins acting on voltage-gated potassium channels. Curr. Pharm. Des. *14*, 2503–2518.

Mourão, C.B.F., and Schwartz, E.F. (2013). Protease inhibitors from marine venomous animals and their counterparts in terrestrial venomous animals. Mar. Drugs *11*, 2069–2112.

Mourier, G., Dutertre, S., Fruchart-Gaillard, C., Ménez, A., and Servent, D. (2003). Chemical synthesis of MT1 and MT7 muscarinic toxins: critical role of Arg-34 in their interaction with M1 muscarinic receptor. Mol. Pharmacol. *63*, 26–35.

Mullier, A., Bouret, S.G., Prevot, V., and Dehouck, B. (2010). Differential distribution of tight junction proteins suggests a role for tanycytes in blood-hypothalamus barrier regulation in the adult mouse brain. J. Comp. Neurol. *518*, 943–962.

Mullis, K., Kay, K., and Williams, D.L. (2013). Oxytocin action in the ventral tegmental area affects sucrose intake. Brain Res. *1513*, 85–91.

Muratspahić, E., Freissmuth, M., and Gruber, C.W. (2019). Nature-Derived Peptides: A Growing Niche for GPCR Ligand Discovery. Trends Pharmacol. Sci. *40*, 309–326.

Namkung, Y., Le Gouill, C., Lukashova, V., Kobayashi, H., Hogue, M., Khoury, E., Song, M., Bouvier, M., and Laporte, S.A. (2016). Monitoring G protein-coupled receptor and β-arrestin trafficking in live cells using enhanced bystander BRET. Nat. Commun. *7*, 12178.

Le Nguyen, D., Heitz, A., Chiche, L., Castro, B., Boigegrain, R.A., Favel, A., and Coletti-Previero, M.A. (1990). Molecular recognition between serine proteases and new bioactive microproteins with a knotted structure. Biochimie *72*, 431–435.

Nieh, E.H., Matthews, G.A., Allsop, S.A., Presbrey, K.N., Leppla, C.A., Wichmann, R., Neve, R., Wildes, C.P., and Tye, K.M. (2015). Decoding neural circuits that control compulsive sucrose seeking. Cell *160*, 528–541.

Nieh, E.H., Vander Weele, C.M., Matthews, G.A., Presbrey, K.N., Wichmann, R., Leppla, C.A., Izadmehr, E.M., and Tye, K.M. (2016). Inhibitory Input from the Lateral Hypothalamus to the Ventral Tegmental Area Disinhibits Dopamine Neurons and Promotes Behavioral Activation. Neuron *90*, 1286–1298.

Nijenhuis, W.A.J., Oosterom, J., and Adan, R.A.H. (2001). AgRP(83–132) Acts as an Inverse Agonist on the Human- Melanocortin-4 Receptor. Mol. Endocrinol. *15*, 164–171.

Nikiforovich, G. V., Taylor, C.M., and Marshall, G.R. (2007). Modeling of the complex between transducin and photoactivated rhodopsin, a prototypical G-protein-coupled receptor. Biochemistry *46*, 4734–4744.

Norton, R.S., and Pallaghy, P.K. (1998). The cystine knot structure of ion channel toxins and related polypeptides. Toxicon *36*, 1573–1583.

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., and De Lima, M.E. (2008). Tx2-6 toxin of the Phoneutria nigriventer spider potentiates rat erectile function. Toxicon *51*, 1197–1206.

O'Connor, E.C., Kremer, Y., Lefort, S., Harada, M., Pascoli, V., Rohner, C., and Lüscher, C. (2015). Accumbal D1R Neurons Projecting to Lateral Hypothalamus Authorize Feeding. Neuron *88*, 553–564.

Olds, J., and Milner, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. J. Comp. Physiol. Psychol. *47*, 419–427.

Olivera, B.M., Gray, W.R., Zeikus, R., McIntosh, J.M., Varga, J., Rivier, J., De Santos, V., and Cruz, L.J. (1985). Peptide neurotoxins from fish-hunting cone snails. Science (80-.). *230*, 1338–1343.

Oprea, T.I. (2002). Current trends in lead discovery: are we looking for the appropriate properties? Mol. Divers. *5*, 199–208.

Owen, D.G., Hall, A., Stephens, G., Stow, J., and Robertson, B. (1997). The relative potencies of dendrotoxins as blockers of the cloned voltage-gated K+ channel, mKv1.1 (MK-1), when stably expressed in Chinese hamster ovary cells. Br. J. Pharmacol. *120*, 1029–1034.

Padilla, S.L., Carmody, J.S., and Zeltser, L.M. (2010). Pomc-expressing progenitors give rise to antagonistic neuronal populations in hypothalamic feeding circuits. Nat. Med. *16*, 403–405.

Palczewski, K., Kumasaka, T., Hori, T., Behnke, C.A., Motoshima, H., Fox, B.A., Le Trong, I., Teller, D.C., Okada, T., Stenkamp, R.E., et al. (2000). Crystal structure of rhodopsin: A G protein-coupled receptor. Science *289*, 739–745.

Pape, H.-C., and Pare, D. (2010). Plastic synaptic networks of the amygdala for the acquisition, expression, and extinction of conditioned fear. Physiol. Rev. *90*, 419–463.

Papp, F., Batista, C.V.F., Varga, Z., Herceg, M., Román-González, S.A., Gaspar, R., Possani, L.D., and Panyi, G. (2009). Tst26, a novel peptide blocker of Kv1.2 and Kv1.3 channels from the venom of Tityus stigmurus. Toxicon *54*, 379–389.

Park, S.P., Kim, B.M., Koo, J.Y., Cho, H., Lee, C.H., Kim, M., Na, H.S., and Oh, U. (2008). A tarantula spider toxin, GsMTx4, reduces mechanical and neuropathic pain. Pain *137*, 208–217.

Parkes, S.L., and Balleine, B.W. (2013). Incentive memory: evidence the basolateral amygdala encodes and the insular cortex retrieves outcome values to guide choice between goal-directed actions. J. Neurosci. *33*, 8753–8763.

Peiren, N., Vanrobaeys, F., de Graaf, D.C., Devreese, B., Van Beeumen, J., and Jacobs, F.J. (2005). The protein composition of honeybee venom reconsidered by a proteomic approach. Biochim. Biophys. Acta *1752*, 1–5.

Peng, X.J., Tso, I.M., and Li, S.Q. (2002). Five new and four newly recorded species of jumping spiders from Taiwan (Araneae: Salticidae). Zool. Stud. *41*, 1–11.

Petrel, C., Hocking, H.G., Reynaud, M., Upert, G., Favreau, P., Biass, D., Paolini-Bertrand, M., Peigneur, S., Tytgat, J., Gilles, N., et al. (2013). Identification, structural and pharmacological characterization of τ-CnVA, a conopeptide that selectively interacts with somatostatin sst3 receptor. Biochem. Pharmacol. *85*, 1663–1671.

Petruzzelli, M., Schweiger, M., Schreiber, R., Campos-Olivas, R., Tsoli, M., Allen, J., Swarbrick, M., Rose-John, S., Rincon, M., Robertson, G., et al. (2014). A switch from white to brown fat increases energy expenditure in cancer-associated cachexia. Cell Metab. *20*, 433–447.

Pin, J.-P., Galvez, T., and Prézeau, L. (2003). Evolution, structure, and activation mechanism of family 3/C G-protein-coupled receptors. Pharmacol. Ther. *98*, 325–354.

Pinto, S., Roseberry, A.G., Liu, H., Diano, S., Shanabrough, M., Cai, X., Friedman, J.M., and Horvath, T.L. (2004). Rapid rewiring of arcuate nucleus feeding circuits by leptin. Science *304*, 110–115.

Pogozheva, I.D., Chai, B.X., Lomize, A.L., Fong, T.M., Weinberg, D.H., Nargund, R.P., Mulholland, M.W., Gantz, I., and Mosberg, H.I. (2005). Interactions of human melanocortin 4 receptor with nonpeptide and peptide agonists. Biochemistry *44*, 11329–11341.

Postic, G., Gracy, J., Périn, C., Chiche, L., and Gelly, J.-C. (2018). KNOTTIN: the database of inhibitor cystine knot scaffold after 10 years, toward a systematic structure modeling. Nucleic Acids Res. *46*, D454–D458.

Psichas, A., Reimann, F., and Gribble, F.M. (2015). Gut chemosensing mechanisms. J. Clin. Invest. *125*, 908–917.

Puzziferri, N., Roshek, T.B., Mayo, H.G., Gallagher, R., Belle, S.H., and Livingston, E.H. (2014). Long-term follow-up after bariatric surgery: a systematic review. JAMA *312*, 934–942.

Qin, Y., Dey, A., and Daaka, Y. (2013). Protein s-nitrosylation measurement. Methods Enzymol. *522*, 409–425.

Qu, D., Ludwig, D.S., Gammeltoft, S., Piper, M., Pelleymounter, M.A., Cullen, M.J., Mathes, W.F., Przypek, R., Kanarek, R., and Maratos-Flier, E. (1996). A role for melanin-concentrating hormone in the central regulation of feeding behaviour. Nature *380*, 243–247.

Quinton, L., Gilles, N., and De Pauw, E. (2009). TxXIIIA, an atypical homodimeric conotoxin found in the Conus textile venom. J. Proteomics *72*, 219–226.

Quinton, L., Girard, E., Maiga, A., Rekik, M., Lluel, P., Masuyer, G., Larregola, M., Marquer, C., Ciolek, J., Magnin, T., et al. (2010). Isolation and pharmacological characterization of AdTx1, a natural peptide displaying specific insurmountable antagonism of the alpha1A-adrenoceptor. Br. J. Pharmacol. *159*, 316–325.

Reano, S., Graziani, A., and Filigheddu, N. (2014). Acylated and unacylated ghrelin administration to blunt muscle wasting. Curr. Opin. Clin. Nutr. Metab. Care *17*, 236–240.

Reed, J., and Reed, T.A. (1997). A set of constructed type spectra for the practical estimation of peptide secondary structure from circular dichroism. Anal. Biochem. *254*, 36–40.

Rees, D.C., and Lipscomb, W.N. (1982). Refined crystal structure of the potato inhibitor complex of carboxypeptidase A at 2.5 A resolution. J. Mol. Biol. *160*, 475–498.

Reidelberger, R.D., Hernandez, J., Fritzsch, B., and Hulce, M. (2004). Abdominal vagal mediation of the satiety effects of CCK in rats. Am. J. Physiol. Regul. Integr. Comp. Physiol. *286*, R1005-12.

von Reumont, B.M., Undheim, E.A.B., Jauss, R.-T., and Jenner, R.A. (2017). Venomics of Remipede Crustaceans Reveals Novel Peptide Diversity and Illuminates the Venom's Biological Role. Toxins (Basel). *9*.

Rhea, E.M., Rask-Madsen, C., and Banks, W.A. (2018). Insulin transport across the blood-brain barrier can occur independently of the insulin receptor. J. Physiol. *596*, 4753–4765.

Ritter, R.C. (2004). Gastrointestinal mechanisms of satiation for food. Physiol. Behav. *81*, 249–273.

Robinson, S.D., and Norton, R.S. (2014). Conotoxin gene superfamilies. Mar. Drugs *12*, 6058–6101.

Rodrigues, A.R., Almeida, H., and Gouveia, A.M. (2012). Melanocortin 5 receptor signaling and internalization: role of MAPK/ERK pathway and β-arrestins 1/2. Mol. Cell. Endocrinol. *361*, 69–79.

Rodrigues, A.R., Almeida, H., and Gouveia, A.M. (2013). Alpha-MSH signalling via melanocortin 5 receptor promotes lipolysis and impairs re-esterification in adipocytes. Biochim. Biophys. Acta *1831*, 1267–1275.

Rodrigues, A.R.A., Arantes, E.C., Monje, F., Stühmer, W., and Varanda, W.A. (2003). Tityustoxin-K(alpha) blockade of the voltage-gated potassium channel Kv1.3. Br. J. Pharmacol. *139*, 1180–1186.

Roeland, E.J., Bruggeman, A.R., Kamal, A.H., LeBlanc, T.W., Ma, J.D., and Baracos, V.E. (2016). Cancer cachexia: Beyond weight loss. J. Oncol. Pract. *12*, 1163–1171.

Rogowski, R.S., Krueger, B.K., Collins, J.H., and Blaustein, M.P. (1994). Tityustoxin K alpha blocks voltage-gated noninactivating K+ channels and unblocks inactivating K+ channels blocked by alpha-dendrotoxin in synaptosomes. Proc. Natl. Acad. Sci. U. S. A. *91*, 1475–1479.

Rong, M., Duan, Z., Chen, J., Li, J., Xiao, Y., and Liang, S. (2013). Native pyroglutamation of huwentoxin-IV: a post-translational modification that increases the trapping ability to the sodium channel. PLoS One *8*, e65984.

Rosenbaum, M., and Leibel, R.L. (2010). Adaptive thermogenesis in humans. Int. J. Obes. *34*, S47–S55.

Rosin, D.L., Weston, M.C., Sevigny, C.P., Stornetta, R.L., and Guyenet, P.G. (2003). Hypothalamic orexin (hypocretin) neurons express vesicular glutamate transporters VGLUT1 or VGLUT2. J. Comp. Neurol. *465*, 593–603.

Rouget, C., Quinton, L., Maïga, A., Gales, C., Masuyer, G., Malosse, C., Chamot-Rooke, J., Thai, R., Mourier, G., De Pauw, E., et al. (2010). Identification of a novel snake peptide toxin displaying high affinity and antagonist behaviour for the α2-adrenoceptors. Br. J. Pharmacol. *161*, 1361–1374.

Saez, N.J., Senff, S., Jensen, J.E., Er, S.Y., Herzig, V., Rash, L.D., and King, G.F. (2010). Spider-venom peptides as therapeutics. Toxins (Basel). *2*, 2851–2871.

Sailer, A.W., Sano, H., Zeng, Z., McDonald, T.P., Pan, J., Pong, S.S., Feighner, S.D., Tan, C.P., Fukami, T., Iwaasa, H., et al. (2001). Identification and characterization of a second melanin-concentrating hormone receptor, MCH-2R. Proc. Natl. Acad. Sci. U. S. A. *98*, 7564–7569.

Sakurai, T. (1999). Orexins and orexin receptors: Implication in feeding behavior. Regul. Pept. *85*, 25–30.

Samuels, B.C., Zaretsky, D. V., and DiMicco, J.A. (2002). Tachycardia evoked by disinhibition of the dorsomedial hypothalamus in rats is mediated through medullary raphe. J. Physiol. *538*, 941–946.

Sanz, L., Pla, D., Pérez, A., Rodríguez, Y., Zavaleta, A., Salas, M., Lomonte, B., and Calvete, J.J. (2016). Venomic Analysis of the Poorly Studied Desert Coral Snake, Micrurus tschudii tschudii, Supports the 3FTx/PLA_2 Dichotomy across Micrurus Venoms. Toxins (Basel). *8*.

Sato, S., Nakamura, H., Ohizumi, Y., Kobayashi, J., and Hirata, Y. (1983). The amino acid sequences of homologous hydroxyproline-containing myotoxins from the marine snail Conus geographus venom. FEBS Lett. *155*, 277–280.

Scarlett, J.M., Jobst, E.E., Enriori, P.J., Bowe, D.D., Batra, A.K., Grant, W.F., Cowley, M.A., and Marks, D.L. (2007). Regulation of central melanocortin signaling by interleukin-1 beta. Endocrinology *148*, 4217–4225.

Schaeffer, M., Langlet, F., Lafont, C., Molino, F., Hodson, D.J., Roux, T., Lamarque, L., Verdié, P., Bourrier, E., Dehouck, B., et al. (2013). Rapid sensing of circulating ghrelin by hypothalamic appetite-modifying neurons. Proc. Natl. Acad. Sci. U. S. A. *110*, 1512–1517.

Schechter, I., and Berger, A. (1967). On the size of the active site in proteases. I. Papain. Biochem. Biophys. Res. Commun. *27*, 157–162.

Schweitz, H., Heurteaux, C., Bois, P., Moinier, D., Romey, G., and Lazdunski, M. (2006). Calcicludine, a venom peptide of the Kunitz-type protease inhibitor family, is a potent blocker of high-threshold Ca2+ channels with a high affinity for L-type channels in cerebellar granule neurons. Proc. Natl. Acad. Sci. *91*, 878–882.

Sequeira, A.F., Turchetto, J., Saez, N.J., Peysson, F., Ramond, L., Duhoo, Y., Blémont, M., Fernandes, V.O., Gama, L.T., Ferreira, L.M.A., et al. (2017). Gene design, fusion technology and TEV cleavage conditions influence the purification of oxidized disulphide-rich venom peptides in Escherichia coli. Microb. Cell Fact. *16*, 1–16.

Shah, B.P., Vong, L., Olson, D.P., Koda, S., Krashes, M.J., Ye, C., Yang, Z., Fuller, P.M., Elmquist, J.K., and Lowell, B.B. (2014). MC4R-expressing glutamatergic neurons in the paraventricular hypothalamus regulate feeding and are synaptically connected to the parabrachial nucleus. Proc. Natl. Acad. Sci. U. S. A. *111*, 13193–13198.

Shen, W.-J., Yao, T., Kong, X., Williams, K.W., and Liu, T. (2017). Melanocortin neurons: Multiple routes to regulation of metabolism. Biochim. Biophys. Acta. Mol. Basis Dis. *1863*, 2477–2485.

Shi, Y.C., Lau, J., Lin, Z., Zhang, H., Zhai, L., Sperk, G., Heilbronn, R., Mietzsch, M., Weger, S., Huang, X.F., et al. (2013). Arcuate NPY controls sympathetic output and BAT function via a relay of tyrosine hydroxylase neurons in the PVN. Cell Metab. *17*, 236–248.

Shimada, M., Tritos, N.A., Lowell, B.B., Flier, J.S., and Maratos-Flier, E. (1998). Mice lacking melanin-concentrating hormone are hypophagic and lean. Nature *396*, 670–674.

Shon, K.J., Stocker, M., Terlau, H., Stühmer, W., Jacobsen, R., Walker, C., Grilley, M., Watkins, M., Hillyard, D.R., Gray, W.R., et al. (1998). κ-Conotoxin PVIIA is a peptide inhibiting the Shaker K+ channel. J. Biol. Chem. *273*, 33–38.

Siemens, J., Zhou, S., Piskorowski, R., Nikai, T., Lumpkin, E.A., Basbaum, A.I., King, D., and Julius, D. (2006). Spider toxins activate the capsaicin receptor to produce inflammatory pain. Nature *444*, 208–212.

Singh, A., Haslach, E.M., and Haskell-Luevano, C. (2010). Structure-activity relationships (SAR) of melanocortin and agouti-related (AGRP) peptides. Adv. Exp. Med. Biol. *681*, 1–18.

So, S., Chijiwa, T., Ikeda, N., Nobuhisa, I., Oda-Ueda, N., Hattori, S., and Ohno, M. (2008). Identification of the B subtype of γ-phospholipase A2 inhibitor from Protobothrops flavoviridis serum and molecular evolution of snake serum phospholipase A2 inhibitors. J. Mol. Evol. *66*, 298–307.

Sousa, S.R., Vetter, I., and Lewis, R.J. (2013). Venom peptides as a rich source of Cav2.2 channel blockers. Toxins (Basel). *5*, 286–314.

Di Spiezio, A., Sandin, E.S., Dore, R., Müller-Fielitz, H., Storck, S.E., Bernau, M., Mier, W., Oster, H., Jöhren, O., Pietrzik, C.U., et al. (2018). The LepR-mediated leptin transport across brain barriers controls food reward. Mol. Metab. *8*, 13–22.

Srisai, D., Yin, T.C., Lee, A.A., Rouault, A.A.J., Pearson, N.A., Grobe, J.L., and Sebag, J.A. (2017). MRAP2 regulates ghrelin receptor signaling and hunger sensing. Nat. Commun. *8*, 713.

Stamatakis, A.M., Van Swieten, M., Basiri, M.L., Blair, G.A., Kantak, P., and Stuber, G.D. (2016). Lateral Hypothalamic Area Glutamatergic Neurons and Their Projections to the Lateral Habenula Regulate Feeding and Reward. J. Neurosci. *36*, 302–311.

Stehling, E.G., Sfor??a, M.L., Zanchin, N.I.T., Oyama, S., Pignatelli, A., Belluzzi, O., Polverini, E., Corsini, R., Spisni, A., and Pertinhez, T.A. (2012). Looking over toxin-K + channel interactions. clues from the structural and functional characterization of ??-KTx Toxin Tc32, a Kv1.3 Channel Blocker. Biochemistry *51*, 1885–1894.

Steinert, R.E., Feinle-Bisset, C., Asarian, L., Horowitz, M., Beglinger, C., and Geary, N. (2016). Ghrelin, CCK, GLP-1, and PYY(3–36): Secretory Controls and Physiological Roles in Eating and Glycemia in Health, Obesity, and After RYGB. Physiol. Rev. *97*, 411–463.

Steinert, R.E., Feinle-Bisset, C., Asarian, L., Horowitz, M., Beglinger, C., and Geary, N. (2017). Ghrelin, CCK, GLP-1, and PYY(3-36): Secretory Controls and Physiological Roles in Eating and Glycemia in Health, Obesity, and After RYGB. Physiol. Rev. *97*, 411–463.

Stephens, T.W., Basinski, M., Bristow, P.K., Bue-Valleskey, J.M., Burgett, S.G., Craft, L., Hale, J., Hoffmann, J., Hsiung, H.M., and Kriauciunas, A. (1995). The role of neuropeptide Y in the antiobesity action of the obese gene product. Nature *377*, 530–532.

Sternson, S.M., and Eiselt, A.-K. (2017). Three Pillars for the Neural Control of Appetite. Annu. Rev. Physiol. *79*.

Strachan, R.T., Sun, J., Rominger, D.H., Violin, J.D., Ahn, S., Rojas Bie Thomsen, A., Zhu, X., Kleist, A., Costa, T., and Lefkowitz, R.J. (2014). Divergent transducer-specific molecular efficacies generate biased agonism at a G protein-coupled receptor (GPCR). J. Biol. Chem. *289*, 14211–14224.

Stratford, T.R., and Kelley, A.E. (1999). Evidence of a functional relationship between the nucleus accumbens shell and lateral hypothalamus subserving the control of feeding behavior. J. Neurosci. *19*, 11040–11048.

Stuber, G.D., and Wise, R.A. (2016). Lateral hypothalamic circuits for feeding and reward. Nat. Neurosci. *19*, 198–205.

Takacs, Z., Toups, M., Kollewe, A., Johnson, E., Cuello, L.G., Driessens, G., Biancalana, M., Koide, A., Ponte, C.G., Perozo, E., et al. (2009). A designer ligand specific for Kv1.3 channels from a scorpion neurotoxin-based library. Proc. Natl. Acad. Sci. U. S. A. *106*, 22211–22216.

Tan, C.P., Sano, H., Iwaasa, H., Pan, J., Sailer, A.W., Hreniuk, D.L., Feighner, S.D., Palyha, O.C., Pong, S.-S., Figueroa, D.J., et al. (2002). Melanin-concentrating hormone receptor subtypes 1 and 2: species-specific gene expression. Genomics *79*, 785–792.

Tao, Y.X. (2010). The melanocortin-4 receptor: Physiology, pharmacology, and pathophysiology. Endocr. Rev. *31*, 506–543.

Tarcha, E.J., Olsen, C.M., Probst, P., Peckham, D., Muñoz-Elías, E.J., Kruger, J.G., and Iadonato, S.P. (2017). Safety and pharmacodynamics of dalazatide, a Kv1.3 channel inhibitor, in the treatment of plaque psoriasis: A randomized phase 1b trial. PLoS One *12*, e0180762.

Tate, E.W., Kalesh, K.A., Lanyon-Hogg, T., Storck, E.M., and Thinon, E. (2015). Global profiling of protein lipidation using chemical proteomic technologies. Curr. Opin. Chem. Biol. *24*, 48–57.

Teichert, R.W., López-Vera, E., Gulyas, J., Watkins, M., Rivier, J., and Olivera, B.M. (2006). Definition and characterization of the short alphaA-conotoxins: a single residue determines dissociation kinetics from the fetal muscle nicotinic acetylcholine receptor. Biochemistry *45*, 1304–1312.

Temel, J.S., Abernethy, A.P., Currow, D.C., Friend, J., Duus, E.M., Yan, Y., and Fearon, K.C. (2016). Anamorelin in patients with non-small-cell lung cancer and cachexia (ROMANA 1 and ROMANA 2): results from two randomised, double-blind, phase 3 trials. Lancet. Oncol. *17*, 519–531.

Terlau, H., and Olivera, B.M. (2004). Conus venoms: a rich source of novel ion channel-targeted peptides. Physiol. Rev. *84*, 41–68.

Tolle, V., and Low, M.J. (2008). In vivo evidence for inverse agonism of Agouti-related peptide in the central nervous system of proopiomelanocortin-deficient mice. Diabetes *57*, 86–94.

Tota, M.R., Smith, T.S., Mao, C., MacNeil, T., Mosley, R.T., Van der Ploeg, L.H., and Fong, T.M. (1999). Molecular interaction of Agouti protein and Agouti-related protein with human melanocortin receptors. Biochemistry *38*, 897–904.

Turchetto, J., Sequeira, A.F., Ramond, L., Peysson, F., Brás, J.L.A., Saez, N.J., Duhoo, Y., Blémont, M., Guerreiro, C.I.P.D., Quinton, L., et al. (2017). High-throughput expression of animal venom toxins in Escherichia coli to generate a large library of oxidized disulphide-reticulated peptides for drug discovery. Microb. Cell Fact. *16*, 1–15.

Turner, A.H., Craik, D.J., Kaas, Q., and Schroeder, C.I. (2018). Bioactive Compounds Isolated from Neglected Predatory Marine Gastropods. Mar. Drugs *16*.

Tytgat, J., Chandy, K.G., Garcia, M.L., Gutman, G.A., Martin-Eauclaire, M.F., van der Walt, J.J., and Possani, L.D. (1999). A unified nomenclature for short-chain peptides isolated from scorpion venoms: alpha-KTx molecular subfamilies. Trends Pharmacol. Sci. *20*, 444–447.

Upert, G., Mourier, G., Pastor, A., Verdenaud, M., Alili, D., Servent, D., and Gilles, N. (2014). High-throughput production of two disulphide-bridge toxins. Chem. Commun. *50*, 8408–8411.

Utkin, Y.N. (2018). Last decade update for three-finger toxins: Newly emerging structures and biological activities. World J. Biol. Chem. *10*, 17–27.

Van Valen, L. (1973). Evolution Theory. In 1, pp. 1–30.

Venyaminov, Sy., Baikalov, I.A., Shen, Z.M., Wu, C.S., and Yang, J.T. (1993). Circular dichroic analysis of denatured proteins: inclusion of denatured proteins in the reference set. Anal. Biochem. *214*, 17–24.

Verma, D., Wood, J., Lach, G., Herzog, H., Sperk, G., and Tasan, R. (2016). Hunger Promotes Fear Extinction by Activation of an Amygdala Microcircuit. Neuropsychopharmacology *41*, 431–439.

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. (1999). The T-superfamily of conotoxins. J. Biol. Chem. *274*, 30664–30671.

Wang, F., Yan, Z., Liu, Z., Wang, S., Wu, Q., Yu, S., Ding, J., and Dai, Q. (2016). Molecular basis of toxicity of N-type calcium channel inhibitor MVIIA. Neuropharmacology *101*, 137–145.

Wang, H., Liu, T., and Malbon, C.C. (2006). Structure-function analysis of Frizzleds. Cell. Signal. *18*, 934–941.

Warkentin, T.E., Greinacher, A., and Koster, A. (2008). Bivalirudin. Thromb. Haemost. *99*, 830–839.

Waterson, M.J., and Horvath, T.L. (2015). Neuronal Regulation of Energy Homeostasis: Beyond the Hypothalamus and Feeding. Cell Metab. *22*, 962–970.

Waxman, L., Smith, D.E., Arcuri, K.E., and Vlasuk, G.P. (1990). Tick anticoagulant peptide (TAP) is a novel inhibitor of blood coagulation factor Xa. Science *248*, 593–596.

Wen, K.-C., Chang, C.-S., Chien, Y.-C., Wang, H.-W., Wu, W.-C., Wu, C.-S., and Chiang, H.-M. (2013). Tyrosol and its analogues inhibit alpha-melanocyte-stimulating hormone induced melanogenesis. Int. J. Mol. Sci. *14*, 23420–23440.

West, D.B., Fey, D., and Woods, S.C. (1984). Cholecystokinin persistently suppresses meal size but not food intake in free-feeding rats. Am. J. Physiol. *246*, R776-87.

Westerterp, K.R. (2017). Control of energy expenditure in humans. Eur. J. Clin. Nutr. *71*, 340–344.

Wilkie, T.M., Scherle, P.A., Strathmann, M.P., Slepak, V.Z., and Simon, M.I. (1991). Characterization of G-protein alpha subunits in the Gq class: expression in murine tissues and in stromal and hematopoietic cell lines. Proc. Natl. Acad. Sci. U. S. A. *88*, 10049–10053.

Williams, E.K., Chang, R.B., Strochlic, D.E., Umans, B.D., Lowell, B.B., and Liberles, S.D. (2016). Sensory Neurons that Detect Stretch and Nutrients in the Digestive System. Cell *166*, 209–221.

Wing, R.R., and Hill, J.O. (2001). Successful weight loss maintenance. Annu. Rev. Nutr. *21*, 323–341.

Wright, H.T. (1991). Nonenzymatic deamidation of asparaginyl and glutaminyl residues in proteins. Crit. Rev. Biochem. Mol. Biol. *26*, 1–52.

Wu, Y., Wu, X., Yu, J., Zhu, X., Zhangsun, D., and Luo, S. (2014). Influence of disulfide connectivity on structure and bioactivity of α-conotoxin TxIA. Molecules *19*, 966–979.

Yang, L.K., and Tao, Y.X. (2017). Biased signaling at neural melanocortin receptors in regulation of energy homeostasis. Biochim. Biophys. Acta - Mol. Basis Dis.

Yang, J., Xu, J., Gonzalez, R., Lindner, T., Kratochwil, C., and Miao, Y. (2018). 68Ga-DOTA-GGNle-CycMSHhex targets the melanocortin-1 receptor for melanoma imaging. Sci. Transl. Med. *10*, 1–12.

Yang, Y., Atasoy, D., Su, H.H., and Sternson, S.M. (2011). Hunger states switch a flip-flop memory circuit via a synaptic AMPK-dependent positive feedback loop. Cell *146*, 992–1003.

Yi, C.-X., Meyer, C.W., and Jastroch, M. (2013). Leptin action in the brain: How (and when) it makes fat burn. Mol. Metab. *2*, 63–64.

Yi, C.X., Gericke, M., Krüger, M., Alkemade, A., Kabra, D.G., Hanske, S., Filosa, J., Pfluger, P., Bingham, N., Woods, S.C., et al. (2012). High calorie diet triggers hypothalamic angiopathy. Mol. Metab. *1*, 95–100.

Young, A.A. (2012). Brainstem sensing of meal-related signals in energy homeostasis. Neuropharmacology *63*, 31–45.

Zhan, C., Zhou, J., Feng, Q., Zhang, J.-E., Lin, S., Bao, J., Wu, P., and Luo, M. (2013). Acute and long-term suppression of feeding behavior by POMC neurons in the brainstem and hypothalamus, respectively. J. Neurosci. *33*, 3624–3632.

Zhang, C., Zhang, Z., Lin, K.S., Lau, J., Zeisler, J., Colpo, N., Perrin, D.M., and Bénard, F. (2018). Melanoma Imaging Using18F-Labeled α-Melanocyte-Stimulating Hormone Derivatives with Positron Emission Tomography. Mol. Pharm. *15*, 2116–2122.

Zhao, Y., Sun, R., You, L., Gao, C., and Tian, Z. (2003). Expression of leptin receptors and response to leptin stimulation of human natural killer cell lines. Biochem. Biophys. Res. Commun. *300*, 247–252.

Zhu, S., Darbon, H., Dyason, K., Verdonck, F., and Tytgat, J. (2003). Evolutionary origin of inhibitor cystine knot peptides. FASEB J. *17*, 1765–1767.

Zhu, S., Gao, B., and Tytgat, J. (2005). Phylogenetic distribution, functional epitopes and evolution of the CSalphabeta superfamily. Cell. Mol. Life Sci. *62*, 2257–2269.

Zhu, S., Peigneur, S., Gao, B., Umetsu, Y., Ohki, S., and Tytgat, J. (2014). Experimental conversion of a defensin into a neurotoxin: Implications for origin of toxic function. Mol. Biol. Evol. *31*, 546–559.

Ziegler, D.R., Cullinan, W.E., and Herman, J.P. (2002). Distribution of vesicular glutamate transporter mRNA in rat hypothalamus. J. Comp. Neurol. *448*, 217–229.

ANNEXES

Annexe 1. Article « A Venomics approach coupled to High-Throughput toxin production strategies for the identification of the first melanocortin toxins" soumis au Journal of Medicinal Chemistry.

A Venomics approach coupled to high-throughput toxin production strategies identifies the first venom-derived melanocortin receptor agonists

Steve Reynaud[†,⊥], Justyna Ciolek[⊥], Michel Degueldre[ǁ,ε], Natalie J. Saez[ɠ,Ψ], Ana Filipa Sequeira[ϕ,∫], Yoan Duhoo[ɠ,Ψ], Joana L. A. Brás[ϕ,∫], Hervé Meudal[¶], Miguel Cabo Díez[ϫ], Victoria Fernández Pedrosa[ϫ], Marion Verdenaud[⊥], Julia Boeri[⊥], Oscar Pereira Ramos[⊥], Frédéric Ducancel[ǂ], Margot Vanden Driessche[⊥], Rudy Fourmy[δ], Aude Violette[δ], Grégory Upert[⊥], Denis Servent[⊥], Annette G. Beck-Sickinger[Ŧ], Karin Mörl[Ŧ], Céline Landon[¶], Carlos M. G. A. Fontes[ϕ,∫], Rebeca Miñambres Herráiz[ϫ], Ricardo C. Rodríguez de la Vega[§], Steve Peigneur[ϑ], Jan Tytgat[ϑ], Loïc Quinton[ǁ], Edwin De Pauw[ǁ], Renaud Vincentelli[ɠ], Gilles Mourier[⊥], Venomics consortium, Nicolas Gilles[⊥]*

[†]Université Paris-Sud, 15 Rue Georges Clemenceau, Orsay, FR 91405.

[⊥]Commissariat à l'énergie atomique et aux énergies alternatives, Direction de la recherche fondamentale, Institut des Sciences du Vivant Frédéric JOLIOT, Service d'Ingénierie Moléculaire des Protéines (SIMOPRO), CEA Saclay, Gif-sur-Yvette, FR 91191.

[ǁ]Université de Liège, Mass Spectrometry Laboratory, Allée du six Aout 11, Quartier Agora, Liege, BE 4000.

[ε]UCB, Department of Analytical Science Biologicals, Chemin du Foriest, Braine L'Alleud , BE 1420.

[ɠ]Centre National de la Recherche Scientifique, Architecture et Fonction des Macromolécules Biologiques, Campus de Luminy, Marseille, FR 13288.

[Ψ]The University of Queensland, Institute for Molecular Bioscience, Institute for Molecular Bioscience, The University of Queensland, St Lucia, QLD, AUS 4072.

[ϕ]Universidade de Lisboa, CIISA - Faculdade de Medicina Veterinária, Avenida da Universidade Técnica, Lisboa, PT 1300-477.

[∫]NZYTech Lda, Genes & Enzymes, Estrada do Paço do Lumiar, Campus do Lumiar, Edifício E - R/C, Lisboa, PT 1649-038.

[ϫ]Sistemas Genómicos Ltd, Next-Generation Sequencing Laboratory, Ronda de Guglielmo Marconi, 6, Paterna, ES 46980.

[ǂ]Commissariat à l'énergie atomique et aux énergies alternatives, Direction de la recherche fondamentale, Institut des Sciences du Vivant Frédéric JOLIOT, Service de Pharmacologie et Immunoanalyse, CEA Saclay, Gif-sur-Yvette, FR 91191.

[δ] Alphabiotoxine Laboratory sprl, Barberie 15, Montroeul-au-bois, BE 7911.

[Ŧ]Universitat Leipzig, Institute of Biochemistry, Leipzig, DE 04103.

[¶]Centre National de la Recherche Scientifique, Centre de Biophysique Moléculaire, rue Charles Sadron, CS 80054
Orléans, FR 45071.

[§]Universite Paris-Sud, Ecologie, Systematique et Evolution, 15 rue Georges Clémenceau, Orsay, FR 91405.

[ϑ]KU Leuven, Toxicology and Pharmacology, Herestraat 49, Leuven, BE 3000.

Venomics consortium: Venomics is an FP7 project 278346 (2011-2015) with the participation of a mix of public institutions: CEA (Gif sur Yvette, France), Université d'Aix Marseille (Marseille, France), Liege University (Liège, Belgium) and private companies: Sistemas Genómicos Ltd, (Paterna, Valencia, Spain), NZYTech, (Lisbon, Portugal), Zealand Pharma (Glostrup, Denmark), under the management of Absiskey (Grenoble, France).

ABSTRACT: Animal venoms are rich in hundreds of toxins with extraordinary biological activities. Their exploitation is difficult due to their complexity and the small quantities of venom available from most venomous species. We developed a Venomics approach combining transcriptomic and proteomic characterization of 191 species and identified 20,206 venom toxin sequences. Two complementary production strategies based on solid phase synthesis and recombinant expression in *E. coli* generated a physical bank of 3,597 toxins. Screened on four GPCRs, this bank gave an incredible hit rate of 3%. Here, we focus on two novel toxins: N-TRTX-Preg1a, exhibiting an inhibitory cystine knot (ICK) motif, and N-BUTX-Ptr1a, a short scorpion-CSαβ structure. Neither N-TRTX-Preg1a nor N-BUTX-Ptr1a affect ion channels, the known targets of their toxin scaffolds, but bind to four melanocortin receptors with low micromolar affinities and activate the hMC1R/Gs pathway. Phylogenetically, these two toxins form new groups within their respective families and represent novel hMC1R agonists, structurally unrelated to the natural agonists.

Introduction

It is estimated that more than 200,000 venomous species exist on Earth, found in almost all animal subgroups across a range of diverse habitats[1]. Many are terrestrial, such as some vertebrates (e.g., frogs, snakes, platypus, lizards, toads) or arthropods (e.g., scorpions, spiders, centipedes, ticks, ants), some marine, such as cnidarians (e.g., sea anemones and jellyfish) or mollusks (e.g., cuttlefish, octopus, marine cone snails, terebrids, squids), and some aerial including insects (e.g., bugs, wasps, bees) and vampire bats. Venoms are used for defense and/or predation and are composed, among other components, of peptide toxins. Toxins are small proteins up to one hundred residues in size, reticulated by disulfide bridges. Number of toxins present in venoms depends of the species and of the method used to identify them. In the 1990's, publications mentioned few tens of toxins in *Conus* venom. Recently, almost 6,000 toxins were reported in the species *Conus lividus*[2]. All in all, we can estimate that at least 40 million toxins exist on the planet and of these, less than 5,000 have been pharmacologically characterized. Despite this very small number, a plethora of activities have been reported. Toxins interacting with membrane receptors like voltage- and ligand-gated ion channels[3,4] and G-Protein Coupled receptors (GPCR)[5–10] have been frequently described whereas other toxins inhibit enzymes or bind membranes to exert hemolytic and/or antimicrobial actions[11]. Toxins are very often selective (with the exception of some lytic toxins) and make excellent pharmacological tools to better understand the physiological or pathophysiological roles of their targets. In some cases, the toxins themselves or their derivatives have applied clinical uses, such as Byetta® (exenatide from the beaded lizard) a treatment against type-1 diabetes, Prialt® (ziconotide from a cone snail) a treatment for chronic pain, Capoten® (captopril) derived from a snake toxin that is widely used to treat hypertension, but also some other snake toxins e.g. Defibrase® (batroxobin) with clotting properties used to stop bleeding[12]. In addition, more than 20 other toxins are currently investigated in clinical trials as biopharmaceutical and therapeutic tools.

Venoms are a very promising source of selective bioactive molecules and considerable effort is being undertaken to discover new toxins with exploitable activities[13,14]. Classically, toxins are identified by bioguidage, i.e. successive cycles of venom deconvolution and screening tests against a given target. After defining its sequence, the novel toxin of interest must be produced to allow in-depth characterization. This strategy is long, costly and risky and requires large amounts of venoms, unavailable for most venomous animals which are very small in size. This strategy is not adapted for the vast majority of this fantastic biodiversity.

New sequencing strategies that appeared in the 21st century are revolutionizing the toxinological field. These technologies were commonly applied to whole genome sequencing,

however, eukaryote genomes are typically large with tandem repeats and only few genomes from venomous species were reported (several snakes, two scorpions, a handful of spiders and dozens of Hymenopteras) but none from cone snails. By contrast, many studies have taken advantage of these same strategies for the sequencing of various venom gland transcriptomes. Toxins are produced as precursor proteins by specific cells and undergo several maturation steps before being secreted[15]. The formation of disulfide bonds is the most important post-translational modification (PTM) and their arrangement is commonly used to classify toxins[16]. Sequenced transcripts are translated into their cognate amino acid sequences allowing identification of the endoplasmic reticulum signal peptide, the propeptide and mature peptide regions. While tools like SignalP often accurately predict the endoplasmic reticulum signal peptide boundary, relying on trancriptomics alone, it is far more complicated to obtain robust information regarding further processing to the mature peptide due to the lack of information on most pro-protein convertase specific cleavage sites. Moreover, it is important to note that using a purely transcriptomic strategy, the physical presence in the venoms of most of the sequences described in databases is almost never demonstrated. Alternatively, proteomic strategies study only secreted toxins and provide the opportunity to characterize post-translational modifications (PTMs). The combination of these complementary proteomic and transcriptomic strategies leads to a higher level of confidence for mature peptide sequences. Consequently, it is expected that the toxin sequences present in these types of databases will become more and more accurate.

The last and most complicated challenge remains the biological characterization of these natural peptides. This step cannot be done by a simple sequence analysis, especially for peptides with a low sequence similarity with previously characterized toxins. It is necessary to build a physical bank of peptides ready to be screened. The FP7 Venomics Project is part of the biotechnology program "Small-to-Medium Enterprise (SME)-targeted research for developing tools and technologies for high-throughput research". The total project budget was close to 10 million euros (equivalent to 11 million USD) shared by eight partners from five countries. The vision associated with the Venomics Project is not only to investigate the diversity of venom peptides through the development of "omics" technologies but also to replicate toxin diversity *in vitro* to generate novel peptide banks to be used in drug discovery programs[17]. This unique toxin bank was generated and screened on G-Protein Coupled Receptors (GPCRs) for which no toxin modulators have previously been described.

Results

A large venomous animal collection. In total, 323 animal samples from 191 different species were gathered (Table 1). Almost every terrestrial and marine venomous clade was represented by at least one species such as ray (*Potamotrygon motoro*) or venomous lizard (*Heloderma horridum exasperatum*). Small terrestrial arthropods represented 95 species *i.e.* half of the sourcing. 55 species (29%) were studied for the first time, including 44 new species from the most infamous venomous clades (Araneae, Conus, Scorpiones and Serpentes). For instance, 12 previously-unexplored species of Neogastropods were sourced including *Conus canonicus* from the Conidae family, the most studied group of Neogastropoda. The Turridae and Terebridae groups also belong to the Neogastropoda order but have been largely neglected compared to their counterpart[18]. For instance, among over 4,000 described species of Turridae, venom secretions of only twelve species have been studied so far. One of them, *Unedogemmula bisaya* was included in the sourcing. Concerning the Terebridae family, only seven of around 350 described species has been subject of venom analysis. In Venomics sourcing, two of them, *Terebra subulata* and *Terebra guttata*, were chosen. A third species, *Terebra maculata*, never explored before, completed the Terebridae species of the sourcing.

Sourcing also included totally unexplored groups such as Mitridae. Two species considered non-venomous, the *Silurus glanis* fish and the *Trachycephalus resinifictrix* frog, but known to produce toxic skin secretions, were also part of the collection. The diverse collection of source animals encompassed various habitats, morphologies and animal sizes ranging from millimeters (*Manica rubida* ant) to several meters (*Ophiophagus hannah* snake).

Table 1. Selected taxa for sourcing.

Taxa	Total N° of species	Never studied species	Main groups
Neogastropoda			
Conus	38	10	*Conidae*
Other Neogastropoda	5	2	*Turridae, Terebridae & Mitridae*
Octopoda	3	0	*Octopodidae*
Arthropoda			
Chilopoda	6	2	*Scolopendridae*
Scorpiones	40	7	*Buthidae, Scorpionidae, Vaejovidae, Scorpionidae, Hemiscorpiidae & Euscorpiidae*
Araneae	41	23	*Theraphosidae, Ctenidae, Agelenidae, Araneidae, Dipluridae, Lycosidae, Nephilidae, Sparassidae & Theridiidae*
Hymenoptera	8	4	*Apidae, Vespida, Formicidae & Myrmicinae*
Chondrichthyes	1	0	*Potamotrygonidae*
Teleostei	8	2	*Synanceiidae, Plotosidae & Siluridae*
Squamata			
Serpentes	39	4	*Elapidae, Viperidae & Atractaspididae*
Anguimorpha	1	0	*Helodermatidae*
Anura	1	1	*Hylidae*

A huge sequence toxin diversity in animal venoms. *De novo* transcriptomics opens the door to the study of Nature without any previous knowledge of it and with a very high resolution, allowing the unbiased identification and quantification of transcripts in a broad spectrum, even for those of low expression. Total RNA was extracted from venom glands. In order to include species diversity, sample requirements were decreased from 2 µg to 100 ng of RNA, being even as low as 60 ng in specific samples. Illumina specific 2x100 pair-end libraries were constructed from each RNA following manufacturer´s protocols. Over 60 million reads were obtained per sample using Hiseq 2000 sequencing platform with an average length of 500 nucleotides.

Two proteomic strategies were used according to the size of the toxins[19]. The first one, for venoms presenting most of toxins with less than 25 amino acid residues (like Neogastropoda, Hymenoptera, Anura), analyzed the peptides post-disulfide-reduction by liquid chromatography coupled to tandem mass spectrometry. The second strategy, dedicated to the sequencing of the long venom toxins (> 25 amino acid residues (like Chilopoda, Scorpiones, Araneae, Squamata) required additional alkylation and trypsin digestion (short time and low enzyme concentration) steps after disulfide treatment leading to the generation of overlapping peptides (+/- 15 amino acids)[19]. These resulting peptides were then analyzed with the same mass spectrometry platform. The number of sequences varied between fifty to several hundred, depending on the venom. A high number of sequences are observed in Conidae venoms by proteomics. This diversity relies on the presence of post-translational modifications at different sites in a single toxin and has been discussed in previous publications[2,20]. A total of 21,300,000 sequence tags were obtained.

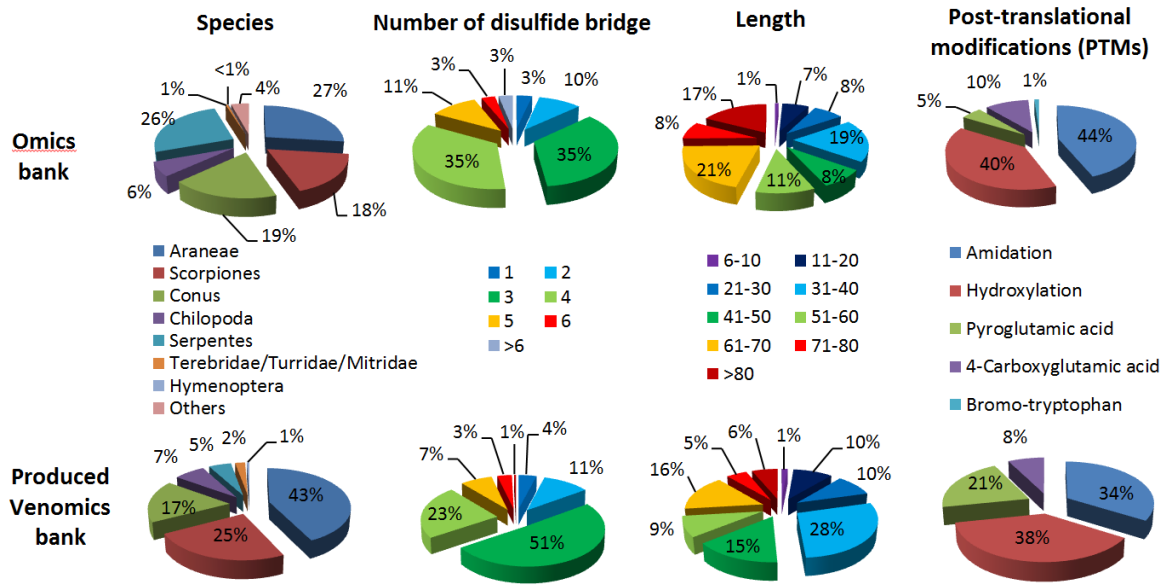


Figure 1. Statistical representation of the sequence banks built from omics analysis (upper panel) and produced Venomics bank (lower panel).

Proteomic and transcriptomic data generated from the same animal were cross-validated to extract only sequences present in both the venom and venom gland transcripts, resulting in 20,206 validated toxin sequences (defined here as peptide up to 110 residues with at least one disulfide bridge). Figure 1 describes these sequences in term of origins, disulfide bridges (1 to 9), sizes (6 to 110) and PTMs. The quasi-totality of Conus sequences (98%) bears at least one PTM, in addition to disulfide bridges. Amidation is found in 2,889 sequences (14%) from all groups studied except Terebridae, Turridae, Mitridae and Teleostei. Hydroxylation on Pro, Met, Lys or Val residues was found in 2,630 sequences (13%) from various groups with a great majority from Conus (70%). In our study, 4-carboxyglutamic acid, pyroglutamic acid and bromo-tryptophan appeared to be Conus specific. Strangely, we did not find pyroglutamic acid in Arachnid venoms (Scorpiones and Araneae), although its presence has been described previously in many other studies[21,22]. The 20,206 toxin sequences were divided into 137 different cysteine patterns. Among the 97 patterns having an even number of cysteines, 38 do not correspond to any known peptide frameworks in any venomous species. Regarding the 59 known patterns, C-C-CC-C-C framework is the most represented comprising 19% of the bank (Table 2). It is related to Inhibitory Cystine Knot Toxins (ICK), which is known to be particularly abundant in spider, snail and scorpion venoms[23–26]. For the first time, ICK peptides were identified at protein level from Mitridae, Turridae, *Manica rubida* ant, Chilopoda and the octopus *Hapalochlaena lunulata.*

Table 2. Cysteine pattern diversity for toxins containing an even number of cysteines identified in the Venomics bank. Newly described patterns are highlighted in grey.

Cysteine framework	N° of Cys	N° of sequences in: Data base	Peptide bank	Taxa in produced bank (N° of species represented)
C-C	2	637	142	Conus (112), Araneae (11), Serpentes (8), Chilopoda (6), Scorpiones (4), Terebridae (1)
CC-CC	4	247	61	Conus (52), Araneae(5), Serpentes (4)
CC-C-C	4	611	106	Conus (92), Araneae (11), Scorpiones (2), Serpentes (1)
C-CCC	4	2	2	Conus (2)
C-CC-C	4	9	1	Conus (1)
C-C-CC	4	63	35	Conus (29), Araneae (5), Serpentes (1)
C-C-C-C	4	764	182	Conus (81), Chilopoda (49), Scorpiones (25), Araneae (21), Terebridae (3), Serpentes (2), Hymenoptera (1)
CCCC-CC	6	10	2	Conus (2)
CCC-C-C-C	6	1	0	
CC-CC-C-C	6	3	0	
CC-C-CC-C	6	2	1	Araneae (1)
CC-C-C-CC	6	450	99	Conus (70), Araneae (26), Serpentes (3)
CC-C-C-C-C	6	10	1	Serpentes (1)
C-CC-CCC	6	1	0	
C-CC-C-C-C	6	20	5	Araneae (4), Conus (1)
C-C-CCC-C	6	1	0	
C-C-CC-CC	6	5	1	Serpentes (1)
C-C-CC-C-C	6	3,739	1,047	Araneae (887), Conus (82), Scorpiones (25), Serpentes (20), Chilopoda (14), Terebridae (12), Hymenoptera (7)
C-C-C-CC-C	6	244	78	Chilopoda (67), Scorpiones (7), Araneae (3), Conus (1)
C-C-C-C-CC	6	83	20	Chilopoda (11), Conus (5), Scorpiones (4)
C-C-C-C-C-C	6	2,199	539	Scorpiones (367), Araneae (76), Conus (33), Serpentes (26), Chilopoda (20), Terebridae (15), Hymenoptera (2)
CC-C-C-C-C-C-C	8	10	2	Conus (1), Scorpiones (1)
C-CCC-C-C-C-C	8	1	0	
C-CC-CC-C-C-C	8	3	0	
C-CC-C-CC-C-C	8	15	7	Araneae (7)
C-CC-C-C-C-C-C	8	11	3	Scorpiones (2), Araneae (1)
C-C-CCCC-C-C	8	1	0	
C-C-CCC-CC-C	8	1	0	
C-C-CC-CC-C-C	8	94	28	Conus (17), Araneae (8), Terebridae (2), Serpentes (1)
C-C-CC-C-CC-C	8	12	3	Araneae (2), Conus (1)
C-C-CC-C-C-CC	8	1	0	
C-C-CC-C-C-C-C	8	438	112	Araneae (68), Terebridae (13), Conus (11), Scorpiones (8), Serpentes (6), Hymenoptera (3), Chilopoda (3)
C-C-C-CCC-C-C	8	29	4	Conus (2), Serpentes (1), Araneae (1)
C-C-C-CC-C-CC	8	2	0	
C-C-C-CC-C-C-C	8	1,544	257	Araneae (190), Scorpiones (27), Chilopoda (17), Conus (10), Hymenoptera (6), Serpentes (5), Terebridae (2)
C-C-C-C-CC-CC	8	4	1	Chilopoda (1)
C-C-C-C-CC-C-C	8	101	43	Scorpiones (23), Serpentes (8), Conus (5), Araneae (4), Chilopoda (2), Terebridae (1)
C-C-C-C-C-CC-C	8	2,128	161	Serpentes (70), Chilopoda (44), Scorpiones (33),

				Araneae (12),Conus (1), Hymenoptera (1)
C-C-C-C-C-C-CC	8	1	0	
C-C-C-C-C-C-C-C	8	2,560	213	Scorpiones (171), Terebridae (23), Chilopoda (11),Conus (4), Serpentes (2), Araneae (2)
CC-C-C-CC-C-C-C-C	10	338	0	
CC-C-C-C-CC-C-C-C	10	2	0	
CC-C-C-C-C-CC-C-C	10	1	1	Scorpiones (1)
CC-C-C-C-C-C-C-C-C	10	1	1	Scorpiones (1)
C-CC-C-CC-C-C-C-C	10	54	7	Conus (4), Araneae (3)
C-CC-C-C-C-C-C-C-C	10	482	0	
C-C-CC-CC-CC-C-C	10	2	1	Scorpiones (1)
C-C-CC-CC-C-C-CC	10	1	0	
C-C-CC-CC-C-C-C-C	10	13	6	Conus (3), Araneae (1), Serpentes (1), Terebridae (1)
C-C-CC-C-CC-C-C-C	10	57	22	Araneae (20), Serpentes (2)
C-C-CC-C-C-C-CC-C	10	4	0	
C-C-CC-C-C-C-C-C-C	10	224	87	Araneae (46), Scorpiones (26), Chilopoda (8), Serpentes (4), Terebridae (2), Hymenoptera (1)
C-C-C-CCC-C-C-C-C	10	12	4	Araneae (2), Conus (1), Serpentes (1)
C-C-C-CC-C-C-CC-C	10	1	0	
C-C-C-CC-C-C-C-C-C	10	86	17	Araneae (14), Scorpiones (2), Terebridae (1)
C-C-C-C-CCCCC-C	10	1	0	
C-C-C-C-CCC-CC-C	10	1	0	
C-C-C-C-CC-C-C-C-C	10	3	1	Araneae (1)
C-C-C-C-C-CC-C-C-C	10	5	0	
C-C-C-C-C-C-CCC-C	10	1	0	
C-C-C-C-C-C-CC-C-C	10	1	1	Serpentes (1)
C-C-C-C-C-C-C-CC-C	10	605	9	Serpentes (7), Scorpiones (2)
C-C-C-C-C-C-C-C-C-C	10	253	107	Scorpiones (51), Araneae (41), Conus (7), Serpentes (4), Terebridae (4)
CC-C-C-C-CC-C-C-C-C	12	5	1	Conus (1)
CC-C-C-C-C-C-CC-C-C	12	2	1	Serpentes (1)
CC-C-C-C-C-C-C-C-C-C	12	1	1	Serpentes (1)
C-CC-C-C-C-C-CC-C-C	12	1	0	
C-CC-C-C-C-C-C-C-C-C	12	4	0	
C-C-CC-CC-CC-CC-C-C	12	2	1	Scorpiones (1)
C-C-CC-CC-C-C-C-C-C	12	2	0	
C-C-CC-C-C-CC-C-C-C	12	5	0	
C-C-CC-C-C-C-CC-C-C	12	15	4	Araneae (4)
C-C-CC-C-C-C-C-C-C	12	3	0	
C-C-C-CC-C-C-C-C-C	12	144	7	Araneae (7)
C-C-C-C-CCC-C-C-	12	3	0	

C-C-C				
C-C-C-C-CC-CC-CC-CC	12	1	1	Scorpiones (1)
C-C-C-C-CC-C-C-C-C-CC	12	1	1	Scorpiones (1)
C-C-C-C-CC-C-C-C-C-C-C	12	33	1	Hymenoptera (1)
C-C-C-C-C-CC-C-C-C-C-C	12	245	99	Scorpiones (90), Araneae (7), Chilopoda (2)
C-C-C-C-C-C-C-C-C-CC-C	12	1	0	
C-C-C-C-C-C-C-C-C-C-CC	12	1	0	
C-C-C-C-C-C-C-C-C-C-C-C	12	22	3	Conus (2), Scorpiones (1)
CC-C-C-C-C-C-CC-C-C-C-C-C	14	3	0	
C-CC-C-C-C-C-C-CC-C-C-C-C	14	1	0	
C-CC-C-C-C-C-C-CC-C-C-C	14	2	1	Conus (1)
C-C-CC-CC-C-C-C-C-C-C-C-C	14	11	0	
C-C-C-CC-C-C-C-C-C-CC-C-C	14	13	3	Conus (2), Araneae (1)
C-C-C-CC-C-C-C-C-C-C-C-C	14	574	7	Araneae (6), Terebridae (1)
C-C-C-C-CC-C-C-CC-C-C-C-C	14	1	1	Araneae (1)
C-C-C-C-C-CC-C-CC-C-C-C-C	14	3	0	
C-C-C-C-C-CC-C-C-C-C-C-C	14	2	0	
C-C-C-C-C-C-CC-C-C-C-C-C	14	4	0	
C-C-C-C-C-C-C-C-C-C-C-C-C	14	14	1	Scorpiones (1)
C-CC-C-C-C-C-C-CC-C-C-C-C-C-C	16	1	0	
C-C-C-C-CC-C-C-C-C-C-C-CC-C	16	2	0	
C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C	16	4	2	Scorpiones (1) , Terebridae (1)
C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C	18	1	1	Scorpiones (1)

Interestingly, 874 sequences covering 40 cysteine patterns display an odd number of cysteine residues (Table 3). Only thirteen were previously described. Two patterns (both lacking vicinal cysteine residues) are particularly abundant: C-C-C-C-C-C-C with 130 sequences (Scorpiones, Serpentes, Octopoda, Conus) and C-C-C-C-C specific to the Araneae taxon with 179 sequences. We did not identify any sequences from Conus sharing the latter pattern, despite it having been previously described[27].

Table 3. Cysteine pattern diversity for toxins containing an odd number of cysteine residues identified in the Venomics bank. Newly described patterns are highlighted in grey.

Cysteine framework	N° of Cys	N° of sequences in : Data base	Peptide bank	Taxa in produced bank (N° of species represented)
CC-C	3	27	3	Araneae (3)
C-CC	3	3	0	
C-C-C	3	37	3	Scorpiones (1), Conus (1), Serpentes (1)
CC-CCC	5	2	0	
CC-C-CC	5	4	1	Conus (1)
CC-C-C-C	5	11	0	
C-CC-C-C	5	17	0	
C-C-CC-C	5	42	1	Araneae (1)
C-C-C-CC	5	3	0	
C-C-C-C-C	5	179	3	Araneae (3)
CC-C-C-CC-C	7	1	0	
C-CC-C-C-C-C	7	3	2	Araneae (2)
C-C-CC-CC-C	7	10	2	Araneae (2)
C-C-CC-C-C-C	7	97	20	Araneae (20)
C-C-C-CC-C-C	7	15	3	Araneae (3)
C-C-C-C-CC-C	7	55	0	
C-C-C-C-C-CC	7	3	0	
C-C-C-C-C-C-C	7	130	2	Chilopoda (1), Scorpiones (1)
CC-C-CC-C-C-C-C	9	2	0	
CC-C-C-CC-C-C-C	9	12	0	
CC-C-C-C-C-CC-C	9	2	0	
C-CC-C-C-C-CC-C	9	9	1	Serpentes (1)
C-CC-C-C-C-C-C-C	9	5	0	
C-C-CC-CCC-C-C	9	1	0	
C-C-CC-CC-C-C-C	9	1	0	
C-C-CC-C-C-C-C-C	9	23	0	
C-C-C-CC-C-C-C-C	9	3	0	
C-C-C-C-CC-C-C-C	9	12	0	
C-C-C-C-C-C-CC-C	9	5	0	
C-C-C-C-C-C-C-CC	9	1	0	
C-C-C-C-C-C-C-C-C	9	14	0	
C-CC-C-C-C-C-CC-C-C	11	1	0	
C-CC-C-C-C-C-C-C-C-C	11	8	0	
C-C-CC-C-C-C-C-C-C-C	11	1	0	
C-C-C-CC-C-C-C-C-C-C	11	1	0	
C-C-C-C-C-C-C-C-CC-C	11	3	0	
C-C-C-C-C-C-C-C-C-CC	11	1	0	
C-C-C-CC-C-C-C-C-C-C-C	13	121	0	
C-C-C-C-C-C-C-C-C-CC-C	13	6	0	
C-C-C-CC-C-C-C-CC-C-C-C-C	15	3	0	

Among these 20,206 sequences, 782 (4%) display a new cysteine pattern and 4,393 (22%) share less than 50% sequence identity with any described peptide or protein in the UniprotKB/SwissProt database using default settings (e-value max =1). It must be noted that the cysteine connectivity of the native toxins can not be explicitly determined using this pipeline.

From sequences to a high throughput screening compatible toxin bank. In order to allow screening and characterization of peptide activity, we developed two high-throughput production protocols. 5,923 sequences representing the broad diversity of the sequence database were selected and produced in the dedicated time of 18 months. Coming from 64 different species, they cover 91 different cysteine patterns, including 14 with an odd number of cysteine residues (Tables 1 and 2).
We selected 931 peptides shorter than 35 residues, with or without PTMs, to be produced by solid phase chemical synthesis[28]. The crude synthetic product was cleaved from the resin, desalted and subjected to an optimized oxidation process, whatever the number or pattern of cysteines[28]. After a final desalting step, liquid chromatography and mass analysis were performed to validate the presence of the expected oxidized mass (Figure S1). 893 toxins from 9 to 34 residues with 1 to 3 disulfide bridges were successfully produced, representing a success rate of 96%.
The remaining 4,992 sequences containing more than 34 residues were selected for heterologous expression in *E. coli*[29–31]. All synthetic genes encoding toxin sequences were subcloned into an optimized periplasmic expression plasmid containing an N-terminal *E. coli* Disulfide bond Isomerase C (DsbC) chaperone fusion tag, a 6xHIS tag for purification and a tobacco etch virus (TEV) protease cleavage site to release the native peptide from the fusion protein. Expressed fusion proteins are purified on a Nickel resin, subjected to overnight TEV cleavage and finally desalted. The correct mass was confirmed by mass analysis (Figure S2). 2,704 toxins, from 35 to 110 residues with 1 to 9 disulfide bridges were successfully produced representing a success rate of 54%.
Determining disulfide connectivity in native toxins is a challenge and the Venomics project did not have the resources to carry out this task. In the absence of this information, cysteine residues were randomly oxidized and the vast majority of toxins presented a unique peak in LC, suggesting that they adopted only one conformation, hopefully the natural one. In the case where LC profiles presented two or three peaks with the good expected oxidized mass, we assumed that it could be due to the generation of peptide isomers[32] (Figures S1 and S2).
Together, these two strategies resulted in a bank of 3,597 peptides composed of 71 different cysteine frameworks (Table 1). In addition, 158 toxins (4%) display a novel cysteine pattern and 1,026 (29%) share less than 50% sequence identity with any previously described peptide or protein in the UniprotKB/SwissProt database. All together, we have successfully translated the high complexity of the omics bank into a physical toxin bank via high throughput methods.

A rich source of GPCRs ligands. Animal toxins are mostly known for their bioactivity on ion channels. Few but very interesting toxins active on GPCRs have also been identified[33]. For example, the commercially available type 2 diabetes treatment, Byetta (derived from a venomous lizard) targets a class A GPCR. We selected four GPCRs linked to unmet medical needs in therapeutic areas such as diabetes, obesity, inflammation and allergy. A competitive binding strategy against a radio-labelled ligand was used for screening. For each receptor, binding conditions were optimized to obtain a Z'-factor above 0.5. The Venomics bank was tested once and 468 hits were obtained. The two first targets provided an amazing high hit rate of 4 and 9%. At the opposite, the two others gave very poor results (Table 4).

Table 4. Hit rate following the Venomics bank screening on different GPCRs.

Targets	Hit rate Global (%)	Origin of hits (%)	
		chemical synthesis	recombinant expression
GPCR1	9	57	43
GPCR2	4	11	89
GPCR3	0.3	64	36
GPCR4	0	0	0

First animal toxins on melanocortin receptors. We concentrated our study on the best-hit receptor, the human melanocortin 4 receptor (hMC4R). This receptor regulates food intake and its genetic invalidation is responsible of morbid obesity. Hit diversity is extraordinarily high in terms of toxin origin, number of disulfide bridges, length (Figure 2) and cysteine pattern (Table 5). Of the 286 hMC4R hits, 80 representative sequences were tested by dose-response characterization (data not shown), displaying affinities from nM to low µM range (Figure 2). Activity of eight hits could not be confirmed, giving a success rate of 90%. Half of the hMC4R-active peptides have an unknown cysteine pattern (Table 5) or share less than 50% sequence identity with toxins in other databases. The ICK-related and the CSαβ-related cysteine patterns are particularly well represented. Toxins with these patterns are renowned for their activities on ion channels. We selected one member of each of these structural families to perform a detailed structural and pharmacological characterization.

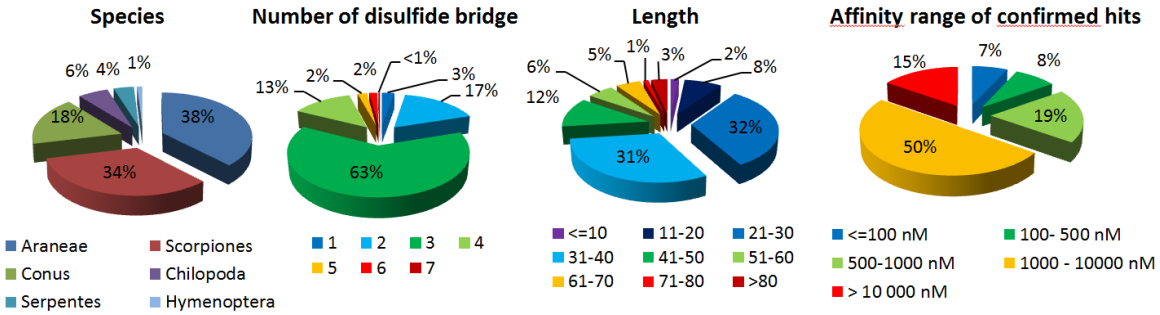


Figure 2. Statistical representation of hits following Venomics bank screening on hMC4R. Affinities of 80 toxins representative of cysteine pattern diversity were measured by binding inhibition of ^{125}I-NDP-α-MSH on CHO cells stably expressing hMC4R.

Table 5. Cysteine pattern diversity of hMC4R hits. Newly described patterns are highlighted in grey.

Cysteine frameworks	N° of Cys	N° of seq	% of total hit	Taxa (number of sequences)
C-C	2	8	3%	Conus (5), Araneae (1), Scorpiones (1), Serpentes (1)
CC-CC	4	16	6%	Conus (13), Araneae (1), Serpentes (1), Undefined (1)
CC-C-C	4	4	1%	Conus (3), Araneae (1)
C-C-CC	4	4	1%	Conus (4)
C-C-C-C	4	25	9%	Conus (14), Chilopoda (5), Scorpiones (5), Araneae (1)
CC-C-C-CC	6	7	2%	Araneae (4), Conus (3)
C-C-CC-C-C	6	97	34%	Araneae (86), Conus (6), Scorpiones (3), Hymenoptera (1), Serpentes (1)
C-C-C-C-CC	6	4	1%	Chilopoda (3), Conus (1)
C-C-C-C-C-C	6	71	25%	Scorpiones (62), Araneae (5), Chilopoda (2), Conus (1), Serpentes (1)
C-C-CC-CC-C	7	1	<1%	Araneae (1)
C-C-CC-C-C-C	7	1	<1%	Araneae (1)
C-C-C-CC-C-C	7	1	<1%	Araneae (1)
C-C-CC-C-C-C-C	8	5	2%	Scorpiones (4), Araneae (1)
C-C-C-CCC-C-C	8	1	<1%	Serpentes (1)
C-C-C-CC-C-C-C	8	9	3%	Scorpiones (5), Chilopda (3), Serpentes (1)
C-C-C-C-C-CC-C	8	10	3%	Serpentes (5), Chilopoda (3), Scorpiones (2)
C-C-C-C-C-C-C-C	8	10	3%	Scorpiones (9), Terebridae (1)
C-C-CC-C-C-C-C-C-C	10	2	1%	Araneae (2)
C-C-C-C-C-C-C-C-C-C	10	4	1%	Scorpiones (2), Araneae (1), Serpentes (1)
CC-C-C-C-C-C-C-C-C-C	12	1	<1%	Serpentes (1)
C-C-C-C-C-CC-C-C-C-C	12	4	1%	Scorpiones (3), Chilopoda (1)
C-C-C-CC-C-C-C-C-C-C-C	14	1	<1%	Araneae (1)

The ICK-related N-TRTX-Preg1a comes from the spider *Poecilotheria regalis* while the CSαβ-related N-BUTX-Ptr1a from the scorpion *Parabuthus transvaalicus*. N letter (uppercase nu greek letter) has been chosen as the activity descriptor for modulators of hMCRs. All the following work was done with synthetically produced toxins of > 95% purity (Figures 3 and S3). N-TRTX-Preg1a and N-BUTX-Ptr1a display low micromolar affinities on every studied subtype of melanocortin receptors stably expressed in CHO-K1 cells for hMC1R, hMC4R, hMC5R and HEK293 cells for hMC3R (Table 6, Figure 4). Agonistic and antagonistic functions on the signaling pathways hMC1R/Gs and hMC4R/Gs were assessed on HEK293 cells. To control the experiments, it was demonstrated that neither toxin induced any effect on control hMCR-free HEK293. Both toxins tested at 100 µM induced cAMP production on hMC1R, similar to 1 nM of α-MSH, while no effect was observed on hMC4R (Figure 4B). In the presence of 1 nM of α-MSH, the toxins slightly enhance hMC1R related cAMP production but did not reduce the cAMP production related to hMC4R (Figure 4B). Both toxins were tested at 100 µM on several sodium channels ($Na_V1.1$, $Na_V1.2$, $Na_V1.3$, $Na_V1.4$, $Na_V1.5$, $Na_V1.6$, $Na_V1.7$ and $Na_V1.8$), potassium channels ($K_V1.1$, $K_V1.2$, $K_V1.3$, $K_V1.4$, $K_V1.5$, $K_V1.6$, $K_V2.1$, $K_V3.1$, $K_V4.3$, $K_V7.1$, $K_V7.3$, $K_V10.1$, $K_V11.1$(hERG), GIRK1), a voltage-gated calcium channel ($Ca_V3.3$) and ligand-gated ion channels nAChR $\alpha1_2\beta1\gamma\delta$ and nAChR α7). No activity could be detected (Figure S4 and S5).

Table 6. Inhibition constant (K_i) values of N-TRTX-Preg1a, N-BUTX-Ptr1a on melanocortin receptors, mean (µM) ± SEM (n=3-10).

	hMC1R	hMC3R	hMC4R	hMC5R
N-TRTX-Preg1a	**1.8 ± 0.3**	**19.8 ± 3.5**	**7.1 ± 1.5**	**10.0 ± 4.4**
N-BUTX-Ptr1a	**2.9 ± 0.6**	**3.9 ± 0.7**	**2.6 ± 0.4**	**2.2 ± 0.8**

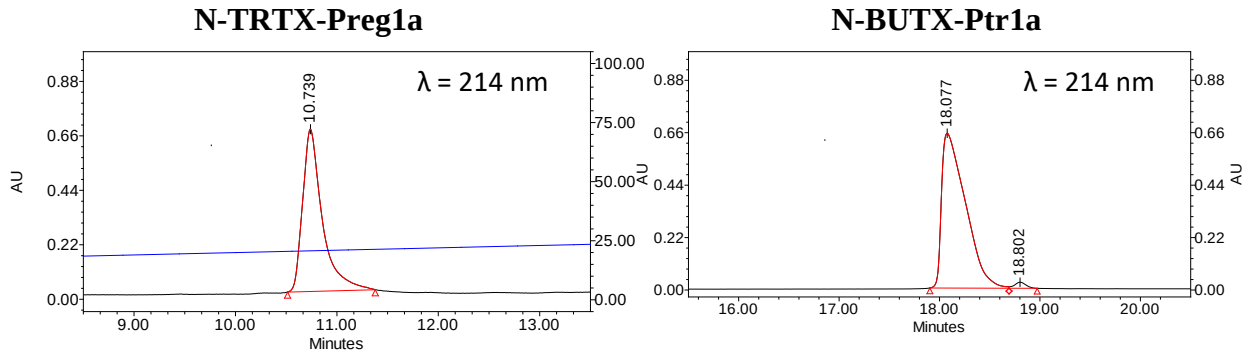


Figure 3. Analytical HPLC chromatogram of the final purified products of N-TRTX-Preg1a and N-BUTX-Ptr1a. Between 10 and 15 µg of each toxin were analyzed on a X Bridge BEH 300Å 5 µm 4.6 x 250 mm with a gradient of 10-50% CH_3CN/TFA 0.1% in 40 min at a flow rate of 1 mL/min. A spontaneous transformation of the N-terminal glutamine (Q form) residue of the peptide N-BUTX-Ptr1a into a pyroglutamic acid (Z form) can be detected after crude peptide deprotection. According to LC-MS analysis performed on Elute Sp Pump HPG 700 (Brucker) coupled with AmaZon SL LC/MS (Bruker), the peak at 18.077 min corresponds to the Q form (average mass of 3795.70 m/z) and represents 96.5% of the total area. The peak at 18.802 min with an average mass of 3778.70 m/z seems to correspond to the Z form of the peptide and represents 3.5% of the total area.

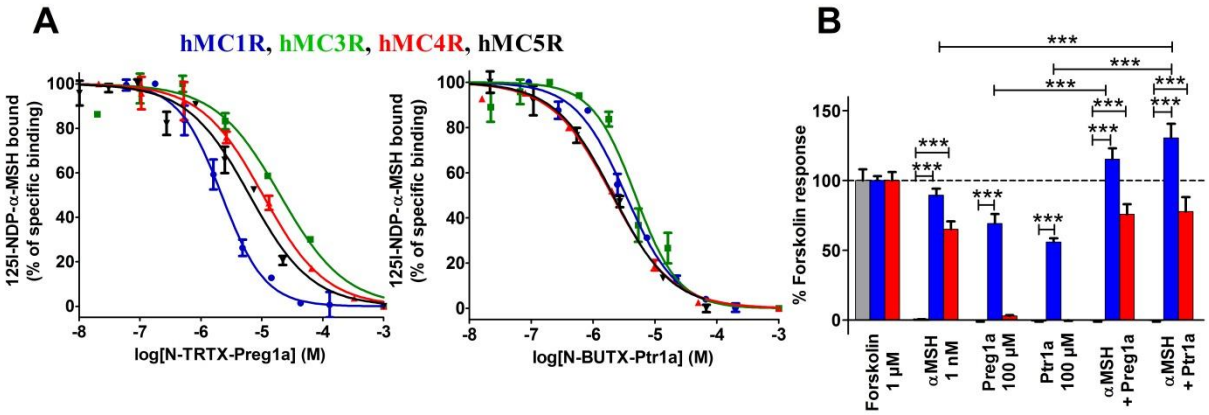


Figure 4. Pharmacological characterization of N-TRTX-Preg1a and N-BUTX-Ptr1a. (A) Representative curves of binding of ^{125}I-NDP-α-MSH on hMC1R, hMC4R and hMC5R (stably expressed in CHO cells) and on hMC3R (stably expressing in HEK293 cells) in the presence of N-TRTX-Preg1a (left panel) or N-BUTX-Ptr1a (right panel). Mean (%) ± SD (n=2-3). (B) Functional assays on HEK293 (grey), HEK293 stably expressing hMC1R (blue) or HEK293 transitory expressing hMC4R (red). cAMP production expressed as % of the positive control (1 µM forskolin). Preg1a: N-TRTX-Preg1a. Ptr1a: N-BUTX-Ptr1a. Mean (%) ± SEM (n = 3), ***: $p<0.001$; one-way analysis of variance (ANOVA) followed by Dunnett's multiple comparison test.

Are N-TRTX-Preg1a and N-BUTX-Ptr1a real ICK-related and CSαβ-related scaffold toxins ? The two N-TRTX-Preg1a and N-BUTX-Ptr1a toxins have a cysteine pattern corresponding to the ICK-related and the CSαβ-related structure with the following pairing C1-C4, C2-C5, C3-C6. As both toxins display original pharmacological activities, we characterized both structures by NMR and determined experimentally their cysteine connectivity. Both peptides are well-folded, illustrated by their good dispersion of resonances. All the ^{1}H (except for the methyl groups of Met2 and Met28 of N-BUTX-Ptr1a), the ^{15}N chemical shifts and the ^{13}C resonances of Cα and Cβ were assigned. Depositions in the BioMagResBank were done with the entry code 34117 for N-TRTX-Preg1a and 34118 for N-BUTX-Ptr1a.

A set of 546 NOEs for N-TRTX-Preg1a were used in the final ARIA run, which includes 454 unambiguous restraints (Table S1). Additional constraints were introduced during the iterative ARIA procedure: 38 dihedral angle restraints, 10 H-bond constraints observed within the β-strands and 3 covalent bonds between sulphur atoms (C2-C19, C9-C22, C18-C27). Indeed, when enough experimental data had been introduced into the first runs of the iterative ARIA process, the C18-C27 disulfide bond is formed in two-thirds of the N-TRTX-Preg1a structures. The fine analysis of NOE constraints and the evaluation of distance between sulfur atoms allowed discrimination between the two remaining possibilities: C2-C19, C9-C22, C18-C27 (favorable) or C2-C22, C9-C19, C18-C27 (excluded). Therefore, the corresponding covalent bonds were added between sulfur atoms to refine the N-TRTX-Preg1a structure in the last runs of the calculations. The 10 models representative of the solution structure (Figure 5A) display a very low RMSD, small potential energy values, respect all the experimental data and are in perfect agreement with standard covalent geometry with 64.8% of the residues in the most favored regions of the Ramachandran plot, and 35.2% in additional allowed regions (Table S1). The N-TRTX-Preg1a structure is very well defined, organized into the ICK motif without ambiguities regarding cysteine pairing. It is composed of a short antiparallel β-sheet made up of two short strands (22–23 and 26–27). The very short turn between these two strands, containing two arginine residues (R24 and R25), is highly constrained. Note the presence of an extended region (7–9), which faces the two-stranded β-sheet, resembling a third quasi-β-strand. Proline residue 12 is in a majority trans configuration (about 80%). N-TRTX-Preg1a sequence produced only a handful of significant hits in public databases and shares only 61% identity with the closest toxin with a known activity, Hs1a from *Cyriopagopus schmidti* (Figure 5C), a calcium channel blocker[36].

Regarding structural determination of N-BUTX-Ptr1a, 839 NOEs (including 558 unambiguous distance restraints) and 4 hydrogen bonds present within the β-sheet were used in the final ARIA run (Table S2). 92% of the 250 final structures display a C1-C4, C2-C5, C3-C6 disulfide bridge array, without adding any constraint on the bonds. The 10 models representative of the solution structure (Figure 5B) display low root mean square deviations (RMSD), small potential energy values, and respect all the experimental data, as attested by the absence of residual violations (Table S2). The Ramachandran plot exhibits 53% of the (ϕ,ψ) angles in the most favored regions, and 40% in the additional allowed regions. The N-BUTX-Ptr1a structure scaffold is a well-known motif within scorpion toxins[48], plant defensins[49] and insect defensins[50], and was designated as a cysteine-stabilized α-helix β-sheet motif (CSαβ motif)[51]. The structure is composed of an α-helix (11–19) and a 2-stranded short β-sheet (26–28; 31–33), stabilized by three disulfide bridges (C6-C27, C12-C32, C16-C34) unambiguously assigned. N-BUTX-Ptr1a shares the same sequence pattern as the majority of α-KTx toxins known to block potassium channels: C-x(4,5)-C-x(3)-C-x(6,9)-[GAS]-K-C-[IMQT]-x(3)-C-x-C (Prosite PS01138). Yet, N-BUTX-Ptr1a shares only 59% identity with the closest toxin with a known activity, Parabutoxin-3 from the same species *Parabuthus transvaalicus* (Figure 5D), a potassium channel blocker[40]. The coordinates of structures have

been deposited in the Protein Data Bank (http://www.pdb.org) with the entry codes 6SAA and 6SAB for N-TRTX-Preg1a and N-BUTX-Ptr1a, respectively.

While N-TRTX-Preg1a and N-BUTX-Ptr1a display the canonical ICK and CSαβ structure, respectively, their sequences diverged so much from most described toxins that homology searches or phylogenetic reconstruction failed to assign them into known toxin groups. On one hand, N-TRTX-Preg1a matched only a handful of spider toxins that could not be linked to the main group of spider ICK toxins. On the other hand, phylogenetic reconstruction with all known alpha-KTx (UniProt) placed N-BUTX-Ptr1a as sister to a clade consisting of several monophyletic alpha-KTx subfamilies (alpha-KTx 2, 4 and 12).

A

B

C

Name	Sequence	%Id	Targets
N-TRTX-Preg1a	RCLHAGAACSGPIQKIPCCG-TCSRRK--CT		MCRs
ω-TRTX-Hs1a	KCLPPGKPCYGATQKIPCCG-VCSHNK--CT	61%	$Ca_v2.2$[37]
U2-TRTX-Pc1a	RCLPAGKTCVRGPMRVPCCG-SCSQNK--CT	54%	Unknown
U6-TRTX-Cg1a	CEPSGKPC-RPLMRIPCCG-SCVRGK--CA	46%	Unknown
ω-CTX-GVIA	CKSPGSSCSP--TSYNCCR-SCNPYTKRCY	29%	$Ca_v2.2$[38]
ω-CTX-MVIIA	CKGKGAKCSR--LMYDCCTGSCRSGK--C-	29%	$Ca_v2.2$[39]

D

Name	Sequence	%Id	Targets
N-BUTX-Ptr1a	QMDMRCSASVECKQKCLKAIGSIFGKCMNKKCKC		MCRs
Parabutoxin-3	EVDMRCKSSKECLVKCKQATGRPNGKCMNRKCKCYPR	59%	$K_v1.1$, $K_v1.2$, $K_v1.3$[40]
Hongotoxin-2	HFIDVKCTTSKECWPPCKAATGKAAGKCMNKKCKCQ	53%	Unknown
Rj1m1	IVIDVPCTVSKECWSACKKAVGTDRGKCMGKKCKCYP	53%	Unknown
Tst26	VFINAKCRGSPECLPKCKQAIGKAAGKCMNGKCKCYP	53%	$K_v1.2$, $K_v1.3$ [41]
Ts7	VFINAKCRGSPECLPKCKEAIGKAAGKCMNGKCKCYP	53%	$K_v1.1$, $K_v1.2$, $K_v1.3$, $K_v1.6$, Shaker K_v [42–44]
Tco30	WCSTCLDLECGASRECYDPCFKAFGRAHGKCMNNKCRCYT	50%	Shaker B KCh[45]
TtBut	WCSTCLDLACGASRECYDPCFKAFGRAHGKCMNNKCRCYT	50%	Unknown
Charybdotoxin	ZFTNVSCTTSKECWSVCQRLHNTSRGKCMNKKCRCYS	41%	$K_v1.2$, $K_v1.3$, $K_v1.5$, $K_v1.6$, KCa1.1, Shaker K_v[22,46]
Agitoxin-2	GVPINVSCTGSPQCIKPCKDA-GMRFGKCMNRKCHCTPK	38%	$K_v1.1$, $K_v1.2$, $K_v1.3$, $K_v1.6$[46,47]

Figure 5. Structural characterization of N-TRTX-Preg1a and N-BUTX-Ptr1a. Primary structure and overlay of the 10 NMR models representative of the solution structure of N-TRTX-Preg1a (A) and N-BUTX-Ptr1a (B). N and C termini are indicated. Comparison of amino acid sequences between N-TRTX-Preg1a (C) or N-BUTX-Ptr1a (D) and the most similar toxins obtained with NCBI Protein Blast program (default settings). Cysteine residues are in bold and orange, conserved residues from N-TRTX-Preg1a or N-BUTX-Ptr1a are in green. Sequence alignment was performed with Clustal Omega (version 1.2.4 from Emboss programs, EBlosum62 matrix for two pair alignment).

Discussion and conclusions

To our knowledge, this is the first time that a thorough '-omic' analysis of venoms has been carried out on such a wide number and diversity of venomous species from various habitats (128 terrestrial, 55 aquatic and 8 aerial) including two European species (*Manica rubida* ant and *Buthus occitanus* scorpion). The *Manica rubida* ant was the smallest species of the project, studied for the first time and 35 sequences were validated. The numbers of validated sequences ranges from 0, due to experimental failures (some collecting conditions were not compatible with proper mRNA extraction), to 1,163, in the venom of the Mediterranean scorpion *Buthus occitanus.* Finally, no validated toxin sequences could be detected in the skin secretions of *Silurus glanis* fish and the *Trachycephalus resinifictrix* frog. An average of 106 toxin sequences per venom were validated, much less than what could be expected according to the assumed richness of venoms. This difference highlights the difficulties in obtaining full and unambiguous toxin sequences. The gotten sequence diversity also highlights our very limited state of knowledge of the toxin world. Indeed, almost one quarter of the toxins share less than 50% sequence identity with known toxin sequences, suggesting unpredictable and original biological activities. Furthermore, 65 new naturally occurring cysteine patterns have been found, suggesting 65 new structural organizations. This is an opportunity to discover peptides with attractive biophysical properties in terms of engineering and thermodynamic stability for optimal drugability. Another level of complexity, not investigated here, is the dimerization (homo- and/or hetero-) of toxins containing an odd number of cysteines. A specific proteomic analysis should be designed to address this point.

Converting a sequence database into physical peptides is challenging, in particular for the oxidation of disulfide bridges. Our recombinant expression strategy was able to directly produce oxidized peptide and has proven to be effective in terms of high throughput production. The two-step chemical synthesis strategy was also extremely efficient. This is mainly thanks to the use of a universal folding strategy applied on all toxins regardless of size or number of cysteines. These production strategies provided peptides with impurities, as no high-efficiency purification step was performed. While the presence of these contaminations could have caused numerous false positive results during the screening tests[52] we shown this was not the case as 90% of the tested hits were confirmed.

GPCRs constitute one of the largest and most versatile families of receptors. They are involved in a wide spectrum of physiological functions. Since the late 20th century, pharmaceutical companies have significantly diminished their natural product programs, partly because of the advances in both HTS and combinatorial synthesis, thereby the creation of large synthetic libraries of small but artificial molecules. Natural product libraries consist of extracts and partially purified fractions and are perceived to be incompatible with HTS. Unfortunately, this loss of interest in natural products correlated with the overall reduction in new leads in the drug development pipeline and the substantial decline in new drug approval[53]. The Venomics bank reconciles venoms, a particularly difficult exploitable resource of natural peptides, with HTS. The screening of the Venomics bank provided an extraordinary hit rate of 3%, much higher than the one reported in literature concerning artificial molecules and estimated at around 0.05%[54,55]. To our knowledge, this is the first large-scale demonstration of the power of natural peptide toxins to provide GPCR ligands. This strategy opens up an unexplored avenue for the discovery of new ligands on receptors linked to diseases with unmet medical needs.

MC4R is mainly expressed in the paraventricular nucleus of the hypothalamus. Its activity is finely regulated by both anorexigenic agonists, the melanocyte stimulating hormones (MSH), and the orexigenic antagonist, Agouti-related protein (AgRP)[56]. Loss-of-function mutations on hMC4R are the primary cause of monogenic obesity (http://www.orpha.net/) making it a therapeutic target in eating disorders. Currently, only one MC4R ligand is in clinical trials

(Setmelanotide, https://clinicaltrials.gov/). Subtype 1 (MC1R) is also of therapeutic interest with its protective role against melanoma. It elicits eumelanin production, a dark, UV radiation-absorbing, photoprotective pigment[57] and promotes nucleotide excision repair of DNA damaged by UV radiation[58]. There is an urgent need for development of innovative ligands to better understand MCR functions and to generate new medicines. N-TRTX-Preg1a and N-BUTX-Ptr1a are the first MCR venom-peptide ligands, structurally unrelated to natural endogenous MCR agonists. Very interestingly, despite having two different structures, they both activate hMC1R without exhibiting the HFRW pharmacophore, previously described in endogenous agonists as the minimal active sequence[59,60]. In addition, they promote a synergistic effect with endogenous agonist α-MSH. This activity deserves to be finely studied as both ligands may activate hMC1R by a new molecular mechanism. N-TRTX-Preg1a and N-BUTX-Ptr1a inhibit ^{125}I-NDP-α-MSH binding on hMC4R but are not able to modulate activate or antagonize the hMC4R/Gs pathway, under our experimental conditions. This discrepancy can comes from our experimental procedure or from a biased function of the two toxins. A deeper exploration of the functions of the two toxins on various MC4R signaling pathways (Gq, β-arrestin, Kinases, etc.) will be performed to better understand their modes of action.

N-TRTX-Preg1a exhibits a common ICK fold only described so far for activity on ion channels. Its sequence is close to calcium channel toxins but does not possess the tyrosine residue in position 16 (N-TRTX-Preg1a numbering), described to be essential for blocking $Ca_V2.2$ channel conductivity[61]. N-BUTX-Ptr1a, close in sequence to potassium channel toxins, possesses a dyad composed of a basic and an aromatic residue[62,63] (F24 and K26, N-BUTX-Ptr1a numbering). This dyad, stated as essential for potassium activity[62,63] but we shown that is clearly not sufficient for N-BUTX-Ptr1a to bind potassium channels. N-TRTX-Preg1a and N-BUTX-Ptr1a sequences diverge so much within their own group that it was even impossible to phylogenetically anchor them into known toxin families. Thus, they constitute new clades for ICK and CSαβ families. Several other toxins of the produced bank sharing more than 80% of sequence identity with these two toxins are present in the Venomics database, suggesting that N-TRTX-Preg1a and N-BUTX-Ptr1a are not the only members of these new families. Production and pharmacological characterization of these natural variants should provide interesting insights with regard to the mode of action of these toxins on hMCRs. This may aid the generation of ligands with improved affinity and selectivity for development of new hMC1R agonists with applications in the treatment and/or prevention of melanoma.

To our knowledge, the Venomics bank is the largest one ever built. We demonstrated that it represents a real opportunity to identify ligands with unusual structures and pharmacological properties. Its high hit rate shows that quality can beat quantity. Two toxins, with classical toxin structures but original pharmacological activities highlight the diversity of toxin properties and open the door to the development of new hMCR modulators.

Experimental Section

Sourcing. We paid important caution to respect the natural resources to which we had access. Thus, we limited the number of collected specimens to the strict number necessary to ensure the amounts and volumes of biological samples were compatible with the experiments to be carried out within the Venomics consortium. For instance, during the Mayotte and Polynesia expeditions, supernumerary cone snails collected during the day, were systematically dropped back to water. In the case of the venomous lizard specimen (*Heloderma horridum exasperatum*), the lizard was anaesthetized while a veterinary surgeon performed surgery to remove only one of the venomous salivary glands of the animal. The animal fully recovered. This demonstrates the applicability of such a strategy of tissue extraction that could certainly be applied with success to rare snakes too.

Transcriptomic analysis. The Illumina paired-end NGS platform was selected to assure sequence and data quality. 203 RNA samples from different venomous species were sequenced through a low-input method. Quality and quantity of total RNA was determined using a Bio-analyzer 2100 (Agilent, California), and Qubit 2.0 (ThermoFisher, Massachussetts). The poly(A)+ mRNA fraction was isolated from total RNA and cDNA libraries were obtained following Illumina´s recommendations. Briefly, poly(A)+ RNA was isolated on poly-T oligo-attached magnetic beads and chemically fragmented prior to reverse transcription and cDNA generation. cDNA fragments then went through an end repair process, the addition of a single 'A' base to the 3' end and then ligation of the adapters. Finally, the products were purified and enriched with PCR to create the indexed final double stranded cDNA library. Library quality was analyzed via a Bioanalyzer 2100 High Sensitivity assay and quantity was determined by real-time PCR using a LightCycler 480 (Roche, Basilea, Suiza). The pool of libraries was sequenced by paired-end sequencing (100 x 2) on an Illumina HiSeq 2000 sequencer (Illumina, California). Bioinformatic analysis was a crucial step of the process to identify peptide sequences. This analysis started with NGS raw data quality control, followed by exhaustive cleaning of primers, adapters, indexes and low-quality bases. The next step consisted of joining the paired-end reads and dereplicating all reads before entering the assembly step itself. To assure the quality of the assembly two different assemblers were used; Velvet (v1.2.10)[64] and Trinity (v2.1.1)[65], creating a robust and comprehensive *de novo* assembly pipeline.

Proteomic analysis. Venoms predominantly containing toxins with less than 25 amino acid residues such as cone snails and wasps, were treated with dithiothreitol (10 mM final concentration, 1 hour at 56 °C) and alkylated with iodoacetamide (20 mM final concentration, 1 hour at room temperature). Reduced and alkylated peptidic toxins were separated by reverse phase chromatography (nanoLC-UPLC Acquity, Waters, Milford, Massachusetts, United Stated) and analyzed by tandem mass spectrometry (Q-Exactive, ThermoFisher, Waltham, Massachusetts ,United States). This "top-down" proteomics approach has the advantage of measuring the mass of the intact full-size toxins but also allows for sequence determination using collision-based fragmentation mass spectrometry experiments. Venoms containing large toxins, such as snake, spider and scorpion venoms, were, after reduction and alkylation, enzymatically cleaved (by using a ratio of 1/50 protein/trypsin, with incubation at 37°C over 1 hour). These cleavage conditions yield truncated peptides cleaved at different positions, with several overlapping peptides covering the full toxin sequence. The peptides were then analyzed with the same nanoUPLC-MS/MS platform, in the same way as the smaller toxins permitting the full toxin sequences to be determined. As large toxins have been transformed into a collection of smaller peptides, the fragmentation yield is then much better and MS/MS spectra can provide useful sequence information.

Sequence validation. For the transcriptomics data, different databases were generated: (1) the first one including genes coding for peptidic toxins by using predictive bioinformatics tools

and (2) the rest of the venom gland transcriptome translated into the six possible open reading frames. The bioinformatics data analysis has been supported by Peaks software (Bioinformatics Solutions Inc., Waterloo, Canada). In addition to classical data analysis, Peaks was also used to *de novo* sequence peptides not identified from the databases, and to predict mutations with Spider algorithms. All the peptides discovered was classified and only those with a false discovery rate of less than 0.1% were considered for inclusion in the Venomics database.

High-throughput peptide production by solid phase synthesis. The high-throughput chemical production of toxins up to 34 residues with or without post-translational modifications (PTMs) was undertaken by solid phase peptide synthesis on an automated Prelude synthesizer (Protein Technologies, USA) equipped with six reactions vessels using the Fmoc strategy. Amino acids were coupled twice for 10 minutes using five or twenty equivalents for toxins up to 25 residues or between 26 and 34 residues, respectively, on a ChemMatrix® or pre-loaded Wang resins using HCTU as a coupling reagent. Both strategies lead to toxins with excellent purities higher than 50%. After a desalting step on homemade C18 columns and a drying step, the refolding of the different toxins was performed directly on the crude material. A universal folding protocol was applied for each toxin as already described[28] using 50 mM HEPES buffer pH 7.5, 20% acetronitrile, 1 mM cysteine, 0.1 mM cystine, over 48 hours at 4 °C. After a desalting step and analytical characterizations by analytical HPLC on a Waters e2695 separation module equipped with a XBridge BEH300 C18 column (5µm , 4.6x250mm, Waters) and by mass spectrometry on a MALDI-TOF spectrometer (4800 MALDI TOF-TOF Analyser from Applied Biosystems, USA), the toxins were distributed in 96 well plates at a concentration between 2-120 µM.

High-throughput peptide production by recombinant expression. Optimization of gene synthesis and expression clone constructs for venom-peptide production have been published previously[31], in addition to the final protocols used in this project[29]. Briefly, genes encoding animal venom peptides were designed with codon optimization for expression in *E. coli* using the ATGenium codon optimization algorithm (NZYTech genes & enzymes, Portugal). Synthetic genes were produced in a high-throughput pipeline and purified PCR products were subsequently cloned into the pHTP4 expression vector using the NZYEasy Cloning kit (NZYTech genes & enzymes, Portugal), according to ligation independent cloning (LIC) technology protocols. In pHTP4, peptide genes are placed under the control of a T7 promoter, and are produced as fusion proteins in the *E. coli* periplasmic space. Fusions contain an N-terminal Disulphide bond isomerase C (DsbC) fusion, an internal hexa-histidine tag for purification and a TEV protease recognition sequence to allow cleavage and isolation of the native peptides. Optimization of expression and purification conditions[30,31] and the final recombinant production protocols used herein[29] have been described previously. All steps were carried out in 24-deep-well plates or in 96-deep-well plates on a Tecan Freedom EVO 200 robot (Switzerland)[66]. pHTP4 expression plasmids were transformed into BL21 (DE3) pLysS *E. coli* and cultured overnight at 25 °C before harvesting. Cell pellets were lysed enzymatically and by sonication. DsbC-His-peptide fusions were isolated from the soluble cell lysates by nickel affinity purification. Concentrations of the 96 purified DsbC-His-fusion proteins were estimated by UV absorbance at 280 nm in a microtiter plate reader (GENios Plus, TECAN, Switzerland). TEVsh protease was added (1/10 w/w) directly to the eluted fusion proteins. The mixture was supplemented with a final concentration of 0.1 mM DTT then incubated overnight at room temperature in a Microtron shaking incubator (INFORS-HT, Switzerland) (200 rpm). Finally, samples were acidified for 1 h with a final concentration of 5% acetonitrile and 0.1% formic acid (FA) at room temperature with mild agitation. TEV protease were precipitated during this acidification process and the 96 recombinant toxins per plate were subjected to a final purification step on C18 chromatographic resin by Solid Phase

Extraction (SPE). The target peptides were eluted from the C18 resin, and analyzed by UHPLC and MS to confirm the identity and the quantity of the toxin (Figure S2). The toxins were distributed in 96 well plates at a concentration between 1-10 µM.

Production of N-TRTX-Preg1a and N-BUTX-Ptr1a toxins. Linear peptide was obtained by solid phase chemical synthesis following the same protocol described above using Fmoc-Thr(tBu)-Wang resin LL (Novabiochem®) for N-TRTX-Preg1a and H-Cys(Trt)-Trityl NovaPEG resin (Novabiochem®) for N-BUTX-Ptr1a. After resin cleavage and ether precipitation, linear peptide was then purified on Waters Preparative HPLC 600 with an XBridge BEH 300Å Prep C18 column (5 µm, 19 x 250 mm, Waters). The oxidation step was performed using Venomics oxidation conditions[28]. Resulting oxidized peptide was finally purified on a Waters 600 controller using an XBridge BEH 300Å Prep C18 column (5 µm, 10 x 250 mm, Waters). Purity of the final product was assessed separately by MALDI-TOF spectrometry (4800 MALDI TOF-TOF Analyser (Applied Biosystems, USA)) and analytical HPLC on Waters e2695 separation module with an XBridge BEH 300Å C18 column (5 µm, 4.6 x 250 mm, Waters). Final quantity was measured by amino-acid quantification using the PicoTag kit (Waters Associates, Milford, MA).

Sequence analysis and phylogenetic reconstruction. Sequences were queried against UniProt (release 2019_02) and refseq_protein (release-96) with blastp 2.29.1. Searching parameters were adjusted for short-length sequences. Spider and scorpion venom proteins in UniProt were annotated as in VenomZone (https://venomzone.expasy.org/ release May, 2019). Multiple sequence alignments were performed with mafft 7.310 --auto option. Phylogenetic trees were obtained with RAxML 8.2.11 --PROTGAMMAAUTO --# 100 options, and MrBayes v3.2.0 ngen=1000000 nchains=4 nruns=2.

NMR structures. Solutions of 0.5 mM N-TRTX-Preg1a, and 1.2 mM N-BUTX-Ptr1a were prepared in 3 mm tubes (respectively, 315 µg of N-TRTX-Preg1a and about 1 mg of N-BUTX-Ptr1a were dissolved in 220 µL of H_20/D_20 (95/5), at pH 4.5). All NMR spectra were acquired on a BRUKER 700 MHz NMR spectrometer equipped with a 5 mm TCI cryoprobe, at 298K. The standard set of 2D ^{1}H-NMR experiments (COSY, 80 ms TOCSY, and 160 ms NOESY) were completed with natural abundance ^{15}N- and ^{13}C-HSQC spectra. Processing, analyses, and calculations were successively performed with Bruker's TopSpin3.2, CcpNMR[67] and Aria2.3[68] software. Chemical shift assignments were carried out following the classical procedure, distance constraints were deduced from the volume integration of NOE correlations and introduced into the standard iterative calculation process of the Aria (ambiguous restraints for iterative assignment) software. For each Aria run, 500 structure calculations were performed per iteration, and 250 structures were calculated in the final iteration in water. For both samples, the first runs were performed allowing all combinations of disulfide pairing (ambiguous intersulfur constraints option). This option was kept for the N-BUTX-Ptr1a, for which the calculations converged towards a unique array of disulfide bridges, and the only additional constraints were the H-bonds observed in the β-sheet already formed in the first runs. For the N-TRTX-Preg1a sample, (φ,ψ) angular restraints deduced by the TALOS software[69] from ^{13}C and ^{1}H chemical shifts were added to NOE constraints. Covalent bonds between sulphur atoms involved in each defined bridge were introduced in the last runs of the calculation. Finally, for each peptide, ten of the structures refined in water in the final iteration were selected according to their agreement with experimental data. The final sets of structures were displayed and analyzed using PyMol (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC), and their quality were evaluated using Promotif[70] and PROCHECK3.5[71] software.

Screening. Radioligands and membranes expressing receptors were purchased from Perkin Elmer (Courtaboeuf, France). For each receptor, screening conditions were optimised in order to get a Z'-factor value superior to 0.5. Z'-factor is defined as: $1 - \frac{3MAD\ max + 3MAD\ min}{Median\ max - Median\ min}$,

where MAD equals Median Absolute Deviation. Following screening, peptides with an associated binding value less than Median – $3xMAD_{median}$[72] were classified as hits. Peptides were screened at concentrations between 0.1 to 10 µM according to their production yields.

Radioligand binding assays. The ^{125}I-NDP-α-MSH and melanocortin receptors were purchased from Perkin Elmer (Courtaboeuf, France). Binding experiments were performed using radioligand at 50 pM in a 100-µL reaction mixture at room temperature in buffer composed of 25 mM Hepes pH 7.4, 1 mM $MgCl_2$, 1 mM $CaCl_2$, 0.2 g/L BSA with various concentrations of competitors. Nonspecific binding was measured in the presence of 1 µM NDP-α-MSH (Tocris Bioscience). After 3 hours of incubation time, samples were filtered through 96 well GF/C filter plates preincubated with 0.5% polyethylenimine. 25 µL of MicroScint 0 was added onto each dry filter and the radioactivity was quantified on a TopCount beta counter (Perfinelmer, Courtaboeuf, France) with 45% yield. Competition binding data were fitted to a one-site/state inhibition mass action curve using Prism software, v5.0.0 (GraphPad, La Jolla,CA, USA). IC_{50} values were converted to K_i, using the Cheng–Prusoff equation[73]. Kd values were 0.34 nM, 0.47 nM, 0.25 nM and 0.88 nM, respectively, for hMC1R, hMC3R, hMC4R and hMC5R. Data represent at least three independent experiments performed in duplicate and are presented as mean ± SEM.

Functional assays. Human embryonic kidney cells (HEK293, DSMZ, ACC305, authentication: full-matching STR reference profile of 293 in the database, authentic) and HEK293-MC1R (human embryonic kidney cells expressing the human melanocortin 1 receptor with an N-terminal HA-tag and C-terminal FLAG-tag due to stable transfection, selection marker: hygromycin, AG Beck-Sickinger, University Leipzig, Leipzig, Germany) were grown in Dulbecco's modified Eagle's medium (DMEM) with 4.5 g/L glucose and L-glutamine and Ham's F12 (1:1, v/v; Lonza) supplied with 15% (v/v) heat-inactivated fetal calf serum (FCS; Lonza) in a humidified atmosphere at 37 °C and 5% CO_2. HEK293 or stably transfected HEK293-HA_MC1R_FLAG cells were transfected in solution with 0.15 µl Lipofectamine 2000 (Invitrogen) per well and 30 ng Plasmid pGL4.29[luc2P/CRE/Hygro] (Promega) per well and additionally 30 ng pcDNA3-hMC4R per well when needed. Lipofectamine and plasmid solutions were prepared in 6 µl Opti-MEM medium (Gibco) per well. After 5 min incubation at RT the Lipofectamine solution was slowly added to the DNA solution and incubated for 30 min at RT. Transfection of 20,000 cells/well in 30 µl medium (DMEM/Ham's F12/ 15% FCS) without antibiotics was performed in solution by slowly adding the DNA/Lipofectamine solution to the cells. After careful mixing, 42 µl per well was plated on a 384-well microplate (Greiner bio-one, F-Bottom, white) and incubated overnight. The day after transfection, medium was removed and cells were stimulated with 15 µl peptide solution in serum free medium (1 nM α-MSH (AG Beck-Sickinger, University Leipzig, Leipzig, Germany) corresponding to ~EC_{85} at hMC1R or ~EC_{75} at hMC4R, 100 µM N-TRTX-Preg1a or 100 µM N-BUTX-Ptr1a) for 3 h at 37 °C, 5% CO_2. As a positive control 1 µM forskolin (Sigma-Aldrich) was used, while serum-free medium was used as the negative control (baseline). For Luciferase detection 15 µl OneGlo Substrate (Promega) was added at RT to each well. The 384-well plate was transferred directly to the plate reader (Tecan Infinite 200 PRO) and luminescence was measured after shaking for 5 sec (double orbital, amplitude: 3.5 mm) and 5 min incubation with an integration time of 1000 ms at room temperature. Data analysis was performed with GraphPad Prism. After baseline subtraction, data were normalized to stimulation with 1 µM Forskolin. Data represent the mean ± SEM from three independent experiments performed in triplicates. Statistical analysis was performed with GraphPad Prism 5.03 (GraphPad Software, Inc., San Diego, California, USA). Statistical significance was determined by one-way analysis of variance (ANOVA) followed by Dunnett's multiple comparison test. *** $p < 0.001$.

Electrophysiology. For the expression of Na_V channels (human (h)$Na_V1.1$, rat (r)$Na_V1.2$, r$Na_V1.3$, r$Na_V1.4$, h$Na_V1.5$, mouse (m)$Na_V1.6$, h$Na_V1.7$, r$Na_V1.8$ channels and the auxiliary subunits rβ1, hβ1), K_V channels (r$K_V1.1$, r$K_V1.2$, h$K_V1.3$, r$K_V1.4$, r$K_V1.5$, r$K_V1.6$, r$K_V2.1$, h$K_V3.1$, r$K_V4.3$, h$K_V7.2$, h$K_V7.3$, h$K_V10.1$, h$K_V11.1$), Ca_V channels (r$Ca_V3.3$), hGIRK1 and nAChRs (hα1β1γδ and hα7) in *Xenopus laevis* oocytes, the linearized plasmids were transcribed using the T7 or SP6 mMESSAGE-mMACHINE transcription kit (Ambion, Carlsbad, CA). The harvesting of stage V–VI oocytes from anesthetized female *Xenopus laevis* frogs was previously described[74]. Oocytes were injected with 50 nl of cRNA at a concentration of 1 ng/nl using a microinjector (Drummond Scientific, Broomall, PA). The oocytes were incubated in a solution containing 96 mM NaCl, 2 mM KCl, 1.8 mM $CaCl_2$, 2 mM $MgCl_2$, and 5 mM HEPES pH 7.4, supplemented with 50 mg/L gentamycin sulfate.
Two-electrode voltage-clamp recordings were performed at room temperature (18–22 °C) using a Geneclamp 500 amplifier (Molecular Devices, Downingtown, PA) controlled by a pClamp data acquisition system (Axon Instruments, Union City, CA). Whole cell currents from oocytes were recorded 1–4 days after injection. Bath solution composition was 96 mM NaCl, 2 mM KCl, 1.8 mM $CaCl_2$, 2 mM $MgCl_2$, and 5 mM HEPES pH 7.4. Toxins were applied directly to the bath. Resistances of both electrodes were kept between 0.8 and 1.5 MΩ. The elicited currents were filtered at 2 kHz (Na_V), 0.5 kHz (K_V & Ca_V), 0.05 kHz (GIRK & nAChR) and sampled at 20 kHz (Na_V), 2 kHz (K_V & Ca_V) and 0.1 kHz (GIRK & nAChR) using a four-pole low-pass Bessel filter. Leak subtraction was performed using a -P/4 protocol. For the electrophysiological analysis, a number of protocols were applied from a holding potential of −90 mV. $K_V1.1$–$K_V1.6$ currents were evoked by 250 ms depolarizations to 0 mV followed by a 250 ms pulse to −50 mV. $K_V2.1$, $K_V3.1$, and $K_V4.3$ currents were elicited by 250 ms pulses to +20 mV from a holding potential of −90 mV. K_V7 and $K_V10.1$ currents were evoked by 2 s depolarizing pulses to 0 mV from a holding potential of −90 mV. Current traces of $K_V11.1$ channels were elicited by applying a +40 mV prepulse for 2 s followed by a step to −120 mV for 2 s. Sodium current traces were evoked by 100 ms depolarization to the voltage corresponding to maximal sodium current in control conditions. For GIRK measurements, the oocytes were clamped at -70 mV and superfused with a high K+ solution (96 mM KCl, 2 mM NaCl, 1 mM $MgCl_2$, 1.8 mM $CaCl_2$, 5 mM HEPES pH 7.5). Ca_V channel currents, carried by 10 mM Ba^{2+}, were elicited by 200 ms test depolarizations applied from -90 mV to voltages ranging from -50 mV or +30 mV. For nAChRs, oocytes were clamped at a holding potential of −70 mV throughout all recordings. To minimize interference of endogenous Ca^{2+}-activated Cl^- channels during the experiments, a Ca^{2+}-free bath perfusion solution was used, containing 82.5 mM NaCl, 2.5 mM KCl, 1 mM $MgCl_2$, 2.5 mM $BaCl_2$, and 5 mM HEPES pH 7.4. Current responses were evoked by 10 s applications of increasing concentrations of acetylcholine (starting at 200 μM). All data were analyzed using pClamp Clampfit 10.1 (Molecular Devices) and Origin 7.5 software (OriginLab, Northampton, MA).
Toxins nomenclature. Newly characterized toxins were named according to the rational nomenclature proposed by King *et al.*[75] with a greek letter associated with a biological function followed by the generic name indicating from which taxon the toxin is derived, genus and species indicated respectively by uppercase and lowercase letters, a number to distinguish different toxins with similar activity and a lowercase letter for isoforms. In this study, uppercase nu greek letter (N) was chosen as new descriptor of toxins active on melanocortin receptors.

ASSOCIATED CONTENT

The Supporting Information is available free of charge on the ACS Publications website at DOI:

AUTHOR INFORMATION
Corresponding Author
*E-mail: Nicolas.gilles@cea.fr

ACKNOWLEDGMENT

ABBREVIATIONS

ANOVA, one-way analysis of variance; BUTX, butantoxin; Cav, voltage-gated calcium channel; CHO, chinese hamster ovary cells; COSY, correlation spectroscopy; CSαβ, cysteine-stabilized α-helix β-sheet; DsbC, Disulfide bond Isomerase C; DMEM, Dulbecco's modified Eagle's medium; Fmoc, fluorenylmethoxycarbonyl; GIRK, G protein-coupled inwardly-rectifying potassium channel; GPCR, G-Protein Coupled receptors; HCTU, H-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate; HEK, human embryonic kidney cells; HEPES, hydroxyethyl piperazineethanesulfonic acid; HPLC, high-performance liquid chromatography; HT, high-throughput; HT, high-throughput screening; ICK, inhibitory cystine knot; Ki, inhibitory constant; Kv, voltage-gated potassium channel; LC, liquid chromatography; hMC1R, human melanocortin 1 receptor; hMC3R, human melanocortin 3 receptor; hMC4R, human melanocortin 4 receptor; hMC5R, human melanocortin 5 receptor; MALDI-TOF, Matrix Assisted Laser Desorption Ionisation - Time of Flight; MAD, median absolute deviation; MS, mass spectrometry; MSH, melanocyte stimulating hormone; N°, number; Nav, voltage-gated sodium channel; NDP-α-MSH, [Nle4,D-phe7]-alpha-melanocyte stimulating hormone; NGS, next-generation sequencing; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; PTM, post-translational modification; RMSD, root-mean-square deviation; RT, room temperature; SD, standard deviation; SEM, standard error of the mean; Seq, sequence; TOCSY, total correlation spectroscopy; TRTX, theraphotoxin; UPLC, ultra-performance liquid chromatography; UV, ultraviolet.

Funding
This work was supported by the European FP7 project 278346 2011-2015 and by the French Infrastructure for Integrated Structural Biology (FRISBI) ANR-10-INSB-05-01. A PhD grant (S.R.) was supported by the University of Paris-Sud.

References

1. Jenner, R. & Undheim, E. *Venom: The Secrets of Nature's Deadliest Weapon*. (Smithsonian Books, 2017).
2. Degueldre, M. *et al.* 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* **130**, 116–125 (2017).
3. Beraud, E. & Chandy, K. G. Therapeutic potential of peptide toxins that target ion channels. *Inflamm. Allergy Drug Targets* **10**, 322–42 (2011).
4. Lewis, R. J., Dutertre, S., Vetter, I. & Christie, M. J. Conus Venom Peptide Pharmacology. *Pharmacol. Rev.* **64**, 259–298 (2012).
5. Servent, D. & Fruchart-Gaillard, C. Muscarinic toxins: tools for the study of the pharmacological and functional properties of muscarinic receptors. *J. Neurochem.* **109**, 1193–202 (2009).
6. Ciolek, J. *et al.* Green mamba peptide targets type-2 vasopressin receptor against polycystic kidney disease. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 7154–7159 (2017).
7. Petrel, C. *et al.* Identification, structural and pharmacological characterization of τ-CnVA, a conopeptide that selectively interacts with somatostatin sst3 receptor. *Biochem. Pharmacol.* **85**, 1663–71 (2013).
8. Rouget, C. *et al.* Identification of a novel snake peptide toxin displaying high affinity and antagonist behaviour for the α2-adrenoceptors. *Br. J. Pharmacol.* **161**, 1361–74

(2010).
9. Quinton, L. *et al.* Isolation and pharmacological characterization of AdTx1, a natural peptide displaying specific insurmountable antagonism of the α 1A- adrenoceptor. *Br. J. Pharmacol.* **159**, 316–325 (2010).
10. Jolkkonen, M., Adem, A., Hellman, U., Wernstedt, C. & Karlsson, E. A snake toxin against muscarinic acetylcholine receptors: amino acid sequence, subtype specificity and effect on guinea-pig ileum. *Toxicon* **33**, 399–410 (1995).
11. Cologna, C. T. *et al.* Peptidomic comparison and characterization of the major components of the venom of the giant ant Dinoponera quadriceps collected in four different areas of Brazil. *J. Proteomics* **94**, 413–22 (2013).
12. Kumar, V. A., Wickremasinghe, N. C., Shi, S. & Hartgerink, J. D. Nanofibrous Snake Venom Hemostat. *ACS Biomater. Sci. Eng.* **1**, 1300–1305 (2015).
13. Pennington, M. W., Czerwinski, A. & Norton, R. S. Peptide therapeutics from venom: Current status and potential. *Bioorg. Med. Chem.* **26**, 2738–2758 (2018).
14. Saez, N. J. & Herzig, V. Versatile spider venom peptides and their medical and agricultural applications. *Toxicon* **158**, 109–126 (2019).
15. Mebs, D. Toxicity in animals. Trends in evolution? *Toxicon* **39**, 87–96 (2001).
16. Kaas, Q., Westermann, J.-C. & Craik, D. J. Conopeptide characterization and classifications: an analysis using ConoServer. *Toxicon* **55**, 1491–509 (2010).
17. Gilles, N. & Servent, D. The European FP7 Venomics Project. *Futur. Med. Chem* **6**, 1611–1612 (2014).
18. Turner, A. H., Craik, D. J., Kaas, Q. & Schroeder, C. I. Bioactive Compounds Isolated from Neglected Predatory Marine Gastropods. *Mar. Drugs* **16**, (2018).
19. Morsa, D. *et al.* Multi-Enzymatic Limited Digestion: The Next-Generation Sequencing for Proteomics? *J. Proteome Res.* **18**, 2501–2513 (2019).
20. Dutertre, S. *et al.* Deep venomics reveals the mechanism for expanded peptide diversity in cone snail venom. *Mol. Cell. Proteomics* **12**, 312–29 (2013).
21. Rong, M. *et al.* Native pyroglutamation of huwentoxin-IV: a post-translational modification that increases the trapping ability to the sodium channel. *PLoS One* **8**, e65984 (2013).
22. Lucchesi, K., Ravindran, A., Young, H. & Moczydlowski, E. Analysis of the blocking activity of charybdotoxin homologs and iodinated derivatives against Ca2+-activated K+ channels. *J. Membr. Biol.* **109**, 269–81 (1989).
23. Saez, N. J. *et al.* Spider-venom peptides as therapeutics. *Toxins (Basel).* **2**, 2851–71 (2010).
24. Gao, B. *et al.* Functional evolution of scorpion venom peptides with an inhibitor cystine knot fold. *Biosci. Rep.* **33**, 513–527 (2013).
25. Robinson, S. D. & Norton, R. S. Conotoxin gene superfamilies. *Mar. Drugs* **12**, 6058–101 (2014).
26. Imperial, J. S. *et al.* The augertoxins: biochemical characterization of venom components from the toxoglossate gastropod Terebra subulata. *Toxicon* **42**, 391–8 (2003).
27. Lavergne, V. *et al.* Optimized deep-targeted proteotranscriptomic profiling reveals unexplored Conus toxin diversity and novel cysteine frameworks. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E3782-91 (2015).
28. Upert, G. *et al.* High-throughput production of two disulphide-bridge toxins. *Chem. Commun.* **50**, 8408–8411 (2014).
29. Duhoo, Y. *et al.* High-Throughput Production of Oxidized Animal Toxins in Escherichia coli. *Methods Mol. Biol.* **2025**, 165–190 (2019).
30. Sequeira, A. F. *et al.* Gene design, fusion technology and TEV cleavage conditions

influence the purification of oxidized disulphide-rich venom peptides in Escherichia coli. *Microb. Cell Fact.* **16**, 1–16 (2017).
31. Turchetto, J. *et al.* High-throughput expression of animal venom toxins in Escherichia coli to generate a large library of oxidized disulphide-reticulated peptides for drug discovery. *Microb. Cell Fact.* **16**, 1–15 (2017).
32. Massonnet, P. *et al.* Disulfide Connectivity Analysis of Peptides Bearing Two Intramolecular Disulfide Bonds Using MALDI In-Source Decay. *J. Am. Soc. Mass Spectrom.* **29**, 1995–2002 (2018).
33. Muratspahić, E., Freissmuth, M. & Gruber, C. W. Nature-Derived Peptides: A Growing Niche for GPCR Ligand Discovery. *Trends Pharmacol. Sci.* **40**, 309–326 (2019).
34. Pallaghy, P. K., Nielsen, K. J., Craik, D. J. & Norton, R. S. A common structural motif incorporating a cystine knot and a triple-stranded beta-sheet in toxic and inhibitory polypeptides. *Protein Sci.* **3**, 1833–9 (1994).
35. Zhu, S., Darbon, H., Dyason, K., Verdonck, F. & Tytgat, J. Evolutionary origin of inhibitor cystine knot peptides. *FASEB J.* **17**, 1765–1767 (2003).
36. Liu, Z. *et al.* 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). doi:10.1074/jbc.M513542200
37. Liu, Z. *et al.* 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.* **281**, 8628–8635 (2006).
38. Favreau, P. *et al.* A new omega-conotoxin that targets N-type voltage-sensitive calcium channels with unusual specificity. *Biochemistry* **40**, 14567–75 (2001).
39. Wang, F. *et al.* Molecular basis of toxicity of N-type calcium channel inhibitor MVIIA. *Neuropharmacology* **101**, 137–45 (2016).
40. Huys, I. *et al.* Purification, characterization and biosynthesis of parabutoxin 3, a component of Parabuthus transvaalicus venom. *Eur. J. Biochem.* **269**, 1854–65 (2002).
41. Papp, F. *et al.* Tst26, a novel peptide blocker of Kv1.2 and Kv1.3 channels from the venom of Tityus stigmurus. *Toxicon* **54**, 379–89 (2009).
42. Rogowski, R. S., Krueger, B. K., Collins, J. H. & Blaustein, M. P. Tityustoxin K alpha blocks voltage-gated noninactivating K+ channels and unblocks inactivating K+ channels blocked by alpha-dendrotoxin in synaptosomes. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 1475–9 (1994).
43. Cerni, F. A. *et al.* Electrophysiological characterization of Ts6 and Ts7, K+ channel toxins isolated through an improved Tityus serrulatus venom purification procedure. *Toxins (Basel).* **6**, 892–913 (2014).
44. Rodrigues, A. R. A., Arantes, E. C., Monje, F., Stühmer, W. & Varanda, W. A. Tityustoxin-K(alpha) blockade of the voltage-gated potassium channel Kv1.3. *Br. J. Pharmacol.* **139**, 1180–6 (2003).
45. Diego-García, E. *et al.* The Brazilian scorpion Tityus costatus Karsch: genes, peptides and function. *Toxicon* **45**, 273–83 (2005).
46. Garcia, M. L., Garcia-Calvo, M., Hidalgo, P., Lee, A. & MacKinnon, R. Purification and characterization of three inhibitors of voltage-dependent K+ channels from Leiurus quinquestriatus var. hebraeus venom. *Biochemistry* **33**, 6834–9 (1994).
47. Takacs, Z. *et al.* A designer ligand specific for Kv1.3 channels from a scorpion neurotoxin-based library. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 22211–6 (2009).
48. Kuzmenkov, A. I., Grishin, E. V. & Vassilevski, A. A. Diversity of Potassium Channel Ligands: Focus on Scorpion Toxins. *Biochem.* **80**, 1764–1799 (2016).
49. Carvalho, A. de O. & Gomes, V. M. Plant defensins--prospects for the biological functions and biotechnological properties. *Peptides* **30**, 1007–20 (2009).

50. Koehbach, J. Structure-Activity Relationships of Insect Defensins. *Front. Chem.* **5**, 45 (2017).
51. Cornet, B. *et al.* Refined three-dimensional solution structure of insect defensin A. *Structure* **3**, 435–48 (1995).
52. Feng, B. Y. & Shoichet, B. K. Synergy and antagonism of promiscuous inhibition in multiple-compound mixtures. *J. Med. Chem.* **49**, 2151–4 (2006).
53. Li, J. W.-H. & Vederas, J. C. Drug discovery and natural products: end of an era or an endless frontier? *Science* **325**, 161–5 (2009).
54. Bender, A. *et al.* Which aspects of HTS are empirically correlated with downstream success? *Curr. Opin. Drug Discov. Devel.* **11**, 327–337 (2008).
55. Oprea, T. I. Current trends in lead discovery: are we looking for the appropriate properties? *Mol. Divers.* **5**, 199–208 (2002).
56. Andermann, M. L. & Lowell, B. B. Toward a Wiring Diagram Understanding of Appetite Control. *Neuron* **95**, 757–778 (2017).
57. Kobayashi, N. *et al.* Supranuclear melanin caps reduce ultraviolet induced DNA photoproducts in human epidermis. *J. Invest. Dermatol.* **110**, 806–10 (1998).
58. Kadekaro, A. L. *et al.* Melanocortin 1 receptor genotype: an important determinant of the damage response of melanocytes to ultraviolet radiation. *FASEB J.* **24**, 3850–60 (2010).
59. Hruby, V. J. *et al.* alpha-Melanotropin: the minimal active sequence in the frog skin bioassay. *J. Med. Chem.* **30**, 2126–30 (1987).
60. Singh, A., Haslach, E. M. & Haskell-Luevano, C. Structure-activity relationships (SAR) of melanocortin and agouti-related (AGRP) peptides. *Adv. Exp. Med. Biol.* **681**, 1–18 (2010).
61. Sousa, S. R., Vetter, I. & Lewis, R. J. Venom peptides as a rich source of Cav2.2 channel blockers. *Toxins (Basel).* **5**, 286–314 (2013).
62. Gasparini, S. *et al.* Delineation of the functional site of alpha-dendrotoxin. The functional topographies of dendrotoxins are different but share a conserved core with those of other Kv1 potassium channel-blocking toxins. *J. Biol. Chem.* **273**, 25393–403 (1998).
63. Dauplais, M. *et al.* On the convergent evolution of animal toxins. Conservation of a diad of functional residues in potassium channel-blocking toxins with unrelated structures. *J. Biol. Chem.* **272**, 4302–9 (1997).
64. Zerbino, D. R. & Birney, E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* **18**, 821–9 (2008).
65. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–52 (2011).
66. Saez, N. J., Nozach, H., Blemont, M. & Vincentelli, R. High throughput quantitative expression screening and purification applied to recombinant disulfide-rich venom proteins produced in E. coli. *J. Vis. Exp.* e51464 (2014). doi:10.3791/51464
67. Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–96 (2005).
68. Rieping, W. *et al.* ARIA2: automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* **23**, 381–2 (2007).
69. Cornilescu, G., Delaglio, F. & Bax, A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR* **13**, 289–302 (1999).
70. Hutchinson, E. G. & Thornton, J. M. PROMOTIF--a program to identify and analyze structural motifs in proteins. *Protein Sci.* **5**, 212–20 (1996).
71. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R. & Thornton, J. M.

AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–86 (1996).
72. Zhang, Chung & Oldenburg. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J. Biomol. Screen.* **4**, 67–73 (1999).
73. Cheng, Y. & Prusoff, W. H. Relationship between the inhibition constant (K1) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochem. Pharmacol.* **22**, 3099–108 (1973).
74. Peigneur, S. *et al.* Where cone snails and spiders meet: design of small cyclic sodium-channel inhibitors. *FASEB J.* **33**, 3693–3703 (2019).
75. 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* **52**, 264–276 (2008).

Web Sites:
Clinical Trial. https://clinicaltrials.gov (accessed on Sep 21, 2019).
Orphanet. http://www.orpha.net (accessed on Sep 21, 2019).
Protein Data bank. http://www.pdb.org (accessed on Sep 21, 2019).
UniProt. https://www.uniprot.org/ (accessed on Oct 21, 2019).
VenomZone. https://venomzone.expasy.org/ (accessed on Oct 21, 2019).

Annexe 2. Informations supplémentaires de l'article « A Venomics approach coupled to High-Throughput toxin production strategies for the identification of the first melanocortin toxins" soumis au Journal of Chemical Chemistry.

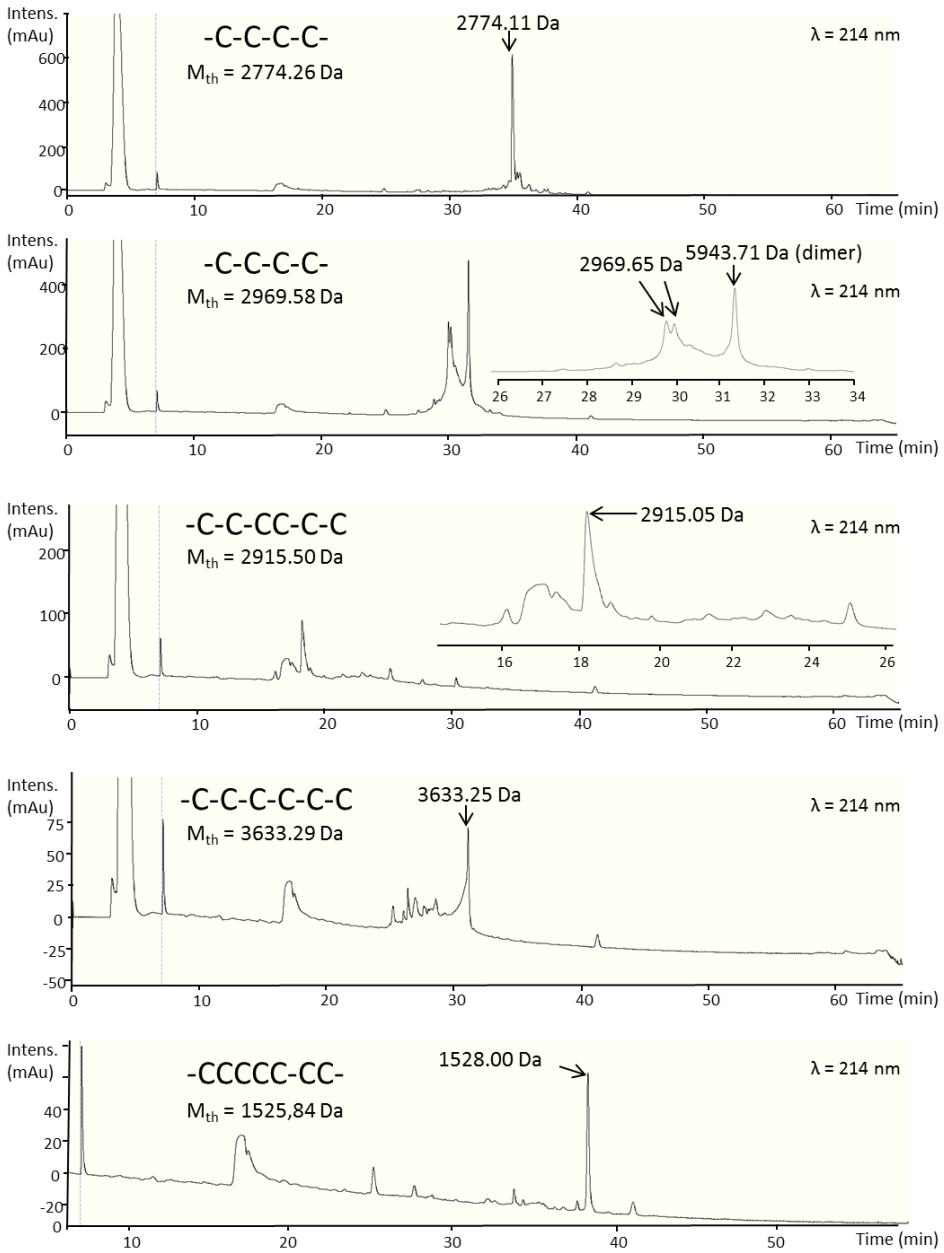


Figure S1. LC-MS analysis of five representative toxins with their cystein patterns, produced by solid phase chemical synthesis. For chemically produced peptides, analysis was performed on Elute Sp Pump HPG 700 (Brucker) coupled with AmaZon SL LC/MS (Bruker) at 214 nm. Column : XBridge® Peptide BEH C18 300Å 3.5µm, 4.6x150 mm (Waters); Gradient : 0-50% CH3CN/TFA0.1% in 50min ; flow rate : 1mL/min. Oxidation of a two disulfide bridge peptide can lead to one or several oxidized forms included dimers whereas oxidation of a three disulfide bridge peptide often lead to only one oxidized form. For –CCCC-CC- previously never described pattern, produced peptide contains only four of the six possible disulfide bridges. The surface of the peak corresponding to the correct mass was integrated for toxin quantification.

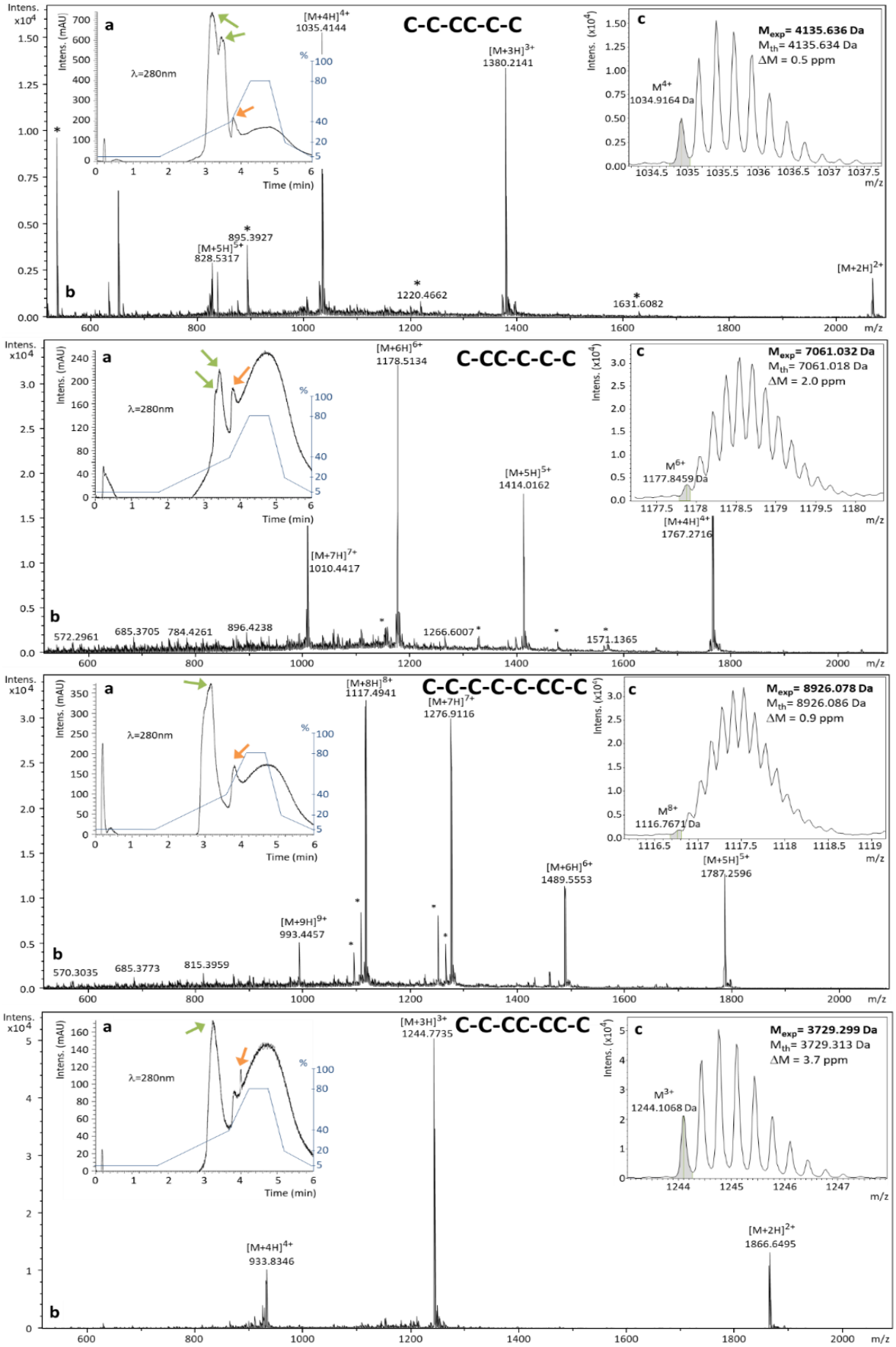

C-C-CC-C-C
[M+4H]4+ 1035.4144
[M+3H]3+ 1380.2141
[M+5H]5+ 828.5317
895.3927
1220.4662
1631.6082
[M+2H]2+
λ=280nm
Time (min)
Mexp= 4135.636 Da
Mth= 4135.634 Da
ΔM = 0.5 ppm
M4+ 1034.9164 Da
C-CC-C-C-C
[M+6H]6+ 1178.5134
[M+5H]5+ 1414.0162
[M+7H]7+ 1010.4417
[M+4H]4+ 1767.2716
572.2961
685.3705
784.4261
896.4238
1266.6007
1571.1365
Mexp= 7061.032 Da
Mth= 7061.018 Da
ΔM = 2.0 ppm
M6+ 1177.8459 Da
C-C-C-C-C-C-C
[M+8H]8+ 1117.4941
[M+7H]7+ 1276.9116
[M+9H]9+ 993.4457
[M+6H]6+ 1489.5553
[M+5H]5+ 1787.2596
570.3035
685.3773
815.3959
Mexp= 8926.078 Da
Mth= 8926.086 Da
ΔM = 0.9 ppm
M8+ 1116.7671 Da
C-C-CC-CC-C
[M+3H]3+ 1244.7735
[M+4H]4+ 933.8346
[M+2H]2+ 1866.6495
Mexp= 3729.299 Da
Mth= 3729.313 Da
ΔM = 3.7 ppm
M3+ 1244.1068 Da

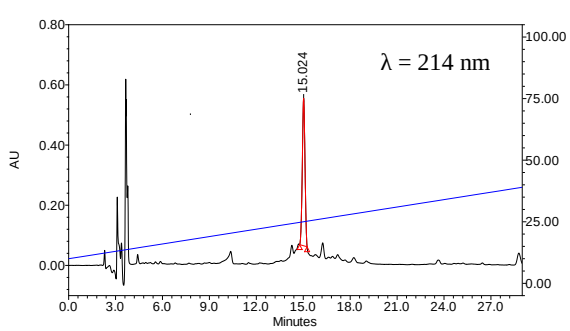
Figure S2. Analysis of 4 representative toxins with their cystein patterns, produced by recombinant expression. A) Analysis on a UHPLC (Accela High Speed LC) on a Hypersil GOLD column C18 175 Å 1.9 µm, 1x50 mm (ThermoScientific) at 280 nm. The acetronitril gradient is drawn on the graph. The surface of the peak corresponding to the correct mass was integrated for toxin quantification. Green arrows indicate the correct mass. Orange arrows indicates a ghost peak corresponding to the 80% acetonitrile flush. B) mass analysis on a MicroTOF (Bruker, accuracy <5ppm, resolution >10,000). An external calibration was performed using NaI clusters. C) Zoom on the most intense charge state. ppm: parts per million.

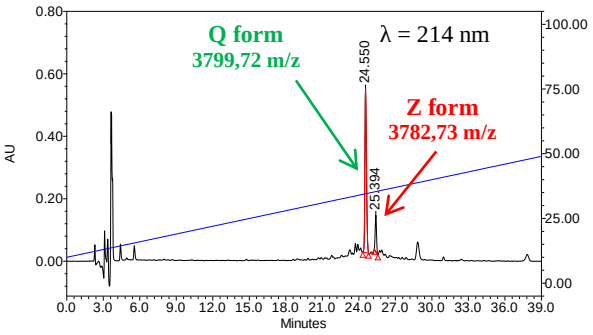
A

N-TRTX-Preg1a
RCLHAGAACSGPIQKIPCCGTCSRRKCT

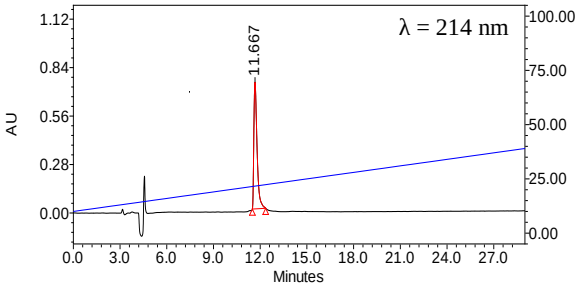
N-BUTX-Ptr1a
QMDMRCSASVECKQKCLKAIGSIFGKCMNKKCKC

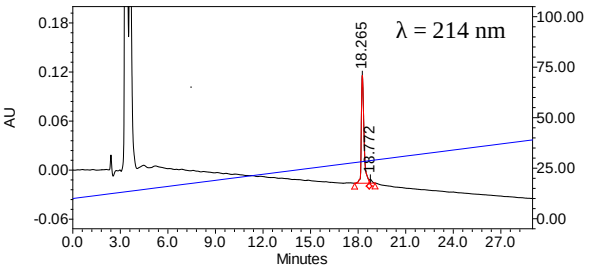
B Crude peptides



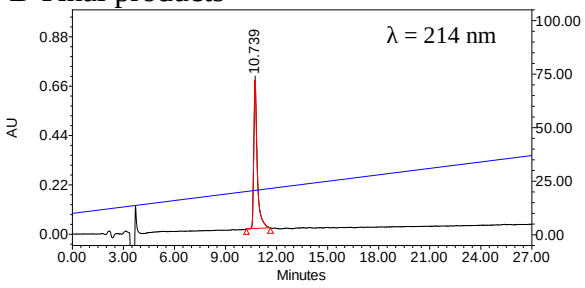


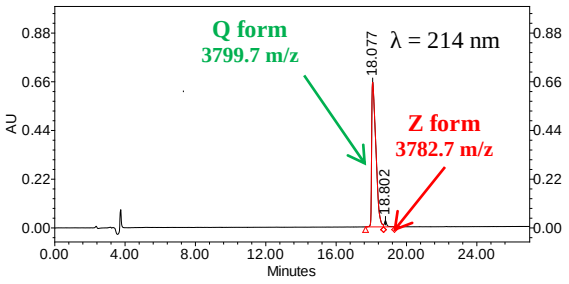
C After 22h of oxidation



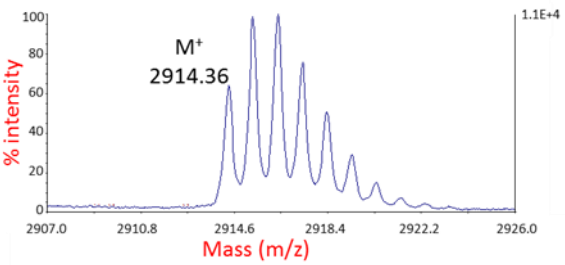


D Final products





E Mass analysis



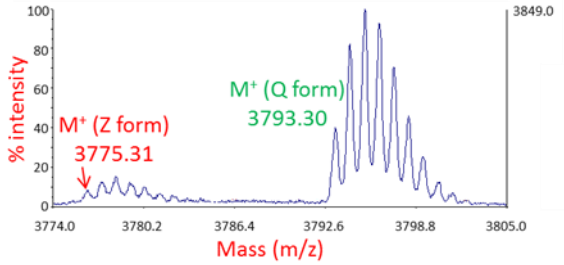


Figure S3. Quality assessment at each stage of M-TRTX-Preg1a (left panel) and M-BUTX-Ptr1a (right panel) production in Fmoc solid-phase chemical synthesis. (A) Peptide sequences with cysteine residues highlighted in red. (B) Analytical HPLC chromatogram of the crude peptide, (C) after 22h of oxidation, and (D) of the final purified product at 214 nm. Column: X-Bridge BEH 300Å 5 µm 4.6 x 250 mm; Gradient: 10-50% CH_3CN/TFA 0.1% in 40 min; flow rate: 1 mL/min. (E) MALDI-TOF spectra of the final pure products. Noted masses correspond to monoisotopic masses. A spontaneous transformation of the N-terminal glutamine (Q form) residue of the peptide N-BUTX-Ptr1a into a pyroglutamic acid (Z form) seems to occur during crude peptide deprotection and can be detected in the final purified product in spite of purification.

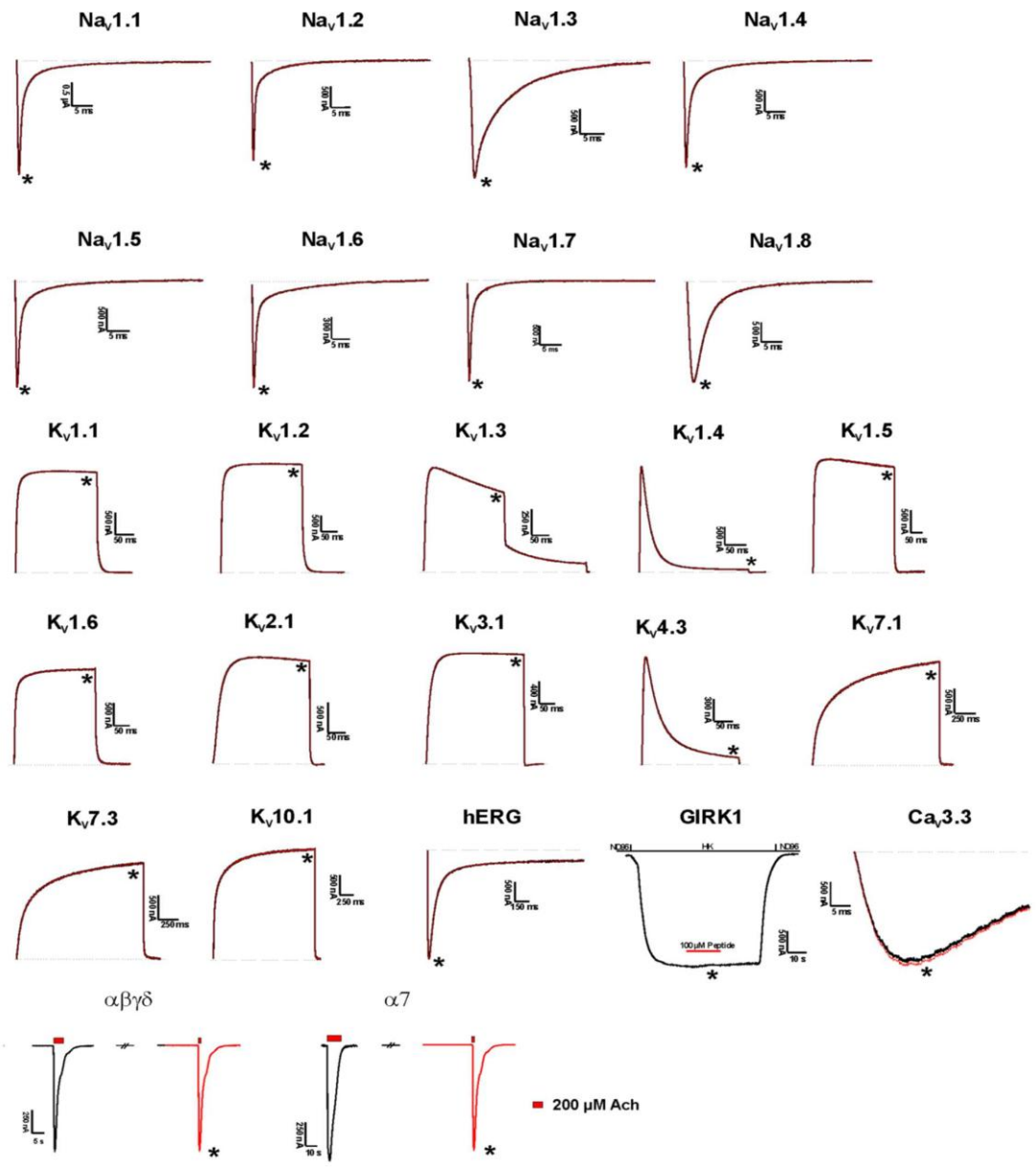


Figure S4. Activity of 100 µM N-TRTX-Preg1a on ion channels. Representative whole cell traces of oocytes expressing cloned Na_v isoforms ($Na_v1.1$-$Na_v1.6$, $Na_v1.8$), K_V isoforms ($K_V1.1$-1.6, $K_V2.1$, $K_V3.1$, $K_V4.3$, $K_V7.1$, $K_V7.3$, $K_V10.1$, hERG), $Ca_V3.3$, GIRK1 or nAChRs (αβγδ and α7). The dotted line indicates the zero-current level. The red * lines indicate the steady-state current peak amplitude in the presence of 100 µM N-TRTX-Preg1a.

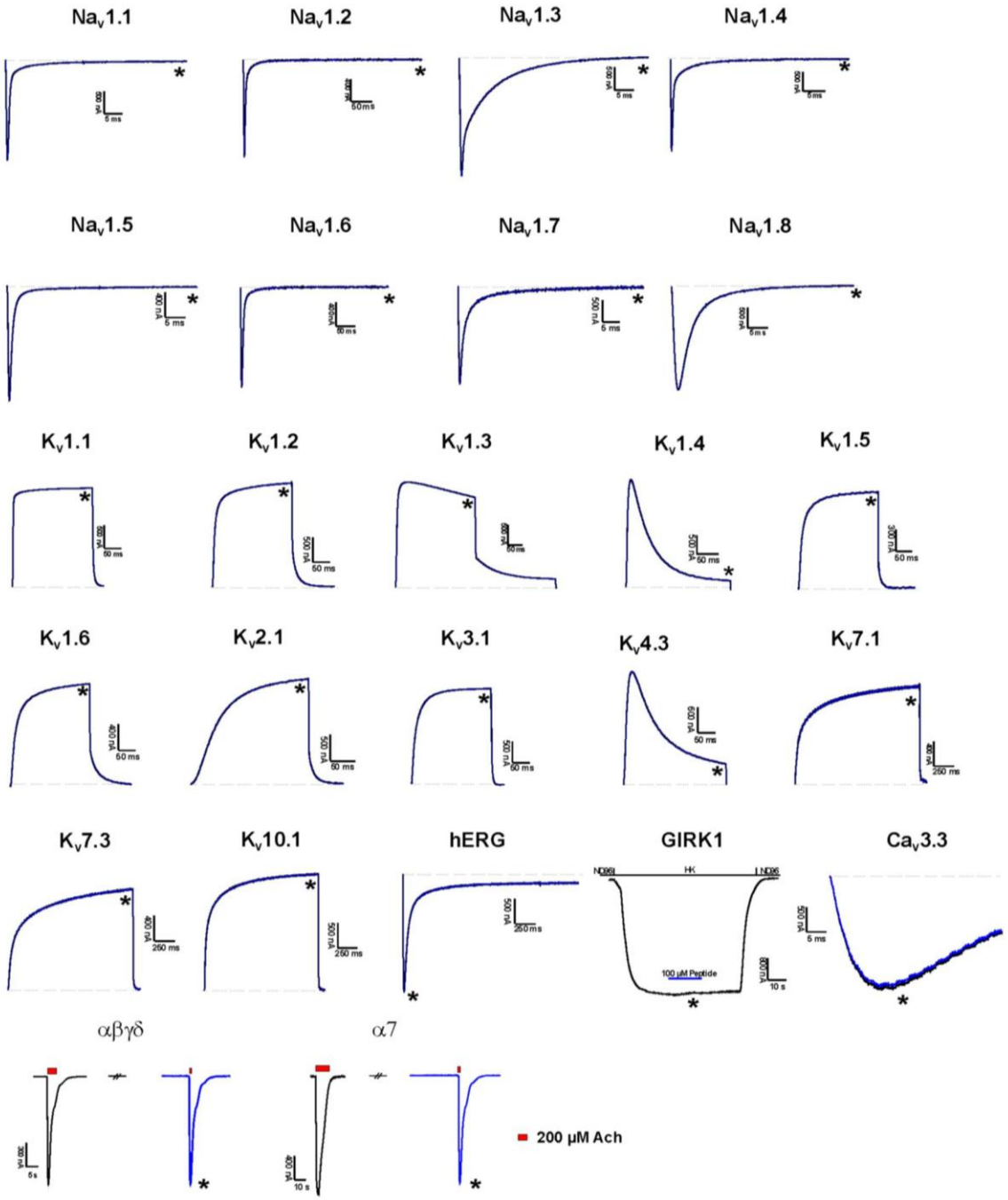


Figure S5. Activity of 100 µM N-BUTX-Ptr1a on ion channels. Representative whole cell traces of oocytes expressing cloned Nav isoforms (Nav1.1-Nav1.6, Nav1.8), KV isoforms (KV1.1-1.6, KV2.1, KV3.1, KV4.3, KV7.1, KV7.3, KV10.1, hERG), CaV3.3, GIRK1 or nAChRs (αβγδ and α7). The dotted line indicates the zero-current level. The blue * lines indicate the steady-state current peak amplitude in the presence of 100 µM N-BUTX-Ptr1a.

Table S1. Experimental NMR data used for the calculations, and final structural statistics for the ten models representative of the solution structure of N-TRTX-Preg1a.

NMR constraints	
Distance restraints	
Total NOE	546
Unambiguous	454
Ambiguous	92
Hydrogen bonds	10
Dihedral Angle Restraints	38
Disulfide bridges[a]	(C2-C19, C9-C22, C18-C27)
Structural Statistics (6SAA.pdb)	
Average violations per structure	
NOEs ≥ 0.3Å	0
Hydrogen bonds ≥ 0.5Å	0
Dihedrals ≥ 25°	0
Dihedrals ≥ 15°	3
Average pairwise rmsd (Å)	0.18
Ramachandran Analysis	
Most favored region	64.8%
Allowed region	35.2%
Generously allowed	0%
Disallowed	0%
Energies (kcal.mol^{-1})[b]	
Electrostatic	-811 ± 8
Van der Walls	-169 ± 3
Total energy	-608 ± 15
Residual NOE energy	26 ± 5

a : Introduced as covalent bonds; b : values are given as mean ± standard deviation (n=10)

Table S2. Experimental NMR data used for the calculations, and final structural statistics for the ten models representative of the solution structure of N-BUTX-Ptr1a.

NMR constraints	
Distance restraints	
Total NOE	839
Unambiguous	558
Ambiguous	281
Hydrogen bonds	4
Disulfide bridges	ambiguous
Structural Statistics (6SAB.pdb)	
Average violations per structure	
NOEs ≥ 0.3Å	0
Hydrogen bonds ≥ 0.5Å	0
Average pairwise rmsd (Å)	0.40
Ramachandran Analysis	
Most favored region	53.0%
Allowed region	40.3%
Generously allowed	5.0%
Disallowed	1.7%
Energies (kcal.mol^{-1})[a]	
Electrostatic	-1252 ± 46
Van der Walls	-116 ± 79
Total energy	-1056 ± 96
Residual NOE energy	22 ± 6

a : values are given as mean ± standard deviation (n=10)

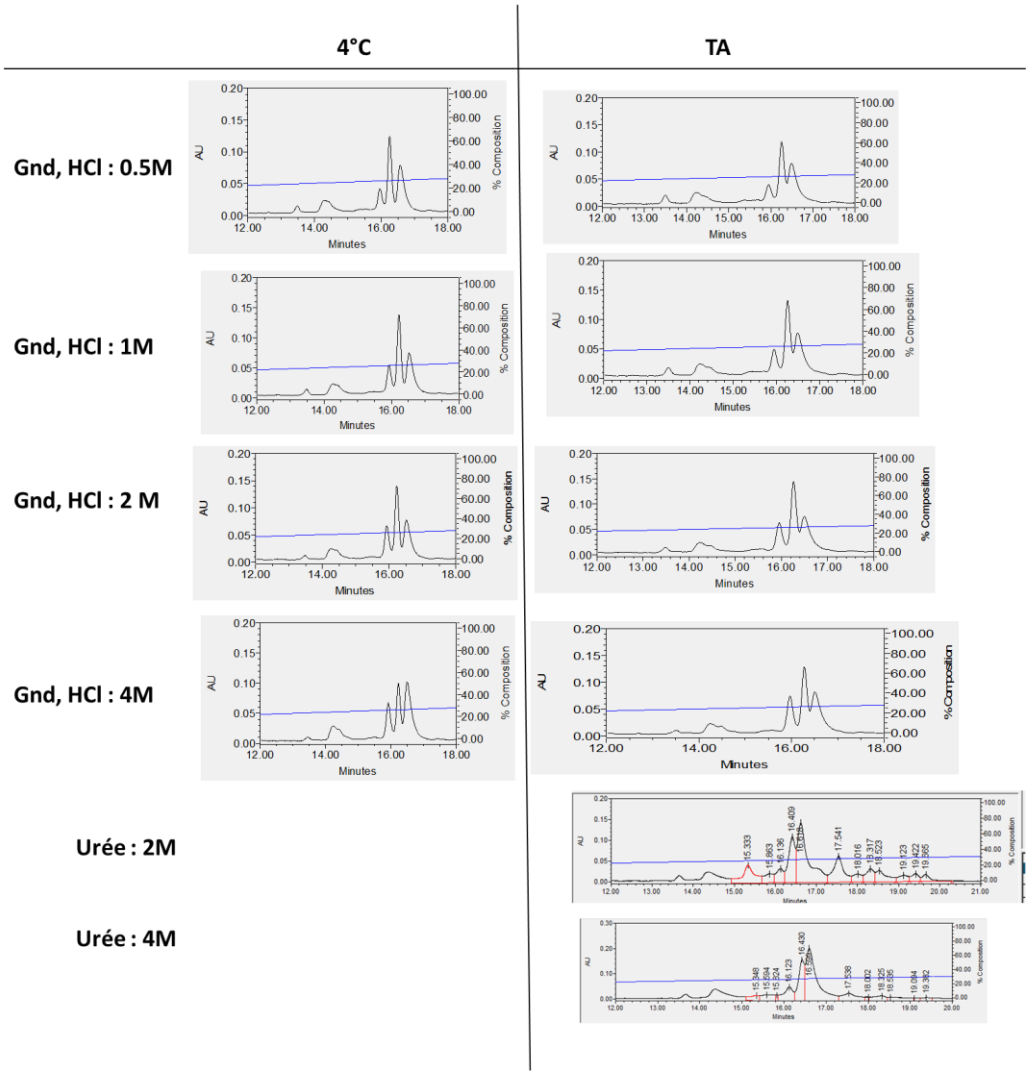
Annexe 3. Criblage de conditions d'oxydation de N-CTX-Can1a. Pour chaque condition, le mélange réactionnel est analysé au bout de 24h. Colonne : X Bridge BEH 300 3.5 µm 4.6 x 150 mm ; Gradient : 10% ACN puis 1% d'ACN/min ; débit : 1 mL/min. λ = 214 nm

- **Variation de la température et de la concentration en agent chaotropique**

Deux couples oxydo-réducteur ont été testés : Cys/CC et GSH/GSSG. Gnd, HCl : chlorure de guanidinium ; GSH : glutathion réduit ; GSSG : disulfure de glutathion ; TA : température ambiante

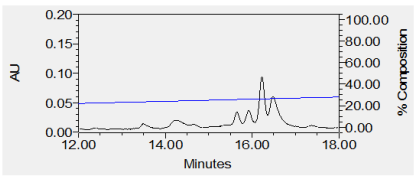
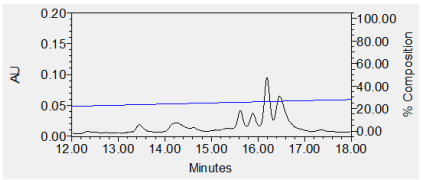
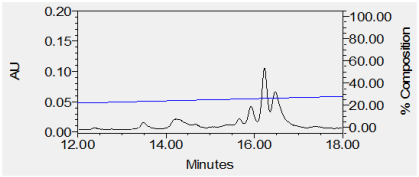
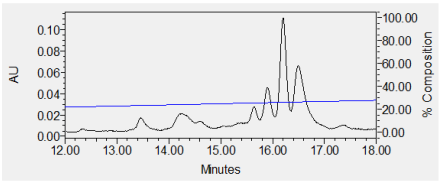
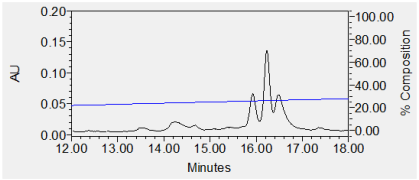
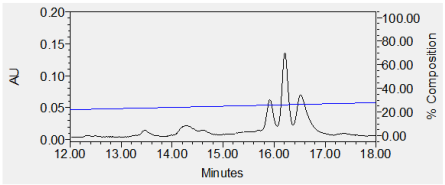
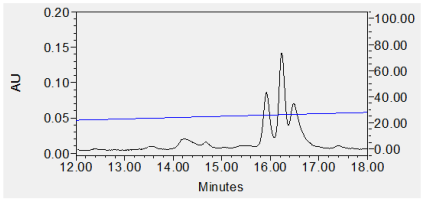
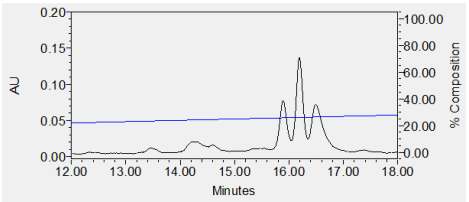
Conditions

	C finale
	mM
HEPES pH 7,5	200
Cys	**1**
CC	**1**
Linéaire	0,04



Conditions

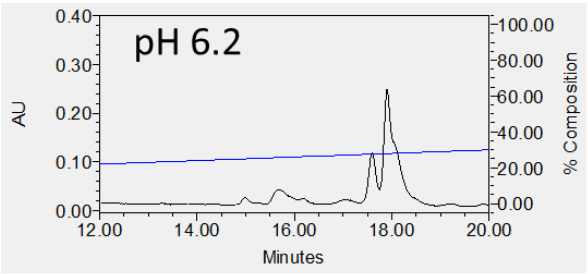
	C finale
	mM
TRIS pH8	200
GSH	**0,5**
GSSG	**0,5**
Linéaire	0,04

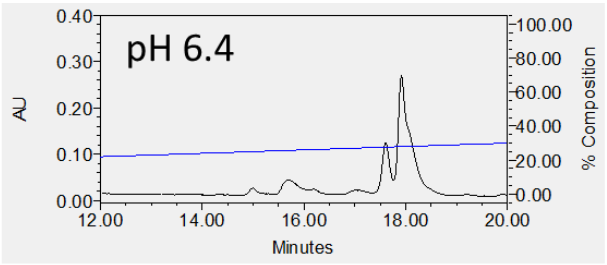
	4°C	TA
Gnd, HCl : 0.5M		
Gnd, HCl : 1M		
Gnd, HCl : 2 M		
Gnd, HCl : 4M		

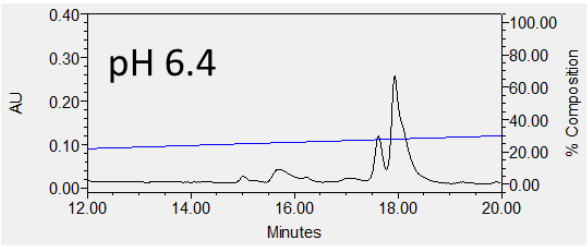
- **Variation du pH (à température ambiante)**

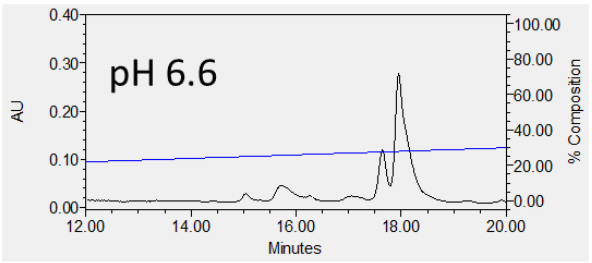
Conditions pH 6.2-7.5

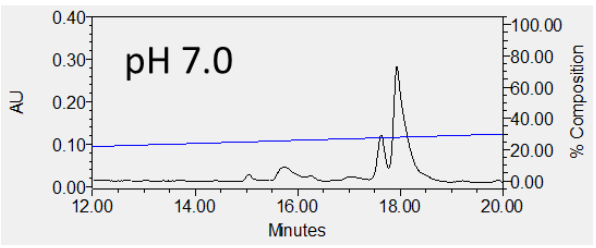
	C finale
HEPES	200 mM
Gnd, HCl	2000
Cys	**1**
CC	**1**
Linéaire	0.04

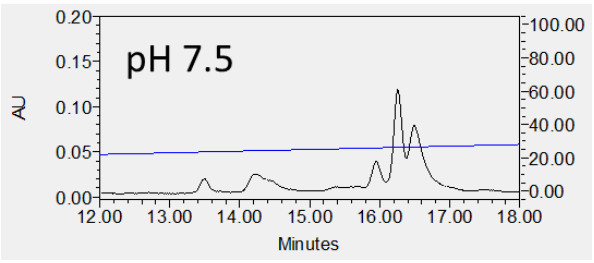






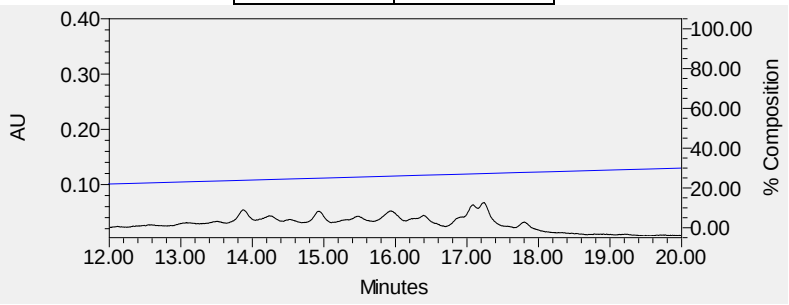






Conditions pH 11

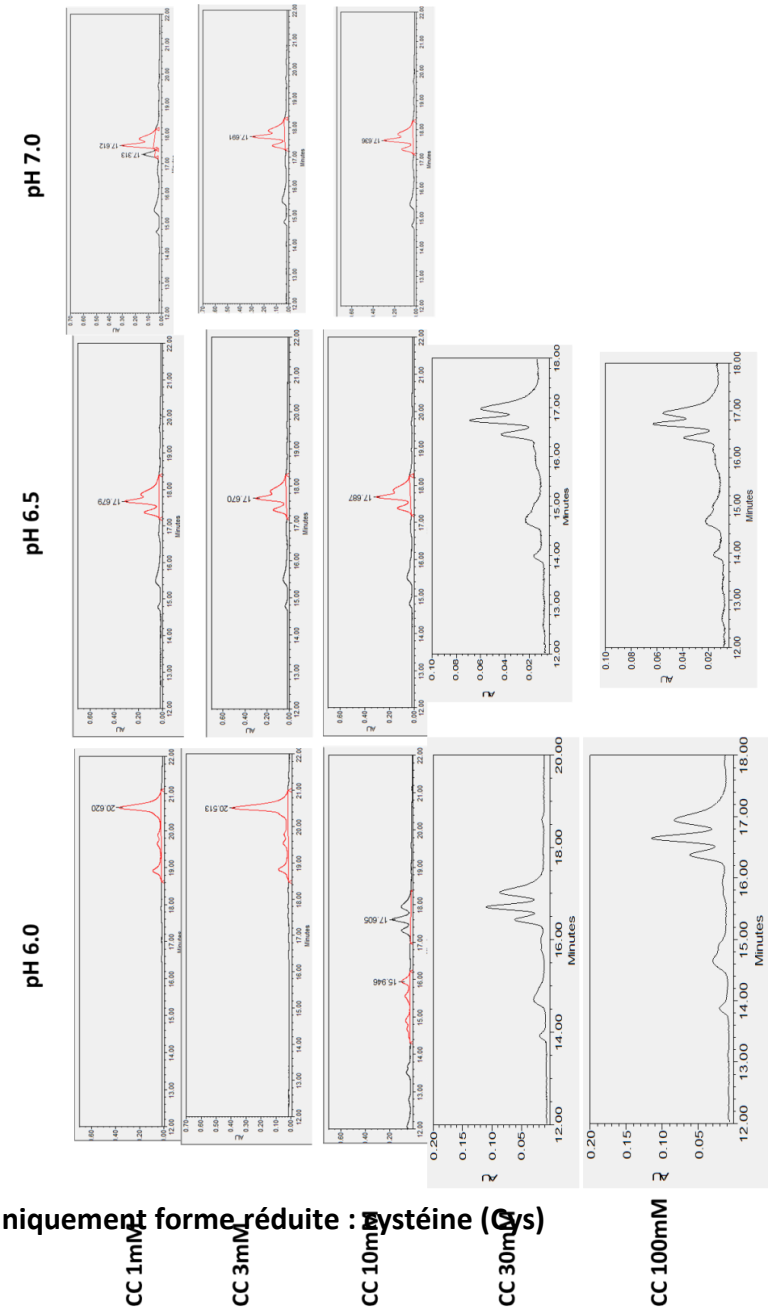
H2O	
TRIS pH 11	200
Gnd, HCl	2000
GSH	**1**
GSSG	**1**
Linéaire	0.04



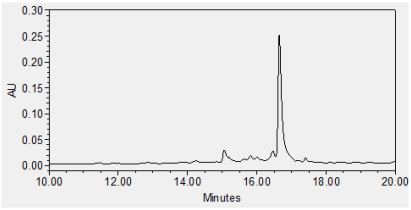
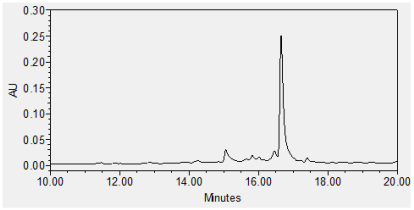
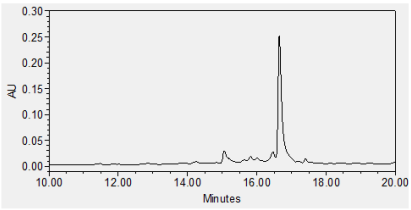
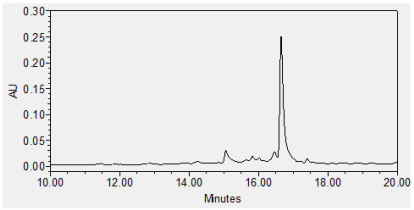
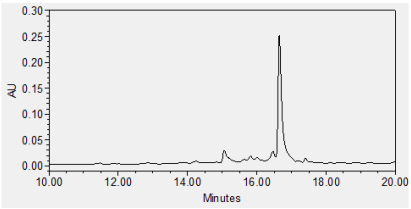
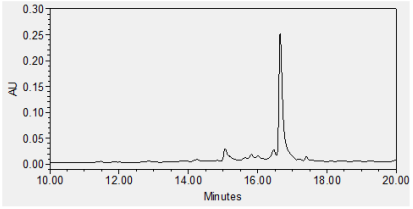
- **Utilisation uniquement forme oxydée : cystine (CC)**

Conditions

	C finale
	mM
HEPES pH 7	200
Gnd, HCl	2000
Linéaire	0.04



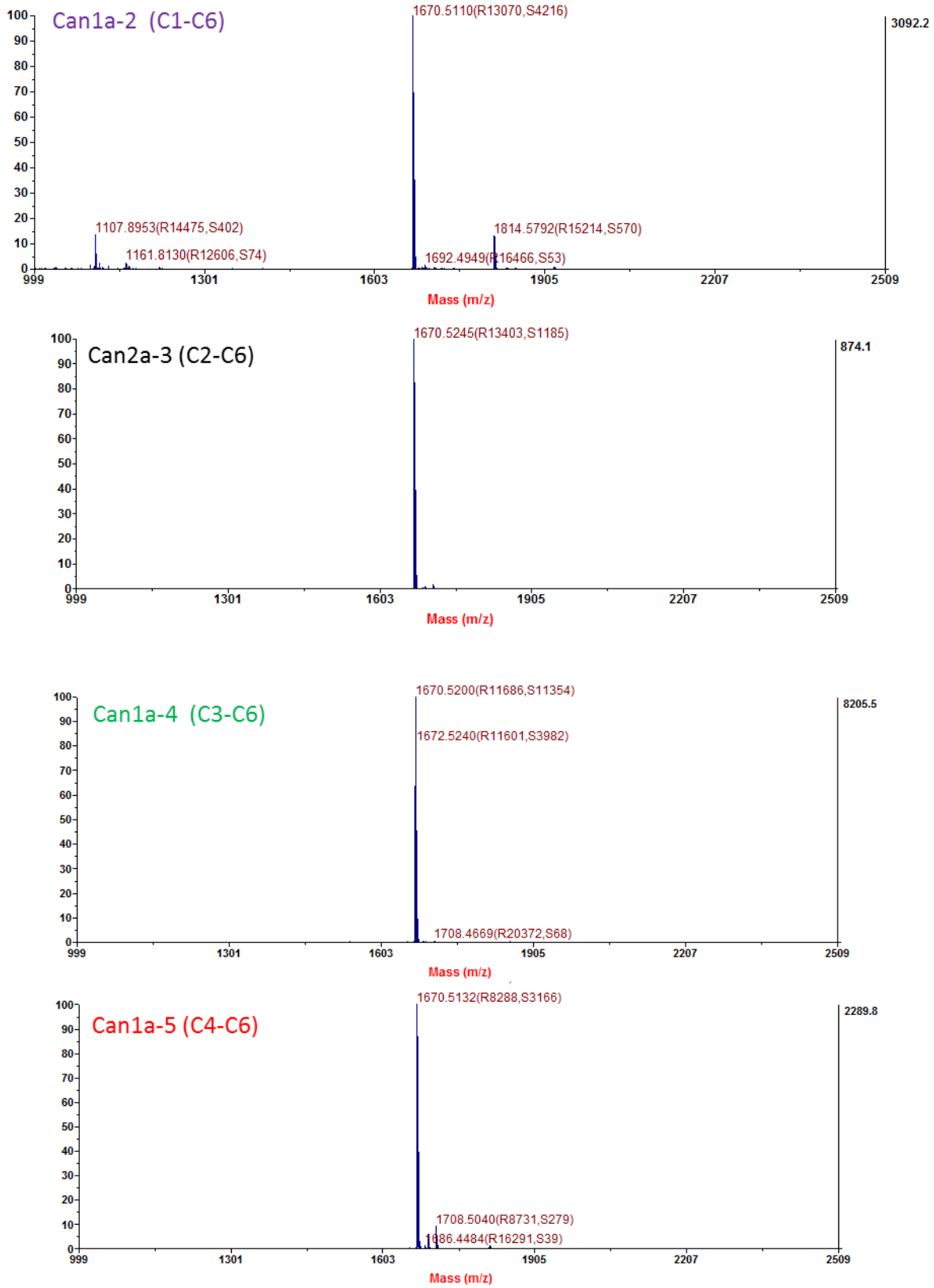
- **Utilisation uniquement forme réduite : cystéine (Cys)**

	pH 6.0	pH 6.5
Cys 1mM		
Cys 3mM		
Cys 10mM		

Annexe 4. Rendements des différentes étapes de production des différentes formes de N-CTX-Can1a, N-CTX-Mil1a et N-CTX-Ltg1a produites en pontage dirigé. Le rendement de synthèse est calculé par dosage des Fmoc. Les rendements des deux étapes d'oxydation sont calculés après purification en HPLC du produit.

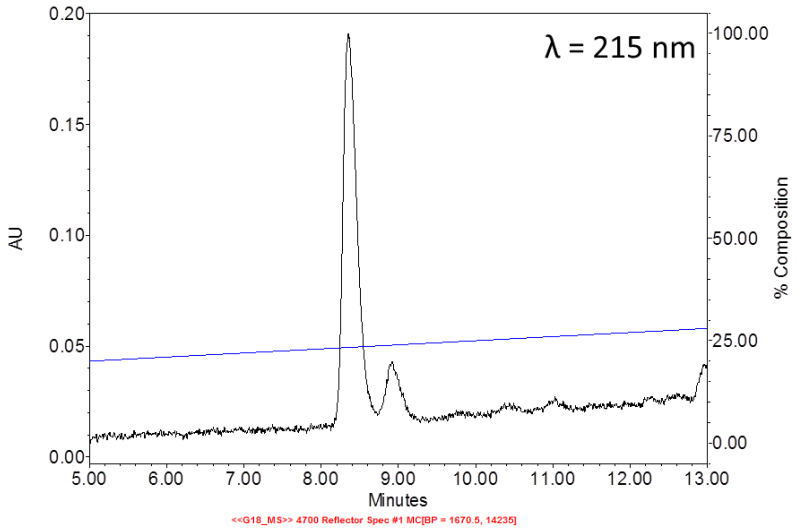
	Pont fixé	Rdt de synthèse	Rdt de la 1ère oxydation	Rdt du pontage dirigé	Rdt global
Can1a-2	C1-C6	65%	17%	29%	3%
Can1a -3	C2-C6	69%	15%	54%	6%
Can1a -4	C3-C6	63%	21%	36%	5%
Can1a -5	C4-C6	65%	21%	71%	10%
Mil1a Bead	C1-C2	51%	39%	24%	5%
Mil1a Globular	C2-C4	52%	29%	28%	4%
Mil1a Ribbon	C2-C3	56%	45%	67%	17%
Ltg1a Bead	C1-C2	56%	83%	53%	25%
Ltg1a Globular	C2-C4	52%	29%	33%	5%
Ltg1a Ribbon	C2-C3	51%	31%	40%	6%

Annexe 5. Spectres de masse MALDI-TOF des différentes formes oxydées en pontage dirigé de N-CTX-Can1a. Le pont fixé est indiqué entre paranthèses. Les masses notées correspondent aux masses monoisotopiques.

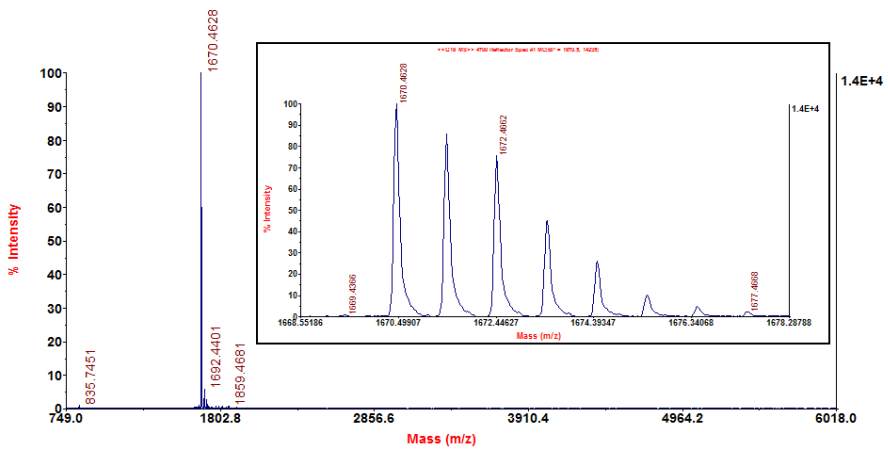


Annexe 6. Oxydation aléatoire de P014846. A) Profil HPLC après 5 jours d'oxydation à 4°C. Colonne : X Bridge BEH 300 5 µm 4.6 x 250 mm ; Gradient : 15% ACN puis 1% d'ACN/min. B) Spectres de masse MALDI-TOF du pic majoritaire purifié. Les masses notées correspondent aux masses monoisotopiques. L'encadré est un agrandissement du massif isotopique de la masse majoritaire.

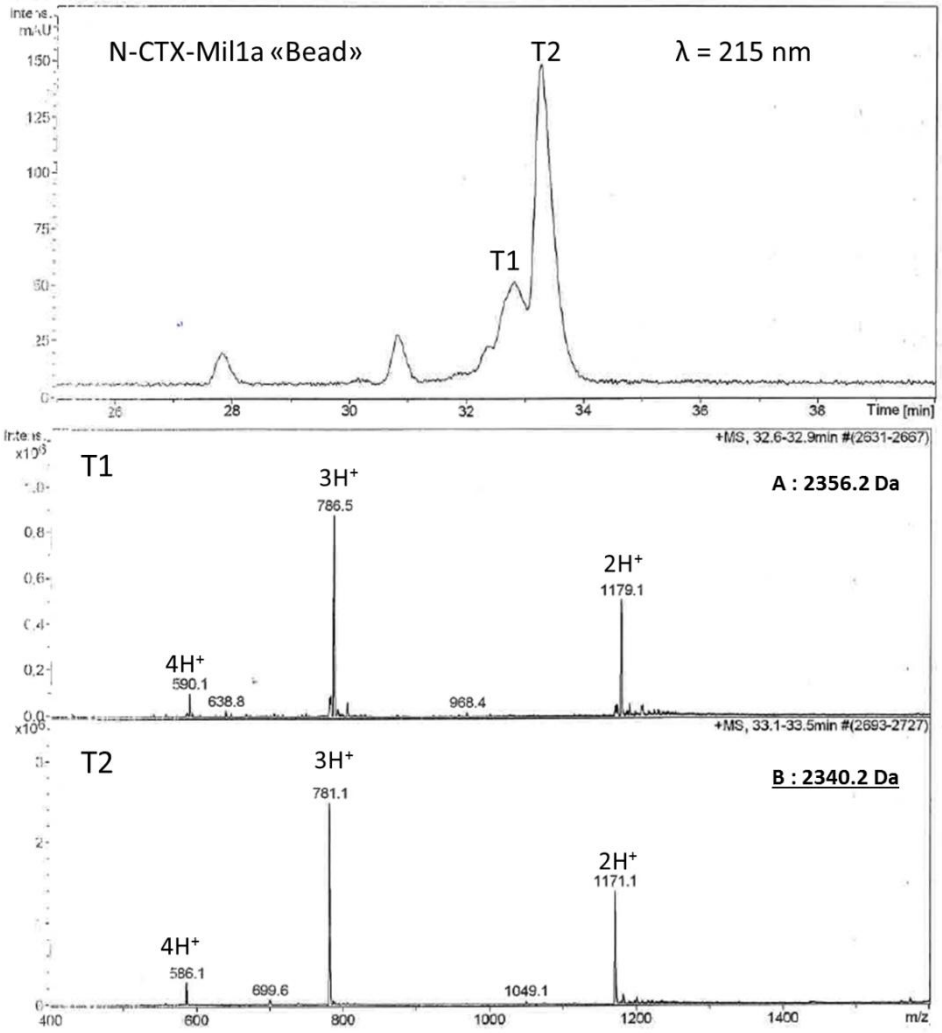
A)

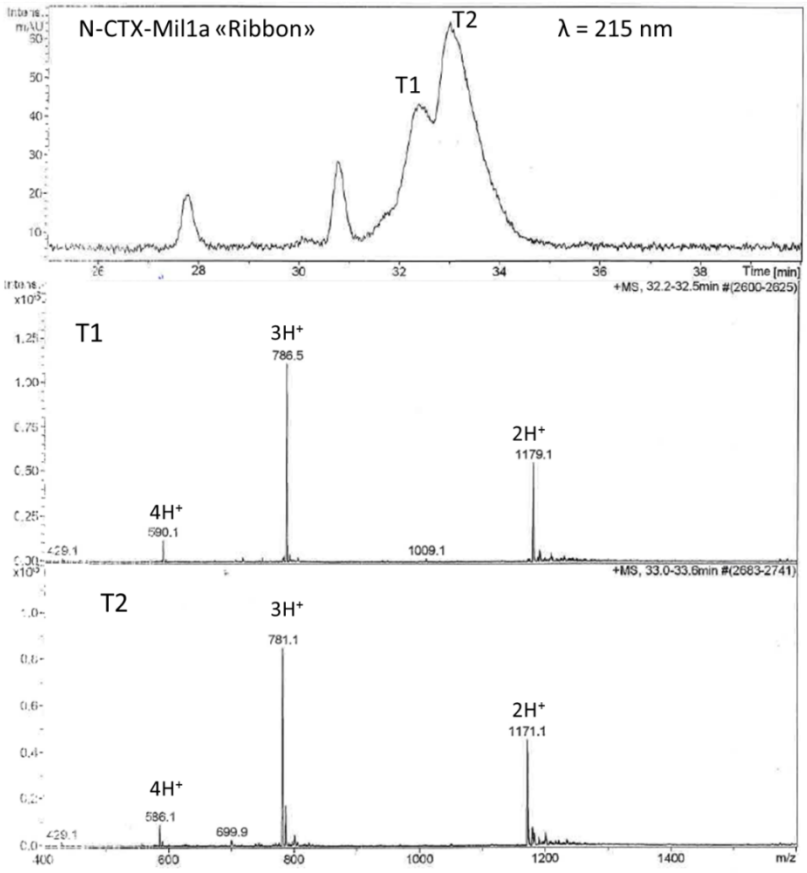


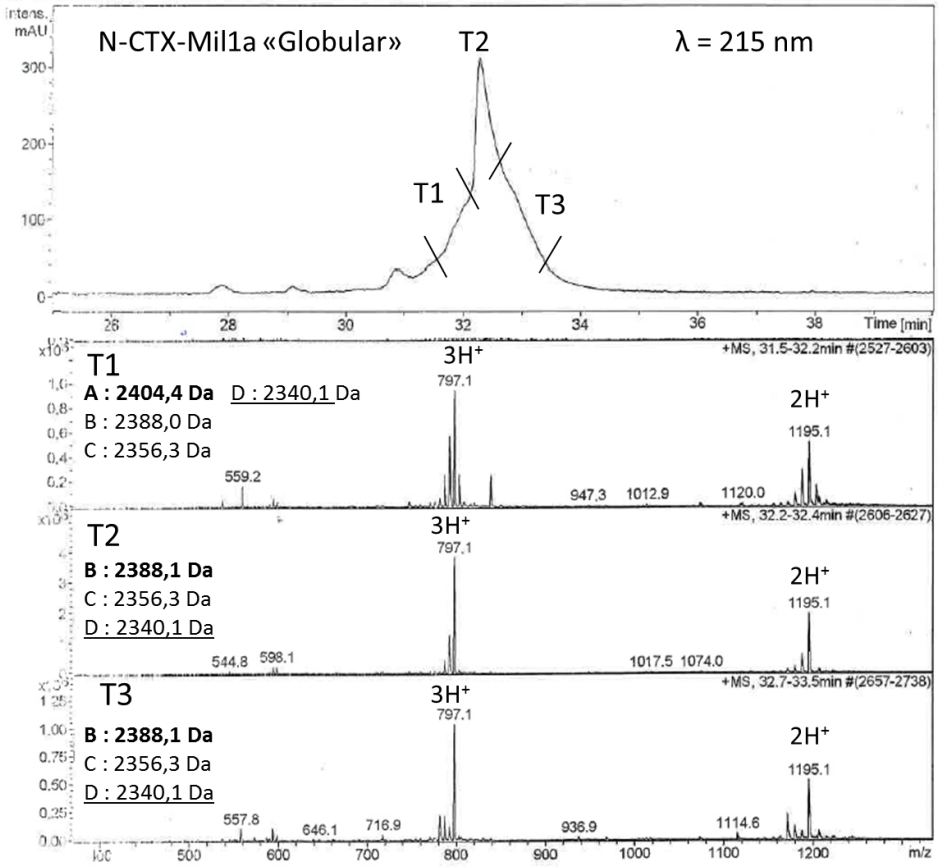
B)



Annexe 7. LC-MS (ESI-IT) des formes oxydées «Bead », «Ribbon » et « Globular » de N-CTX-Mil1a. La masse majoritaire de chacun des spectres est signalée en gras. La masse de la toxine N-CTX-Mil1a est soulignée lorsqu'elle est détectée. Colonne : XBridge® Peptide BEH C18 300Å 3.5 µm, 4.6x150 mm ; Gradient : 0% ACN/TFA0.1% puis 1% CH_3CN/TFA0.1% /min.




N-CTX-Mil1a «Ribbon»
λ = 215 nm
T1
T2
Time [min]
+MS, 32.2-32.5min #(2600-2625)
T1
3H+
786.5
2H+
1179.1
4H+
590.1
429.1
1009.1
+MS, 33.0-33.6min #(2683-2741)
T2
3H+
781.1
2H+
1171.1
4H+
586.1
429.1
699.9
m/z


N-CTX-Mil1a «Globular»
λ = 215 nm
T1
T2
T3
Time [min]
+MS, 31.5-32.2min #(2527-2603)
T1
A : 2404,4 Da
D : 2340,1 Da
B : 2388,0 Da
C : 2356,3 Da
3H+
797.1
2H+
1195.1
559.2
947.3
1012.9
1120.0
+MS, 32.2-32.4min #(2606-2627)
T2
B : 2388,1 Da
C : 2356,3 Da
D : 2340,1 Da
3H+
797.1
2H+
1195.1
544.8
598.1
1017.5
1074.0
+MS, 32.7-33.5min #(2657-2738)
T3
B : 2388,1 Da
C : 2356,3 Da
D : 2340,1 Da
3H+
797.1
2H+
1195.1
557.8
646.1
716.9
936.9
1114.6
m/z

Titre : Des toxines animales comme ligands innovants des récepteurs aux mélanocortines

Mots clés : RCPG, MC4R, Toxines animales, HT1-0, obésité, cachexie

Résumé : Le sous-type 4 des récepteurs aux mélanocortines (MC4R) est une cible thérapeutique pour le traitement des troubles de l'homéostasie énergétique. Aucun traitement efficace n'est pour l'instant sur le marché. Dans ce contexte, une banque de toxines animales synthétiques a été criblée, aboutissant à 286 hits. Au vue de ce chiffre exceptionnel, cinq des toxines les plus originales au niveau séquence ont été sélectionnées et caractérisées pharmacologiquement. N-TRTX-Preg1a de l'araignée *Poecilotheria regalis* et N-BUTX-Ptr1a du scorpion *Parabuthus transvaalicus* sont les premières toxines agonistes du sous-type 1 des récepteurs aux mélanocortines (MC1R). Ce récepteur stimule la pigmentation de la peau et joue un rôle clé dans la prévention du mélanome. Les deux toxines forment chacune un nouveau groupe au sein de leur famille de toxines, respectivement les knottin et les toxines CSαβ. L'étude des conotoxines N-CTX-Can1a et N-CTX-Mil1a, a été abandonnée du fait de la complexité de leur synthèse et de leur faible affinité. N-CTX-Ltg1a, issue du cône *Conus lithoglyphus*, est une toxine prometteuse. Sa séquence ne ressemble à aucune famille connue de toxine et elle présente une affinité nanomolaire sur MC4R. Avec 4 cystéines, elle peut être oxydée de trois façons différentes, rendant son étude complexe. Nous avons simplifié progressivement cette toxine pour aboutir à un peptide linéaire simple à produire et plus affin d'un facteur 3 que la toxine de départ avec une sélectivité d'un facteur 3 à 13 par rapport aux autres MCRs. La Hunger-Toxin (HT1-0) antagonise l'activation des voies de signalisation G_s, G_{15} et β-arrestine 2 du MC4R. Des études de relation structure-activité ont permis de délimiter les zones importantes dans l'interaction avec MC4R et serviront à l'optimisation future du peptide. Prochainement, la confirmation *in vivo* de l'effet anorexigène pourrait ainsi ouvrir la voie vers un nouveau traitement de la cachexie.

Title : Animal toxins as innovative ligands on melanocortin receptors

Keywords : GPCR, MC4R, animal toxin, HT1-0, obesity, cachexia

Abstract: Melanocortin 4 receptor (MC4R) is a therapeutic target for the treatment of energy homeostasis disorders. No effective treatment is currently available on the market. In this context, screening of a synthetic toxin bank resulted in 286 hits. In light of this exceptional figure, five of the most original sequences have been selected and pharmacologically characterized. N-TRTX-Preg1a from the spider *Poecilotheria regalis* and N-BUTX-Ptr1a from the scorpion *Parabuthus transvaalicus* are the first toxins agonist of the melanocortin 1 receptor (MC1R). This receptor stimulates skin pigmentation and plays a key role in preventing melanoma. Each forms a new group within its family, respectively knottin and CSαβ toxins. The study of N-CTX-Can1a and N-CTX-Mil1a, two conotoxins, was discontinued due to the complexity of their synthesis and their low affinity for MC4R. N-CTX-Ltg1a, from the cone snail *Conus lithoglyphus*, is a very promising toxin. It is not linked to any known toxin family and exhibits nanomolar affinity on MC4R. With four cysteines, the toxin can be oxidized into three different forms, making its study complex. We gradually simplified it to a linear peptide that is easy to produce and 3-fold more potent on MC4R than the original toxin with a 3-13-fold selectivity compared to other MCRs. Hunger-Toxin (HT1-0) antagonizes the activation of the signaling pathways G_s, G_{15} and β-arrestin 2 of MC4R. Structure-activity relationship studies have identified important areas in the interaction with MC4R and will be used for future peptide optimization. In the near future, the in vivo confirmation of the anorexigenic effect could thus pave the way for a new treatment of cachexia.

Route de l’Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France