



**HAL**  
open science

# Études par dynamique moléculaire de l'interaction de Récepteurs Couplés aux Protéines-G avec leurs partenaires extra et intra-cellulaires

Bartholomé Delort

► **To cite this version:**

Bartholomé Delort. Études par dynamique moléculaire de l'interaction de Récepteurs Couplés aux Protéines-G avec leurs partenaires extra et intra-cellulaires. Biologie moléculaire. Université Montpellier, 2018. Français. NNT : 2018MONT113 . tel-02181478

**HAL Id: tel-02181478**

**<https://theses.hal.science/tel-02181478>**

Submitted on 12 Jul 2019

**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.

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

En Ingénierie Moléculaire

École Doctorale 459 - Sciences Chimiques Balard

Unité de recherche IBMM - Institut des Biomolécules Max Mousseron

Etudes par dynamique moléculaire de l'interaction de  
Récepteurs Couplés aux Protéines-G avec leurs  
partenaires extra et intra-cellulaires

Présentée par Bartholomé DELORT  
Le 19 Novembre 2018

Sous la direction de Nicolas FLOQUET

Devant le jury composé de

M. Jean MARTINEZ	Professeur	Université de Montpellier	Président du Jury
M. Nicolas FLOQUET	C.R. CNRS	IBMM, Montpellier	Directeur de Thèse
Mme Esther KELLENBERGER	Professeur	Université de Strasbourg	Rapporteur
M. Marc BAADEN	D.R. CNRS	IBPC, Paris	Rapporteur
M. Patrick FUCHS	MCF	Université Paris Diderot	Examineur
M. Jean-Philippe PIN	D.R. CNRS	IGF, Montpellier	Examineur



UNIVERSITÉ  
DE MONTPELLIER



- « Qu'est-ce que vous faites ?
- Oh, je ne fais que... Vous savez, je ne fais que... que regarder. Une petite...  
expérience. Après tout, on ne sait jamais.
- On sait jamais quoi ?
- Rien qu'on ne sait jamais, vous savez.
- Des fois, on sait quand même, fit Ridculle. Je crois que j'sais beaucoup de trucs que  
j'savais pas avant. C'est incroyable ce qu'on finit par savoir, je m'dis des fois. Je  
m'demande souvent quels nouveaux trucs je vais savoir.
- Ben, on ne sait jamais.
- Pour sûr.»

Dialogue extrait du « **Père Porcher** » illustrant la méthode scientifique

*A Terry Pratchett, père de trop nombreuses citations*

# REMERCIEMENTS

Je souhaite remercier en premier lieu Monsieur le Professeur Jean MARTINEZ qui m'a accueilli au sein de son laboratoire l'Institut des Biomolécules Max Mousseron, et pour avoir soutenu ma candidature me permettant de poursuivre mon sujet de stage en thèse. Je remercie également Monsieur le Professeur Pascal DUMY qui par la suite prit la relève à la direction du laboratoire.

Je souhaite mes remerciements les plus sincères à l'intention de Monsieur le Docteur Nicolas FLOQUET qui m'a permis de découvrir avec plaisir la modélisation moléculaire après des années de chimie. Je le remercie pour l'encadrement et la formation lors de ma thèse de doctorat dans ce domaine, et sur bon nombres d'autres aspects de la recherche publique. Je le remercie également pour toutes les critiques et améliorations de mon travail. Enfin, je le remercie pour avoir permis la réalisation de ce manuscrit et de présenter aujourd'hui ces travaux représentant un peu plus de trois ans de recherche.

J'adresse également mes remerciements à Madame la Professeure Esther KELLENBERGER et à Monsieur le Docteur Marc BAADEN pour avoir accepté d'être rapporteurs de mes travaux de thèse, mais également à Monsieur le Docteur Jean-Philippe PIN, à Monsieur le Professeur Patrick FUCHS et à Monsieur le Professeur Jean MARTINEZ pour avoir accepté de juger mon travail.

Je remercie également les autres membres de l'équipe modélisation Monsieur le Docteur Maxime LOUET et Monsieur le Docteur Pedro RENAULT pour les discussions et les conseils sur mon travail, pour la qualité scientifique de vos travaux et pour la bonne humeur du bureau. Je remercie également nos anciens colocataires de l'équipe informatique du laboratoire Boris RIGAULT, Jeremy BERRE et surtout Florent RAUSSIN pour son aide lors de mon apprentissage Linux et de mes problèmes techniques. Enfin, je remercie la climatisation, membre essentiel de l'équipe, toujours présente au quotidien et lors nos plus terribles étés.

J'aimerais également remercier tous les membres de mes deux équipes au sein du laboratoire, pour les discussions et encore une fois la bonne humeur de ces dernières années. Je remercie plus particulièrement le Docteur Jean-Louis BANERES pour m'avoir permis de réaliser les tests biologiques de mes peptides, et au Docteur Pascal VERDIE et Manon MAUREL de m'avoir formé à la synthèse peptidique. Je remercie aussi toutes les personnes avec qui nous avons collaborés.

Pour finir, j'adresse mes plus francs remerciements à mes amis qui n'auront malheureusement pas le droit aux remerciements qu'ils méritent à cause de phrases beaucoup trop longues et d'un manque de place évident malgré une police 11. Je remercie donc alphabétiquement Alexandre, Alice, Christophe, Julie et Julie, Kévin, Laurent, Laurine, Lisa, Manon, Titouan et tous ceux que je n'écris pas, mais que je n'oublie pas. Enfin, je remercie ma famille pour devoir faire avec ma notion si particulière du temps qui passe, et pour conclure Caroline, qui a la chance/fardeau de me supporter au quotidien. **Merci.**

# TABLE DES MATIERES

<b>1 INTRODUCTION .....</b>	<b>3</b>
<b>1.1 Les Récepteurs Couplés aux Protéines-G.....</b>	<b>3</b>
1.1.1 Historique.....	3
1.1.2 Structures des RCPGs .....	7
1.1.3 Les différentes familles de RCPGs .....	9
1.1.3.a Rhodopsine ou classe A .....	11
1.1.3.b Sécrétine ou classe B1 .....	13
1.1.3.c Adhésion ou classe B2 .....	14
1.1.3.d Glutamate ou classe C.....	15
1.1.3.e Frizzled ou classe F.....	15
1.1.3.f Taste2 ou classe T.....	16
1.1.4 Partenaires principaux des RCPGs .....	17
1.1.4.a Partenaires extra-cellulaires .....	17
1.1.4.b Partenaires intra-cellulaires.....	19
1.1.4.c Dimérisation / oligomérisation.....	22
1.1.5 Activation des RCPGs .....	23
1.1.5.a Formation de l'hétéro-trimère .....	23
1.1.5.b Complexation avec un RCPG .....	24
1.1.5.c Activation du couple RCPG:pG.....	26
1.1.5.d Relargage du GDP et dissociation .....	27
1.1.5.e Après l'activation de la pG .....	29
<b>1.2 Les champs de forces.....</b>	<b>30</b>
1.2.1 Champs de forces classiques.....	31
1.2.2 Champs de forces polarisables .....	32
1.2.3 Champs de forces réactifs .....	33
1.2.4 Champs de forces gros-grains .....	33
<b>1.3 Modélisation de la dynamique intrinsèque et des interactions impliquant les RCPGs .....</b>	<b>35</b>
1.3.1 Liaison du ligand.....	36
1.3.2 Docking .....	38

1.3.3	Dynamique intrinsèque du récepteur .....	39
1.3.4	Simulations de systèmes de taille plus importante.....	41
1.3.5	Dimérisation / oligomérisation des RCPGs .....	43
<b>1.4</b>	<b>Objectifs de la thèse.....</b>	<b>45</b>
<b>2</b>	<b>MATERIELS ET METHODES.....</b>	<b>51</b>
<b>2.1</b>	<b>Mécanique moléculaire .....</b>	<b>51</b>
2.1.1	Champ de forces MARTINI E1NeDyn .....	51
2.1.2	Schéma de conversion tout-atomes vers gros-grains .....	52
2.1.3	Equations d'énergie.....	54
2.1.4	Minimisation d'énergie .....	57
<b>2.2</b>	<b>Dynamique moléculaire .....</b>	<b>59</b>
2.2.1	Théorie .....	59
2.2.2	Conditions périodiques aux limites.....	59
2.2.3	REMD .....	60
2.2.4	Protocoles de simulations.....	63
<b>2.3</b>	<b>Analyses des résultats.....</b>	<b>64</b>
2.3.1	RMSD .....	64
2.3.2	Clustering .....	64
2.3.3	Analyse en composantes principales.....	65
<b>2.4</b>	<b>Structures initiales.....</b>	<b>66</b>
2.4.1	Construction par homologie .....	66
2.4.2	Récepteurs de la Neurotensine .....	66
2.4.3	Récepteur de la Chimiokine C-X-C de type-4 .....	68
2.4.4	Récepteurs de la Vasopressine .....	70
2.4.5	Récepteur de la Ghréline .....	72
2.4.6	Récepteur de la Dopamine .....	74
2.4.7	Récepteurs de l'Adénosine.....	75
2.4.8	Peptides C-terminaux des Protéines-G.....	77
2.4.9	Construction de « systèmes doubles » pour les simulations CG-REMD.....	78

<b>2.5 Moyens informatiques utilisés .....</b>	<b>80</b>
2.5.1 Clusters de calcul .....	80
2.5.2 Logiciels utilisés .....	81
<b>2.6 Synthèse et test biologiques de peptides inhibiteurs des interactions protéine-G:protéine-G .....</b>	<b>82</b>
2.6.1 Synthèse .....	82
2.6.1.a Préparation des résines .....	82
2.6.1.b Robot synthétiseur .....	83
2.6.1.c Décrochage et déprotection des peptides linéaires .....	83
2.6.1.d Cyclisation, décrochage et déprotection du peptide cyclique .....	84
2.6.1.e Purification et analyses .....	84
2.6.2 Tests biologiques .....	85
2.6.2.a Préparation des systèmes .....	85
2.6.2.b Mise en place du test .....	87
<b>3 RESULTATS.....</b>	<b>91</b>
<b>3.1 Prédiction de la liaison de peptide aux RCPGs .....</b>	<b>91</b>
3.1.1 Validation de la méthode .....	91
3.1.1.a Contexte .....	91
3.1.1.b Article .....	93
3.1.1.c Commentaires .....	111
3.1.1.d Le cas des peptides « scrambled » .....	113
<b>3.2 Prédiction de la sélectivité de peptides à différents récepteurs.....</b>	<b>116</b>
3.2.1 Design de peptides sélectifs de NTR2 par dynamique moléculaire.....	116
3.2.1.a Contexte .....	116
3.2.1.b Article .....	117
3.2.1.c Commentaires .....	129
3.2.2 Étude de la sélectivité de peptides dérivés de la Neurotensine .....	130
3.2.2.a Contexte .....	130
3.2.2.b Article .....	131
3.2.2.c Commentaires .....	145

<b>3.3 Prédiction des interactions RCPG:RCPG.....</b>	<b>147</b>
3.3.1 Dimère de RCPGs.....	147
3.3.1.a Contexte.....	147
3.3.1.b Article.....	149
3.3.1.c Commentaires.....	160
3.3.2 Synthèse et test de peptides inhibiteurs des interactions protéine-G:protéine-G.....	161
3.3.2.a Choix des peptides.....	161
3.3.2.b Synthèse des peptides.....	162
3.3.2.c Test biologiques d'inhibition.....	165
3.3.3 Prédiction du mode de dimérisation / oligomérisation des RCPGs.....	168
3.3.3.a Contexte.....	168
3.3.3.b Résultats.....	169
3.3.3.c Perspectives.....	174
<b>3.4 Autres applications.....</b>	<b>176</b>
3.4.1 Prédiction de la liaison d'une toxine aux récepteurs de la Vasopressine.....	176
3.4.1.a Contexte.....	176
3.4.1.b Résultats.....	176
3.4.1.c Comparaison avec la Vasopressine.....	182
3.4.1.d Perspectives.....	183
3.4.2 Prédiction de la liaison de peptides au récepteur de la Ghreline.....	185
3.4.2.a La Ghreline.....	185
3.4.2.b Le LEAP2.....	191
3.4.2.c Perspectives.....	192
3.4.3 Prédiction de la liaison et de la sélectivité de peptides issus de l'hélice C-terminale de Protéines-G.....	193
3.4.3.a Contexte.....	193
3.4.3.b Résultats.....	193
3.4.3.c Perspectives.....	196
<b>4 CONCLUSIONS ET PERSPECTIVES.....</b>	<b>197</b>
<b>5 REFERENCES.....</b>	<b>200</b>

# LISTE DES FIGURES

Figure 1 : Structure générale d'un Récepteur Couplé aux Protéines-G.....	5
Figure 2 : Modèle d'un RCPG complexé à une protéine-G hétéro-trimérique .....	8
Figure 3 : Evolution du nombre de structures de RCPGs disponibles dans la PDB .....	9
Figure 4 : Classification des récepteurs de la sous-famille Rhodopsine .....	12
Figure 5 : Classification des récepteurs de la sous-famille Sécrétine .....	14
Figure 6 : Classification des récepteurs de la sous-famille Adhésion .....	14
Figure 7 : Classification des récepteurs de la sous-famille Glutamate.....	15
Figure 8 : Classification des récepteurs de la sous-famille Frizzled .....	16
Figure 9 : Classification des récepteurs de la sous-famille Taste2.....	16
Figure 10 : Représentation des effets provoqués par la liaison d'un ligand à un RCPG.....	18
Figure 11 : Représentation de l'effet modulateur d'un ligand allostérique.....	19
Figure 12 : Structure d'une Kinase-RCPG.....	21
Figure 13 : Structure de la $\beta$ -arrestine liée à un RCPG .....	21
Figure 14 : Modèle de dimère DRD2/G <sub>i</sub> :GHSR-1a/G <sub>q</sub> .....	23
Figure 15 : Comparaison de l'orientation de deux structures Cryo-EM de protéines-G.....	25
Figure 16 : Comparaison de structures cristallographiques de $\beta$ 2AR .....	27
Figure 17 : Schématisation des modèles de sorties du GDP .....	28
Figure 18 : Illustration des objectifs de la thèse .....	46
Figure 19 : Représentation d'une particule d'eau dans le champ de forces MARTINI .....	52
Figure 20 : Représentation des acides aminés naturels dans le champ de forces MARTINI.....	53
Figure 21 : Interactions atomiques dans le champ de forces MARTINI EIneDyn.....	55
Figure 22 : Représentation de l'énergie d'un système .....	57
Figure 23 : Représentation de la minimisation par la méthode de la plus grande pente .....	58
Figure 24 : Représentation des conditions périodiques aux limites .....	60
Figure 25 : Représentation de la distribution de l'énergie potentielle en REMD .....	61
Figure 26 : Schématisation du demultiplexage ou « demux ».....	62
Figure 27 : Représentation schématique de l'algorithme utilisé pour le clustering .....	65
Figure 28 : Modèles de départ de NTR1r et de NT[8-13] après construction.....	67
Figure 29 : Modèles de départ de CXCR4 et de CVX15 après construction .....	69
Figure 30 : Modèles de départ de V2R et de l'AVP après construction.....	71
Figure 31 : Structure générale et alignement des toxines de type « kunitz » .....	72
Figure 32 : Modèles de départ de GHSR-1a et de la Ghréline après construction.....	73
Figure 33 : Structure RMN du peptide LEAP2.....	74

Figure 34 : Modèle de départ du récepteur DRD2 après construction .....	75
Figure 35 : Modèles de départ des récepteurs A1R et A2AR après construction .....	76
Figure 36 : Construction des peptides C-terminaux $G_{\alpha}$ .....	77
Figure 37 : Schématisation du processus de création des « systèmes doubles » .....	79
Figure 38 : Schématisation du processus de préparation des systèmes pour le REMD .....	79
Figure 39 : Généralisation du gradient d'éluant utilisé pour la purification par HPLC .....	85
Figure 40 : Schématisation de la modification de l'activité de $G_i$ .....	86
Figure 41 : Schématisation de la production de l'hétéro-dimère DRD2:GHSR-1a .....	86
Figure 42 : Comparaison des conformations liées de CVX15 .....	112
Figure 43 : Comparaison des orientations de NT[8-12] liée .....	113
Figure 44 : Evolution de la liaison des peptides « scrambled » de la Neurotensine.....	115
Figure 45 : Schématisation du processus de construction du dimère .....	148
Figure 46 : Structures des segments extraits de $G_i$ à l'interface $G_i:G_q$ .....	161
Figure 47 : Schéma réactionnel de l'accrochage du premier acide aminé à la résine .....	162
Figure 48 : Schéma réactionnel de la synthèse des peptides linéaires.....	163
Figure 49 : Schéma réactionnel de la synthèse du peptide cyclique .....	164
Figure 50 : Mesure de l'effet des peptides synthétisés.....	167
Figure 51 : Nomenclature du tétramère.....	168
Figure 52 : Rotation observées lors des simulations partant du dimère DRD2:GHSR-1a.....	171
Figure 53 : Alignement des homo-dimères DRD2 obtenus sur le dimère DRD2:GHSR-1a .....	173
Figure 54 : Modèle du tétramère proposé par le résultats des simulations CG-REMD .....	174
Figure 55 : Modèles de la MQ-1 liée au récepteur V2R obtenus .....	177
Figure 56 : Dynamique de liaison de MQ-1.....	180
Figure 57 : Evolution de la liaison des différentes toxines.....	182
Figure 58 : Modèles de la Vasopressine liée au V2R obtenus .....	183
Figure 59 : Structure RMN du peptide de la Ghréline liée.....	185
Figure 60 : Représentation des élastiques MARTINI ELNeDyn appliqués à la Ghréline .....	186
Figure 61 : Conformations les plus représentatives de la Ghréline obtenues.....	187
Figure 62 : Représentation des contacts entre la Ghréline et son récepteur GHSR-1a .....	189
Figure 63 : Présence des différents clusters en fonction de la température.....	190
Figure 64 : Etudes structure-activité du LEAP2.....	191
Figure 65 : Modèles liés des peptides LEAP2 obtenus et comparaison avec la Ghréline.....	192
Figure 66 : Flexibilité des récepteurs et des peptides en simulations CG .....	194
Figure 67 : Contacts peptide:récepteur non spécifiques des peptides C-terminaux .....	194
Figure 68 : Résultats des simulations CG-REMD réalisées sur les peptides C-terminaux .....	195

# LISTE DES TABLEAUX

Tableau 1 : Populations et structures résolues pour les sous-familles de RCPGs .....	11
Tableau 2 : Liste des effecteurs des quatre familles de la sous-unité $\alpha$ et du complexe $\beta\gamma$ .....	20
Tableau 3 : Niveaux d'interactions entre les particules gros-grains dans MARTINI .....	53
Tableau 4 : Matrice d'interaction entre les particules gros-grains dans MARTINI .....	54
Tableau 5 : Séquences des différents peptides dérivés de la Neurotensine .....	68
Tableau 6 : Composition des différents échantillons pour les tests d'activité $G_i$ .....	87
Tableau 7 : Séquences du peptide Neurotensine et de ses peptides « scrambled ».....	114
Tableau 8 : Distances mesurées lors de simulations du dimère et validées par LRET .....	149
Tableau 9 : Séquences des peptides synthétisés .....	162
Tableau 10 : Mesure de la charge des résines avec le premier acide aminé greffé .....	163
Tableau 11 : Composition des échantillons préparés pour les tests d'activité de $G_i$ .....	166
Tableau 12 : Représentation simplifiée des principales interfaces de dimérisation .....	170
Tableau 13 : Principales mutations réalisées sur la toxine et le récepteur.....	179
Tableau 14 : Alignement et $K_i$ envers V2R de différentes toxines.....	181

# LISTE DES EQUATIONS

Equation 1 : Energie potentielle du système .....	35
Equation 2 : Intéractions de Van der Waals .....	35
Equation 3 : Intéractions électrostatiques.....	35
Equation 4 : Liaison de deux particules .....	35
Equation 5 : Angle formé par trois particules .....	35
Equation 6 : Angle dièdre formé par quatre particules.....	35
Equation 7 : Angle dièdre impropre formé par quatre particules .....	35
Equation 8 : Force d'un élastique MARTINI E1NeDyn.....	35
Equation 9 : Seconde loi de Newton .....	59
Equation 10 : Dérivée partielle de la fonction d'énergie potentielle .....	59
Equation 11 : Probabilité d'échange des coordonnées en REMD .....	60
Equation 12 : RMSD .....	64
Equation 13 : Relation Absorbance-Charge de la résine .....	83

# LISTE DES ABREVIATIONS

## Biologie

5-HW	5-Hydroxytryptophane
A1R	Récepteur de l'Adénosine A1
A2AR	Récepteur de l'Adénosine A2A
$\beta$ 2AR	Récepteur Adrénérgique $\beta$ 2
BRET	Transfert d'Énergie de Bioluminescence par Résonance
Cryo-EM	Cryo-Microscopie Electronique
CXCR4	Récepteur de Chimiochine C-X-C de type-4
DRD2	Récepteur de la Dopamine D2
DTX-E	Dendrotoxine-E
ECL	Boucle Extra-Cellulaire
Fmoc	Fluorénylméthoxycarbonyl
FRET	Transfert d'Énergie de Fluorescence par Résonance
GDP	Guanosine Diphosphate
GHSR-1a	Récepteur de la Ghreline-1a
GRK	Kinase-RCPG
GTP	Guanosine Triphosphate
H8	Hélice 8
ICL	Boucle Intra-Cellulaire
LEAP2	Peptide Antimicrobien du Foie 2
LRET	Transfert d'Énergie de Luminescence par Résonance
MQ-1	Mambaquarétine-1
NC-IUPHAR	« Union of Pharmacology, Comitee on Receptor Nomenclature and Classification »
nOe	Effet Overhauser Nucléaire
NT	Neurotensine
NTR1	Récepteur de la Neurotensine-1
NTR1r	Récepteur de la Neurotensine-1 du rat
NTR2	Récepteur de la Neurotensine-2
pG	Protéine-G
POPC	1-palmitoyl-2-oléoyl-sn-glycéro-3-phosphocholine
RCPG	Récepteur Couplé aux Protéines-G
RGS	Protéine Régulatrice du Signal des Protéines-G
RMN	Spectrométrie par Résonance Magnétique Nucléaire
TM	Domaine Trans-Membranaire
TROESY	« TRansferred nuclear Overhauser Enhancement SpectroscopY »
V1aR	Récepteur de la Vasopressine de type-1a
V2R	Récepteur de la Vasopressine de type-2

## Simulation

AA (simulation)	Tout-Atomes
ACP	Analyse en Composante Principale
CG	Gros-Grains
CP	Composante Principale
EINeDyn	« Elastic Network in Dynamics »
GPU	Processeur Graphique
GROMACS	« GRONingen MACHine for Chemical Simulations »
MD	Dynamique Moléculaire
MM	Mécanique Moléculaire
PBC	Conditions Périodiques aux Limites
PDB	« Protein Data Bank »
REMD	Dynamique moléculaire « Replica-Exchange »
RMSD	« Root Mean Square Deviation »
SD	Méthode de la Plus Grande Pente
WHAM	Méthode d'Analyse d'Histogrammes Pondérés

## Synthèse

AA (synthèse)	Acide Aminé
AcN	Acétonitrile
Clt	Résine 2-chloro Chlorotrytle
DCM	Dichlorométhane
DIEA	N,N-Diisopropyléthylamine
DMF	N,N-Diméthylformamide
Et <sub>2</sub> O	Ether Diéthylique
HATU	Hexafluorophosphate de (diméthylamino)-N,N-diméthyl(3H-[1,2,3]triazolo[4,5-b]pyridin-3-yloxy)méthaniminium
HCO <sub>2</sub> H	Acide Méthanoïque
HPLC	Chromatographie en Phase Liquide à Haute Performance
LC-MS	Chromatographie en Phase Liquide - Spectrométrie de Masse
MeOH	Méthanol
NaCl	Chlorure de Sodium
Pip	Pipéridine
SPPS	Synthèse Peptidique sur Support Solide
TA	Température Ambiante
TFA	Acide Trifluoroacétique
TIS	Triisopropylamine
TRIS	Trishydroxyméthylaminométhane



# Chapitre 1

## Introduction



© MRAH



# 1 Introduction

## 1.1 Les Récepteurs Couplés aux Protéines-G

Les Récepteurs Couplés aux Protéines-G (RCPGs) sont les membres de la famille de protéines la plus représentée chez l'espèce humaine. En effet, plus de 800 RCPGs ont été identifiés chez l'homme [Lagerström et Schiöth 2008] même si l'on ne connaît pas la fonction réelle d'un certain nombre d'entre eux. Ces récepteurs membranaires, par leur forte implication au niveau physiologique et pathologique, constituent actuellement plus d'un tiers des cibles thérapeutiques de médicaments sur le marché [Sriram et Insel 2018]. Elle représente également la plus large famille génétique chez les mammifères avec plus de 3% des gènes. Leur nom est issu de leur capacité à se lier aux protéines-G (pGs), elles-mêmes nommées ainsi car capables de lier les Guanosine Di et Triphosphate (GDP et GTP). Les RCPGs sont également capables de se lier à une large variété de ligands qui incluent de petites molécules organiques, des peptides, des protéines, mais aussi des lipides [Fredriksson et al. 2003; Venkatakrishnan et al. 2013]. Près de la moitié de ces récepteurs sont d'ailleurs capables de s'activer en réponse à des stimuli extérieurs, comme les photons. Ces récepteurs sensoriels sont impliqués dans la vision, la perception des odeurs et du goût sucré, amer et umami [Julius et Nathans 2012]. Ces récepteurs sont retrouvés de manière ubiquitaire dans toutes les cellules, chaque type cellulaire exprimant néanmoins un set de récepteurs qui lui est propre. En novembre 2017, il est estimé qu'un peu plus d'une centaine de RCPGs sont la cible d'un médicament sur le marché [Hauser et al. 2017]. Au vu du nombre de celles-ci, les RCPGs forment donc encore un vivier de cibles potentielles particulièrement riche pour le développement futur de drogues.

### 1.1.1 Historique

C'est au début du 20<sup>ème</sup> siècle que le pharmacologiste Langley décrivit précisément les notions de substance réceptrice et d'effecteur [Langley 1905]. Longtemps, ce concept ne fit pas l'unanimité, et ce malgré une forte progression du domaine des années 1920 à 1970 par des pionniers tels qu'Ariens, Clark, Dale, Stephenson, Black, Gaddum, Schild et Furchgott [Limbird 1996]. Dans les années soixante, Sutherland montra l'existence d'un second messenger, l'Adénosine Monophosphate Cyclique (AMPC), comme résultat de l'activation de certains récepteurs [Robison, Butcher, et Sutherland 1967]. Au même moment Krebs découvrait les kinases AMPC-dépendantes, un effecteur direct de l'AMPC [Walsh, Perkins, et

Krebs 1968]. En 1971, Rodbell proposa l'existence d'une protéine régulatrice du nucléotide Guanine qui serait un transducteur entre les récepteurs d'hormones et l'Adénylyl cyclase [Rodbell et al. 1971], cette protéine sera nommée quelques années plus tard  $G_s$  [Gilman 1987]. Ces découvertes permettront une forte progression du domaine car il était désormais possible de suivre l'activation d'un récepteur par le suivi de cette voie de signalisation, au lieu des méthodes à l'époque classiques mais moins précises comme la mesure de la contraction des myocytes.

Dans les années soixante-dix, l'existence de molécules couplées aux enzymes produisant les seconds messagers restait à prouver. Pour se faire, des techniques de liaison de radio ligand furent développées. Par exemple le laboratoire de Snyder développa des ligands pour des récepteurs de neurotransmetteurs comme les récepteurs opiacés [Pert et Snyder 1973]. En une décennie, on vit l'explosion du nombre de radio ligands disponibles pour étudier les phénomènes de liaisons aux RCPGs. Ceci permit d'obtenir des informations importantes sur la régulation dynamique de nombreux récepteurs, mais également sur le couplage entre ceux-ci et les protéines-G [De Lean, Stadel, et Lefkowitz 1980; Kent, De Lean, et Lefkowitz 1980]. Une autre conséquence fut la capacité à identifier spécifiquement un RCPG en utilisant les radios ligands, ce qui permit ensuite de les purifier sélectivement et donc de les isoler. Aussi, il fut possible de découvrir un certain nombre de nouveaux sous-types de RCPGs. Cette purification, le laboratoire de Lefkowitz y participa activement, ce qui amena à l'isolement du récepteur Adrénergique  $\beta_2$  ( $\beta_2AR$ ) [Shorr, Lefkowitz, et Caron 1981], puis des récepteurs Adrénergiques  $\alpha_{1B}$  et  $\alpha_{2A}$  [Lomasney et al. 1986; Regan et al. 1986].

En 1983, la séquence de la Rhodopsine fut déterminée par dégradation d'Edman, technique de séquençage classique des protéines [Hargrave et al. 1983], suggérant la présence de sept domaines hydrophobes, possiblement trans-membranaires. A l'époque, seule la Bactériorhodopsine également impliquée dans la réponse aux signaux lumineux, était séquencée et possédait également les sept domaines hydrophobes [Khorana et al. 1979]. Il fut émis l'hypothèse qu'il s'agissait d'une caractéristique des récepteurs sensibles aux photons. Toutefois en 1986, il fut possible de cloner le gène du récepteur  $\beta_2AR$ , facilité par le fait qu'il ne soit codé que par un seul exon [Dixon et al. 1986]. Ces différents travaux de recherche permirent déjà à l'époque de mettre en évidence des caractéristiques communes à tous les RCPGs (**Figure 1**) : sept domaines transmembranaires (TM) hydrophobes connectés par des boucles intra et extra-cellulaires (ICL et ECL respectivement) possédant un site de glycosylation conservé et des sites de phosphorylation dans la partie C-terminale.

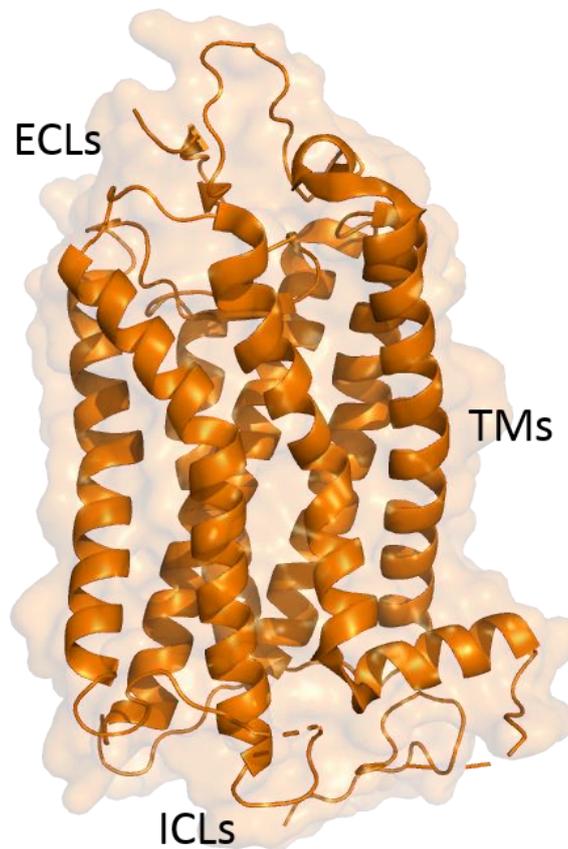


Figure 1 : Structure générale d'un Récepteur Couplé aux Protéines-G ; ici la structure de la Rhodopsine (code PDB : 1F88) [Palczewski et al. 2000].

Depuis ces travaux fondateurs, le nombre de RCPGs isolés a considérablement augmenté, même si de nombreux récepteurs, appelés récepteurs orphelins, n'ont pas encore de ligand endogène identifié. L'augmentation de la sensibilité des méthodes de tests pharmacologiques *in vitro* a également permis de mettre en évidence une activité constitutive (ou activité basale) pour nombre de ces récepteurs [Leurs et al. 2000; Seifert et Wenzel-Seifert 2002], des effets divers de leurs ligands selon leur capacité à stimuler ou inhiber leur activité : agoniste, antagoniste, agoniste inverse [Latorraca, Venkatakrishnan, et Dror 2017] ou encore les effets allostériques liés à leur dimérisation/oligomérisation [Ferré et al. 2014]. Des études récentes ont également montré que leur activité pouvait être régulée par la composition même de la membrane [Inagaki et al. 2012; Dawaliby et al. 2016; Casiraghi et al. 2016; Ramírez-Anguita et al. 2018]

Ces dernières années, la détermination de la structure des RCPGs par cristallographie ou autres méthodes de résolution structurale a connu un véritable essor. Depuis la résolution de la structure de la Rhodopsine en 2000 [Palczewski et al. 2000], de nombreuses autres structures

ont été déposées dans la « Protein Data Bank » (PDB) montrant ces récepteurs seuls ou complexés à différents ligands agonistes, antagonistes ou agonistes inverses, ou encore montrant ces récepteurs couplés à leurs partenaires intra-cellulaires favoris : protéines-G [Rasmussen et al. 2011; García-Nafría et al. 2018; Koehl et al. 2018] ou  $\beta$ -Arrestine [Zhou et al. 2017].

Ces structures constituent autant de modèles de départ possibles pour des études de modélisation moléculaire, avec néanmoins quelques précautions à prendre. En effet, la plupart du temps, la cristallisation de ces récepteurs nécessite des modifications structurales non négligeables comme par exemple l'insertion du côté intra-cellulaire du Lysozyme T4 entre les TM5 et TM6, ou encore la troncation de ces récepteurs du côté N ou C-terminal, quelquefois sur des segments de plus de 30 résidus. Même si ces structures apportent des informations clés à l'échelle moléculaire, elles décrivent un nombre encore assez limité de récepteurs (environ 50). Il est aussi à noter que la plupart de ces récepteurs semblent cristalliser dans un même état conformationnel, et ce quelque-soit le ligand. Il y a donc une barrière technologique qui doit encore être franchie afin d'élargir ces données structurales et décrire la grande variabilité conformationnelle de ces récepteurs à l'échelle moléculaire, comme montré par les études biophysiques récentes [Preininger, Meiler, et Hamm 2013; Latorraca, Venkatakrishnan, et Dror 2017; Vaidehi, Bhattacharya, et Larsen 2014]. Ces deux dernières années, les avancées en Cryo-Microscopie Electronique (Cryo-EM) ont permis d'obtenir de nombreuses nouvelles structures, et d'apporter une vue un peu plus dynamique de ces récepteurs. Malgré tout, et ce quelque-soit la méthode de résolution structurale utilisée, le résultat reste une vue figée de ces systèmes qui pourtant subissent des mouvements de large amplitude lors de leur activation, comme montré par des études biophysiques récentes [V. V. Gurevich et Gurevich 2017; Latorraca, Venkatakrishnan, et Dror 2017; Schihada et al. 2018]. Ces changements de conformation apparaissent à des échelles de temps qui restent trop courtes pour être caractérisées à l'échelle structurale (de l'ordre de la  $\mu$ s) [Dror et al. 2012]. C'est pourquoi l'étude par modélisation moléculaire des RCPGs apparaît comme une alternative qui vient compléter cette vue rigide de ces systèmes et permettre la prédiction de leurs mouvements à l'échelle moléculaire. Ces 10 dernières années, les avancées informatiques permettant notamment l'utilisation de GPUs pour le calcul scientifique mais aussi le développement de modèles (champs de forces) simplifiés comme les modèles gros-grains, ont rendu presque facile l'étude de ces systèmes à une échelle de temps bien supérieure à la  $\mu$ s.

### 1.1.2 Structures des RCPGs

Les Récepteurs Couplés aux Protéines-G sont caractérisés par sept domaines transmembranaires (TM) structurés en hélice et qui traversent de part en part la bicouche lipidique de la membrane cellulaire. Ces segments sont reliés entre eux par trois boucles extra-cellulaires (ECL) et trois boucles intra-cellulaires (ICL). La partie N-terminale, de longueur, rôle et composition très variés d'une famille de récepteur à l'autre, se trouve à l'extérieur de la cellule alors que la partie C-terminale se trouve à l'intérieur. Chez les récepteurs de la sous-classe A (Rhodopsine), le segment C-terminal adopte souvent une structure en hélice et se positionne à plat sur la face intra-cellulaire de la membrane, ce segment est également appelé « hélice 8 » [Cherezov et al. 2007].

Du côté extra-cellulaire, les sept segments membranaires forment ensemble une cavité servant de poche de liaison au ligand. Selon les récepteurs, cette poche est plus ou moins grande, et plus ou moins positionnée en profondeur [Katritch, Cherezov, et Stevens 2012]. Du côté intra-cellulaire les RCPGs se lient à différents partenaires, notamment (comme leur nom l'indique) aux protéines-G (**Figure 2**). De nombreuses structures sont disponibles décrivant ces partenaires privilégiés des RCPGs [Rasmussen et al. 2011; García-Nafría et al. 2018; Koehl et al. 2018]. Il s'agit d'un hétéro-trimère composé des sous-unités  $\alpha$ ,  $\beta$  et  $\gamma$ . La sous unité  $\alpha$  se compose d'une part d'un domaine dit « ras-like », retrouvé dans les petites protéines-G [J. Cherfils et al. 1997] et composé d'un feuillet  $\beta$  à six brins et de cinq hélices  $\alpha$ , et d'autre part d'un domaine hélicoïdal, très mobile [Preininger, Meiler, et Hamm 2013] et difficile à observer en cristallographie en l'absence du GDP [Rasmussen et al. 2011]. La sous unité  $\beta$  se compose d'une hélice dans sa partie N-terminale et d'un ensemble de sept feuillets  $\beta$  à quatre brins antiparallèles. Elle interagit avec la sous-unité  $\gamma$ , structurée en hélice. Il est à noter ici que les différentes sous-unités des protéines-G sont la plupart du temps ancrées à la membrane *via* des résidus modifiés par des chaînes grasses.

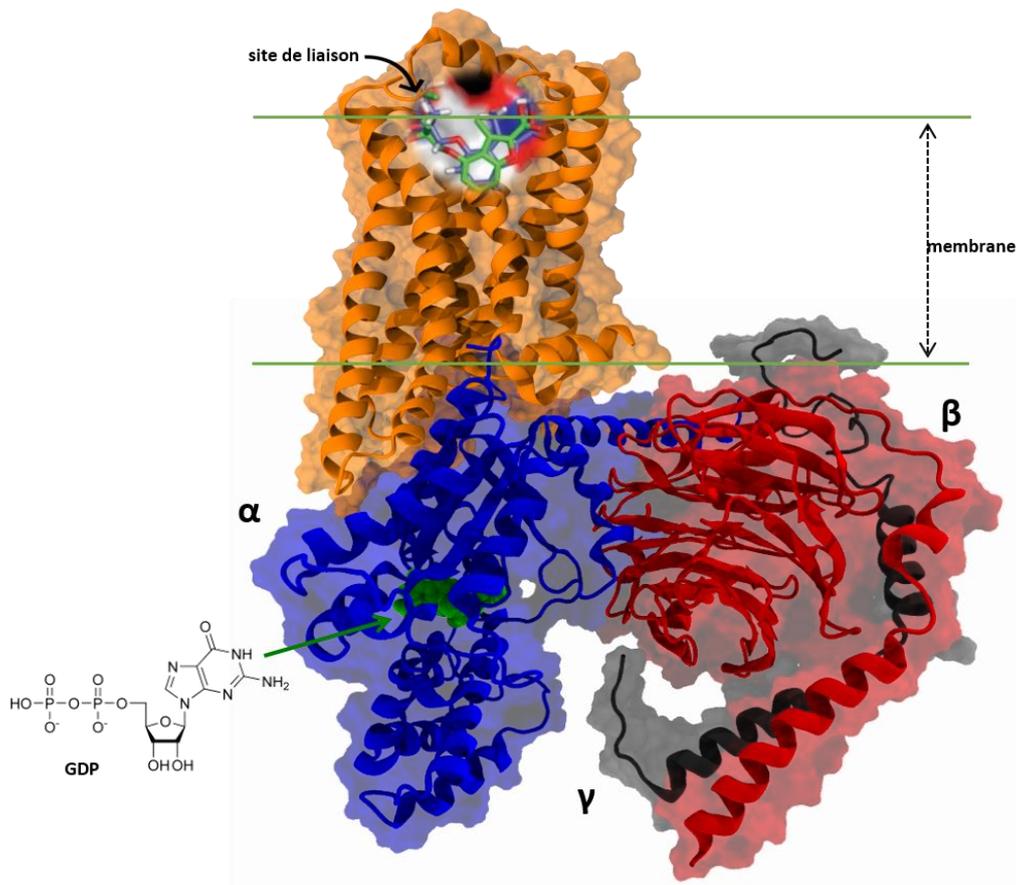


Figure 2 : Modèle d'un RCPG (en orange) complexé à une protéine-G hétéro-trimérique; ici le récepteur  $\beta 2AR$  (code PDB : 3SN6) complexé à  $G_i$  [Rasmussen et al. 2011], construite d'après sa structure cristallographique (code PDB : 1GP2) [Wall et al. 1995].

C'est en 2000 qu'est publiée la première structure d'un RCPG. Il s'agit de la structure cristallographique de la Rhodopsine bovine [Palczewski et al. 2000]. Cette structure restera unique jusqu'en 2007 et la publication de la structure du récepteur Adrénérique  $\beta 2$  ( $\beta 2AR$ ) [Rasmussen et al. 2007]. Les années suivantes, de nombreuses structures seront publiées décrivant de nombreux autres récepteurs liés à des ligands de natures très diverses (**Figure 3**), mais n'apportant finalement que peu de nouvelles informations clés quant à la structure et/ou la dynamique des RCPGs. La grande avancée suivante réside dans la publication de la première structure d'un complexe RCPG:pG, à savoir le complexe  $\beta 2AR:G_s$  [Rasmussen et al. 2011].

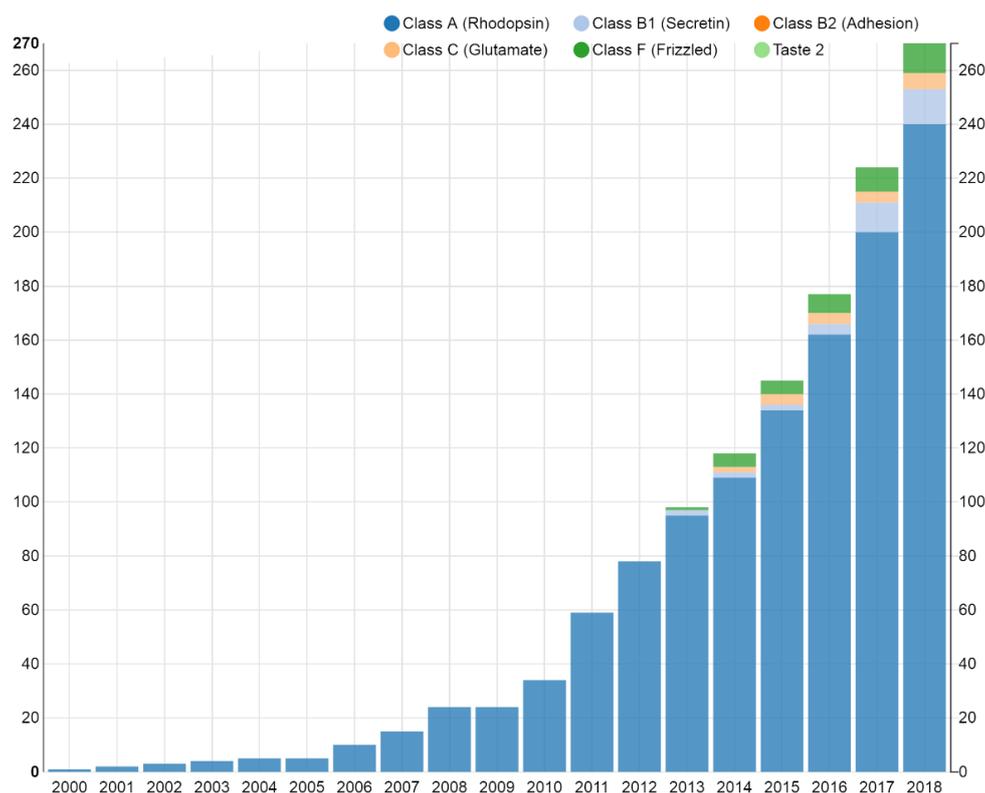


Figure 3 : Evolution du nombre de structures de RCPGs disponibles dans la PDB ; d'après la base de données GPCRdb.org au 1<sup>er</sup> Septembre 2018.

Dans tous les cas en revanche, la dynamique de ces récepteurs a été limitée en vue de leur cristallisation, *via* la stabilisation de leur structure par plusieurs biais : la co-cristallisation avec un ligand de bonne affinité et stabilisant le récepteur dans un état pseudo-inactif (antagoniste ou agoniste inverse), la production de récepteurs mutés et thermo-stables par mutagenèse dirigée et/ou l'ajout d'une protéine de fusion dans la boucle ECL3 ou d'un anticorps intracellulaire. À l'heure actuelle, les structures de 52 RCPGs différents ont été résolues, majoritairement présentant une conformation inactive [GPCRdb.org].

### 1.1.3 Les différentes familles de RCPGs

Il y a trente-cinq ans, Hargrave publiait la première structure primaire (séquence) d'un RCPG : le récepteur Rhodopsine [Hargrave et al. 1983]. Cette protéine située dans la membrane plasmique était capable de transférer l'énergie d'un signal lumineux externe vers l'intérieur de la cellule. Depuis, de nombreux RCPGs ont été identifiés et caractérisés, nécessitant de les regrouper en sous-familles [Lefkowitz 2004], selon leur séquence, leur capacité à lier différents types de ligands ou également en fonction des réponses biologiques qu'ils initient.

Au cours du temps, de nombreuses classifications ont vu le jour. Puisque toutes les séquences de RCPGs ont été identifiées, la plus exhaustive est basée sur une classification phylogénétique de celles-ci. Attwood et Finland ont en 1993 classé les RCPGs en se basant sur les empreintes de séquence des sept domaines TMs, spécificité structurale des RCPGs [Attwood et Findlay 1993]. En 1994, Kolakowski proposa une classification comportant toutes les protéines pouvant se lier aux protéines-G et non seulement les récepteurs de type Rhodopsine [Kolakowski 1994]. Ce système nomme les groupes ainsi formés par une lettre allant de A à F, et O pour ceux possédant bien les sept TMs mais ne se liant pas aux protéines-G. Parmi ces classes, les récepteurs de phéromones (classe D) et les récepteurs de l'Adénosine monophosphate cyclique (classe E) sont des RCPGs qui ne sont pas présents chez l'homme. Une variation de cette nomenclature est utilisée par l' « Union of Pharmacology, Comitee on Receptor Nomenclature and Classification » (NC-IUPHAR). Bockaert et Pin proposèrent en 1999 une classification similaire dans laquelle les RCPGs sont regroupés dans des familles numérotées de 1 à 5 en fonction de leur structure mais également de la nature de leur ligand [Bockaert et Pin 1999].

En 2001, l'accès au génome humain a permis de récupérer la majeure partie des séquences complètes des RCPGs [Venter et al. 2001]. Indépendamment, Fredriksson se basa sur la phylogénie des 802 RCPGs connus pour les séparer en famille [Fredriksson et al. 2003]. Il s'agit de la classification GRAFS, acronyme du nom des 5 principales familles : Glutamate, Rhodopsine, Adhésion, Frizzled/Taste2 et Sécrétine. On retrouve ici une classification proche de celle de Kolakowski. Le groupe A devint la famille Rhodopsine, accueillant la majeure partie des protéines, et se vit alors divisé en quatre sous-familles. Le groupe B quant à lui, se sépara pour devenir les groupes Sécrétine et Adhésion (respectivement B1 et B2). Il rajoute également un nouveau groupe de récepteurs découvert peu de temps avant : Taste2 [Adler et al. 2000].

La principale nomenclature utilisée aujourd'hui est celle de Fredriksson même si les différentes familles sont toujours désignées par leur nom de classe utilisé par NC-IUPHAR et issu de la classification de Kolakowski. Pour désigner les familles dans ce manuscrit, les deux seront utilisées. Le **Tableau 1** présente la répartition des RCPGs dans ces différentes sous-familles et leurs structures disponibles dans la PDB.

Classe	Famille GRAFS	Membres chez l'homme	Structures totales	Récepteurs uniques	Complexés à un ligand	Complexés avec une protéine-G
A	Rhodopsine	689	240	45	233	8
B1	Sécrétine	15	13	4	12	3
B2	Adhésion	33	0	0	0	0
C	Glutamate	22	6	2	6	0
F	Frizzled	11	11	1	11	0
T	Taste2	25	0	0	0	0
Total		795	270	52	262	11

Tableau 1 : Populations et nombre de structures résolues pour chacune des sous-familles de RCPGs ; les 6 RCPGs de la classe O (Autres) humains ont été exclus de ce tableau ; d'après GPCRdb.org au 1<sup>er</sup> Septembre 2018.

### 1.1.3.a Rhodopsine ou classe A

La famille des récepteurs Rhodopsine ou classe A correspond à la plus vaste famille avec près de 700 membres et, de par le nombre de récepteurs qu'elle contient, est la principale cible des molécules à visée thérapeutique. Cette sous-classe de récepteurs possède une grande diversité, à la fois au niveau structural mais aussi de par la nature des ligands qu'elle englobe : peptides, lipides ou odorants. Toutefois, cette diversité n'est pas contenue dans la partie N-terminale de ces récepteurs, généralement assez courte, mais dans les TMs. Il y a tout de même des exceptions comme le récepteur de la Thrombine qui possède un site de clivage intrinsèque dans sa partie N-terminale, et qui peut servir de ligand tant qu'il n'est pas clivé par la Thrombine [Vu et al. 1991]. Cette hétérogénéité des récepteurs de la classe A a conduit à une subdivision en 4 sous-groupes :  $\alpha$ ,  $\beta$ ,  $\gamma$ , et  $\delta$ .

La structure de la Rhodopsine publiée en 2000 a permis de montrer un réseau de connexions inter-hélice essentiel pour la stabilisation et l'activation du récepteur et commun à la grande partie des récepteurs de la classe A [Palczewski et al. 2000]. Quelques années plus tard, la structure du  $\beta$ 2AR apporte un second modèle de RCPG « Rhodopsine-like » [Cherezov et al. 2007]. La comparaison de ces deux structures montre déjà toute la problématique de se baser sur un modèle unique pour prédire la structure de toute cette famille de RCPGs, avec des différences drastiques dans la partie extra-cellulaire, mais aussi dans l'orientation de certains résidus des TMs.

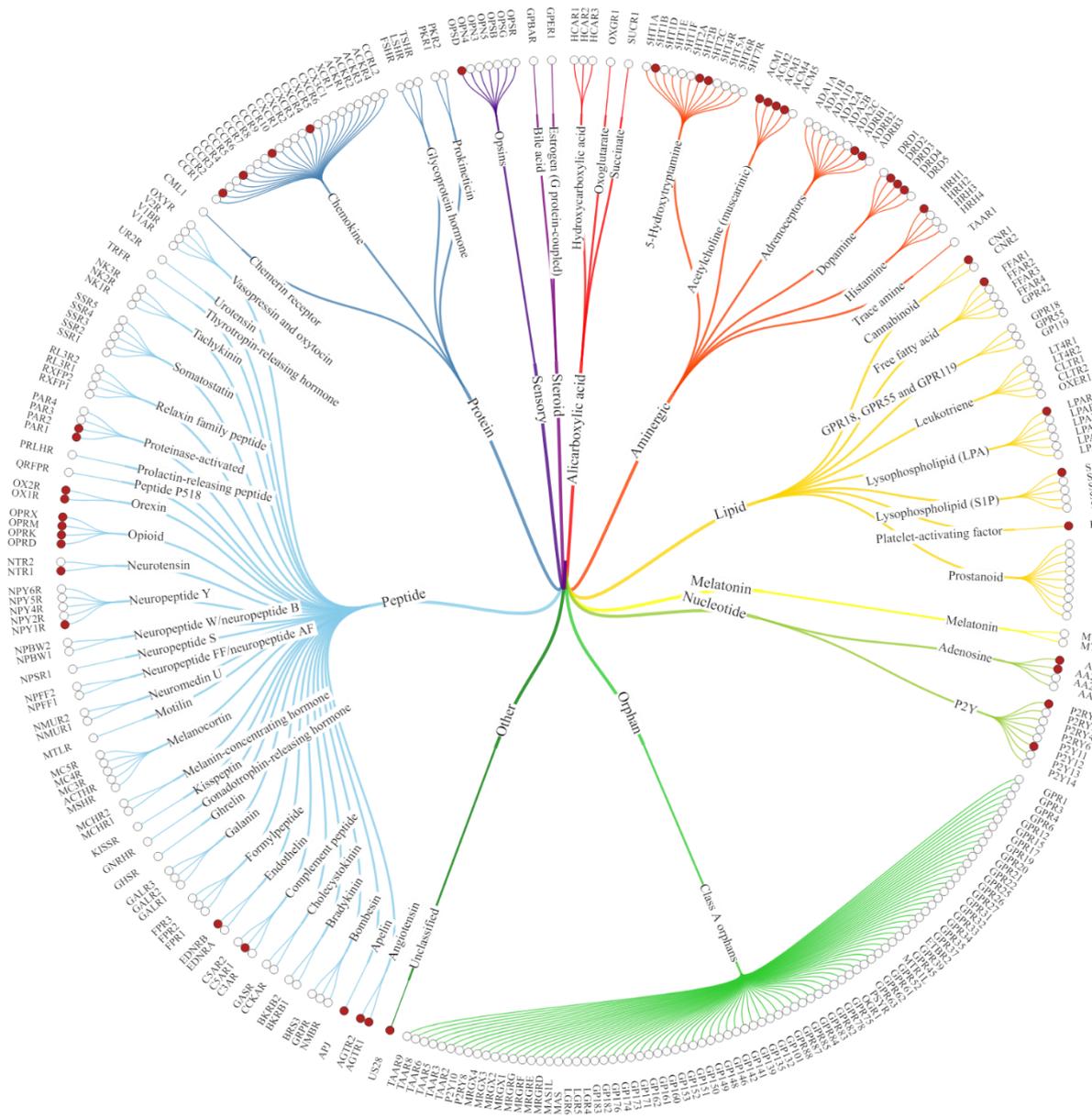


Figure 4 : Classification des récepteurs de la sous-famille Rhodopsine (classe A); les points rouges correspondent aux récepteurs dont la structure (ou partie de la structure) est résolue ; d’après GPCRdb.org au 1<sup>er</sup> Septembre 2018.

Le sous-groupe  $\alpha$  comporte d’importantes cibles thérapeutiques comme les récepteurs de la Dopamine, les récepteurs de l’Histamine, les récepteurs de la Sérotonine ou encore les récepteurs Adrénérgiques. Ces ligands, essentiellement amines, se lient dans le site de liaison orthostérique délimité par les TMs. Il s’agit de cibles thérapeutiques très importantes avec par exemple les récepteurs Muscariniques impliqués dans le développement des cellules tumorales lors de cancers, mais également la cible de traitements traitant la schizophrénie à cause de son rôle de régulateur de la neurotransmission de la Dopamine et du Glutamate [Spindel 2012; Bymaster et al. 2002].

Le sous-groupe  $\beta$  est composé principalement de récepteurs liant des peptides et dont le site de liaison se trouve entre les TMs avec la participation des boucles extra-cellulaires et de la partie N-terminale. Parmi ceux-ci on retrouve les récepteurs de l'Endothéline, de la Gonadotropine, de l'Ocytocine, mais également le récepteur du Neuropeptide Y. Ce dernier est d'ailleurs une cible thérapeutique intéressante. En effet, le Neuropeptide Y est impliqué dans un certain nombre de maladie comme l'athérosclérose, l'hypertension, l'obésité, l'épilepsie et des dérèglements métaboliques [M. Yi et al. 2018]. De plus, celui-ci est surexprimé dans de nombreux types de cancers, ce qui en fait un biomarqueur pour le diagnostic et le traitement de cette maladie [Li, Tian, et Wu 2015].

Le sous-groupe  $\gamma$  regroupe des récepteurs se liant à des peptides et des lipides et contient les récepteurs Opioïdes, les récepteurs de l'Angiotensine ou encore les récepteurs de Chimiochine, notamment impliqués dans la propagation du Virus de l'Immunodéficience Humain (VIH) et certains cancers [Fätkenheuer et al. 2005; Keane et al. 2004].

Le sous-groupe  $\delta$  contient des récepteurs se liant à des glycoprotéines et à des nucléotides, et d'autres récepteurs activés par les protéases. Mais ce groupe est majoritairement constitué des récepteurs olfactifs avec un nombre estimé à 400 chez l'homme, et qui sont divisés en deux classes I et II. Leur site de liaison se trouve sur les parties les plus extra-cellulaires des TM3, TM5, TM6, TM7 et l'ECL2 [Man, Gilad, et Lancet 2004]. Ils peuvent être activés par une grande variété de molécules odorantes (ou odorants), peu spécifiques. En effet, un même odorant est capable la plupart du temps d'activer plusieurs récepteurs olfactifs [Malnic et al. 1999]. Dans ce champ d'application, il y a une notion de code combinatoire qui pourrait permettre à la cellule de reconnaître un grand nombre d'odorants. Il est à noter que la majorité des récepteurs olfactifs demeurent orphelins.

### **1.1.3.b Sécrétine ou classe B1**

La famille des récepteurs Sécrétine ou classe B1 comprend 15 membres. Ces récepteurs ont tous un domaine extra-cellulaire qui lie des hormones peptidiques dont la Calcitonine, le Glucagon, la Sécrétine ou des peptides intestinaux vasoactifs. Leurs sites de liaison se trouvent dans leur partie N-terminale très variable d'un récepteur à l'autre mais possédant toute trois ponts disulfures, et les trois boucles extra-cellulaires. Il a été proposé que la liaison de l'hormone permettrait de faire un pont physique (mais non covalent) entre ces deux sites de liaison, stabilisant ainsi le récepteur dans un ensemble de conformations actives [Unson et al. 2002; Grace et al. 2004].

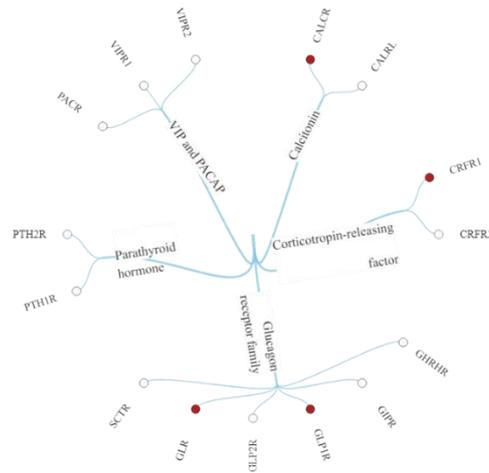


Figure 5 : Classification des récepteurs de la sous-famille Sécrétine (classe B1) ; les points rouges correspondent aux récepteurs dont la structure (ou partie de la structure) est résolue ; d'après GPCRdb.org au 1<sup>er</sup> Septembre 2018.

### 1.1.3.c Adhésion ou classe B2

La famille des récepteurs Adhésion ou classe B2 comporte 33 récepteurs. Premièrement regroupés avec les récepteurs Sécrétine, ces récepteurs se distinguent par une partie N-terminale très longue et contenant de nombreux sites fonctionnels spécifiques, ainsi qu'un domaine protéolytique. Ces domaines de la partie N-terminale sont identiques à d'autres protéines avec des domaines lectine, laminine, ou immuglobuline. Les récepteurs de cette famille se lient aux protéines de la matrice extra-cellulaire [Liebscher, Schöneberg, et Prömel 2013].

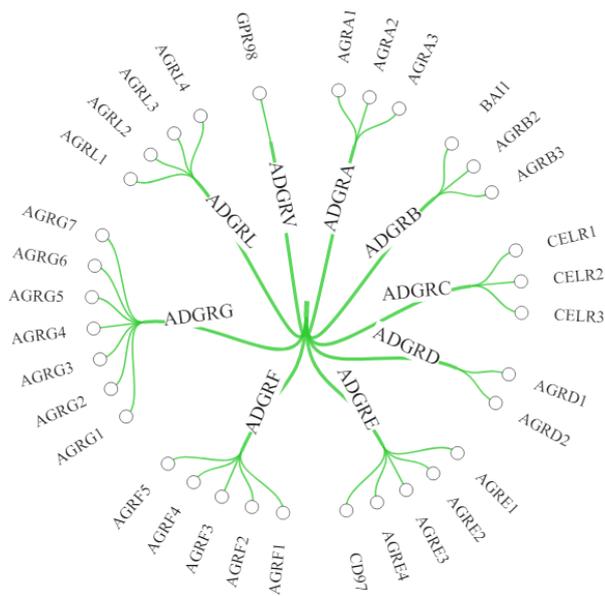


Figure 6 : Classification des récepteurs de la sous-famille Adhésion (classe B2) dont aucune structure n'est connue ; d'après GPCRdb.org au 1<sup>er</sup> Septembre 2018.

### 1.1.3.d Glutamate ou classe C

La famille des récepteurs Glutamate ou classe C comporte 22 membres avec les deux récepteurs GABA, le récepteur senseur de calcium, les huit récepteurs métabotropes du Glutamate et les récepteurs du goût sucré et umami. Le site de liaison se situe dans la région N-terminale repliée en deux domaines dont la structure tertiaire est maintenue par des ponts disulfures intra-protomériques [Kunishima et al. 2000; O'Hara et al. 1993].

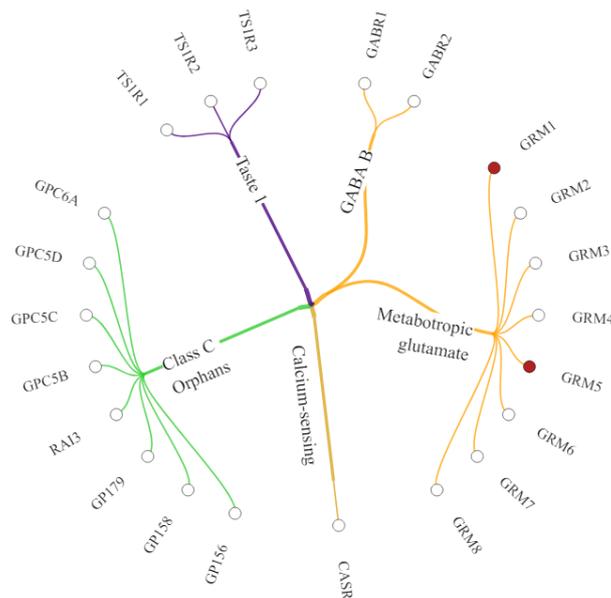


Figure 7 : Classification des récepteurs de la sous-famille Glutamate (classe C) ; les points rouges correspondent aux récepteurs dont la structure (ou partie de la structure) est résolue ; d'après GPCRdb.org au 1<sup>er</sup> Septembre 2018.

### 1.1.3.e Frizzled ou classe F

La famille des récepteurs Frizzled ou classe F regroupe les récepteurs Frizzled (FZD), au nombre de 10 chez l'homme, ainsi que le récepteur Smoothened (SMO). Ils possèdent une région N-terminale riche en cystéines dont huit sont conservées dans tous les récepteurs. Les principaux ligands des FZD sont les glycoprotéines Wnt [Dijksterhuis, Petersen, et Schulte 2014], impliquées dans l'embryogénèse et le cancer, alors que le SMO semble fonctionner sans ligand comme unité de signalisation [Murone, Rosenthal, et de Sauvage 1999].

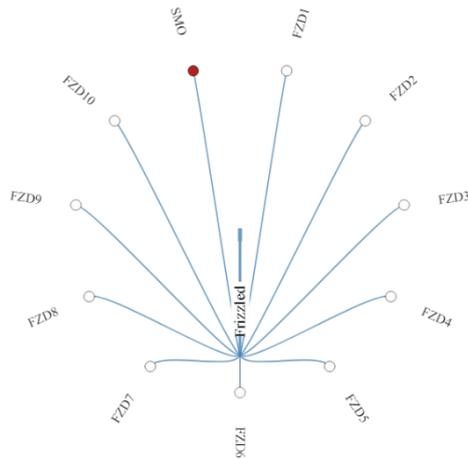


Figure 8 : Classification des récepteurs de la sous-famille Frizzled (classe F) ; les points rouges correspondent aux récepteurs dont la structure (ou partie de la structure) est résolue ; d'après GPCRdb.org au 1<sup>er</sup> Septembre 2018..

### 1.1.3.f Taste2 ou classe T

Les récepteurs de la famille Taste2 ou classe T ont d'abord été regroupés par classification phylogénétique avec la famille des récepteurs Frizzled malgré leurs grandes différences, avant de devenir une classe à part entière [Nordström et al. 2011]. La famille comporte 25 membres qui sont principalement orphelins. Les récepteurs Taste2 sont des RCPGs relativement courts avec des extrémités N et C-terminales très courtes, et ne possédant pas de pont disulfure entre les boucles extra-cellulaires [Adler et al. 2000]. On les retrouve impliqués dans les mécanismes de perception du goût amer [Chandrashekar et al. 2000].

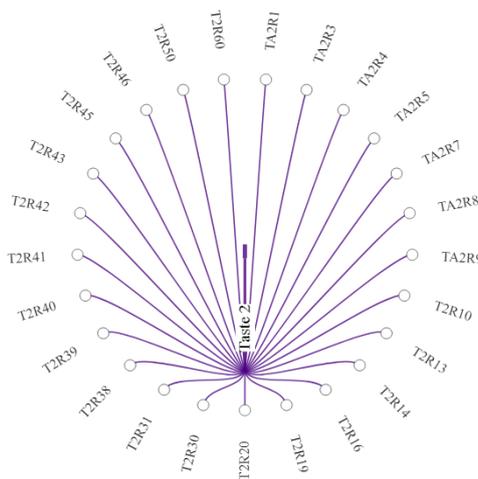


Figure 9 : Classification des récepteurs de la sous-famille Taste2 (classe T) dont aucune structure n'est connue ; d'après GPCRdb.org au 1<sup>er</sup> Septembre 2018.

## 1.1.4 Partenaires principaux des RCPGs

Les Récepteurs Couplés aux Protéines-G disposent d'un large panel de partenaires, extra et intra-cellulaires. Ces partenaires possibles incluent des ligands organiques ou peptidiques de tailles assez diverses ou bien encore des protéines. Leurs interactions avec leurs différents partenaires ont pour effet de moduler l'activité de ces récepteurs, en agissant sur leurs équilibres conformationnels intrinsèques [Manglik et Kobilka 2014]. L'activité de certains RCPGs, notamment de la classe A, peut être modulée par la liaison d'un simple ion sodium [Katritch et al. 2014]. Plus récemment, le rôle de la membrane a aussi été pointé du doigt, montrant que sa composition pouvait directement agir sur le degré d'activation de ces récepteurs [Luchetti et al. 2016; Manna et al. 2016; Bolivar et al. 2016; Casiraghi et al. 2016]. Enfin, il ne faut pas oublier ici de rappeler que ces récepteurs sont aussi capables d'interagir entre eux (homo ou hétéro-dimérisation/oligomérisation) [Ferré et al. 2014], ouvrant des possibilités de cross-régulation quasi-infinies.

### 1.1.4.a Partenaires extra-cellulaires

Comme décrit déjà plus haut, chaque récepteur peut accueillir de manière spécifique des ligands dont la nature et la taille sont très variées : médiateurs aminergiques, peptides, protéines, lipides, nucléotides, stéroïdes, acides alicarboxyliques, ions, ou encore être activés par des stimuli sensoriels externes comme des photons dans le cas de la Rhodopsine ou des odorants pour d'autres récepteurs de la classe A. La liaison de ligands dans le site orthostérique des RCPGs peut, selon les cas, stimuler (ligand agoniste) ou bloquer (antagoniste) l'activité du récepteur. Du fait que nombre de ces récepteurs sont également caractérisés par une activité basale, c'est à dire une activité en l'absence de ligand, une molécule capable de diminuer cette activité sera alors désignée comme agoniste inverse. Dans tous les cas, un ligand peut s'avérer plus ou moins efficace, on parle alors par exemple d'agoniste plein pour un ligand capable d'augmenter la réponse biologique à son maximum connu et d'agoniste partiel pour un ligand moins efficace (**Figure 10**). Cette terminologie peut également être utilisée pour les agonistes inverses. Dans ce contexte en revanche, l'antagoniste est une molécule capable de se lier au récepteur sans provoquer d'effet sur l'activité basale, mais qui peut, par compétition, bloquer l'accès d'autres ligands à la poche orthostérique.

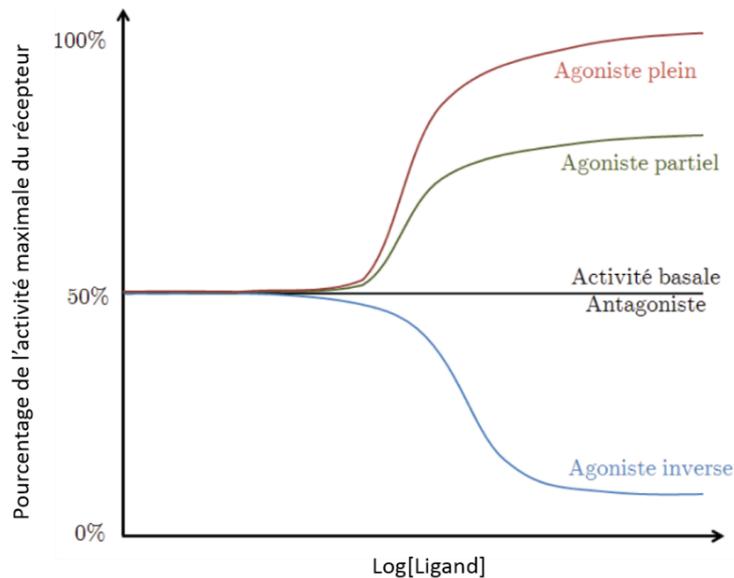


Figure 10 : Représentation des effets possibles provoqués par la liaison d'un ligand à un RCPG sur l'activité de celui-ci ; d'après [Louet 2012].

Etant donné qu'un même récepteur peut être associé à différentes voies de signalisation, il est très important de faire le lien entre cette activité et une voie de signalisation bien définie. Ainsi, un même ligand peut s'avérer être agoniste d'une certaine voie et antagoniste d'une autre en se liant au même récepteur [Galandrin et Bouvier 2006]. Les données issues de la biophysique, et notamment de la RMN montrent que ces récepteurs oscillent entre différentes conformations actives et inactives, en équilibre [Okude et al. 2015; Latorraca, Venkatakrishnan, et Dror 2017]. Cet équilibre explique l'activité basale de certains de ces récepteurs. Aussi, l'activation ou l'inactivation d'un récepteur par un ligand peut être vue comme le déplacement de cet équilibre vers une conformation particulière, permettant le couplage plus ou moins facile aux différents partenaires intra-cellulaires, et donc l'activation préférentielle de certaines voies de signalisation. Cette sélectivité de voie de signalisation est encore mal comprise mais capitale pour le développement de composés aux effets plus contrôlés (ligands « biaisés »). Il a été suggéré que cette sélectivité pourrait être liée à une position intermédiaire du ligand (ou « vestibulaire »), avant d'adopter son orientation finale dans la poche orthostérique [Dror, Pan, et al. 2011; Kofuku et al. 2012].

En plus de ce site de liaison « orthostérique » bien identifié, il a été suggéré que l'activité des RCPGs pourrait être régulée par la fixation de ligands ou partenaires dans d'autres sites dits « allostériques ». On parle de modulateur allostérique positif (MAP) si la liaison de ce ligand augmente l'activité du récepteur ou de modulateur allostérique négatif (MAN) s'il la diminue (**Figure 11**). De ce point de vue, et comme déjà suggéré plus haut, les ions, la membrane et

autres partenaires des RCPGs peuvent être tous vus comme des modulateurs allostériques dès lors qu'ils ne se lient pas dans la poche orthostérique.

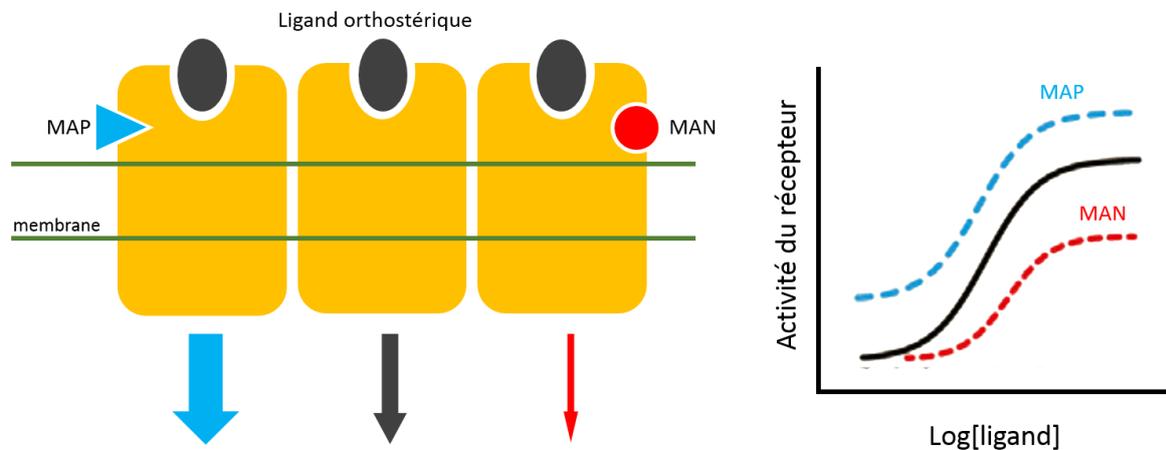


Figure 11 : Représentation de l'effet modulateur positif (MAP) et négatif (MAN) d'un ligand allostérique sur l'activité d'un RCPG.

### 1.1.4.b Partenaires intra-cellulaires

Les RCPGs possèdent de nombreux partenaires possibles du côté intra-cellulaire, dont évidemment les protéines-G comme déjà discuté plus haut. Les protéines-G sont ubiquitaires et jouent un rôle très important dans un grand nombre de processus biologiques, car étant les premières actrices de l'activation d'une pléiade de voies de signalisation. Il existe différents types de protéines-G hétéro-trimériques ayant des fonctions cellulaires spécifiques. C'est principalement la sous-unité  $\alpha$  qui va définir les propriétés de la protéine-G. Il existe quatre familles de sous-unités  $\alpha$  :  $\alpha_s$ ,  $\alpha_i$ ,  $\alpha_q$  et  $\alpha_{12}$ . Les sous-unités  $\beta$  et  $\gamma$  quant à elles, forment un complexe très stable généralement impossible à dénaturer [Handbook of Cell Signaling 2003]. Enfin, il est à noter que les différentes sous-unités de la protéine-G présentent très peu de diversité conformationnelle, notamment les sous-unités  $\beta$  et  $\gamma$ .

Les interactions à l'échelle moléculaire des protéines-G avec les RCPGs, ainsi que leur mécanisme d'activation seront décrits plus en détails dans la partie suivante (voir partie 1.1.5). Il a été montré récemment qu'en se couplant à leurs récepteurs, les protéines-G sont directement capables d'affecter leur dynamique intrinsèque, que ce soit avec ou sans ligand [Renault et al. 2018]. Récemment, notre équipe a montré que dans le cas du récepteur de la Ghréline, il existait un complexe récepteur:protéine-G inactif [Damian et al. 2015].

## 1 Introduction

Les protéines-G, après activation, interagissent avec différentes protéines appelées effecteurs. Ces effecteurs provoquent alors soit une cascade d'interactions amenant à la réponse biologique, on parle alors de seconds messagers, soit être directement impliqués dans cette réponse biologique, comme c'est le cas par exemple des canaux ioniques. Lors de l'activation des protéines-G, leurs sous-unités  $\alpha$  et  $\beta\gamma$  ont alors différents partenaires possibles qui dépendent directement de leur nature. Ces effecteurs sont présentés dans le **Tableau 2**.

$G_s\alpha$	adenylate cyclases $\uparrow$ ; canaux potassiques $\uparrow$ Src tyrosine kinases (c-Src, Hck) $\uparrow$ ; tubuline GTPase $\uparrow$
$G_i\alpha$	adenylate cyclase $\downarrow$ ; Rap1GAP $\uparrow$ ; cGMP-PDE $\uparrow$ activation des ERK/MAPkinase par voie Rap1GAPII $\uparrow$ canaux calciques $\downarrow$ ; activation de Cdc42 par GRIN1 $\uparrow$ canaux potassique $\uparrow$ ; tubuline GTPase $\uparrow$ ; tyrosine kinases Src (c-Src, Hck) $\uparrow$
$G_q\alpha$	Phospholipase C $\beta$ $\uparrow$ ; p63-RhoGEF $\uparrow$ ; tyrosine kinase de Bruton $\uparrow$ canaux potassiques $\uparrow$
$G_{12/13}\alpha$	Phospholipase D $\uparrow$ ; Phospholipase C $\epsilon$ $\uparrow$ ; NHE-1 $\uparrow$ ; iNOS $\uparrow$ ; HSP90 $\uparrow$ adhésion cellulaire par voie E-cadherine $\uparrow$ ; p115RhoGEF $\uparrow$ ; PDZ-RhoGEF $\uparrow$ Radixin $\uparrow$ ; Protein phosphatase 5 (PP5) $\uparrow$ ; activation de PKA par voie AKAP110 $\uparrow$
$G\beta/\gamma$	Phospholipase C $\beta$ s $\uparrow$ ; adenylylase cyclase I $\downarrow$ ; adenylylase cyclases II, IV, VII $\uparrow$ PI-3 kinases $\uparrow$ ; canaux potassiques (GIRK1,2,4) $\uparrow$ ; P-Rex1 $\uparrow$ ; JNK $\uparrow$ canaux calciques (type- N, P/Q, R) $\downarrow$ ; tubuline GTPase $\uparrow$ kinases Src $\uparrow$ ; recrutement de GRK2 à la membrane $\uparrow$ ; protéine kinase D $\uparrow$ tyrosine kinase de Bruton $\uparrow$ ; p114-RhoGEF $\uparrow$

Tableau 2 : Liste des effecteurs des quatre familles de la sous-unité  $\alpha$  et du complexe  $\beta\gamma$  ; d'après [Milligan et Kostenis 2006].

Après leur activation, les RCPGs sont désensibilisés *via* différents partenaires intracellulaires. Les kinases-RCPG (GRK) (**Figure 12**) phosphorylent ainsi les sérines et thréonines présentes sur la face intra-cellulaire du récepteur, notamment son segment C-terminal, afin de provoquer le départ de la protéine-G [Bouvier et al. 1988; Nakata et al. 1994]. Les kinases sont des enzymes de type transférase catalysant les réactions de phosphorylation par ajout d'un ion phosphate de l'Adénosine triphosphate (ATP) vers certains acides aminés. Chez l'homme, les GRKs sont au nombre de sept.



Figure 12 : Structure d'une Kinase-RCPG, ici la GRK5 (code PDB : 4TNB) [Komolov, Bhardwaj, et Benovic 2015].

Toutefois cette phosphorylation n'est pas suffisante pour produire une désensibilisation complète du récepteur, ce qui nécessite l'implication d'autres partenaires comme l'Arrestine [Pitcher et al. 1992]. Les arrestines (**Figure 13**) sont une famille de protéines régulatrices ayant une haute affinité pour le récepteur phosphorylé [Lohse et al. 1990]. Ces protéines sont capables de découpler la protéine-G du récepteur par compétition [Krupnick et Benovic 1998], et jouent également un rôle dans l'internalisation du récepteur et le trafic intra-cellulaire des RCPGs. Elles peuvent également jouer un rôle varié dans la signalisation cellulaire en se complexant à d'autres protéines [E. V. Gurevich et Gurevich 2006]. Ces arrestines sont au nombre de quatre chez les vertébrés. Deux des arrestines sont impliquées dans la vision : l'Arrestine-1 dans les bâtonnets et les cônes, et l'Arrestine-4 présente uniquement dans les cônes. Les Arrestine-2 et Arrestine-3, aussi appelées  $\beta$ -Arrestine et  $\beta$ -Arrestine-2, sont ubiquitaires, mais principalement exprimées dans le cerveau.

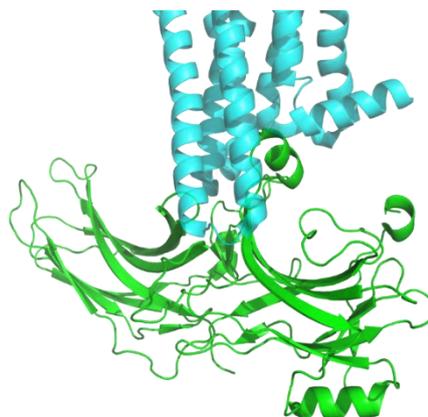


Figure 13 : Structure de la  $\beta$ -Arrestine liée à un RCPG (code PDB : 5W0P) [Kang et al. 2015].

### 1.1.4.c Dimérisation / oligomérisation

La plupart des études visant à comprendre la signalisation *via* les RCPGs se basent sur l'hypothèse d'un fonctionnement en monomère. Or, de plus en plus de travaux montrent la capacité des RCPGs à se dimériser en formant des homo-dimères, des hétéro-dimères ou encore des assemblages (oligomères) encore plus importants [A. A. Kaczor et Selent 2011]. Ainsi, l'activité d'un RCPG n'est pas liée uniquement aux changements conformationnels provoqués par la liaison de son ligand, mais aussi possiblement modulée de manière allostérique par ses interactions avec d'autres récepteurs. Il a été montré que cette dimérisation permet de moduler positivement ou négativement la liaison de ligands aux différents protomères impliqués [Ferré et al. 2014; Haack et McCarty 2011; Sleno et Hébert 2018].

De nombreuses méthodes expérimentales ont été développées pour identifier le degré d'oligomérisation des RCPGs qu'il s'agisse de méthodes biochimiques, biophysiques ou physiologiques [Guo et al. 2017]. On note d'ailleurs plusieurs structures cristallographiques d'homo-dimères de RCPGs pour les récepteurs de la Rhodopsine (codes PDB : 2Z73 et 3CAP) [Murakami et Kouyama 2008; Park et al. 2008], le récepteur CXCR4 (codes PDB : 3OE0 et 3ODU) [B. Wu et al. 2010], le récepteur  $\mu$ -Opioides (code PDB : 4DKL) [Manglik et al. 2012] ou encore le récepteur  $\beta$ 1AR (code PDB : 4GPO) [J. Huang et al. 2013]. L'étude de ces complexes présente un fort intérêt pharmacologique car la complexation des RCPGs entre eux crée de « nouvelles cibles » potentielles. Avec ce gain d'intérêt, on a vu de nombreuses études de dynamique moléculaire et autres méthodes de modélisation utilisées pour prédire ces interactions. Celles-ci seront discutées dans un autre chapitre (voir partie 1.3.5).

Malgré tout, ces mécanismes restent encore mal compris. Entre autres interrogations, la question du mécanisme de couplage à leurs partenaires des RCPGs impliqués dans de tels assemblages reste encore à élucider. Dans un article récent publié au sein de notre équipe (voir partie 3.4.1) nous avons proposé et validé l'existence d'un modèle (**Figure 14**) dans lequel les deux récepteurs de la Ghreline et de la Dopamine forment un hétéro-dimère tout en étant couplés aux protéines  $G_q$  et  $G_i$ , respectivement, qui se retrouvent parfaitement imbriquées [Damian et al. 2018]. Dans cet assemblage, chaque récepteur forme également un homo-dimère, aboutissant ainsi à un modèle comprenant quatre récepteurs et deux protéines-G. On retrouve un assemblage assez similaire avec les récepteurs  $GABA_B$  où deux hétéro-dimères  $GABA_{B1}$ : $GABA_{B2}$  sont associés à l'inverse *via* une interface homo-dimérique des deux  $GABA_{B1}$  [Comps-Agrar et al. 2011].

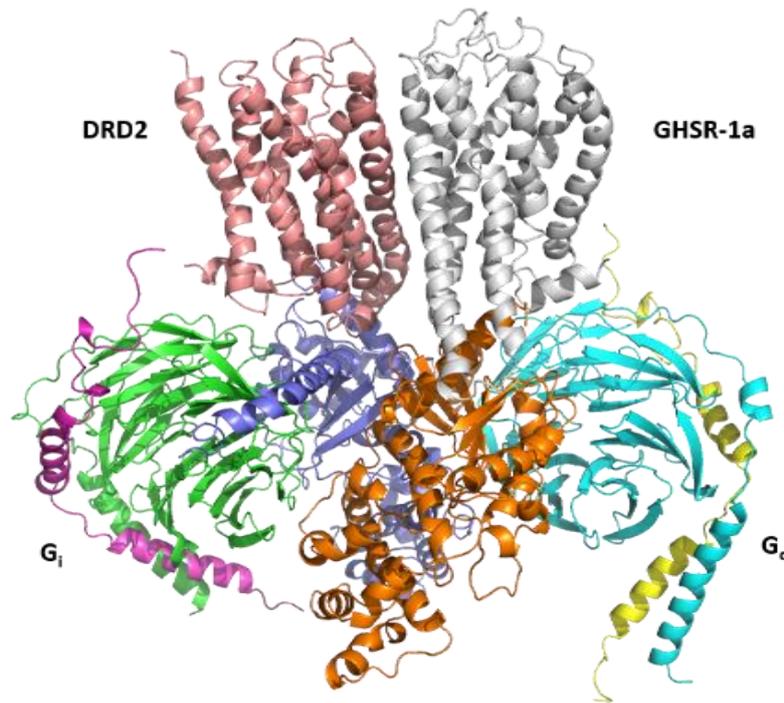


Figure 14 : Modèle de dimère impliquant les récepteurs de la Ghréline et de la Dopamine et les protéines  $G_i$  et  $G_q$ , ce dimère sera décrit plus précisément dans la partie 3.3.1.

L'oligomérisation des RCPGs permettrait peut-être d'expliquer le rôle de certains récepteurs orphelins. En effet, au cours du temps certains récepteurs se sont vu perdre leur titre d'orphelin par identification de leur ligand endogène avec en exemples les récepteurs de la Ghréline, de la Métastatine ou de l'Oréxine, [Civelli et al. 2006]. Toutefois, certains récepteurs pourraient demeurer orphelins, ne possédant pas du tout de ligand endogène, mais seraient utilisés comme modulateur dans des macro-complexes avec des récepteurs non-orphelins. On peut citer comme exemple les couples TAS1R1/ TAS1R3 et TAS1R2/ TAS1R3 [G. Nelson et al. 2001; Greg Nelson et al. 2002].

## 1.1.5 Activation des RCPGs

### 1.1.5.a Formation de l'hétéro-trimère

On peut considérer que le cycle d'activation d'un RCPG commence bien avant l'arrivée du ligand sur son récepteur avec la formation de la protéine-G : complexation d'une sous-unité  $\alpha$  inactive avec le complexe  $\beta\gamma$ . L'ensemble est maintenu par une zone de contact au cœur très hydrophobe et par le contact entre l'hélice N-terminale de  $\alpha$  et des feuillets de  $\beta$ . On a découvert à ce jour 21 sous-unités  $\alpha$  pour 6 sous-unités  $\beta$  et 12 sous-unités  $\gamma$ , représentant alors une

quantité conséquente de complexes possibles. Il n'existe toutefois qu'un nombre restreint de ces combinaisons. On retrouve très peu de diversité dans le cœur hydrophobe formé par  $\alpha$  et  $\beta$  dans ces différentes sous-unités, suggérant que la spécificité de couplage se ferait sur la seconde interface. De plus, les structures disponibles de  $\beta\gamma$  montrent une forte conservation structurale, montrant que cette partie est très peu rigide, comme discuté dans les travaux antérieurs de l'équipe [Louet 2012; Wall et al. 1995; Lambright et al. 1996; Nishimura et al. 2010]. On note tout de même une légère flexibilité au niveau des hélices N-terminales et d'une boucle de  $\beta$  ; cette flexibilité a été montrée comme importante lors de l'activation du complexe récepteur:protéine-G [Preininger, Meiler, et Hamm 2013; Damian et al. 2015].

L'hétéro-trimère de protéine-G est formé dans l'appareil de Golgi et adressé à la membrane, notamment grâce aux résidus d'ancrage issus de modifications post-traductionnelles qui lui permet de rester ancré à la membrane et de faciliter sa diffusion. Il existe aussi des pGs dont l'association a lieu directement au contact de la membrane. Toutefois, des études ont montré qu'un simple fragment de la sous-unité  $\alpha$ , notamment l'hélice C-terminale, était suffisant pour se lier sélectivement aux RCPGs et modifier leur dynamique intrinsèque [Carpenter et Tate 2016; Semack et al. 2016; Csépanyi-Kömi, Lévy, et Ligeti 2012]. Le complexe  $\beta\gamma$  viserait donc principalement à séquestrer la sous-unité  $\alpha$  pour l'amener à la membrane grâce notamment à ses modifications lipidiques. Après activation, le complexe  $\beta\gamma$  interagit également avec un certain nombre d'effecteurs, comme déjà décrit plus haut dans ce chapitre.

### 1.1.5.b Complexation avec un RCPG

La spécificité d'interaction entre RCPGs et protéines-G permet d'obtenir une grande variété de réponses biologiques. Comme montré par les différentes structures désormais disponibles, la pG [Rasmussen et al. 2011; García-Nafría et al. 2018; Koehl et al. 2018] s'insère *via* son hélice C-terminale de la sous-unité  $\alpha$  sur la face intra-cellulaire du récepteur en interagissant avec une crevasse formée grâce à l'écartement des TM5 et TM6. Comme évoqué précédemment, de nombreuses études ont montré que cette hélice C-terminale était suffisante pour se lier au récepteur, sans autre contact nécessaire, et pouvait mimer au niveau structural l'effet de la protéine-G complète [Scheerer et al. 2008]. Il est à noter une seconde zone d'interaction entre les TM6 et TM7 du récepteur et l'hélice  $\alpha 4$  et le brin  $\beta 6$  de la sous-unité  $\alpha$ . Les études récentes de microscopie ont permis de caractériser différents couples récepteur:protéine-G, suggérant une différence d'orientation significative de l'hélice C-

terminale selon la nature de la sous-unité  $\alpha$  (**Figure 15**) [Draper-Joyce et al. 2018; García-Nafría et al. 2018]. Ces différences pourraient suggérer une sélection variable de conformations du récepteur en fonction de la protéine-G qui lui est liée. L'alignement de ces différentes structures permet aussi pour la première fois de deviner les mouvements intrinsèques du complexe récepteur:protéine-G, jusqu'alors uniquement suggérés par le calcul [Louet et al. 2011, 2012].

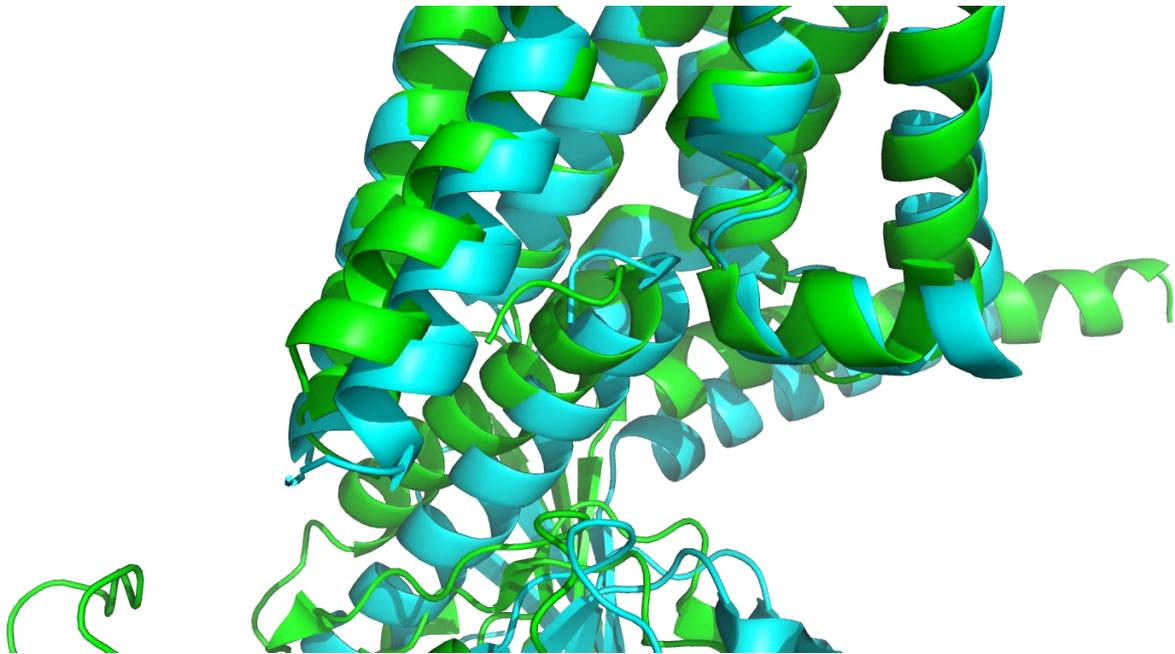


Figure 15 : Comparaison de l'orientation des deux protéines-G  $G_s$  et  $G_i$ , dans deux structures résolues par Cryo-EM ; le complexe A1R: $G_i$  est représenté en bleu (code PDB : 6D9H) [Draper-Joyce et al. 2018] ; le complexe A2AR: $G_s$  est représenté en vert (code PDB : 6GDG) [García-Nafría et al. 2018].

S'il était autrefois admis que le récepteur devait être activé par un ligand pour pouvoir adopter la conformation nécessaire à son couplage avec la pG, de nombreuses données de la littérature suggèrent l'existence d'un pré-couplage récepteur:protéine-G. Ce pré-couplage permettrait notamment d'expliquer un temps de réponse très rapide après l'arrivée du ligand du côté extra-cellulaire. Des effecteurs pourraient également faire partie de ce pré-couplage, préparant ainsi une voie d'activation spécifique [Qin et al. 2011].

Comme déjà discuté plus haut, de plus en plus d'études montrent l'effet de la protéine-G sur la sélection de conformations du récepteur [Louet et al. 2013; Flock et al. 2015; Saleh, Saladino, et al. 2017; DeVree et al. 2016]. Un travail réalisé cette année au laboratoire et soumis pour publication [Renault et al. 2018] montre notamment l'effet allostérique de la mini- $G_s$ , simplification de l'hétéro-trimère en ne conservant de la sous-unité  $\alpha$  que le domaine GTPase [Nehmé et al. 2017], sur la poche de liaison du ligand du récepteur de l'Adénosine.

### 1.1.5.c Activation du couple RCPG:pG

L'activation du récepteur par un ligand se manifeste sous la forme d'un changement conformationnel, ou plutôt comme une sélection spécifique de conformations adoptées par le récepteur ou le couple récepteur:protéine-G. Les boucles extra-cellulaires participent directement à la liaison du ligand mais aussi à son maintien en position, une fois dans la poche de liaison orthostérique [Wheatley et al. 2012]. Des mutations de ces boucles entraînent une diminution de la liaison du ligand au récepteur, montrant leur rôle essentiel dans l'affinité et la sélectivité de liaison [Arora, Sakai, et Catt 1995; Noda et al. 1994; Nguyen et al. 2016]. Même si de nombreuses structures existent désormais montrant les RCPGs couplés à différents types de ligands, ou à différents partenaires (notamment intra-cellulaires), la cristallographie semble plus ou moins toujours capturer un même état conformationnel de ces récepteurs. En effet, les structures cristallographiques disponibles montrent peu de réarrangement conformationnel des RCPGs en fonction de la nature du ligand co-cristallisé (**Figure 16**). Trop souvent dans la littérature, un lien a été fait entre structure active et structure couplée. Dans un travail récent effectué au laboratoire, l'existence d'une conformation couplée inactive a été montrée pour le récepteur de la Ghréline [Damian et al. 2015]. Pourtant, on sait que cette liaison engendre des changements de conformation importants, comme montré par les études récentes de RMN [Casiraghi et al. 2016], et impliquant possiblement des mouvements de large amplitude [Louet et al. 2013]. Ces réarrangements locaux et globaux se répercutent sur la face intra-cellulaire et donc sur la pG entraînant son activation. Même sans ligand, il y a un équilibre entre la (les) conformation(s) active(s) et inactive(s) du récepteur, expliquant l'activité basale de certains RCPGs. Là encore, la RMN permet d'observer de tels équilibres conformationnels [Casiraghi et al. 2016]. On suppose que la liaison du ligand dans la poche orthostérique va concourir au déplacement de cet équilibre vers une direction ou une autre selon sa nature: agoniste ou agoniste inverse. En dépit de la grande conservation de structures, notamment les sept TMs, les RCPGs présentent une très grande diversité dans leur boucles extra-cellulaires, leur segment N-terminal et leur poche de liaison.

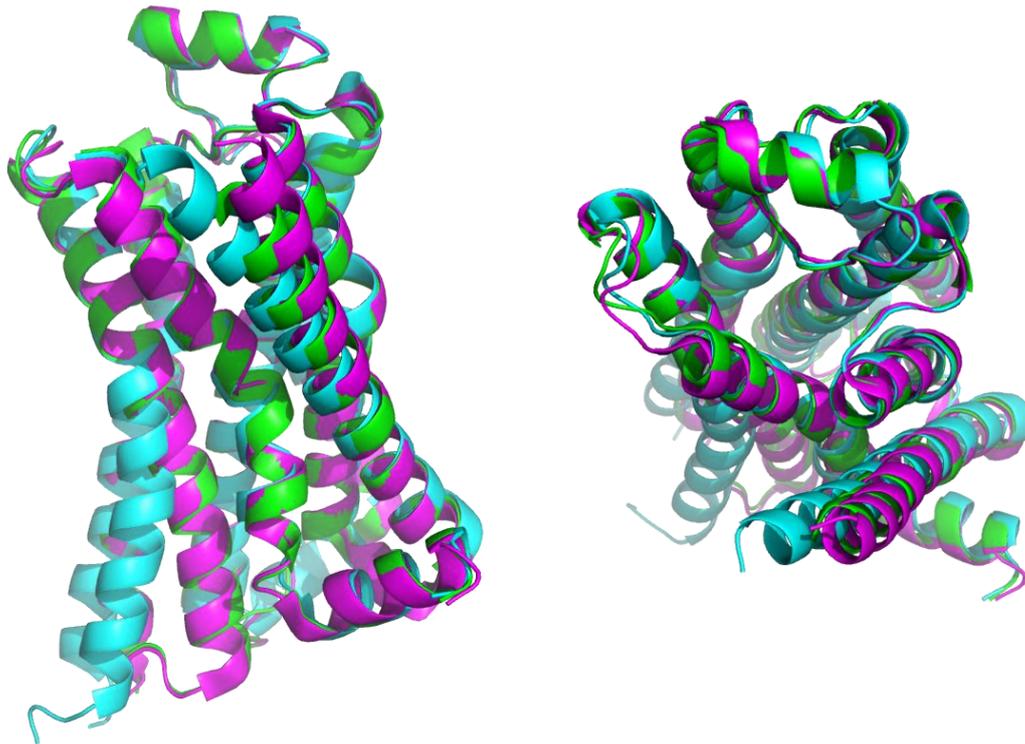


Figure 16 : Comparaison de structures cristallographiques de B2AR complexé à un ligand agoniste en bleu (code PDB : 4LDE) [Ring et al. 2013], à un ligand antagoniste en magenta (code PDB : 3NYA) [Wacker et al. 2010], et à un ligand agoniste inverse en vert (code PDB : 5D5A) [C. Y. Huang et al. 2016].

#### 1.1.5.d Relargage du GDP et dissociation

L'activation de la pG entraîne la sortie du GDP présent dans la sous-unité  $\alpha$ . Cette sortie implique des mouvements de larges amplitudes, le nucléotide étant situé au cœur des deux sous-domaines de la sous-unité  $\alpha$ . Plusieurs modèles d'ouverture ont alors été proposés à ce jour. Dans le premier, appelé « bras de levier » (**Figure 17A**) [Iiri, Farfel, et Bourne 1998], les sous-unités  $\beta\gamma$  s'écartent afin de laisser sortir le GDP. Dans le second, appelé « levier de vitesse » (**Figure 17B**) [Jacqueline Cherfils et Chabre 2003], les sous-unités  $\beta\gamma$  s'appuient au contraire sur la sous-unité  $\alpha$  pour créer la voie de sortie du nucléotide. Un autre modèle plus récent montre un rôle plus important de la sous-unité  $\alpha$ , et repose sur l'hélice C-terminale en contact à la fois avec le récepteur et le GDP, et dont le mouvement entraîne la sortie du nucléotide. La seule structure cristallographique présentant la sous-unité  $\alpha$  sans le GDP montre bien l'écartement entre les deux domaines de celle-ci [Rasmussen et al. 2011].

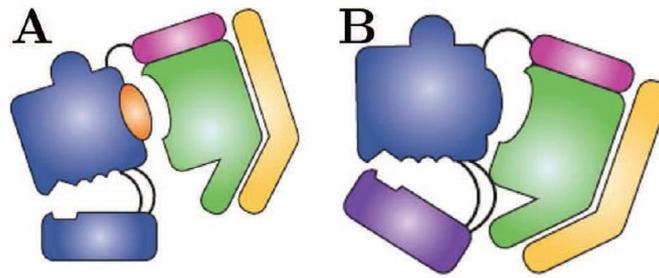


Figure 17 : Schématisation des modèles de sorties du GDP, d'après [Oldham et Hamm 2008] ; A) Modèle « bras de levier » ; B) Modèle « levier de vitesse ».

Au laboratoire des études de modélisation ont permis de simuler les mouvements de sortie du GDP. Dans la première étude, qui combine dynamique moléculaire et analyse de modes normaux, un des modes observés montre une large ouverture entre les domaines Ras-like et hélicoïdal [Louet et al. 2011]. Dans la seconde étude, qui combine dynamique moléculaire dirigée et reconstruction de profils d'énergie libre par échantillonnage en parapluie (« umbrella sampling »), il est suggéré que le relargage du GDP se fait plus favorablement par son côté phosphate, toujours avec l'écartement des deux domaines comme précédemment discuté [Louet, Martinez, et Floquet 2012]. Il est à noter qu'une sortie du côté base ne serait effectivement pas possible dans les oligomères mis en évidence récemment au laboratoire [Damian et al. 2018].

Après la sortie du GDP, celui-ci est remplacé par une molécule de GTP. Là aussi, le mécanisme de liaison du GTP reste encore mal connu, mais il est raisonnable de suggérer un chemin inverse à la sortie du GDP, les deux nucléotides ne se différenciant que d'un groupement phosphate. Les deux molécules ayant la même affinité pour la sous-unité  $\alpha$  vide, l'activation des pGs repose donc sur une plus grande concentration en GTP qu'en GDP dans la cellule [Oldham et Hamm 2008; Zwingmann, Leibfritz, et Hazell 2003].

L'entrée du GTP a pour conséquence l'écartement de la sous-unité  $\alpha$  du dimère  $\beta\gamma$ . Dans une autre étude menée au laboratoire, la dynamique moléculaire dirigée (« targeted MD » ou TMD) a été utilisée pour forcer la sous-unité  $\alpha$  à passer dans sa forme active liée au GTP. Les résultats de ces calculs montrent de larges mouvements domaine-domaine avec une dissociation presque complète du dimère  $\beta\gamma$  [Louet et al. 2012]. Des données obtenues par FRET ou BRET montrent également que la dissociation de la protéine-G n'est pas complète car un signal est toujours perçu après activation de la pG [T.-M. Yi, Kitano, et Simon 2003; Azpiazu et Gautam 2004; Galés et al. 2005]. La vitesse de dissociation est également dépendante de la nature du

trimère [Digby, Sethi, et Lambert 2008]. Là aussi, une dissociation incomplète pourrait expliquer le temps de réponse rapide de ces systèmes membranaires, comme déjà discuté plus haut pour le pré-couplage.

### 1.1.5.e Après l'activation de la pG

Une fois activées, les Protéines-G interagissent avec des protéines dites effectrices. Les différentes sous-unités  $\alpha$  et  $\beta\gamma$  possèdent une grande variété de partenaires qui engendrent la réponse cellulaire en induisant une cascade d'interactions, notamment en produisant des molécules intermédiaires ou seconds messagers.

Afin de compléter le cycle d'activation des RCPGs, le retour à l'état inactif nécessite de reformer le complexe lié au GDP. Pour cela, la sous-unité  $\alpha$  comporte une activité GTPasique intrinsèque. Deux mécanismes d'hydrolyse du GTP en GDP ont été proposés sur la base de calculs semi-empiriques ou quantiques. Le premier repose sur un modèle associatif où l'accumulation de charges négatives sur un oxygène non liant du groupement  $\gamma$  du GTP provoquerait la rupture entre le phosphate  $\beta$  et le phosphate  $\gamma$  [Coleman et al. 1994; Kleuss et al. 1994]. Dans ce mécanisme, seraient impliquées une arginine, une thréonine et une glutamine de la sous-unité  $\alpha$ , qui ensemble permettraient de stabiliser l'état de transition. Le second modèle dit dissociatif voit les charges négatives du phosphate  $\beta$  s'accumuler sur l'oxygène le liant au phosphate  $\gamma$  [Maegley, Admiraal, et Herschlag 1996]. Un ion  $Mg^{2+}$  est également impliqué dans ce mécanisme d'hydrolyse en participant à l'élongation de la liaison entre les deux phosphates grâce à un stockage des électrons qu'il restituerait ensuite aux deux groupements partants résultants [Rudack et al. 2012].

Toutefois, cette activité GTPasique est très faible. Afin de parvenir à la rapide désactivation des pGs, d'autres protéines sont impliquées, les protéines régulatrices du signal des protéines-G (RGS) qui augmentent significativement cette hydrolyse [Lambert et al. 2010]. Les RGS se fixent sur la protéine-G, à proximité de la poche de liaison du nucléotide et à l'interface entre les domaines Ras-like et hélicoïdal [Soundararajan et al. 2008]. Il a été proposé que cette interaction permettrait la réorientation d'une glutamine afin qu'elle interagisse de façon plus stable avec le phosphate en  $\gamma$  du GTP, stabilisant ainsi la sous-unité  $\alpha$  dans un état de transition et permettant à une thréonine à proximité de stabiliser la position du cation  $Mg^{2+}$  [Kimple et al. 2009].

### 1.2 Les champs de forces

Un champ de forces est un ensemble de paramètres permettant d'évaluer l'énergie potentielle d'une molécule à partir de ses coordonnées cartésiennes. L'objectif des différents algorithmes utilisés en modélisation moléculaire est d'échantillonner les conformations d'un même système moléculaire afin d'identifier celles présentant une énergie minimum. Dans ce travail nous avons utilisé principalement la dynamique moléculaire. La dynamique moléculaire permet de simuler le comportement d'un système moléculaire au cours du temps. En dynamique moléculaire, la fonction d'énergie potentielle est utilisée (dérivée) pour calculer, à chaque temps  $t$ , les forces appliquées à chacune des particules du système. Ces forces permettent ensuite, *via* les équations du mouvement de Newton, de prédire les nouvelles positions de ces particules à un temps  $t + \Delta t$ . Plus le niveau de détail utilisé pour décrire le système sera important et plus le pas d'intégration ( $\Delta t$ ) sera petit. Par exemple, l'utilisation d'un champ de forces dans lequel tous les atomes y compris les hydrogènes sont représentés (champ de forces « tout-atomes »), nécessitera d'utiliser un pas d'intégration de 1 femtoseconde (fs) permettant de décrire la fréquence d'oscillation des liaisons impliquant ces hydrogènes. Les champs de forces tout-atomes sont de loin les plus utilisés, car ils permettent généralement de simuler une variété de molécules assez larges. De nombreux champs de forces ont été optimisés pour simuler le comportement dynamique des peptides, protéines, acides nucléiques, lipides. Il existe aussi des champs de forces spécialisés dans la représentation de certains types de molécules.

Les champs de forces classiquement utilisés pour les protéines ne permettent généralement pas de créer ou briser une liaison chimique et ne permettent pas non plus de prédire des modifications de distributions de charges (polarisation). La taille des systèmes impliqués dans les processus biologiques, en particulier de ceux étudiés dans ce manuscrit (récepteurs dans la membrane), limite énormément leur échantillonnage conformationnel par dynamique moléculaire car requiert alors des moyens de calcul très importants. Par exemple, la production d'une microseconde ( $\mu s$ ) de dynamique moléculaire pour un récepteur dans la membrane nécessite environ une semaine à 10 jours sur une station de travail équipée de GPUs dès lors que l'on utilise un champ de forces tout-atomes. Ces contraintes ont encouragé les chercheurs à développer des champs de forces dans lesquels d'autres simplifications ont été introduites, notamment au niveau structural, comme dans les champs de forces gros-grains, largement utilisés au cours de ce travail, rendant presque facile l'étude de larges systèmes moléculaires à l'échelle de la  $\mu s$ .

### 1.2.1 Champs de forces classiques

C'est en 1930 que sont posées les bases des premiers champs de forces moléculaires avec un modèle de type sphère-élastique utilisé pour décrire la longueur et l'angle de liaison, mais également les interactions des atomes non liés en utilisant les équations de Van der Waals [Andrews 1930]. Mais ce n'est qu'en 1968 que sera décrit le premier champ de forces généralisé [Lifson et Warshel 1968]. Les premiers champs de forces sont alors limités à quelques types d'atomes et sont surtout utilisés pour simuler des systèmes moléculaires organiques. Aujourd'hui il existe de nombreux champs de forces permettant de simuler une grande variété de biomolécules.

Le champ de force « Assisted Model Building with Energy Refinement » (AMBER) est l'un des tous premiers développés pour l'étude de macromolécules biologiques tel que l'ADN ou les protéines [Weiner et Kollman 1981]. Les paramètres du champ de forces sont issus de calculs de mécanique quantique, ou encore de données expérimentales comme les diffusions micro-ondes et de neutrons, ou encore la RMN.

Le champ de forces « Optimized Potentials for Liquid Simulations » (OPLS) inclut des modèles d'atomes unis, où les hydrogènes aliphatiques ne sont pas explicitement représentés, et des modèles tout atomes, et est utilisé pour les peptides et les molécules organiques [Jorgensen et Tirado-Rives 1988]. Les paramètres d'élongation et flexibilité des liaisons sont des modifications d'AMBER. Il est plus spécifiquement dédié au calcul de l'énergie de conformation de molécules en phase gazeuse, au calcul de l'énergie libre de solvation de molécules en phase liquide et au calcul d'autres propriétés thermodynamiques.

Le champ de force « Chemistry at HARvard Molecular Mechanics » (CHARMM) a été développé pour simuler le comportement de molécules biologiques telles que les peptides, protéines, acides nucléiques, mais aussi les acides gras et les sucres [Bernard R. Brooks et al. 1983]. En plus de données expérimentales, la détermination des paramètres a été réalisée par de nombreux calculs quantiques. CHARMM est aussi le nom d'un programme permettant de réaliser des calculs de minimisation d'énergie, calculs de modes normaux, simulations de dynamique moléculaire, Monte-Carlo, etc...

Le champ de force GROMOS porte le nom là aussi d'une suite de logiciels/programmes de simulation moléculaire « GRONingen MOlecular Simulation » (GROMOS) avec lequel il est fourni [van Gunsteren et al. 1996]. Il permet de simuler des systèmes biomoléculaires comportant des protéines, des nucléotides, des sucres ou des polymères, etc... Le champ de

forces est souvent confondu avec GROMACS (« GRONingen Machine for Chemical Simulations ») qui est également un ensemble de logiciels de simulation [Oostenbrink et al. 2004].

### 1.2.2 Champs de forces polarisables

Les champs de forces classiques ne prennent pas en compte la polarisation électrostatique et les transferts de charge inter ou intra-moléculaires. Or on sait que dans un milieu composé d'eau ou en présence d'ions qui possèdent une charge forte, il existe des phénomènes importants de polarisation pouvant impacter l'énergie et le comportement structural et dynamique des protéines. Pour répondre à cette problématique les champs de forces polarisables ont été développés. Plusieurs types de méthodes existent : le modèle « dipôle induit » où comme son nom l'indique un dipôle est induit au centre de tous les contributeurs en réponse au champ électrique total ou encore le modèle de charge fluctuante où la charge atomique varie en réponse à l'environnement selon le principe d'équilibre de l'électronégativité. Des versions polarisables des champs de forces classiques existent tel que CHARMM [Sandeep Patel et III 2006] ou AMBER [Cieplak, W. Caldwell, et A. Kollman 2001], et d'autres ont été directement développés comme champs de forces polarisables.

Par exemple, le champ de force Drude a été développé dans CHARMM [B. R. Brooks et al. 2009] avant d'être intégré à NAMD [Jiang et al. 2011] et d'autres logiciels de simulation. La première étape du développement de ces champs de forces a été la paramétrisation de l'eau, puis d'autres petites molécules, acides nucléiques ou lipides, permettant ainsi la construction de bicouches lipidiques.

Le champ de forces « Atomic Multipole Optimized Energetics for Biomolecular Applications/simulation » (AMOEBA) est un champ de forces initialement développé pour l'eau [Dudek et Ponder 1995], mais qui a fini par intégrer des paramètres pour la simulation de protéines [Y. Shi et al. 2013]. Le champ de forces est capable de prendre en compte des molécules organiques comme les alcanes, les alcools, les amines, les amides ou encore des molécules aromatiques.

### 1.2.3 Champs de forces réactifs

Dans la plupart des champs de forces, il est impossible de créer ou de briser une liaison chimique, les éléments liés et non-liés ne variant pas au cours de la simulation. Ainsi pour pouvoir simuler des réactions chimiques, des champs de forces dits réactifs ont vu le jour. Il existe deux types de champs de forces réactifs. Les premiers, définis comme « méthodes de distance de réaction avec cut-off », se basent sur un critère de distance : si les deux particules compatibles sont assez proches, une liaison se forme. Ces champs de forces sont principalement utilisés pour étudier les phénomènes de polymérisation [Akkermans, Toxvaerd, et Briels 1998]. Les seconds, appelés champs de forces réactifs empiriques, ont pour objectifs de développer des champs de forces prédictifs permettant de reproduire les cinétiques de liaisons.

Le champ de forces réactif (ReaxFF) permet de décrire la formation des liaisons implicites sans nécessiter de calculs de mécanique quantique excessifs [van Duin et al. 2001], permettant ainsi, au dépend de la précision, d'atteindre des échelles de temps de simulation inaccessibles en mécanique quantique. Ce champ de forces a été notamment utilisé pour simuler des réactions chimiques à des interfaces hétérogènes, comme notamment la catalyse hétérogène et le dépôt de couches minces atomiques (« Atomic Layer Déposition » ou ALD) [Senftle et al. 2016].

### 1.2.4 Champs de forces gros-grains

La modélisation de larges systèmes et de processus biologiques apparaissant à l'échelle de la micro voire de la milliseconde est très coûteuse en ressources de calculs. Les champs de forces Gros-Grains (CG) permettent de simplifier le système en représentant un groupe d'atomes par une seule et unique sphère (particule) et donc de diminuer drastiquement le nombre d'interactions liantes et surtout non-liantes à calculer. Par ce biais et cette perte de détails au niveau structural, les simulations notamment de dynamique moléculaire vont pouvoir être beaucoup plus rapides. Il est donc important de comprendre la nature de ce que l'on cherche à étudier pour choisir le champ de force adéquat. Le choix du champ de forces doit être fait en trouvant un équilibre entre précision et efficacité du calcul, en conservant les propriétés principales du système étudié.

Les premiers modèles Gros-Grains avaient pour objectifs l'étude des repliements des protéines. En 1975, l'un des modèles CG précurseur était publié et représentait chaque acide aminé par deux sphères : une pour le squelette et une pour la chaîne latérale [Levitt et Warshel 1975]. Le degré de liberté résiduel était alors la torsion entre deux particules liées. Depuis, de

nombreux modèles ont été développés avec une grande diversité dans les applications et les schémas de conversion utilisés. Le champ de forces OPEP par exemple, utilise une représentation presque atomistique pour le squelette mais une unique particule pour la chaîne latérale [Maupetit, Tuffery, et Derreumaux 2007]. On retrouve également des champs de forces développés pour le docking de protéines où les résidus sont représentés par trois particules et les protéines maintenues rigides [Zacharias 2003]. Il est également possible de réaliser des simulations avec plusieurs niveaux de représentations. Par exemple, il a été possible de simuler un canal ionique dans une représentation tout-atomes, placé dans une membrane représentée dans un modèle gros-grains [Q. Shi, Izvekov, et Voth 2006].

Le champ de force MARTINI a été développé au départ pour la simulation des lipides [Marrink et al. 2007], puis d'autres biomolécules comme les protéines [Monticelli et al. 2008] et plus récemment l'ADN [Uusitalo et al. 2015]. Les paramètres d'interactions utilisés pour les atomes liés sont issus de simulations tout-atomes, alors que les paramètres utilisés pour les interactions non-liées se basent sur des calculs d'énergie libre de partitionnement de nombreux composés entre phases polaire et apolaire. Ce champ de forces a été développé initialement pour le logiciel de simulation GROMACS, avant d'être rendu compatible avec d'autres logiciels de simulation comme NAMD. Ce champ de force sera décrit plus en détail dans la partie 2.1. Ces dernières années, l'utilisation de ce champ de forces a connu un véritable engouement, et a vu son champ d'application s'étendre à des systèmes de plus en plus larges. Entre autres études, ce modèle a ainsi permis de proposer la formation d'oligomères de la Rhodopsine dans une membrane [Periole et al. 2007], ou de simuler des nanoparticules complètes afin de reproduire quantitativement leur capacité à délivrer des composés [Lee et Larson 2008]. Au laboratoire, une des premières études utilisant ce champ de forces a visé à simuler l'ancrage d'une protéine du VIH sur la face interne de la membrane cellulaire, tout en reproduisant sa capacité à recruter un type de lipides particulier, le PIP2 [Charlier et al. 2014].

### 1.3 Modélisation de la dynamique intrinsèque et des interactions impliquant les RCPGs

Comme discuté plus haut dans ce chapitre, les RCPGs sont des protéines membranaires hautement dynamiques qui subissent des changements conformationnels à la fois de faibles et de plus larges amplitudes lors de leur liaison à leurs différents partenaires, et ce depuis la liaison d'un ligand du côté extra-cellulaire jusqu'au couplage avec leurs effecteurs intra-cellulaires. Les données structurales existantes donnent peu d'informations sur ces processus dynamiques. Aussi, la modélisation moléculaire et en particulier la dynamique moléculaire (MD) apparaît comme une alternative intéressante pour compléter cette vue rigide apportée par les structures de ces systèmes aujourd'hui disponibles.

De nombreuses études visant à étudier les changements de conformation des RCPGs ont été réalisées suite à l'apparition de la première structure cristallographique de la Rhodopsine en 2000 [Palczewski et al. 2000]. Parmi ces études, nombre d'entre elles ont également cherché à comprendre le mode d'interaction et l'effet du Rétinal, son ligand endogène, sur sa structure intrinsèque [Hornak et al. 2010; Lemaître, Yeagle, et Watts 2005; Leioatts et al. 2015]. Une autre vague d'études basées notamment sur la dynamique moléculaire suivra la publication de la structure du récepteur  $\beta$ 2AR en 2007 [Cherezov et al. 2007]. Depuis 2011 et la publication de la première structure montrant l'interaction entre un RCPG et une protéine-G [Rasmussen et al. 2011], beaucoup de travaux ont visé à comprendre la dynamique intrinsèque de ce complexe à la membrane [Louet et al. 2013; Flock et al. 2015; Saleh, Saladino, et al. 2017; DeVree et al. 2016]. Parmi ces derniers, ceux effectués au laboratoire, y compris avant la publication de cette structure, restent pionniers en la matière et proposaient déjà à l'époque un effet de la protéine-G sur la dynamique intrinsèque des RCPGs [Louet et al. 2011].

Comme il existe différents niveaux de représentation des systèmes moléculaires (partie 1.2), il existe de nombreuses techniques différentes basées sur la dynamique moléculaire. Le choix de la méthode la plus adaptée à la question posée doit être fait notamment en fonction de l'échelle de temps de simulation qui sera requise pour reproduire le phénomène attendu à l'échelle moléculaire. Aussi, de nombreuses techniques de MD biaisées ont vu le jour et ont été appliquées aux RCPGs. Ces techniques introduisent différents biais, soit directement sur la fonction d'énergie potentielle comme dans le cas de la dynamique accélérée [Kappel, Miao, et McCammon 2015; Miao et al. 2016], soit en forçant l'exploration des conformations le long de coordonnées réactionnelles bien choisies, comme dans le cas de la métadynamique [Saleh, Ibrahim, et al. 2017; Provasi, Bortolato, et Filizola 2009].

### 1.3.1 Liaison du ligand

Comprendre les mécanismes moléculaires mis en jeu lors de la reconnaissance et la liaison de ligands par les RCPGs est primordial pour le développement de molécules à visée thérapeutique. Malheureusement, cette liaison apparaît à l'échelle de la microseconde, avant d'être suivie par une dissociation encore plus lente. A cause des limitations de puissance de calcul, la MD conventionnelle a longtemps été incapable de simuler de tels événements. Pour contourner ces limitations, des techniques de MD ont été développées.

La « Random acceleration MD », aussi appelé « random expulsion MD », a été utilisée pour simuler la sortie des ligands de la Rhodopsine en 2007 et du récepteur  $\beta$ 2AR en 2009 montrant deux chemins très différents [T. Wang et Duan 2007, 2009]. Pour la Rhodopsine, la boucle ECL2 formant un repliement  $\beta$  bouche totalement l'accès à la poche de liaison, obligeant le ligand à passer entre les TM4-TM5 ou TM5-TM6. Pour le récepteur  $\beta$ 2AR, cette même boucle forme une courte hélice  $\alpha$ , permettant d'avoir une ouverture de la poche de liaison du côté extra-cellulaire par laquelle le ligand peut rentrer et/ou sortir. Dans cette même étude, il a été proposé que le départ du ligand pourrait entraîner la rupture d'un pont salin important sur le plan structural entre le TM7 et la boucle ECL2, conduisant à deux chemins de sortie possibles de chaque côté de ce pont, et qui seront retrouvés dans des simulations de « Steered MD » simulant la dissociation de ligands des récepteurs  $\beta$ 1AR et  $\beta$ 2AR [González et al. 2011]. Dans cette étude le chemin préférentiel emprunté par le ligand passe entre la boucle ECL2 et les TM5, TM6 et TM7. Par métadynamique, la liaison d'un ligand au récepteur  $\delta$ -Opioïde a été également simulée en utilisant un récepteur construit par homologie [Provasi, Bortolato, et Filizola 2009]. Dans cette étude le ligand passe par une crevasse formée entre les boucles ECL2 et ECL3. De plus, il a été proposé que la position finale du ligand dans le récepteur présente une constante d'équilibre très proche des valeurs expérimentales, suggérant la possibilité d'étudier quantitativement la liaison de ligands aux RCPGs avec cette méthode.

C'est finalement en 2011 que l'on put observer la liaison d'un ligand à un RCPG en utilisant des techniques classiques de dynamique moléculaire, sans aucune contrainte appliquée. Dans cet article, de très longues simulations (dépassant de loin la milliseconde) ont été produites en utilisant le champ de forces CHARMM [Bernard R. Brooks et al. 1983], et permettant d'observer la liaison de petites molécules au récepteur  $\beta$ 2AR en utilisant le super ordinateur dédié au calcul « Anton » [Dror, Pan, et al. 2011]. Les auteurs montrent une entrée du ligand entre la boucle ECL2 et les TM5, TM6 et TM7, en accord avec les simulations de Steered MD discutées plus tôt. En 2012, une autre étude des mêmes auteurs a suggéré

l'impossibilité pour l'antagoniste Tiotropium (TTP) à pouvoir se lier au fond du site orthostérique du récepteur de l'Acétylcholine M3, en désaccord avec les données cristallographiques disponibles pour ce couple ligand:récepteur [Kruse et al. 2012]. Par ailleurs, ces auteurs ont été les premiers à suggérer que le ligand pourrait se placer dans un site vestibulaire situé à l'entrée de la poche, suggérant une position intermédiaire du ligand agoniste lors de sa liaison au récepteur. En 2013, cette position décrite comme « vestibule » fut identifiée comme site de liaison de plusieurs modulateurs allostériques négatifs du récepteur M2 [Dror et al. 2013], puis d'un modulateur positif avec une structure cristallographique à l'appui montrant une position intermédiaire pour l'un de ces ligands [Kruse et al. 2013].

Ces études employant un champ de forces classique restent néanmoins très coûteuses en temps de calcul et accessibles uniquement aux groupes possédant des machines de calcul démesurées et dédiées à la dynamique moléculaire, limitant de fait la généralisation de cette approche à d'autres couples ligand:récepteur. L'emploi de techniques biaisées permettant d'accélérer l'échantillonnage des conformations de tels complexes s'est ensuite généralisé ces dernières années. Ainsi, le développement de la dynamique accélérée a permis en 2013 d'étudier la liaison de trois ligands différents au récepteur Muscarinique M3 : un antagoniste TTP, un agoniste plein l'Acétylcholine et un agoniste partiel l'Arécoline [Kappel, Miao, et McCammon 2015]. Comme la dynamique classique, cette technique a permis aussi l'identification de possibles positions intermédiaires, notamment pour les antagonistes. Grâce au biais appliqué à la fonction d'énergie potentielle, le temps nécessaire pour observer la liaison du ligand TTP à son récepteur a été réduit d'environ 80 fois.

Dans ce manuscrit nous présentons le développement, la validation et l'utilisation à des fins prédictives, d'un protocole combinant « Replica-Exchange MD » (REMD) et le champ de forces de MARTINI, pour prédire, sans aucune contrainte appliquée, la liaison de différents partenaires des RCPGs, notamment de peptides. Notre approche a été validée sur les peptides [8-13] de la Neurotensine ( $\text{NH}_3^+$ -Arg-Arg-Pro-Tyr-Ile-Leu- $\text{CO}_2^-$ ) et CVX15 ( $\text{NH}_3^+$ -Arg-Arg-Nal-Cys-Tyr-Gln-Lys-DPro-Pro-Tyr-Arg-Cit-Cys-Arg-Gly-DPro- $\text{CO}_2^-$ ) se liant respectivement aux récepteurs NTR1 de rat et CXCR4 [Delort et al. 2017]. En utilisant cette approche combinée, nous montrons qu'il est possible d'obtenir des conformations quasiment identiques à la cristallographie en un temps assez court (plusieurs centaines de  $\mu\text{s}$  de temps de simulation produits en quelques jours) et en ayant accès à des moyens de calculs très raisonnables (cluster de laboratoire avec une dizaine de nœuds de 30~40 cœurs).

### 1.3.2 Docking

La prédiction de la position d'un ligand dans sa poche de liaison se fait le plus généralement par amarrage moléculaire ou « docking ». La littérature montre un nombre incalculable d'études dans lesquelles des composés ont été « dockés » dans un ou plusieurs RCPGs. De nombreux logiciels de docking existent et sont basés le plus souvent sur des fonctions de score empiriques permettant de classer les différentes conformations du ligand au fond de la poche. Contrairement aux techniques de dynamique moléculaire décrites ci-dessus, le docking ne tient compte généralement pas (ou peu) de la flexibilité de la cible. Aussi, les résultats de docking dépendront totalement de la conformation du récepteur utilisée au départ. Une alternative est l'utilisation de la dynamique moléculaire pour produire différentes conformations représentatives de la poche de liaison [Miao et al. 2016]. En docking, seule la conformation du ligand est pleinement explorée, en faisant varier ses angles dièdres, le plus souvent en utilisant des algorithmes génétiques. Un des problèmes majeurs du docking provient donc des limitations de cet échantillonnage qui s'avère difficile au-delà d'une dizaine de degrés de liberté [Meng et al. 2011]. Ainsi, docker un peptide dans un RCPGs n'est pas raisonnable au-delà de quelques résidus. Certains logiciels permettent également de rendre flexibles quelques résidus de la cible, mais le choix de ces résidus s'avère toujours difficile, et limite énormément l'échantillonnage des conformations du complexe [S.-Y. Huang et Zou 2010].

Pourtant, des expériences de criblage *in silico* à haut débit effectués sur la structure du récepteur  $\beta$ 2AR [Sabio, Jones, et Topiol 2008; Kolb et al. 2009] ont permis d'identifier six nouveaux ligands de très bonne affinité de ce récepteur et présentant pour la plupart des caractéristiques (« chémotypes ») proches de ligands déjà connus. En revanche, et comme déjà discuté plus haut, une grande majorité de RCPGs ne disposent pas encore d'une structure cristallographique de bonne résolution, et il faut souvent utiliser la modélisation par homologie pour construire un modèle de départ de son récepteur d'intérêt. Cette étape de modélisation par homologie devrait fortement limiter les chances de succès de telles études de docking. Pourtant, de bons résultats ont pu être décrits dans la littérature en partant de modèles par homologie [Beuming et Sherman 2012; Loo et al. 2018]. De nombreuses méthodes ont été également développées pour limiter cette incertitude au niveau structural, comme par exemple celle appliquée au récepteur TRH-R [Engel et al. 2008]. Dans cette étude, un criblage basé sur une représentation pharmacophorique de la poche de liaison a permis d'identifier des antagonistes (les plus efficace à l'époque) présentant une sélectivité pour le récepteur TRH-R1 *versus* le TRH-R2.

La conception de ligands sélectifs reste un défi majeur dans le développement de nouvelles drogues. Aussi, au sein des sous-familles de RCPGs, il n'est pas rare de trouver de nombreux sous-types de récepteurs avec des sites orthostériques présentant une très forte conservation, rendant difficile l'identification de ligands sélectifs par docking [Allen et Roth 2011]. Une alternative est la recherche de ligands allostériques. Lors d'un criblage, Huang et al. ont identifié un modulateur allostérique positif du GPR68 et produit des modèles par homologie de ce récepteur afin de prédire la position de ce ligand par docking [Huang et al. 2015]. Cette position a ensuite été utilisée comme base pour cribler et identifier de nombreuses autres molécules comme modulateurs de ce récepteur.

Ces approches décrites ci-dessus sont toutes basées sur un modèle du récepteur et sont généralement regroupées sous le terme de « structure-based drug design ». Dans un autre type d'approche cette fois « ligand-based », la conception de ligands peut se faire cette fois en se basant uniquement sur la structure ou les pharmacophores de ligands déjà connus. Par exemple, en se basant sur le modèle pharmacophore d'une molécule se liant à la Rhodopsine, le criblage d'une librairie de molécules a conduit à la découverte de trois nouveaux composés stabilisant la Rhodopsine dans ses conformations actives [Taylor et al. 2010].

Dans notre travail, nous avons cherché à développer une approche dynamique où les conformations des deux partenaires sont explorées, afin de mieux prédire leurs interactions. Cette approche dynamique permet non seulement l'identification de la position finale du ligand au fond de la poche orthostérique du récepteur mais aussi celle de positions intermédiaires, qui pourraient s'avérer importantes pour l'affinité et la sélectivité de ligands, et possiblement allostériques.

### 1.3.3 Dynamique intrinsèque du récepteur

La compréhension de la dynamique intrinsèque des RCPGs est un point clé car elle devrait permettre d'expliquer comment le message porté par un ligand extra-cellulaire est transmis à l'intérieur de la cellule, et donc aussi comment le favoriser ou le bloquer spécifiquement. L'augmentation très nette ces dernières années du nombre de structures disponibles a permis d'observer ces récepteurs sous différentes formes « active » et « inactive ». Ainsi, il a été montré par simulations de dynamique moléculaire classique que la transition du récepteur  $\beta$ 2AR de sa forme « active » couplée à la protéine-G à sa forme « inactive », passait par un état conformationnel intermédiaire [Dror, Arlow, et al. 2011]. Inversement, la MD accélérée permet de simuler l'activation du récepteur Muscarinique M2 en retirant le ligand

antagoniste de la structure cristallographique « inactive », cette activation passant alors par deux intermédiaires conformationnels [Miao et al. 2013]. Toujours en se basant sur le mouvement de couplage aux protéines-G comme élément clé de l'activation, plusieurs états intermédiaires de basse énergie ont pu être mis en évidence le long du chemin d'activation de la Rhodopsine et du récepteur  $\beta$ 2AR en combinant métadynamique et « adiabatic-biasing MD » (ABMD) afin d'étudier les profils d'énergie libre de ce mouvement [Provasi et Filizola 2010]. Dans cette étude, après avoir choisi les conformations « active » et « inactive », les intermédiaires furent alors produits par « clustering » de plusieurs ABMD indépendantes, en utilisant comme critère des valeurs de RMSD calculées sur les résidus appartenant aux TMs. Ces clusters ont été utilisés ensuite pour définir le chemin suivi en métadynamique. D'autres simulations utilisant la même technique ont permis de voir un effet du ligand sur l'équilibre entre différentes conformations du  $\beta$ 2AR [Provasi et al. 2011]. Dans cette étude, des ligands agonistes plein, partiel, inverse et antagoniste, ont été placés dans la poche du récepteur, soit en partant de la position connue en cristallographie (si disponible), soit par des calculs de docking. Des simulations du récepteur avec et sans ligand ont ensuite été réalisées afin de déterminer le profil d'énergie libre le long du chemin entre les formes supposées inactive et active du récepteur. Sans ligand et avec l'antagoniste, deux conformations de basse énergie ont pu être ainsi observées, une correspondant à la forme « inactive » et l'autre proche de la forme « active ». En présence d'un agoniste inverse, seule la forme « inactive » est stable. Pour les agonistes en revanche, même si une sélection de la conformation « active » semble se dégager, les auteurs ont observé la présence d'autres minima pouvant s'expliquer par l'absence de pG dans les simulations. Au final, cette étude semble montrer que des approches biaisées peuvent apporter des informations sur les équilibres conformationnels des RCPGs. Si la métadynamique apparaît comme une méthode de choix pour remonter au profil d'énergie libre le long d'un chemin conformationnel, on sait néanmoins que les résultats obtenus avec cette méthode sont extrêmement dépendants du choix de la ou des coordonnées(s) réactionnelle(s) utilisées pour biaiser le potentiel.

L'autre technique de choix pour limiter le temps de calcul est la dynamique accélérée. En utilisant cette technique, il a été suggéré qu'il était possible d'observer une diminution importante de la flexibilité du récepteur M3 par un effet allostérique de la liaison au cœur du récepteur d'un ion sodium, le bloquant dans sa forme « inactive » [Miao, Caliman, et McCammon 2015]. Le problème principal de la dynamique accélérée est qu'elle biaise l'ensemble du potentiel ; lorsque nous avons essayé cette technique au laboratoire, nous avons observé de nombreux problèmes de déstructuration des RCPGs, en particulier dans les régions

subissant de grandes forces, comme l'hélice 8 accolée à la membrane, conduisant même à son dépliement total lors de simulations à l'échelle de la centaine de nanosecondes. En dynamique accélérée, nous avons pu observer par nous-mêmes également que les boucles extra-cellulaires, non structurées, subissent également de larges changements conformationnels, peu réalistes.

Dans notre travail, nous avons fait le choix de nous orienter vers une méthode non biaisée, sans aucune contrainte appliquée intra ou inter-moléculaire, et sans aucun biais sur le potentiel lui-même. Le seul biais appliqué dans notre approche est le biais cinétique apporté par la température. Comme le biais appliqué au potentiel dans la dynamique accélérée, on pourrait penser que ce biais cinétique pourrait agir de façon non rationnelle sur l'exploration des conformations du récepteur. Dans nos études, l'utilisation du modèle gros-grains MARTINI et de son réseau élastique permet d'éviter ce type de problèmes en maintenant les structures secondaires et tertiaires des protéines tout en décrivant de manière raisonnable leur dynamique intrinsèque.

### 1.3.4 Simulations de systèmes de taille plus importante

Il y a très peu d'études qui décrivent la dynamique d'interaction entre les RCPGs et leurs partenaires intra-cellulaires, probablement parce que ce sont des systèmes plus gros et plus difficiles à simuler à l'échelle moléculaire. Au laboratoire, nous avons été les premiers à publier une étude dynamique d'une protéine-G hétéro-trimérique ( $G_i$ ) ancrée à la membrane [Louet et al. 2012]. En utilisant des analyses de modes normaux, un mécanisme permettant une large ouverture à l'interface des domaines Ras-like et hélicoïdal a été suggéré comme impliqué dans le relargage du GDP [Louet et al. 2011]. En combinant cette fois dynamique moléculaire dirigée (TMD) et reconstruction de profils d'énergie libre, notre équipe a proposé que ce relargage du GDP se faisait plus favorablement par son côté phosphate [Louet, Martinez, et Floquet 2012]. Enfin, une étude de TMD forçant la sous-unité  $\alpha$  à transiter de sa forme liée au GDP à celle liée au GTP a permis d'observer de larges mouvements domaine-domaine avec une dissociation presque complète du dimère  $\beta\gamma$  [Louet et al. 2012]. D'autres études ont suggéré une flexibilité accrue de la sous-unité  $\alpha$  une fois couplée à son RCPG [Ceruso, Periole, et Weinstein 2004; Raimondi et al. 2008]. Toutefois, dans ces études, un certain nombre d'approximations importantes existent, comme l'absence de membrane et de l'hélice N-terminale permettant l'ancrage de la protéine-G et jouant un rôle important dans son activation [Oldham et Hamm 2008].

Suite à ces travaux, notre équipe s'est orienté vers l'étude des RCPGs complexés à leurs partenaires, suggérant que la protéine-G pouvait largement influencer sur le comportement dynamique de ces récepteurs [Louet et al. 2013]. La mise au point de méthodes permettant l'échantillonnage de mouvements à la fois de faible et de plus large amplitude [Costa et al. 2015] a permis d'explorer par dynamique moléculaire, les mouvements collectifs dans un complexe récepteur:protéine-G, toujours ancré à la membrane [Damian et al. 2015]. Dans l'étude de mouvements collectifs de tels complexes, il est néanmoins toujours difficile de comprendre lequel des partenaires contribue le plus à sa dynamique. Dans la plupart des études, la question posée est l'effet du récepteur sur la protéine-G, avec pour objectif de voir la connexion allostérique entre le site de liaison du ligand et la poche de liaison du nucléotide [Flock et al. 2017; Saleh, Ibrahim, et al. 2017].

Pourtant, des études expérimentales récentes effectuées sur le récepteur Adrénérique  $\beta_2$  ont montré que la protéine-G pouvait moduler l'activité du récepteur, même en l'absence de ligand [DeVree et al. 2016]. Dans cette article, il est proposé que la présence de la protéine-G pourrait agir à distance sur la poche de liaison du ligand et sur son degré d'ouverture/fermeture, permettant notamment de stabiliser une conformation plus fermée de la poche, gênant ainsi le relargage du ligand et augmentant son affinité apparente. Dans un travail soumis pour publication, notre équipe a pu montrer un tel effet en utilisant des simulations de dynamique moléculaire et le récepteur de l'Adénosine complexé à la mini- $G_s$  (code PDB : 5G53 [Carpenter et al. 2016] comme système modèle [Renault et al. 2018].

Dans ce travail, nous avons également cherché à utiliser notre protocole pour prédire la dynamique d'interaction de différents récepteurs avec leurs protéines-G favorites, notamment pour tenter de comprendre la sélectivité de couplage qui existe entre ces deux familles de protéines.

### 1.3.5 Dimérisation / oligomérisation des RCPGs

Un autre degré de complexité des RCPGs vient avec leur capacité à interagir entre eux, au sein de la membrane, pour former des dimères ou oligomères. Comme c'est le cas pour prédire le mode de liaison de ligands, le docking peut également être utilisé pour prédire le mode de dimérisation de RCPGs. Au contraire d'autres approches basées sur la dynamique moléculaire et qui sont discutées plus loin dans ce paragraphe, le docking reste peu coûteux en temps de calcul mais n'apporte qu'une vue statique de ces assemblages. La problématique principale d'utiliser le docking pour ce type d'applications réside dans les paramètres de dessolvation usuellement implémentés dans les fonctions de scores et beaucoup plus adaptés à des ligands ou protéines solubles. Ainsi, pour prédire l'assemblage de domaines hydrophobes tels que les TMs dans un milieu lipidique, les méthodes doivent être adaptées comme montré par exemple dans un travail de docking utilisant ZDOCK [F. Fanelli et al. 2013]. Ce type d'approche peut s'avérer utile pour générer des modèles de départ pour d'autres calculs, comme par exemple des simulations de dynamique moléculaire, comme montré pour un certain nombre d'études de dimérisation, du récepteur de la Neurotensine-1 [Casciari, Dell'Orco, et Fanelli 2008] ou de l'Adénosine [F. Fanelli et Felling 2011]. De façon assez surprenante, ces études mettent en avant une interface de dimérisation impliquant les boucles ICL2 et ICL3, ainsi que l'hélice 8.

La dynamique moléculaire a été utilisée pour étudier et prédire l'oligomérisation des RCPGs. Toutefois, la taille imposante de ces systèmes nécessite de puissantes ressources de calcul pour les simuler à un niveau de détail « tout-atomes ». Aussi, des études utilisant les champs de forces gros-grains ont été publiées, montrant la possibilité d'utiliser ce type de modèles pour prédire à l'échelle moléculaire, l'auto-assemblage de RCPGs tel que la Rhodopsine [Periole et al. 2007]. Dans ce travail, plusieurs interfaces de dimérisation furent observées et étudiées plus en détail, notamment par le calcul de potentiels de forces moyennes (PMF). Ces calculs suggèrent une interface préférentielle TM1/TM8/H8 aboutissant à un dimère symétrique. Des résultats très similaires ont été rapportés récemment sur le récepteur  $\beta$ 2AR [Ghosh, Sonavane, et Joshi 2014]. La grande limitation de ces simulations réside dans la tendance du champ de forces MARTINI à agglutiner les récepteurs entre eux pour former des agrégats stables et difficiles à rompre [Javanainen, Martinez-Seara, et Vattulainen 2017]. Aussi, les équilibres conduisant à l'association/dissociation de ces oligomères sont mal représentés, limitant ainsi l'interprétation des résultats. Dans notre travail, nous avons pu constater que l'utilisation du REMD permet (ou pourrait permettre) de discuter de tels équilibres car

## 1 Introduction

---

autorisant grâce à la montée en température de dissocier/réassembler de tels dimères /oligomères de RCPGs. En revanche, la plupart de nos simulations ont été faites avec un modèle très simplifié de membrane, le plus souvent présentant 100% de POPC (1-palmitoyl-2-oléoyl-sn-glycéro-3-phosphocholine). Hors, il a été récemment montré que la composition de la membrane et notamment la présence de cholestérol pouvait avoir un effet sur le mode de dimérisation des RCPGs [Prasanna, Sengupta, et Chattopadhyay 2016]. Ces simulations réalisées avec le champ de forces MARTINI, suggèrent que sans cholestérol, le récepteur  $\beta$ 2AR se dimérise *via* l'interface TM4/TM5-TM4/TM5. Plus la concentration de cholestérol est grande et plus cette interface disparaît progressivement en faveur d'autres interfaces comme TM4/TM5-TM1/TM2, puis TM1/TM2-TM1/TM2 au-delà des 50% en cholestérol. Une possible explication est la complexation du cholestérol sur les différentes zones impliquées dans les interfaces, notamment autour du TM4. Cet effet du cholestérol a également été suggéré par la simulation d'autres récepteurs, comme le récepteur CXCR4 [Pluhackova et al. 2016].

De manière intéressante, des études expérimentales récentes ont montré l'effet potentiel du cholestérol non seulement sur la dimérisation, mais aussi sur le couplage et l'activation de RCPGs [Luchetti et al. 2016; Manna et al. 2016; Bolivar et al. 2016; Casiraghi et al. 2016]. D'autres études computationnelles très récentes ont suggéré également l'effet de certains lipides, notamment du PIP2, sur l'activation et la sélectivité de couplage des RCPGs [Yen et al. 2018]. Ces effets de la membrane sur la dimérisation ou le couplage restent néanmoins encore soumis à controverse comme discuté dans d'autres articles récents combinant expérience et modélisation moléculaire [Meral et al. 2018].

Dans ce travail, nous avons cherché à utiliser notre protocole pour prédire la dimérisation de différents récepteurs en simulant la formation d'homo-dimères et d'hétéro-dimères, pour justifier des interfaces décrites dans un dimère proposé par notre équipe.

## 1.4 Objectifs de la thèse

L'objectif de cette thèse a été de tester, valider, et utiliser à des fins prédictives, un protocole assez peu coûteux en ressources de calcul, basé sur la dynamique moléculaire, et surtout ne présentant aucun biais ou contrainte inter-moléculaire, pour prédire à l'échelle moléculaire, les interactions entre les RCPGs et leurs partenaires principaux (**Figure 22**). Notre choix s'est rapidement orienté vers l'utilisation d'un modèle gros-grains et le champ de forces MARTINI, bien validé pour les lipides et les protéines, et déjà largement appliqué à ce type de récepteurs, comme discuté dans les paragraphes précédents. L'originalité de notre approche réside dans la combinaison de ce champ de forces à une méthode d'échantillonnage augmenté « Replica-Exchange » permettant, comme nous le montrons dans ce manuscrit, une bonne convergence des données et le franchissement de barrières d'énergie sans lequel le système serait piégé dans des conformations particulières (minima locaux).

L'objectif a été d'apporter une vue dynamique de ces interactions venant ainsi compléter celle, plus rigide, apportée par les nombreuses structures cristallographiques désormais disponibles de ces récepteurs et de leurs partenaires.

Nous nous sommes particulièrement intéressés à la prédiction du mode de liaison de ligands, notamment peptidiques, du côté extra-cellulaire, mais aussi à la complexation de ces récepteurs avec leurs partenaires favoris intra-cellulaires, les protéines-G. Un résultat assez inattendu au départ vient de la capacité de ce protocole à prédire également la sélectivité de ces récepteurs pour leurs différents partenaires, et ce malgré les approximations inhérentes au champ de forces utilisé. Nous nous sommes aussi posé la question de l'utilité d'un tel protocole pour prédire les interactions protéine:protéine et notamment l'association de ces récepteurs sous forme de dimères ou oligomères.

La première partie de ce travail montre la validation de notre protocole sur deux récepteurs qui ont été co-cristallisés avec leurs ligands peptidiques : les récepteurs de la Neurotensine (NTR1) et des Chimioquinas C-X-C (CXCR4). Comme discuté plus haut dans ce chapitre, la prédiction du mode d'interaction d'un peptide avec son récepteur est une tâche particulièrement difficile car elle nécessite un bon échantillonnage des conformations des deux partenaires.

## 1 Introduction

---

Dans la deuxième partie, nous montrons que ce même protocole qui nécessite des moyens de calcul très raisonnables, peut de ce fait être utilisé à moyen débit, c'est-à-dire pour prédire le mode d'interaction de plusieurs peptides dans plusieurs récepteurs homologues. Aussi, nous avons appliqué notre méthode à différents peptides dérivés de la Neurotensine, conçus sur la base d'hypothèses issues de simulations de dynamique moléculaire tout-atomes et montrant une sélectivité pour les récepteurs humains NTR1 et NTR2.

La troisième partie et la quatrième partie discutent de résultats préliminaires déjà obtenus avec la même méthode et qui sont confrontés à des résultats expérimentaux pour prédire l'assemblage de dimères/oligomères de récepteurs Dopamine/Ghréline (DRD2:GHSR-1a), l'interaction de toxines de venin de serpent avec le récepteur de la Vasopressine (V2R), l'interaction du peptide de la Ghréline avec son récepteur (GHSR-1a), et enfin la sélectivité de différents récepteurs vis-à-vis des différentes protéines-G ( $G_s$ ,  $G_i$ ,  $G_q$ ).

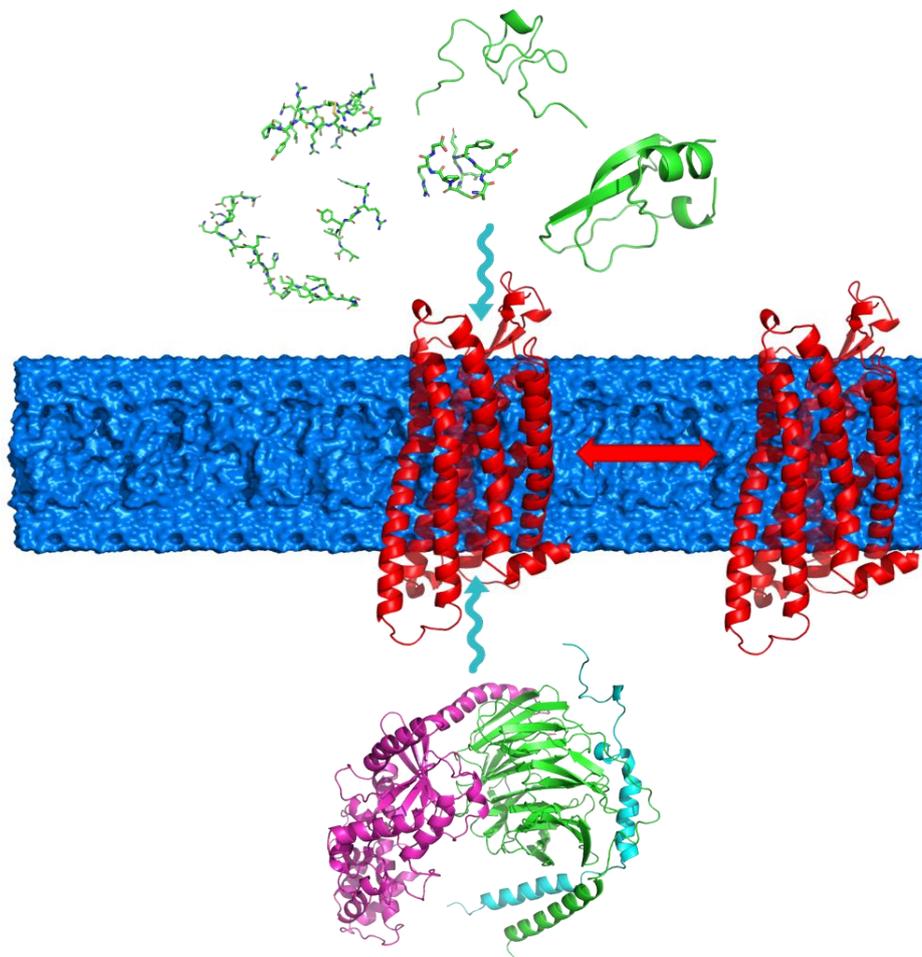


Figure 18 : Illustration des objectifs de la thèse.

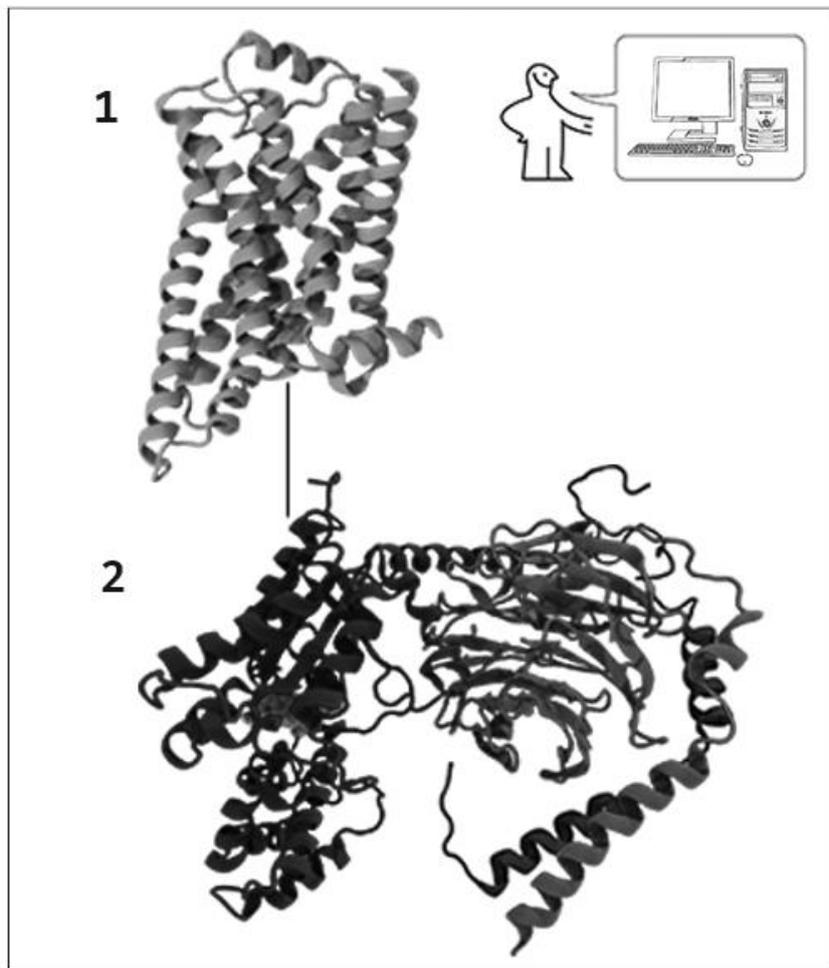




# Chapitre 2

## Matériels et Méthodes

### KOPPLADE





## 2 Matériels et Méthodes

### 2.1 Mécanique moléculaire

En mécanique moléculaire (MM) on utilise une description newtonienne des particules constituant le système d'intérêt pour prédire l'énergie associée à chacune de ses conformations. Cette méthode empirique englobe différents algorithmes mathématiques et surtout un ensemble de paramètres (champ de forces) permettant de calculer l'énergie potentielle associée au système. En MM, les atomes (ou dans certains cas groupes d'atomes), sont considérés comme des sphères solides et incompressibles, avec un rayon de Van der Waals et des charges fixes. Ces sphères sont liées par des liaisons assimilées à des ressorts qui ne peuvent être rompues. Puisque la mécanique moléculaire se base sur le calcul de fonctions mathématiques, son champ d'applications s'est considérablement élargi ces dernières années, grâce aux progrès de l'informatique (augmentation de la puissance de calcul).

#### 2.1.1 Champ de forces MARTINI EINEDyn

Durant mon travail de thèse, j'ai principalement utilisé le champ de force MARTINI. Celui-ci tient son nom de la ville où il a été développé : Groningue aux Pays-Bas, surnommée « *Martinistad* » à cause de ses nombreux lieux associés à Saint-Martin. Il s'agit d'un champ de force Gros-Grains (CG) développé par le laboratoire de Marrink et publié/validé en 2004 pour la simulation des lipides [Marrink, de Vries, et Mark 2004], puis à partir de 2007 étendu à d'autres types de biomolécules comme les protéines [Monticelli et al. 2008] ou plus récemment l'ADN [Uusitalo et al. 2015].

Notre travail portant sur les RCPGs, des protéines de taille moyenne mais fortement structurées, nous avons utilisé la version EINEDyn (« Elastic Network in Dynamics ») [Periole et al. 2009] du champ de forces MARTINI 2.2. Dans ce modèle, un réseau d'élastiques est appliqué entre certaines particules du système (plus de précision dans la partie 2.1.2) permettant de maintenir les éléments de structures secondaires ainsi que la structure tertiaire de la protéine. Dans certains cas, la structure quaternaire de la protéine peut également être maintenue grâce au rajout d'élastiques inter-chaînes. Pour les RCPGs, l'utilisation de ce champ de forces à part entière permet de décrire et contrôler la forme générale de ces récepteurs tout en décrivant de manière raisonnable leur dynamique intrinsèque et d'association.

### 2.1.2 Schéma de conversion tout-atomes vers gros-grains

Le modèle MARTINI repose sur une conversion 4-pour-1, c'est-à-dire qu'en moyenne, quatre atomes lourds sont représentés par une unique sphère. Avec MARTINI, une sphère d'eau englobe par exemple quatre molécules d'eau (**Figure 19**), un schéma de « mapping » qui d'emblée accélère la diffusion des solutés d'un facteur cinq à dix par rapport à un modèle tout-atomes [Marrink, de Vries, et Mark 2004]. Toutefois, pour décrire la complexité structurale de certains groupements annulaires, notamment les cycles aromatiques des tryptophanes, tyrosines, phénylalanines et histidines, une conversion 2 ou 3-pour-1 doit être utilisée pour certaines parties du système.

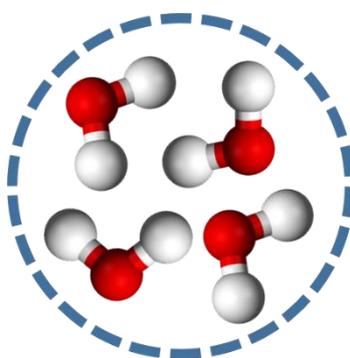


Figure 19 : Représentation d'une particule d'eau en gros grains dans le champ de forces MARTINI.

Le champ de forces comprend quatre types principaux de sphères : les apolaires (C), les intermédiaires (N), les polaires (P), et les chargées (Q). De plus, chaque type est à nouveau divisé : les polaires et les apolaires possèdent un numéro de 1 à 5 selon l'intensité de leur polarité (5 étant le plus polaire) ; les intermédiaires et les chargées possèdent un label manifestant leur capacité ou non (0) à donner un hydrogène (d), à les accepter (a), ou à pouvoir faire les deux (da). Toutes les particules ont un diamètre de collision  $\sigma$  de 0,47nm et une masse de 72amu. Ainsi, lorsque deux sphères se touchent, la distance entre leurs centres respectifs est de presque 5Å. Il est important de garder cela en mémoire pour les analyses de contacts faites lors de nos études et dans lesquelles un critère de 7Å a été utilisé. Dans MARTINI, les différents acides aminés naturels sont représentés comme reporté sur la **Figure 20**. Pour les cycles, les particules utilisées se différencient par la lettre S au début de leur nom (ex : SP1) et présentent un diamètre de collision de 0,43nm et une masse de 45amu. Il est à noter trois autres types de particules : BP4 (« Big-P4 ») qui permet de représenter une particule d'eau aux propriétés antigels (les particules d'eau peuvent spontanément s'ordonner et se figer, vers 280~300K dans MARTINI) grâce à une « taille » plus grande de 0,57nm qui va désordonner le système, et deux

particules d'acides aminés AC1 et AC2 qui sont utilisées pour éviter les gênes stériques entre particules chargées et apolaires (Q-C) liées au sein de la protéine. Enfin, toujours pour les acides aminés, chaque sphère a un type complémentaire en fonction de ses atomes d'origines : BB (« backbone ») pour l'unique sphère représentant le squelette de la molécule et centrée sur le Carbone  $\alpha$ , et SC (« side chain ») suivi d'un nombre pour désigner les sphères représentant la chaîne latérale.

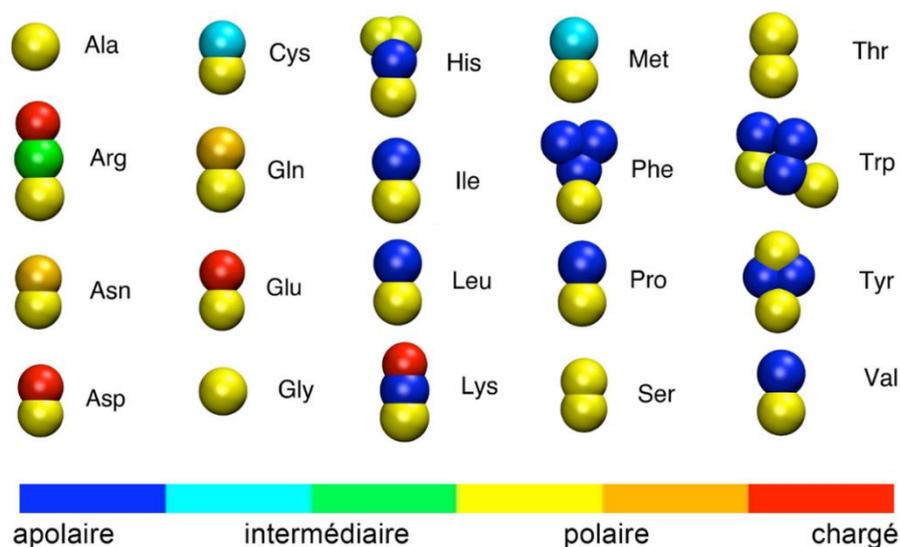


Figure 20 : Schéma de conversion en gros-grains des 20 acides aminés naturels avec le champ de forces MARTINI ; d'après [Monticelli et al. 2008].

La nature de la particule influence la valeur de la profondeur du puit de potentiel  $\epsilon_{ij}$  utilisée dans le calcul des interactions non-liantes de type Van der Waals (voir partie 2.1.3). Toutes les combinaisons possibles de sphères donnent lieu à un total de dix niveaux d'interaction différents (**Tableaux 3 et 4**).

Dénomination	Type	Valeur du potentiel de Lennard-Jones
O	Super attractif	$\epsilon = 5,6\text{kJ.mol}^{-1}$
I	Attractif	$\epsilon = 5,0\text{kJ.mol}^{-1}$
II	Presque attractif	$\epsilon = 4,5\text{kJ.mol}^{-1}$
III	Semi attractif	$\epsilon = 4,0\text{kJ.mol}^{-1}$
IV	Intermédiaire	$\epsilon = 3,5\text{kJ.mol}^{-1}$
V	Presque intermédiaire	$\epsilon = 3,1\text{kJ.mol}^{-1}$
VI	Semi répulsif	$\epsilon = 2,7\text{kJ.mol}^{-1}$
VII	Presque répulsif	$\epsilon = 2,3\text{kJ.mol}^{-1}$
VIII	Répulsif	$\epsilon = 2,0\text{kJ.mol}^{-1}$
IX	Super répulsif	$\epsilon = 2,0\text{kJ.mol}^{-1}$ et $\sigma = 0,62\text{nm}$

Tableau 3 : Niveaux d'interactions entre les différentes particules gros-grains dans MARTINI.

## 2 Matériels et Méthodes

sub	Q				P					N				C					
	da	d	a	0	5	4	3	2	1	da	d	a	0	5	4	3	2	1	
Q	da	O	O	O	II	O	O	O	I	I	I	I	I	IV	V	VI	VII	IX	IX
	d	O	I	O	II	O	O	O	I	I	I	I	III	I	IV	V	VI	VII	IX
	a	O	O	I	II	O	O	O	I	I	I	I	III	IV	V	VI	VII	IX	IX
	0	II	II	II	IV	I	O	I	II	III	III	III	III	IV	V	VI	VII	IX	IX
P	5	O	O	O	I	O	O	O	O	O	I	I	I	IV	V	VI	VI	VII	VIII
	4	O	O	O	O	O	I	I	II	II	III	III	III	IV	V	VI	VI	VII	VIII
	3	O	O	O	I	O	I	I	II	II	II	II	II	IV	V	V	VI	VII	VII
	2	I	I	I	II	O	II	II	II	II	II	II	II	III	IV	IV	V	VI	VII
	1	I	I	I	III	O	II	II	II	II	II	II	II	III	IV	IV	IV	V	VI
N	da	I	I	I	III	I	III	II	II	II	II	II	II	IV	IV	V	VI	VI	VI
	d	I	III	I	III	I	III	II	II	II	II	III	II	IV	IV	V	VI	VI	VI
	a	I	I	III	III	I	III	II	II	II	II	II	III	IV	IV	V	VI	VI	VI
	0	IV	IV	IV	IV	IV	IV	III	III	IV	IV	IV	IV	IV	IV	IV	V	VI	VI
C	5	V	V	V	V	V	V	IV	IV	IV	V	V	V						
	4	VI	VI	VI	VI	VI	VI	V	IV	IV	V	V	V	IV	IV	IV	IV	V	V
	3	VII	VII	VII	VII	VI	VI	V	V	IV	VI	VI	VI	IV	IV	IV	IV	IV	IV
	2	IX	IX	IX	IX	VII	VII	VI	VI	V	VI	VI	VI	V	V	V	IV	IV	IV
	1	IX	IX	IX	IX	VIII	VIII	VII	VII	VI	VI	VI	VI	V	V	IV	IV	IV	IV

Tableau 4 : Matrice d'interaction entre les différentes particules gros-grains dans MARTINI ; d'après [Marrink et al. 2007].

Comme déjà discuté plus haut, en plus de ces interactions, un élastique est ensuite appliqué à toutes les paires de sphères BB, séparées d'au moins 3 résidus (résidus  $i$  et  $i+3$ ), et distants d'au plus 0,9nm. L'ensemble de ces informations : noms et types des sphères, liaisons, élastiques ainsi que les valeurs d'angles et d'angles dièdres de la structure de départ tout-atomes utilisée sont regroupées dans le fichier de topologie, propre à chaque protéine du système.

### 2.1.3 Equations d'énergie

L'énergie potentielle du système est calculée de façon empirique par l'intermédiaire du champ de forces. Celle-ci est dépendante des coordonnées cartésiennes de chaque particule et peut se calculer par la somme des forces liantes et non-liantes exercées sur celle-ci (**Figure 21**). Dans les interactions non-liantes on retrouve les interactions de Van der Waals et les interactions électrostatiques. Dans les interactions liantes on a des termes exprimant la liaison de deux particules, la flexion de l'angle formé par trois particules, la rotation d'un angle dièdre formé par quatre particules, les angles dièdres impropres servant à maintenir plans certains cycles. Enfin, un terme décrit la force exercée par les élastiques placés entre des particules proches mais qui ne sont pas directement liées ( $>i+2$ ).

$$E_{totale} = E_{nonliante} + E_{liante} \quad (1)$$

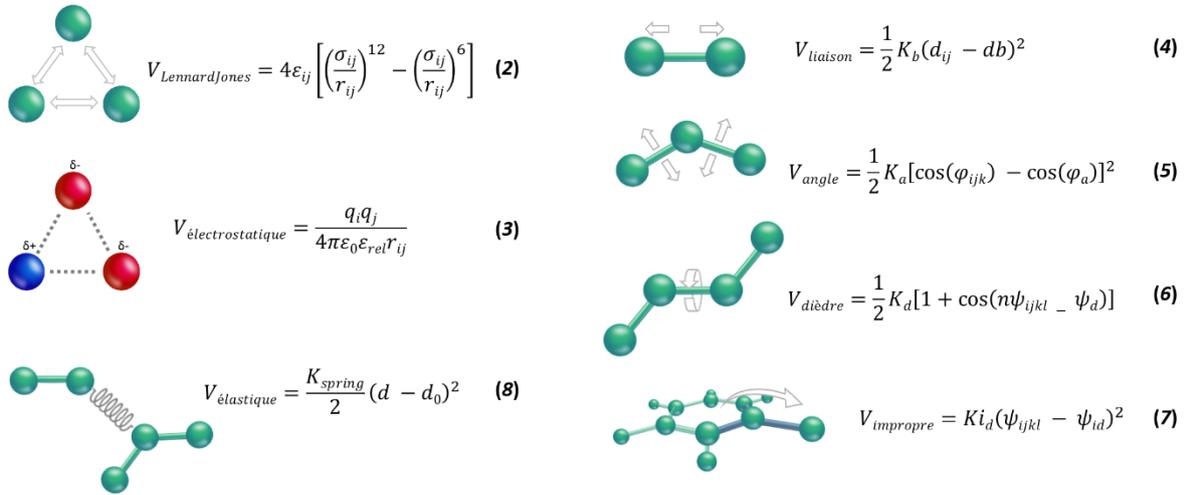


Figure 21 : Schématisation et équations des interactions atomiques dans le champ de forces MARTINI EINEDyn [Periole et al. 2009]

Deux particules  $i$  et  $j$  situées à une distance de  $r_{ij}$  interagissent *via* un potentiel de type Lennard-Jones (**Equation 2**). La force de l'interaction est déterminée par la valeur de la profondeur du puits de potentiel  $\epsilon_{ij}$ . Comme décrit dans la partie précédente (voir partie 2.1.2) l'utilisation de MARTINI conduit à seize principaux types de sphères et seulement dix types d'interactions, avec une échelle de  $\epsilon_{ij}$  allant de 2,0 à 5,6 kJ.mol<sup>-1</sup>. La taille effective des particules est donnée par le diamètre de collision  $\sigma = 0,47$  nm pour toutes les particules, exceptées celles issues d'anneaux pour lesquelles  $\sigma = 0,43$  nm et  $\epsilon_{ij}$  possède 75% de la valeur standard. Le potentiel de Lennard-Jones n'est pas appliqué aux particules liées par une liaison ou un élastique. De plus, ces particules portant une charge  $q$  interagissent *via* une fonction d'énergie Coulombienne (**Equation 3**) avec une constante diélectrique relative  $\epsilon_{rel}$  de 15. Encore une fois, ce terme est exclusif aux particules n'étant pas liées par un élastique.

Le calcul des termes non-liés du système est de loin le plus coûteux en temps de calcul puisque il s'applique en principe à toutes les paires d'atomes. Dans la pratique et pour simplifier les calculs, ces interactions ne sont mesurées que pour les paires de particules dont la distance est inférieure à une valeur définie appelée rayon de coupure (ou « cut-off ») [Verlet 1967]. Toutefois, l'arrêt brutal du calcul des énergies au rayon de coupure peut provoquer des instabilités énergétiques. La transition est donc adoucie par des équations diminuant progressivement la force avec l'augmentation de la distance. Dans notre travail, nous avons utilisé la fonction *shift* de GROMACS [Lindahl, Hess, et van der Spoel 2001]. Pour les

interactions de Van der Waals, la fonction *shift* est appliquée entre  $r_{shift} = 0,9\text{nm}$  et  $r_{cut} = 1,2\text{nm}$ , et pour les interactions électrostatiques, entre  $r_{shift} = 0\text{nm}$  et  $r_{cut} = 1,2\text{nm}$ .

Pour les particules liées, différents termes s'appliquent pour représenter les interactions liées. L'énergie associée à chacune des liaisons « covalentes » du système est calculée par un potentiel harmonique (**Equation 4**) centré sur une distance à l'équilibre de  $d_b = \sigma = 0,47$  ou  $0,43\text{nm}$  pour les différentes particules. Dans l'équation,  $d_{ij}$  est la valeur effective de la distance entre les deux particules et  $K_b$  la constante de raideur associée à chaque paire de particules. Pour décrire les angles associés aux liaisons, un potentiel harmonique du type cosinus est utilisé (**Equation 5**). Ainsi l'énergie potentielle associée à l'angle  $\varphi_{ijk}$  dépend là aussi de la valeur à l'équilibre  $\varphi_a$ , avec  $K_a$  une constante de force associée à chaque triplette de particules. Comme pour les champs de forces classiques, les angles dièdres permettent de définir la structure secondaire du squelette principal du peptide ou de la protéine. L'énergie associée aux dièdres est calculée par une décomposition en séries de Fourier (**Equation 6**) où  $\psi_{ijkl}$  est la valeur observée et  $\psi_d$  la valeur à l'équilibre ;  $K_d$  est la constante de force associée et représente communément la barrière d'énergie à franchir pour pouvoir effectuer la rotation. Les angles impropres sont utilisés pour permettre de reproduire certaines propriétés de déformation des groupements plans, notamment des cycles. L'énergie associée est évalué grâce à l'utilisation d'un potentiel harmonique (**Equation 7**) où  $\psi_{ijkl}$  est l'angle observée,  $\psi_{id}$  la valeur à l'équilibre et  $K_{id}$  est la constante de force.

Avec MARTINI, la force exercée par un élastique appliqué entre deux particules est également calculée par un potentiel harmonique (**Equation 8**) avec  $d$  la distance entre les particules et  $d_0$  la distance séparant les particules lors de l'application de l'élastique. La valeur à l'équilibre dépend donc ici de la structure utilisée au départ.  $K_{spring}$  représente la raideur de l'élastique. Dans notre travail nous avons utilisé généralement la valeur standard définie par MARTINI ( $K_{spring} = 500\text{kJ}\cdot\text{mol}^{-1}\cdot\text{nm}^{-2}$  et cut-off =  $0,9\text{nm}$ ). Néanmoins, nous avons également étudié la possibilité de modifier ces valeurs afin d'obtenir un réseau élastique permettant une meilleure description des fluctuations de nos protéines/peptides d'intérêt, plus proche de celle obtenue avec un champ de forces plus détaillé (« tout-atomes »). Pour ce faire nous nous sommes inspirés d'une étude basée sur la dynamique moléculaire [Globisch et al. 2013].

### 2.1.4 Minimisation d'énergie

L'un des intérêts majeurs de la modélisation moléculaire est d'identifier la (les) conformation(s) la (les) plus stable(s). Un objet constitué de  $N$  particules adopte différentes conformations selon ses degrés de liberté, définissant un espace à  $3N$  dimensions décrit comme l'hyper-surface de l'énergie potentielle. A chaque conformation est associée une énergie potentielle différente. Le but de la minimisation est de trouver la ou les conformations de plus basse énergie d'un système. En ces points appelés minima d'énergie, les dérivées partielles premières de la fonction d'énergie potentielles sont nulles et la moindre variation structurale de la conformation entraînera une augmentation de l'énergie. Il faut toutefois faire la différence entre le minimum d'énergie global où l'énergie est la plus basse, toutes conformations possibles confondues, et les autres appelés minima locaux (**Figure 22**).

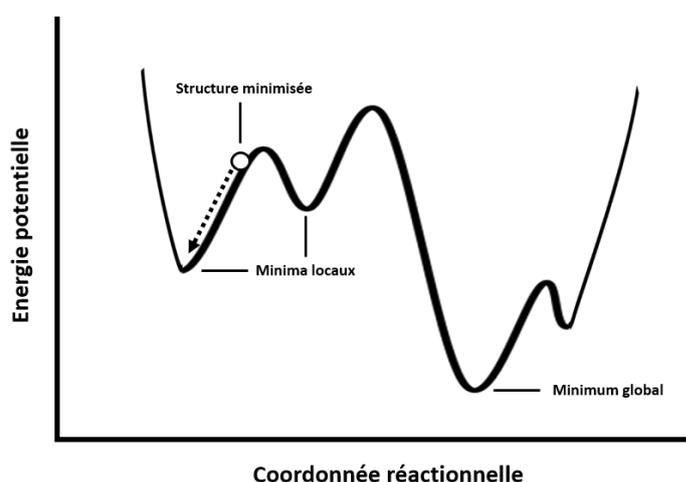


Figure 22 : Représentation à une dimension de l'énergie du système en fonction d'une coordonnée réactionnelle.

Trouver le minimum global d'un système complexe est très difficile et dépend du nombre de degrés de liberté qu'il présente et nécessite l'utilisation d'autres algorithmes. La minimisation d'énergie constitue néanmoins la première étape d'un protocole de modélisation moléculaire, quel qu'il soit, car elle assure de partir d'une conformation de basse énergie, après application des différents paramètres du champ de forces. Les algorithmes de minimisation utilisent selon les cas les dérivées de premier ou de second ordre de la fonction d'énergie potentielle. Le choix de ces différents algorithmes est généralement lié à la complexité du système, à son niveau de détail, mais aussi au type de calcul envisagé après minimisation. Dans notre travail nous avons utilisé uniquement la méthode de la plus grande pente (« Steepest Descent » ou SD) [Cauchy 1847] (**Figure 23**) qui utilise la dérivée de premier ordre de la fonction d'énergie. En premier lieu, l'énergie de la conformation initiale est calculée. Ensuite,

## 2 Matériels et Méthodes

---

les particules sont déplacées dans l'espace à  $3N$  dimensions. La nouvelle valeur de l'énergie est alors calculée ainsi que la dérivée première (gradient). La structure est à nouveau déplacée d'un pas arbitraire dans la direction du gradient négatif (donc vers une énergie plus faible). On réitère l'opération jusqu'à atteindre un minimum d'énergie. Cette méthode a l'avantage d'être très rapide car on se dirige toujours vers une conformation plus stable, mais possède comme défaut une mauvaise convergence. Cette méthode est généralement utilisée pour supprimer assez rapidement les clashes stériques possiblement liés à la construction du modèle et à l'application du champ de forces.

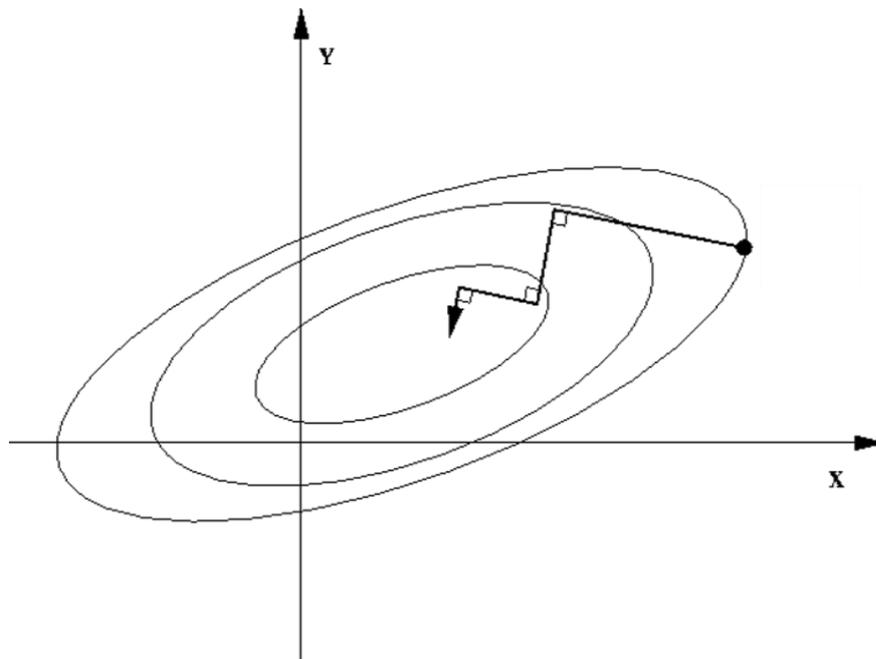


Figure 23 : Représentation de la minimisation par la méthode de la plus grande pente ; d'après [Hjorteland 1999].

## 2.2 Dynamique moléculaire

### 2.2.1 Théorie

En mécanique moléculaire, il est possible de simuler l'évolution d'un système de particules au cours du temps. Cette méthode appelée Dynamique Moléculaire (MD) repose sur l'intégration des lois du mouvement de Newton. La seconde loi de Newton est ainsi exprimée par l'**Equation 9** où  $\vec{F}_i$  est la force exercée sur la particule  $i$  de masse  $m_i$ , de position  $\vec{r}_i(x, y, z)$  avec une accélération de  $\vec{a}_i$ .

$$\vec{F}_i = m_i \vec{a}_i = m_i \frac{\partial^2 \vec{r}_i}{\partial t^2} \quad (9)$$

La force exercée sur cette particule peut aussi être exprimée par l'opposée de la dérivée partielle de la fonction d'énergie potentielle  $U$ , formalisée par l'**Equation 10**.

$$\vec{F}_i = - \frac{\partial U(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)}{\partial \vec{r}_i} \quad (10)$$

En connaissant l'énergie potentielle du système à un temps  $t$ , il est alors possible d'associer un vecteur vitesse à chacune de ses particules. De façon itérative, la connaissance de la position au temps  $t$  de chacune des particules permet ainsi la détermination de leur nouvelle position à un temps  $t + \Delta t$ , où  $\Delta t$  est le pas d'intégration. On parle de trajectoire de dynamique moléculaire pour désigner un ensemble de conformations successives d'un même système moléculaire. La valeur de  $\Delta t$  dépend directement du niveau de détail utilisé pour les simulations. Ainsi, si avec un champ de forces classique un pas de  $\Delta t = 1\text{fs}$  ( $10^{-15}\text{s}$ ) est généralement utilisé, le modèle gros-grains employé lors de cette thèse permet l'utilisation d'un  $\Delta t$  de 20fs.

### 2.2.2 Conditions périodiques aux limites

En dynamique moléculaire, la protéine est le plus souvent placée au centre d'une boîte qui va contenir d'autres molécules (eau, ions, lipides) visant à reproduire son environnement biophysique. Toutefois, si cette boîte de simulation était dans le vide, on observerait rapidement la dispersion des éléments qui la composent. Pour maintenir les propriétés biophysiques du système, il est nécessaire d'appliquer des conditions périodiques aux limites (« Periodic Boundary Conditions » ou PBC). Cette méthode consiste à copier le système étudié dans les trois directions de l'espace (**Figure 24**). Ainsi, une particule qui sortirait par un côté de la boîte

arriverait directement par le côté opposé. En utilisant ce biais, le système simulé paraît avoir des bornes infinies et garde ainsi une densité constante. Le calcul des interactions non-liées tient bien évidemment compte des conditions périodiques et fait intervenir la version la plus proche de chacune des particules impliquées.

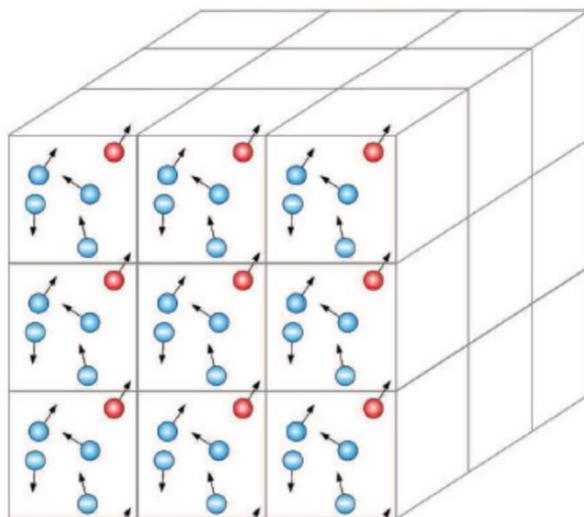


Figure 24 : Représentation des conditions périodiques aux limites ; d'après [Louet 2012].

### 2.2.3 REMD

La dynamique moléculaire de type « Replica-Exchange » (« Replica Exchange Molecular Dynamics » ou REMD) est une méthode d'échantillonnage accélérée utilisant comme biais la température, et qui permet de simuler des transitions de conformation qui ne pourraient être observées à la température classiquement utilisée lors de simulations (300K) car séparées par de trop importantes barrières d'énergie. Cette méthode a été développée initialement pour étudier le repliement de protéines [Sugita et Okamoto 1999], mais a depuis connu bon nombre d'applications [Luitz et Zacharias 2014; Felts et al. 2004; Pitera et Swope 2003; Liu et al. 2015].

Lors de simulations REMD, plusieurs répliques du même système sont simulées à différentes températures (une réplique = une température). A intervalles de temps réguliers (dans nos simulations gros-grains, tous les 1000 pas soit 20ps), le programme (ici GROMACS) va tenter « d'échanger » les coordonnées cartésiennes complètes du système simulé à deux températures adjacentes. La probabilité de cet échange est définie par l'**Equation 11** où  $U_1$  et  $U_2$  désignent l'énergie potentielle du même système aux deux températures  $T_1$  et  $T_2$ .

$$P(1 \leftrightarrow 2) = \min \left( 1, \exp \left[ \left( \frac{1}{k_b T_1} - \frac{1}{k_b T_2} \right) (U_1 - U_2) \right] \right) \quad (11)$$

Cette méthode permet ainsi d'augmenter l'échantillonnage des conformations du système en rendant plus facile le passage de barrières d'énergie à haute température, tout en gardant, à chacune d'entre elles une distribution de type Boltzmann de l'énergie associée au système. Pour des échanges optimaux, on retrouve dans la littérature une probabilité moyenne acceptée entre 20% et 40% [Rathore, Chopra, et de Pablo 2004; Sunita Patel, Vierling, et Tama 2014; Zhang, Wu, et Duan 2005], et pouvant aller jusqu'à 80% [Pitera et Swope 2003; Rhee et Pande 2003]. Cette valeur dépend directement du nombre de degrés de liberté du système étudié. Aussi, les paramètres ajustables en REMD incluent l'échelle de température utilisée ainsi que la fréquence de test des échanges. Dans tous nos calculs, nous avons utilisé le champ de forces gros-grains MARTINI qui réduit énormément le nombre de degrés de liberté. Lors de tests préliminaires, nous avons observé que pour un système classique incluant un récepteur couplé aux protéines-G dans une membrane avec de l'eau de chaque côté, une probabilité d'échange d'environ 20% en moyenne pouvait être obtenue en utilisant un écart de 3K entre deux répliques adjacentes. Dans la plupart de nos études, nous avons utilisé un minimum de 50 répliques entre 300 et 450K. Pour un système modèle de RCPG comme décrit dans le paragraphe sur la construction des systèmes (voir partie 2.4.9), on obtient des distributions d'énergie et le recouvrement représentés sur la **Figure 25**.

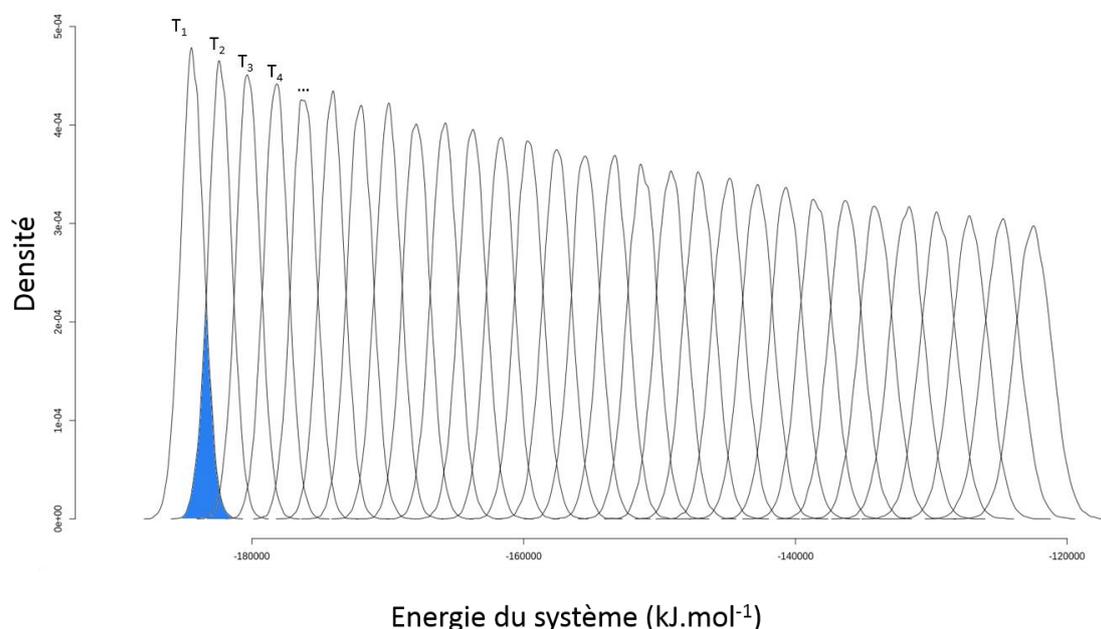


Figure 25 : Représentation de la distribution de l'énergie potentielle d'un de nos systèmes à différentes températures lors de simulations REMD ; la zone en bleu montre le recouvrement pour deux températures adjacentes, c'est-à-dire la zone où un échange de coordonnées du système entre ces deux températures est possible.

## 2 Matériels et Méthodes

Par défaut dans GROMACS, il y a une tentative d'échange tous les 1000 pas. Mais puisqu'on ne peut pas échanger simultanément les coordonnées d'un système de température  $T$  aux températures  $T-i$  et  $T+i$ , les échanges ont lieu en alternance sur les paires 'impaires' et 'paires'. En pratique, avec quatre répliques 1, 2, 3, et 4, on aura aux pas 1000, 3000, ..., des tentatives d'échange entre les réplique 1-2 et 3-4 ; puis aux pas 2000, 4000, ..., des tentatives d'échange entre les répliques 2-3.

A la fin du calcul REMD on obtient des trajectoires non continues. Par exemple, le modèle initialement équilibré à 300K va au cours du temps se déplacer le long des différentes températures. Ainsi, chacune des trajectoires obtenues contient l'ensemble des modèles (conformations) acceptées à la température correspondante. Dans notre travail, nous nous sommes particulièrement intéressés à la dynamique de liaison de différents partenaires des RCPGs. Aussi, pour récupérer le chemin de liaison que suit par exemple un peptide, on réalise un démultiplexage (« demux ») des trajectoires. Des scripts sont déjà prévus dans GROMACS permettant de stocker dans un fichier l'ensemble des informations d'échanges opérés lors d'une simulation. L'outil *gmx trjcat* permet alors de reconstituer les trajectoires continues (**Figure 26**). Bien entendu et à l'inverse, il ne faut pas oublier ici qu'une trajectoire continue contient des conformations issues possiblement de températures différentes.

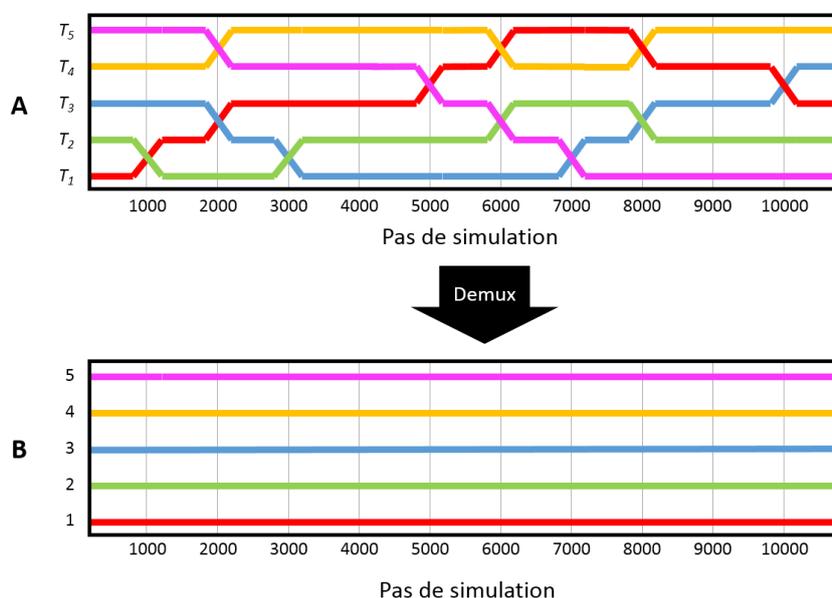


Figure 26 : Schématisation du demultiplexage ou « demux »; A) Trajectoires obtenues en REMD avec un « replex » standard (pas entre chaque tentative d'échange) de 1000 pas de simulation, la trajectoire obtenue est dépendante de la température simulée et n'est donc pas continue; B) Trajectoires obtenues après le « demux » par GROMACS, la trajectoire est dépendante du modèle de départ et la température du système varie.

### 2.2.4 Protocoles de simulations

Pendant cette thèse, toutes les dynamiques moléculaires ont été réalisées avec GROMACS, en utilisant le champ de forces MARTINI 2.2 avec EIneDYn. Pour la préparation des systèmes et les simulations, les paramètres standard définis par les concepteurs du champ de forces ont été utilisés : soit un pas d'intégration de 20fs, un rayon de cut-off de 0,9nm pour l'ajout d'élastiques entre résidus, une constante de raideur  $K_{spring}$  de  $500\text{kJ}\cdot\text{mol}^{-1}\cdot\text{nm}^{-2}$  pour tous les élastiques et un cut-off de 1,2nm pour le calcul des interactions non-liantes. L'énergie de chacun des systèmes étudiés a d'abord été minimisée par la méthode de la plus grande pente (5000 pas, assurant pour ce type de système un bon indice de convergence).

Par la suite, la conformation minimisée a été utilisée comme point de départ pour être équilibrée à chacune des températures d'étude lors de courtes simulations successives (50ns) réalisées dans les conditions NVT puis NPT. L'intégration des équations de Newton a été réalisée avec l'algorithme « leap-frog ». La température et la pression ont été maintenues constantes en utilisant les algorithmes « v-rescale » et « berendsen », respectivement.

Lors de la phase d'équilibrage, les vitesses initiales de toutes les particules ont été assignées aléatoirement selon une loi de distribution de Maxwell-Boltzmann. Lors du REMD, l'énergie du système a été testée tous les 1000 pas soit toutes les 20ps. Au final, pour chaque système étudié on obtient  $N$  trajectoires correspondant à chacune des répliques produites aux  $N$  températures. Les premières trajectoires produites lors de la thèse avaient un temps de simulation de  $2\mu\text{s}$ , puis finalement de  $10\mu\text{s}$  avec environ 50 températures étudiées et un intervalle de température de 3K, soit un temps de simulation total allant de 100 à  $500\mu\text{s}$  pour chaque système étudié. Pour la plupart des systèmes, plusieurs simulations REMD ont été réalisées afin de vérifier la convergence des données. Cette convergence est en particulier discutée pour le cas de la liaison du peptide Ghréline à son récepteur, pour lequel 3 expériences de  $500\mu\text{s}$  ont été effectuées (1,5ms de simulation au total) (voir partie 3.4.2.a).

## 2.3 Analyses des résultats

### 2.3.1 RMSD

Pour comparer deux conformations du même système moléculaire, on utilise assez fréquemment le calcul de RMSD (« Root Mean Square Deviation »). Pour deux conformations  $i$  et  $j$ , présentant toutes deux  $N$  particules de coordonnées position  $\vec{r}_i$  et position  $\vec{r}_j$ , le RMSD est donné par l'Equation 12:

$$RMSD(i, j) = \sqrt{\frac{1}{N} \sum_{n=1}^N (\vec{r}_{i,n} - \vec{r}_{j,n})^2} \quad (12)$$

Le RMSD est également directement lié à la manière dont les deux conformations sont alignées. Dans le cas par exemple d'un complexe protéine:protéine où chacun des partenaires serait flexible, l'alignement peut jouer un rôle primordial sur l'interprétation des résultats. Certains résidus très flexibles peuvent en effet perturber l'alignement et augmenter les valeurs résultantes de RMSD, masquant ainsi des valeurs bien plus convenables qui auraient pu être obtenues avec un alignement optimal. Un calcul de RMSD nécessite donc un choix méthodique des sélections à comparer ainsi qu'un alignement optimal des conformations. Enfin, le RMSD ne s'applique que pour deux objets ayant le même nombre de particules, ce qui rend difficile par exemple le calcul de RMSD entre récepteurs d'une même famille, à moins d'aligner ces modèles uniquement sur une région ou un groupe d'atomes conservés.

### 2.3.2 Clustering

Certaines de nos simulations génèrent plusieurs millions de conformations. Pour les trier et extraire les conformations les plus représentatives, on effectue un « clustering ». Il existe de nombreux algorithmes de clustering. Celui utilisé dans notre travail [Daura et al. 1999] est implémenté dans GROMACS et regroupe les conformations en utilisant un critère de RMSD (Figure 27). La première étape la création d'une matrice contenant les valeurs de RMSD pour chaque paire de conformations obtenue. Il est ensuite considéré que chaque conformation est le centre d'un cluster. Pour chaque centre, on compte alors le nombre de conformations qui présentent un RMSD inférieur à une valeur de cut-off définie au préalable. Le centre qui a le plus de connexions est choisi comme cluster numéro 1 et toutes les conformations de ce cluster sont extraites. Parmi les conformations restantes, on prend alors le nouveau centre avec le plus

de connexions et on réitère la démarche jusqu'à avoir extrait toutes les conformations. Toutefois, le problème de cette méthode est la dispersion des conformations au sein d'un même cluster. En effet, des conformations présentant entre-elles un RMSD supérieur au cut-off peuvent être regroupées au sein d'un même cluster si elles sont suffisamment proches de son centre. Un bon choix de la valeur du cut-off est donc essentiel.

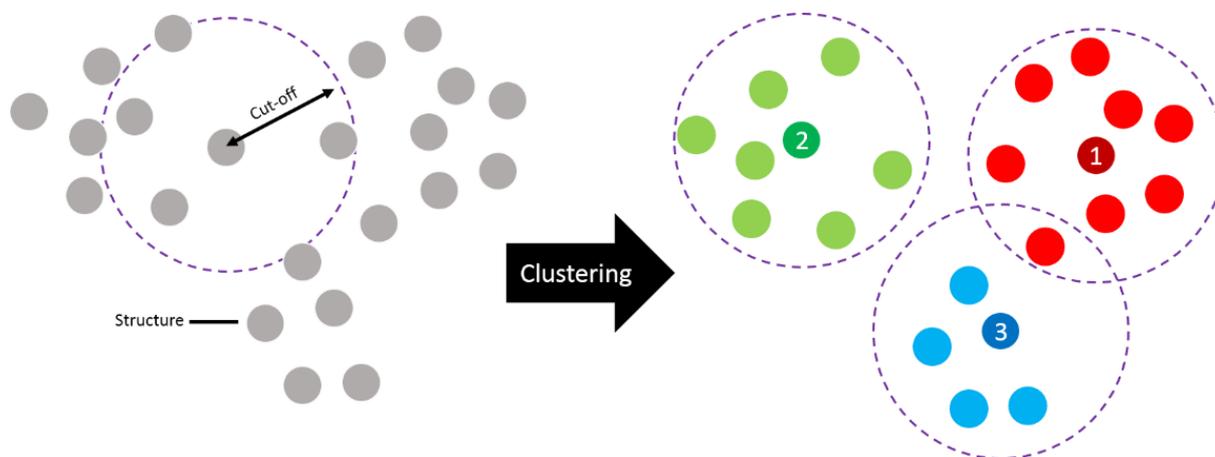


Figure 27 : Représentation schématique de l'algorithme utilisé pour le clustering ; chaque sphère correspond à une conformation ou structure, les couleurs représentant les clusters finaux classés par populations décroissantes.

Après clustering, il est facile d'obtenir une « trajectoire » contenant les centres de chacun des clusters, classés dans l'ordre décroissant de population, ou bien d'obtenir autant de trajectoires qu'il existe de clusters, chacune d'entre elles contenant toutes les conformations qui ont été regroupées ensemble. Lors de notre travail, toutes les analyses de clustering ont été réalisées avec l'outil *gmx cluster* de GROMACS avec l'algorithme de Daura et al [Daura et al. 1999] et un cut-off allant de 2Å pour les plus petites structures à 5Å pour les clusterings de dimères de RCPGs.

### 2.3.3 Analyse en composantes principales

L'Analyse en Composantes Principales (ACP) est une méthode statistique d'analyse de données multi-variables. Le but est d'identifier la relation entre ces données afin d'en obtenir une représentation simplifiée. Pour cela, de nouvelles variables, appelées « composantes principales » (CP), sont créées afin de différencier au mieux les données. On représente généralement le résultat d'une ACP par une projection des données en deux dimensions en choisissant les deux premières CPs comme axes permettant la plus grande dispersion des données dans ce nouveau plan.

### 2.4 Structures initiales

#### 2.4.1 Construction par homologie

Même si de nombreux récepteurs possèdent désormais une structure disponible dans la PDB (voir partie 1.1.2) on ne dispose encore, pour la plupart, que de leur séquence en acides aminés. En l'absence de structure disponible pour modéliser un récepteur d'intérêt, la modélisation par homologie peut être utilisée comme une alternative pour créer un modèle de départ pouvant être ensuite affiné par différents moyens, notamment par dynamique moléculaire. Dans ce travail, nous avons utilisé le logiciel MODELLER [Sali et Blundell 1993]. Pour réaliser la construction d'un récepteur par homologie, il est nécessaire de disposer d'un alignement parfait entre la séquence du récepteur à construire et le « template », récepteur dont on utilise la structure comme référence et qui doit idéalement partager un bon taux d'identité avec le récepteur à construire.

Dans notre travail, nous avons aussi utilisé MODELLER pour construire *de novo* les parties manquantes des structures cristallographiques, notamment les boucles intra-cellulaires, afin d'obtenir une séquence continue et permettre au champ de forces de placer correctement les élastiques (intra-moléculaires). Dans le cas d'ECL3, celle-ci a été systématiquement raccourcie à environ six résidus afin de simplement lier les TM5 et TM6, la boucle n'étant pas impliquée lors de la liaison des ligands extra-cellulaires, simplifiant ainsi le calcul. Une exception toutefois pour l'étude sur les peptides issus des protéines-G où MODELLER, également capable de prolonger des segments en hélice, a été utilisé pour la construction des récepteurs de l'Adénosine et conserver la totalité de la boucle ECL3.

#### 2.4.2 Récepteurs de la Neurotensine

La Neurotensine (NT) est un peptide de 13 acide-aminés, et isolé en 1973 par Carraway et Leeman à partir d'extraits d'hypothalamus de bovin [Carraway et Leeman 1973]. La Neurotensine joue une pléiade de rôles physiologiques et pathologiques, incluant la maladie de Parkinson, la schizophrénie, l'analgésie, l'hypotension et l'hypothermie ainsi que dans la progression du cancer du poumon. Elle possède également de nombreux effets cardiovasculaires et des propriétés antipsychotiques [Dobner 2006; Osadchii 2015]. Elle est exprimée par les neurones du système nerveux central et par les cellules intestinales. La Neurotensine agit comme hormone ou neurotransmetteur agoniste en se liant, avec une affinité

équivalente, à trois récepteurs différents : NTR1 et NTR2 qui sont des RCPGs de classe A et NTR3 qui est de classe B1. En se liant au NTR1, le peptide provoque l'hypothermie ou l'hypotension. En revanche la liaison du même peptide au NTR2 entraîne un effet analgésique rendant les peptides sélectifs du NTR2 très intéressants du point de vue thérapeutique. Des études de structure-activité ont permis de montrer que les résidus [8-13] sont responsables de l'activité du peptide [Granier et al. 1982].

Dans notre travail, les récepteurs NTR1 de rat, NTR1 et NTR2 humains, natifs ou mutés ont été construits à partir de la structure cristallographique du récepteur de la Neurotensine-1 du rat (NTR1r) (code PDB : 4BUO) [Egloff et al. 2014]. Cette structure décrit un variant thermostable exprimé par *Escherichia coli* et résolu à 2,75Å, en présence de son peptide agoniste NT[8-13] ( $\text{NH}_3^+$ -Arg-Arg-Pro-Tyr-Ile-Leu- $\text{CO}_2^-$ ). Cette structure montre néanmoins quelques lacunes, notamment au niveau les boucles ICL1 et ICL3 que nous avons construites *de novo* avec MODELLER, la boucle ICL3 ayant été raccourcie pour faire uniquement la jonction entre les TM5 et TM6.

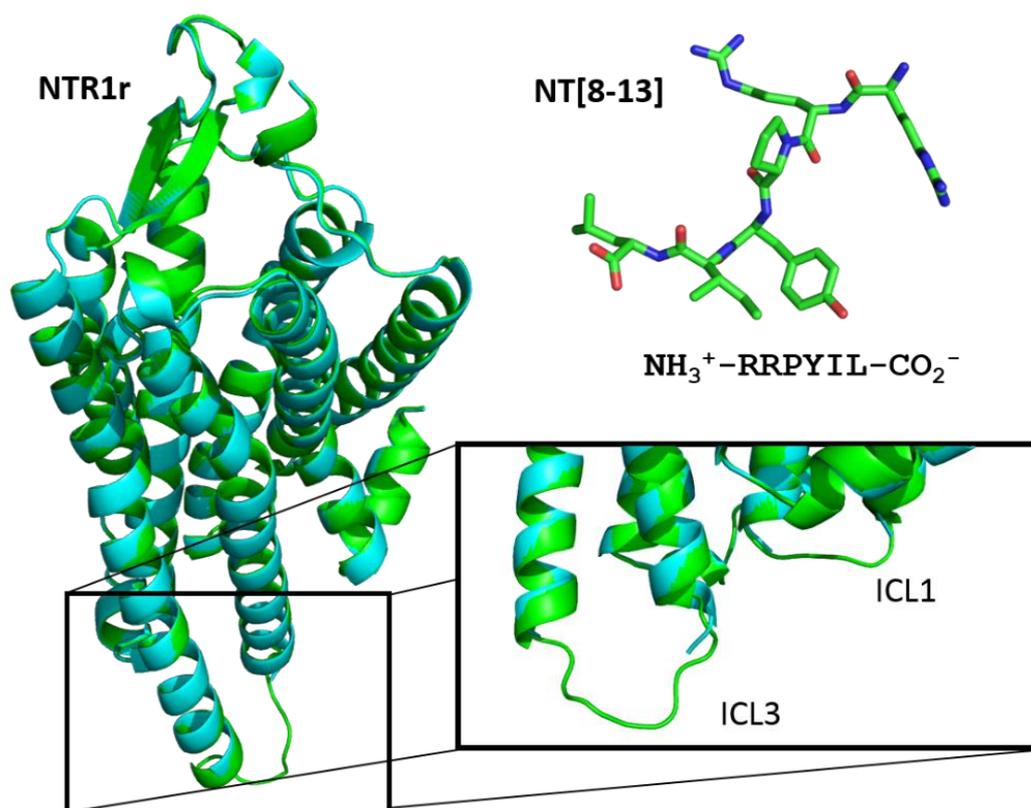


Figure 28 : Modèles de départ de NTR1r et de NT[8-13] après construction en vert, à partir de la structure cristallographique 4BUO en bleu [Egloff et al. 2014].

Les récepteurs humains à la Neurotensine-1 et 2 (NTR1 et NTR2) ont été construits là aussi par homologie en utilisant le récepteur NTR1r (code PDB : 4BUO) [Egloff et al. 2014] comme template. Les séquences utilisées de NTR1 et NTR2 sont issues de la base UniProt (respectivement P30989 et O95665) et présentent un taux d'identité avec le récepteur de rat de 83% et 41%, respectivement.

Lors de nos travaux, les mutations ponctuelles de ces différents récepteurs (NTR1:R212E, NTR2:E179R et NTR2:E179R/E181R) ont été introduites avec PyMOL [Schrödinger], en utilisant comme point de départ les modèles déjà produits pour chacun de ces récepteurs.

Dans toutes nos études, la conformation de départ utilisée pour les peptides natifs ou modifiés est celle décrite dans les structures cristallographiques disponibles, montrant le peptide NT[8-13] dans une conformation quasi-étendue. Les séquences de ces peptides sont présentées dans le **Tableau 5**.

Peptide	Séquence
NT[8-13] (natif)	$\text{NH}_3^+$ -Arg-Arg-Pro- <b>Tyr</b> -Ile-Leu- $\text{CO}_2^-$
11Tyr	$\text{NH}_3^+$ -Lys-Lys-Pro- <b>Tyr</b> -Ile-Leu- $\text{CO}_2^-$
11Lys	$\text{NH}_3^+$ -Lys-Lys-Pro- <b>Lys</b> -Ile-Leu- $\text{CO}_2^-$
11Asp	$\text{NH}_3^+$ -Lys-Lys-Pro- <b>Asp</b> -Ile-Leu- $\text{CO}_2^-$
11Glu	$\text{NH}_3^+$ -Lys-Lys-Pro- <b>Glu</b> -Ile-Leu- $\text{CO}_2^-$

Tableau 5 : Séquences des différents peptides dérivés de la Neurotensine et étudiés lors de notre travail.

### 2.4.3 Récepteur de la Chimiokine C-X-C de type-4

Le récepteur de la Chimiokine C-X-C de type-4 (CXCR4) fait partie d'une famille de vingt récepteurs de Chimiokine, qui sont impliqués dans la migration de nombreux types cellulaires, notamment les leucocytes [Allavena et al. 2000]. Le récepteur CXCR4 est activé exclusivement par la Chimiokine CXCL12 aussi appelé « Stromal cell-Derived Factor » ou SDF-1. Des souris « knock-out » de CXCR4 ou de CXCL12 montrent des défauts dans le développement vasculaire et le développement du système nerveux central et de l'hématopoïèse conduisant à la mort de l'embryon [Zou et al. 1998]. CXCR4 a également été associé à plus de 23 types de cancers, et favorise la formation de métastases et l'angiogenèse tumorale [Chatterjee, Azad, et Nimmagadda 2014]. Un autre rôle clé du CXCR4 est son implication dans l'infection de cellules hôtes par le Virus de l'Immunodéficience Humaine (VIH) en jouant un

rôle de corécepteur [Alkhatib et Berger 2007]. Enfin, le récepteur CXCR4 a été montré comme formant des homo et hétéro-dimères, modulant son activité [B. Wu et al. 2010]. Ce récepteur est un des premiers à avoir été cristallisé sous forme de dimère, présentant une interface de type TM4:TM5.

Nos modèles du récepteur CXCR4 ont été construits directement à partir de sa structure cristallographique (code PDB : 3OE0) [B. Wu et al. 2010]. Ce récepteur a été résolu avec une résolution de 2,9Å en présence d'un lysozyme T4 inséré dans sa boucle ECL3 et de son ligand antagoniste CVX15 ( $\text{NH}_3^+$ -Arg-Arg-**Nal**-Cys-Tyr-Gln-Lys-**DPro**-Pro-Tyr-Arg-**Cit**-Cys-Arg-Gly-**DPro**- $\text{CO}_2^-$ ) dans la poche orthostérique. Une fois encore, les boucles ICL1 et ICL3 ont été raccourcies et construites *de novo* avec MODELLER. Le ligand a lui aussi été construit à partir de sa structure cristallographique. Pour nous éviter le paramétrage de nouveaux résidus dans le champ de forces MARTINI, les acides aminés non naturels présents dans la séquence du peptide ont été modifiés à l'aide de PyMOL. Les D-Proline ont été mutées en L-Proline, la Citrulline remplacée par une Arginine et la 2-Naphthylalanine par un Tryptophane en accord avec des expériences de SAR issues de la littérature [DeMarco et al. 2006].

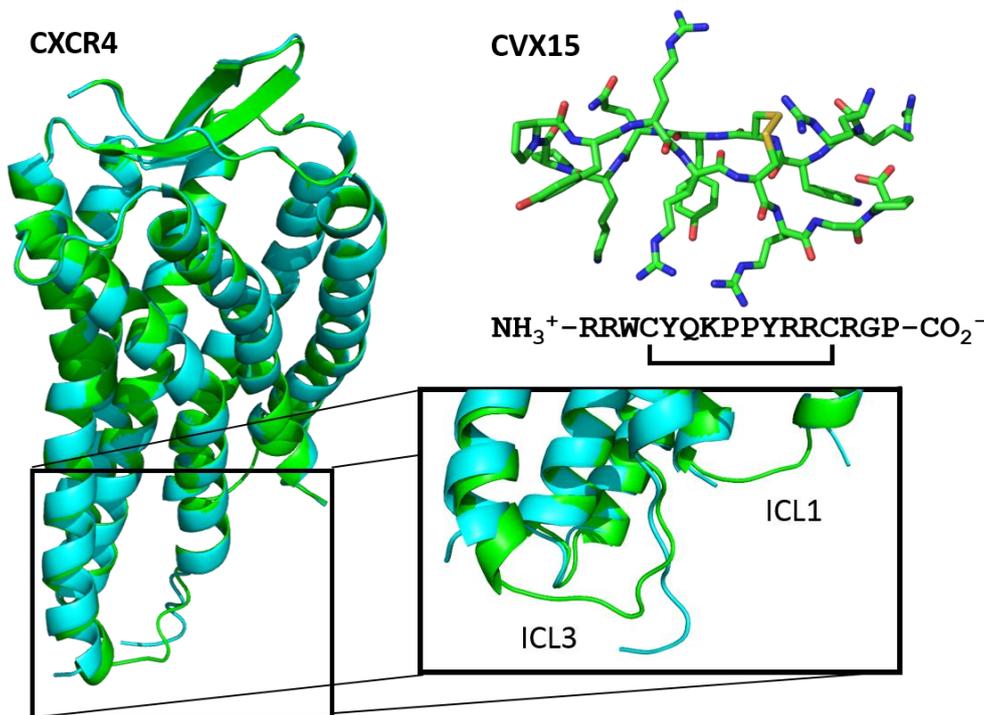


Figure 29 : Modèles de départ de CXCR4 et de CVX15 après construction en vert ; la structure cristallographique utilisée comme point de départ (code PDB : 3OE0) est reportée en bleu [B. Wu et al. 2010].

### 2.4.4 Récepteurs de la Vasopressine

Les récepteurs de la Vasopressine forment un groupe de 4 protéines homologues avec les récepteurs de la Vasopressine de type-1 (V1aR et V1bR), le récepteur de la Vasopressine de type-2 (V2R) et le récepteur de l'Ocytocine (OTR). V1aR participe à la vasoconstriction et à plusieurs fonctions au niveau hépatique. V1bR stimule la sécrétion de l'Hormone Corticotrope et est impliqué dans plusieurs fonctions cognitives [Koshimizu et al. 2012]. V2R régule la réabsorption de l'eau dans les reins [Juul et al. 2014]. OTR quant à lui participe au comportement social et aux fonctions reproductives [Gimpl et Fahrenholz 2001]. Leur ligand naturel, la Vasopressine, Hormone Antidiurétique, ou encore l'Arginine-Vasopressine (AVP), est une hormone de 9 acides aminés. Elle intervient en plus de ces quatre récepteurs dans le cycle cellulaire lors de la mitose et peut jouer le rôle d'inhibiteur de l'apoptose [Miller et al. 2013].

La structure du récepteur V2R n'étant pas connue, nous avons construit un modèle de ce récepteur par homologie en utilisant comme modèle de référence la structure cristallographique du récepteur  $\delta$ -Opioïde de la souris (code PDB : 4EJ4) [Manglik et al. 2012] avec lequel il partage 26% d'identité. La séquence utilisée pour le V2R est issue d'UniProt (P30518). Cette identité de séquence semble assez faible pour utiliser le récepteur  $\delta$ -Opioïde comme template. En effet, pour la construction par homologie des protéines, une conservation d'au moins 35% de la séquence est nécessaire pour garantir la similitude de structure. Toutefois, il a été montré que des protéines membranaires avec une identité de séquence de 20~35% présentaient malgré tout une forte conservation structurale [Olivella et al. 2013].

Le récepteur de la Vasopressine V1aR humain a été construit avec le même protocole en utilisant la même structure de référence. La séquence utilisée pour l'homologie est issue d'UniProt (X5D2B0). Le récepteur V1aR partage 26% d'identité avec le récepteur  $\delta$ -Opioïde, et 40% avec le V2R.

Nous avons réalisé des simulations de dynamique de liaison de deux partenaires du V2R : la Vasopressine et des toxines de venin de serpent présentant un repliement de type « kunitz ».

Le peptide AVP a été construit sous contraintes RMN (collaboration avec Hélène DEMENE, Centre de Biochimie Structurale, Montpellier). Sans trop de surprise, la conformation du peptide obtenue après injection des contraintes est très proche d'une des structures du peptide connue en cristallographie (code PDB : 1JK4 et 1YF4) [C. K. Wu et al. 2001; Syed Ibrahim et Patabhi 2005]. En effet, ce peptide ( $\text{NH}_3^+$ -Cys-Tyr-Phe-Gln-Asn-Cys-

Pro-Arg-Gly-CO<sub>2</sub><sup>-</sup>) présente un pont disulfure maintenant sa conformation qui contraint fortement sa structure.

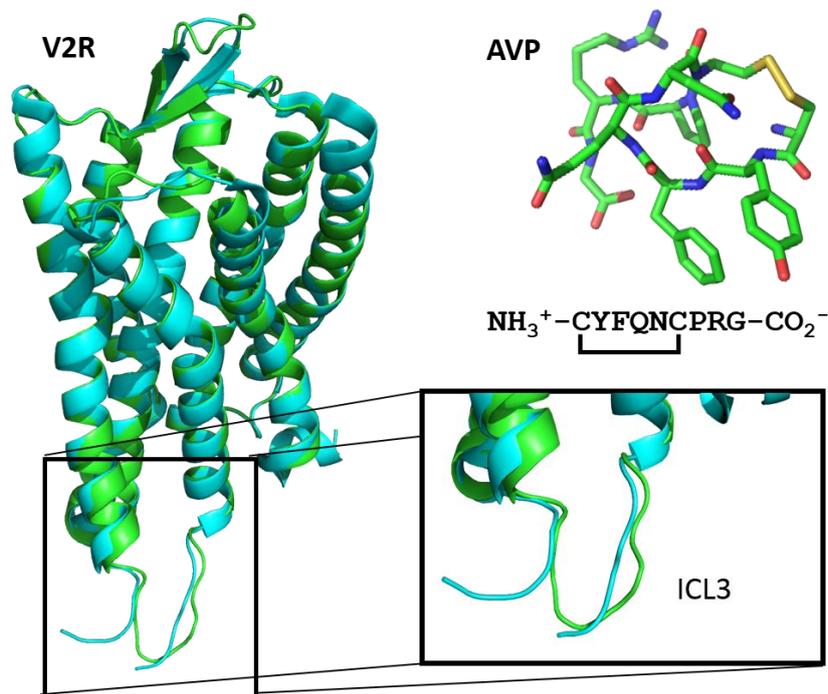


Figure 30 : Modèles de départ de V2R et de l'AVP après construction en vert ; structure cristallographique du récepteur  $\delta$ -Opiöide 4EJ4 en bleu, utilisée comme référence pour la construction [Manglik et al. 2012].

Concernant les toxines (collaboration avec Nicolas GILLES, Laboratoire de Toxinologie Moléculaire et Biotechnologies, CEA Saclay), la Mambaquarétine-1 (MQ-1) issue du serpent Mamba Vert a été construite à partir de sa structure cristallographique (code PDB : 5M4V) [Ciolek et al. 2017]. Elle a été isolée du venin de l'animal par fractionnement multiple et criblage sur le V2R. Cette toxine possède 57 résidus et trois ponts disulfures. Sa structure cristallographique a été résolue par diffraction aux rayons-X avec une résolution de 1,06Å [Droctové 2018].

Par homologie, nous avons construit d'autres toxines homologues de la même famille : E7FL11 produit par la famille de serpent Elapidae et la Dendrotoxine-E (DTX-E) produit par la famille de serpent Mamba Noir, deux toxines testées *in silico* qui présentent respectivement un taux d'identité de 61% et de 63% avec MQ-1 et qui ne se lient pas au V2R [Droctové 2018].

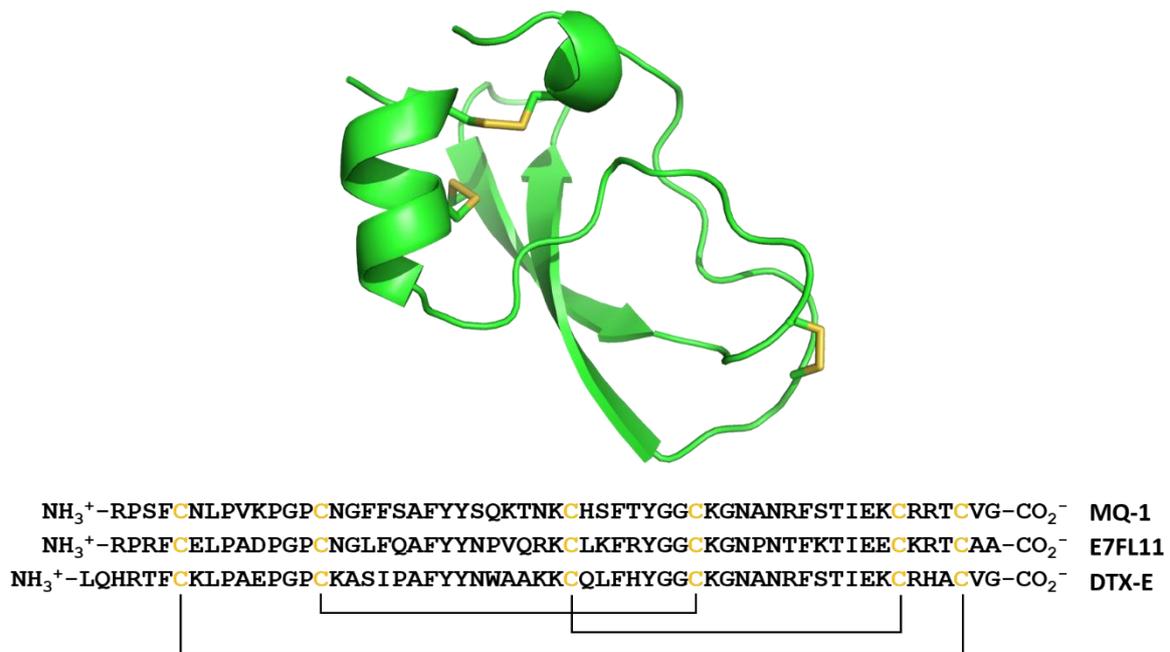


Figure 31 : Structure générale des toxines de type « kunitz » ; alignement des 3 séquences des toxines utilisées au cours de nos calculs.

### 2.4.5 Récepteur de la Ghréline

La Ghréline, est une hormone permettant de stimuler la production de l'Hormone de Croissance (GH) et joue un rôle important dans la régulation de la prise alimentaire. Il s'agit d'une protéine de 28 résidus qui possède la particularité d'être octanoylée sur une sérine à son extrémité N-terminale ( $\text{NH}_3^+$ -GSOFLSPEHQRVQQRKESKKPPAKLQPR-CO<sub>2</sub><sup>-</sup>). Sans cette modification post-traductionnelle, le peptide ne se lie quasiment pas à son récepteur [Bednarek et al. 2000]. Plusieurs études, notamment de RMN et modélisation moléculaire [Großbauer et al. 2010; Staes et al. 2010], ont suggéré un rôle de cet Octanoyl dans l'ancrage à la membrane, pouvant permettre la diffusion du peptide vers son récepteur. Des études de structure-activité ont montré que la Ghréline se liait à son récepteur *via* sa partie N-terminale, les 5 premiers résidus N-terminaux présentant déjà une affinité pour le récepteur de l'ordre de 50nM [Bednarek et al. 2000]. La première fonction identifiée de la Ghréline est d'activer le récepteur sécrétagogue de l'Hormone de Croissance (GHSR-1a) [Kojima et al. 1999]. La GH est une hormone multifonction impliquée dans le métabolisme, la croissance et la reproduction [Devesa, Almengló, et Devesa 2016]. Mais la Ghréline joue aussi un rôle dans l'appétit par la sécrétion du Neuropeptide Y et de l'Oréxine, ainsi que l'inhibition de la Pro-opiomélanocortine [Tschöp, Smiley, et Heiman 2000; Cowley et al. 2003; Toshinai et al. 2003]. La Ghréline est également impliquée dans la stimulation de la fonction cardiaque, la régulation du métabolisme

des carbohydrates *via* un mécanisme d'inhibition de l'insuline, et a une influence sur la prolifération cellulaire [Korbonits et al. 2004].

La structure du récepteur de la Ghréline n'étant pas connue, celle-ci a été construit par homologie en utilisant le récepteur NTR1 de rat comme référence (code PDB : 4BUO) [Egloff et al. 2014] avec lequel il partage 35% d'identité. La séquence utilisée est issue d'UniProt (Q92847). Pour nos simulations, le peptide a été réduit à ses 12 premiers résidus ( $\text{NH}_3^+$ -Gly-Ser-**Ser(oct)**-Phe-Leu-Ser-Pro-Glu-His-Gln-Arg-Val- $\text{CO}_2^-$ ) et construit à partir de résultats de RMN (collaboration avec le groupe d'Alain MILON, Institut de Pharmacologie et de Biologie Structurale, Toulouse) qui ont permis la résolution de la conformation du peptide liée à son récepteur. La méthode TROESY (« TRAnferred nuclear Overhauser Enhancement Spectroscopy ») est la même que celle qui a été utilisée par le même groupe pour résoudre la structure de la Dynorphine liée au récepteur  $\kappa$ -Opioïde [O'Connor et al. 2015]. Cette structure de la Ghréline montre une conformation presque toute étendue du peptide. En ce qui concerne nos calculs, l'information principale issue de cette structure est l'orientation de l'Octanoyl le long de la chaîne principale du peptide. Pour les simulations sans l'Octanoyl en position 3 (Sérine octanoylée), ce résidu a été muté en Sérine *via* PyMOL.

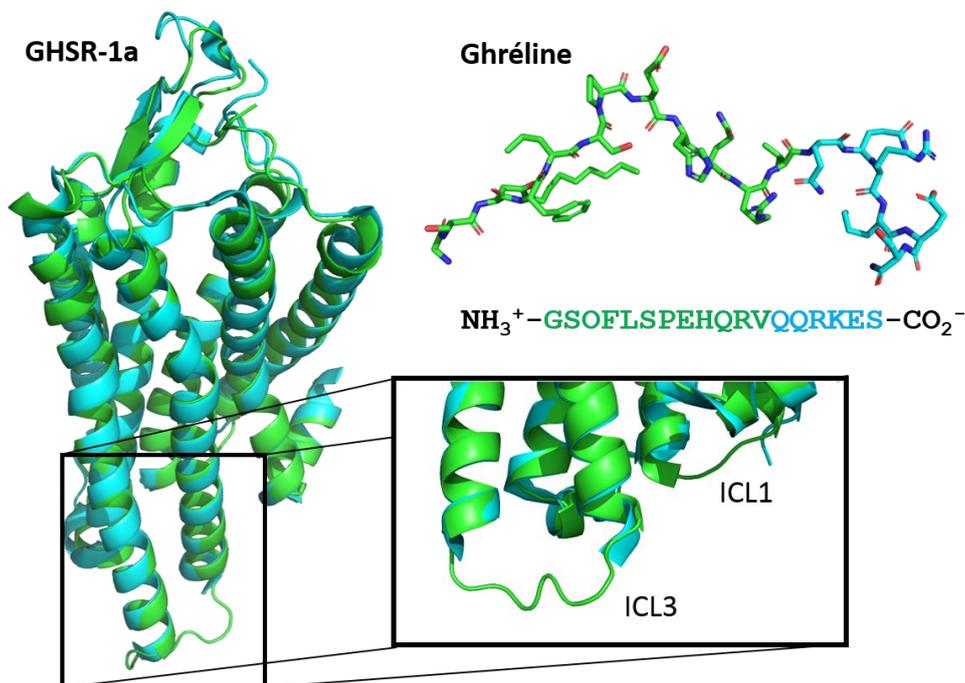


Figure 32 : Modèles de départ de GHSR-1a et de la Ghréline après construction en vert ; structure cristallographique du récepteur Neurotensine-1 (code PDB : 4BUO) en bleu [Egloff et al. 2014].

Ces derniers mois, nous avons également eu l'occasion de tester notre protocole de liaison *in silico* sur un autre peptide se liant au récepteur de la Ghréline, le peptide LEAP2, à l'origine identifié comme peptide antimicrobien exprimé dans le Foie et récemment identifié comme ligand endogène antagoniste du récepteur GHSR1-a [Ge et al. 2018]. Ce peptide de 40 résidus ( $\text{NH}_3^+$ -MTPFWRGVSLRPIGASCRDDSECITRLCRKRRC SL SVAQE-CO<sub>2</sub><sup>-</sup>) jouerait un rôle protecteur lors d'infections virales même si l'on ne connaît pas encore précisément son mécanisme d'action [Krause et al. 2003]. Dans nos simulations, le LEAP2 a été construit à partir de sa structure RMN (code PDB : 2L1Q) [Henriques et al. 2010]. En parallèle de tests menés au laboratoire, ce peptide a été construit entièrement ou bien limité à ses résidus [1-12] ( $\text{NH}_3^+$ -Met-Thr-Pro-Phe-Trp-Arg-Gly-Val-Ser-Leu-Arg-Pro-CO<sub>2</sub><sup>-</sup>) ou [15-40] ( $\text{NH}_3^+$ -ASCRDDSECITRLCRKRRC SL SVAQE-CO<sub>2</sub><sup>-</sup>).

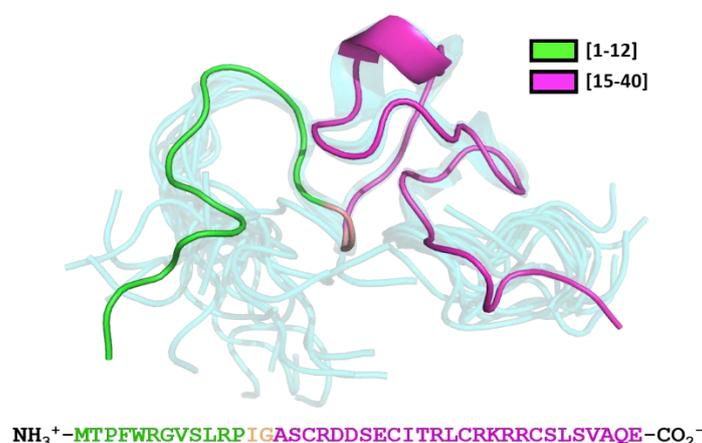


Figure 33 : Structure RMN du peptide LEAP2 (code PDB : 2L1Q) [Henriques et al. 2010] ; la conformation de départ utilisée lors de nos calculs est reportée en couleurs (conformation 19).

### 2.4.6 Récepteur de la Dopamine

La Dopamine est un neurotransmetteur impliqué dans des processus biologiques aussi divers que le système de récompense, l'addiction, de contrôle des mouvements coordonnés, le métabolisme et les sécrétions hormonales [Beaulieu et Gainetdinov 2011]. Un dérèglement du système dopaminergique est impliqué dans différentes maladies comme la dépression, la maladie de Parkinson, les troubles de l'attention, la nausée et la schizophrénie [Barbeau 1970]. La Dopamine interagit avec 5 récepteurs différents parmi lesquels le récepteur de la Dopamine D2 (DRD2) [Missale et al. 1998], capable de dimériser avec le récepteur GHSR-1a [Salahpour et Caron 2012]. Le récepteur DRD2 est la cible principale des traitements antipsychotiques et contre la maladie de Parkinson [Meltzer, Matsubara, et Lee 1989; Roth 2007].

Une structure du récepteur de la Dopamine D2 a été récemment résolue (code PDB : 6CM4) [S. Wang et al. 2018]. Il s'agit d'un variant thermostable exprimé par *Spodoptera frugiperda* et résolu à 2,87Å en présence d'un lysozyme T4 inséré dans la boucle ECL3. Le récepteur a été co-cristallisé avec un antagoniste monoaminergique sélectif : la Rispéridone. Comme pour les autres récepteurs, les parties manquantes incluant les boucles ICL1 et ICL3 ont été construites *de novo* avec MODELLER et la boucle ICL3 raccourcie pour faire uniquement la jonction entre les TM5 et TM6.

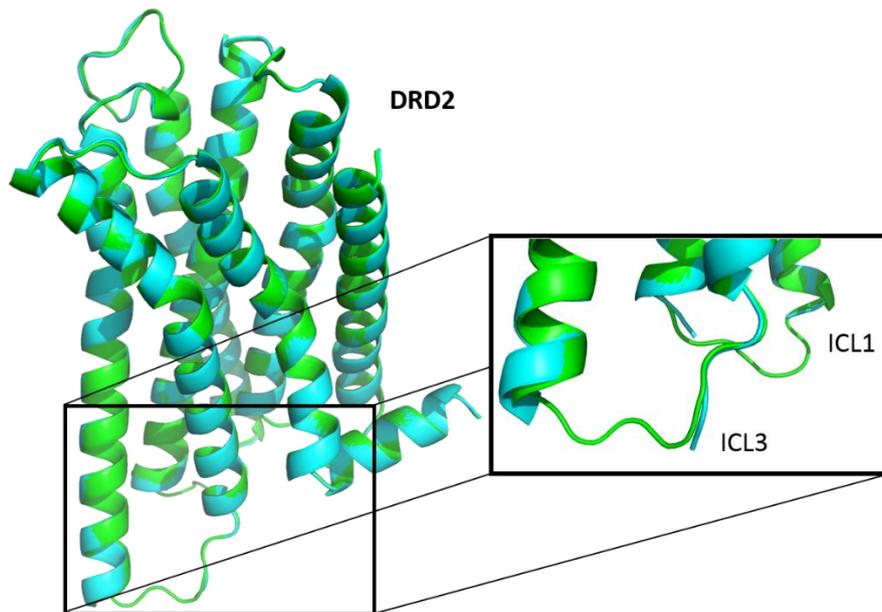


Figure 34 : Modèle du récepteur DRD2 après construction en vert; structure cristallographique (code PDB : 6CM4) en bleu [Wang et al. 2018].

### 2.4.7 Récepteurs de l'Adénosine

L'Adénosine est un nucléoside pouvant se lier à quatre sous-types de récepteurs : A1, A2A, A2B et A3 [Sheth et al. 2014]. Dans le système nerveux central, l'Adénosine est un neuromodulateur (de l'Acétylcholine, du GABA, du Glutamate, ou encore de la Sérotonine), aussi impliqué dans la plasticité synaptique et la neuroprotection lors d'évènements de stress hypotoxique, ischémique et oxydant [Sebastião et Ribeiro 2000; Sperlággh et Vizi 2011; Mendonça et Ribeiro 2001; Wei, Li, et Chen 2011]. Dans le système cardiovasculaire, l'Adénosine permet la vasoconstriction ou la vasodilatation des artères et des veines. Elle régule également la production de cytokine et la prolifération des cellules T. Plus précisément, le récepteur de l'Adénosine A1 (A1R) est une cible privilégiée pour le traitement des douleurs neuropathiques, la fibrillation atriale et les lésions provoquées par une ischémie-réperfusion.

## 2 Matériels et Méthodes

Le récepteur de l'Adénosine A2A (A2AR) quant à lui joue un rôle important dans le système immunitaire, les fonctions cardiaques et le système nerveux central, avec par exemple un rôle dans la libération du Glutamate [Fredholm et al. 2011].

Le modèle utilisé pour la construction du récepteur A1R est sa structure résolue par Cryo-EM (code PDB : 6D9H) [Draper-Joyce et al. 2018], montrant son interaction avec la protéine-G hétéro-trimérique  $G_{i2}$ . Le récepteur résolu est une chimère avec le récepteur Muscarinique M4 exprimé par *Trichoplusia ni* et résolu à 3,6Å en présence de son ligand endogène l'Adénosine. En utilisant MODELLER, les TM5, TM6 ont été prolongés d'un tour d'hélice en accord avec des structures cristallographiques de A2AR (codes PDB : 3RFM, 3UZA et 3VGA) [Dore et al. 2011; Congreve et al. 2012; Hino et al. 2012] et l'hélice 8 de trois tours en accord avec une structure cristallographique de A2AR (code PDB : 2YDV) [Lebon et al. 2011]. La boucle ICL3 a ensuite été construite *de novo*.

L'autre récepteur de la sous-famille utilisé dans nos simulations est le récepteur A2AR dont la structure a été résolue par diffraction aux rayons-X (code PDB : 5G53) [Carpenter et al. 2016], en présence de la protéine mini- $G_s$  hétéro-trimérique. Il est exprimé par *Trichoplusia ni* et *Escherichia coli* et résolu à 4,11Å. De la même manière, MODELLER est utilisé pour prolonger d'un tour d'hélice les TM5 et TM6 et de deux tours l'hélice 8, selon les structures cristallographiques de A2AR citées précédemment. La boucle ICL3 a ensuite été construite *de novo*.

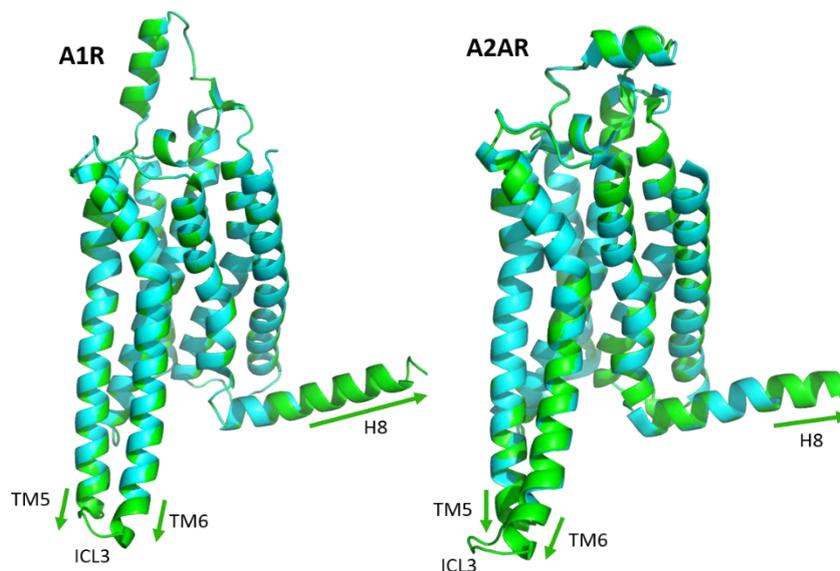


Figure 35 : Modèles de départ des récepteurs A1R et A2AR après construction en vert ; structures résolues (codes PDB : 6D9H et 5G53, respectivement) en bleu [Draper-Joyce et al. 2018; Carpenter et al. 2016].

### 2.4.8 Peptides C-terminaux des Protéines-G

En 2016, il a été montré que seule, la partie C-terminale en hélice de la sous-unité  $\alpha$  de la protéine-G pouvait permettre de conserver sa sélectivité vis-à-vis de différents RCPGs [Semack et al. 2016]. Dans notre travail, nous avons essayé de reproduire cette observation par simulations CG-REMD. Pour ce faire, les peptides C-terminaux de  $G_{\alpha i}$ ,  $G_{\alpha s}$ , et  $G_{\alpha q}$  ont été construits à partir de la structure de  $G_{\alpha s}$  complexée au récepteur A2AR humain (code PDB : 5G53) [Carpenter et al. 2016] et présentant une résolution de 3,4Å. Les 26 derniers résidus de  $G_{\alpha s}$  ont donc été extraits pour servir de template pour les peptides C-terminaux  $G_{\alpha i}$  (27% d'identité), et  $G_{\alpha q}$  (38% d'identité).

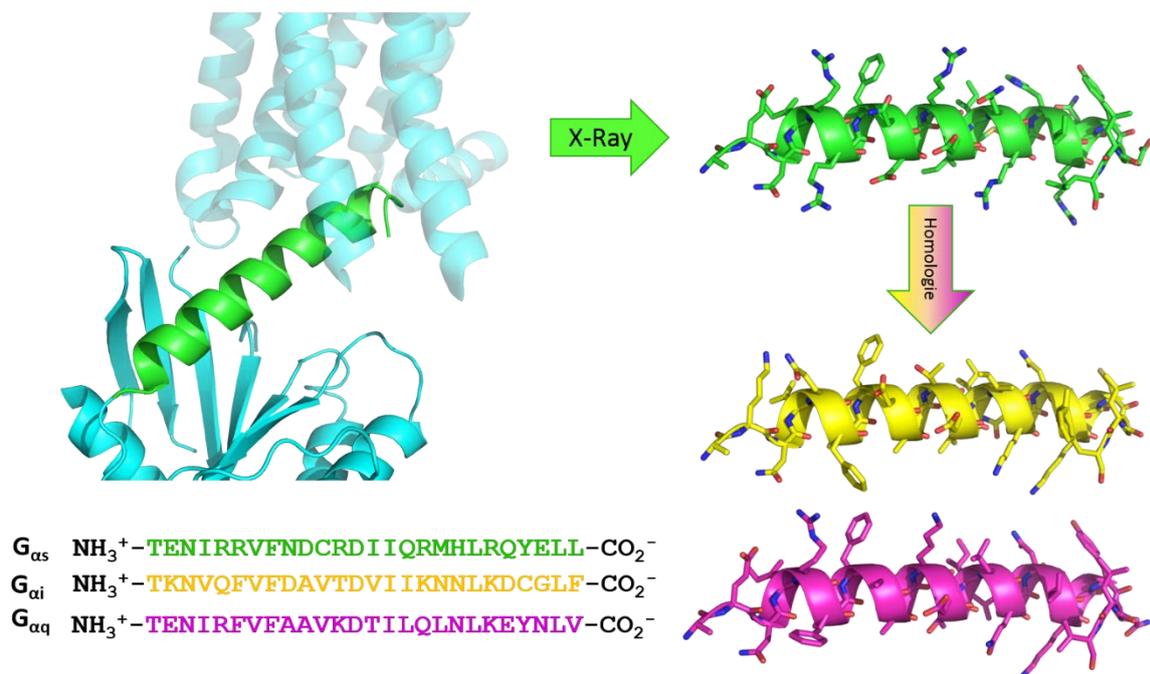


Figure 36 : Construction du peptide C-terminal  $G_{\alpha s}$  issu de la structure cristallographique de A2AR: $G_{\alpha s}$  (code PDB : 5G53) [Carpenter et al. 2016] et modèles construits par homologie des peptides C-terminaux de  $G_{\alpha i}$  et  $G_{\alpha q}$ .

### 2.4.9 Construction de « systèmes doubles » pour les simulations CG-REMD

Lors de simulations de dynamique moléculaire de systèmes membranaires, l'emploi de conditions périodiques aux limites (« Periodic Boundary Conditions » ou PBC) a pour conséquence une communication des compartiments extra et intra-cellulaires. Aussi, au contraire du récepteur dont les mouvements sont limités au plan  $(x,y)$  de la bicouche lipidique, le ligand va pouvoir diffuser librement dans la boîte de simulation, et ainsi interagir indifféremment avec les faces extra ou intra-cellulaires du récepteur. Ceci n'est pas souhaitable car bien entendu irréaliste du point de vue biologique, mais aussi parce que cela constitue une perte de temps au niveau calcul, altérant l'échantillonnage des conformations du complexe peptide:récepteur. Afin de remédier à ce problème, et après discussion avec Romain GAUTIER et Stefano VANNI, nous avons choisi dans toutes nos simulations de travailler sur des systèmes dupliqués sur l'axe  $z$  et dans lesquels deux récepteurs identiques se font face et délimitent un compartiment « extra-cellulaire » où seront piégés le ou les ligands.

La construction de tels systèmes se fait en plusieurs étapes. Le récepteur d'intérêt est tout d'abord converti en gros-grains en utilisant le champs de force MARTINI [Marrink, de Vries, et Mark 2004] ; pour ce faire nous avons utilisé le serveur CHARMM-GUI [CHARMM-GUI; Jo et al. 2008; Qi et al. 2015] (**Figure 37A**). L'axe principal du récepteur est ensuite utilisé pour le placer dans un modèle de membrane 100% POPC en utilisant là aussi les outils disponibles dans CHARMM-GUI [E. L. Wu et al. 2014]. Le système est complété de chaque côté de la membrane par des molécules d'eau et des ions afin de neutraliser la charge globale du système. Dans toutes nos simulations nous avons choisi une même taille de membrane de  $100 \times 100 \text{Å}$  sur les axes  $x$  et  $y$ . En ce qui concerne l'axe  $z$ , une couche d'eau de  $13 \text{Å}$  a été ajoutée de chaque côté du récepteur.

Dans une seconde étape (**Figure 37B**), les sphères d'eau et d'ions sont retirées de ce système simple et la taille de la boîte est augmentée manuellement de  $20 \text{Å}$  sur l'axe  $z$ . Le système obtenu est alors dupliqué, tourné de  $180^\circ$  autour de l'axe  $x$  et translaté, pour générer un système double où les deux récepteurs se font face. Par la suite, le ligand est rajouté dans le compartiment extra-cellulaire ainsi formé, avant finalisation de la construction du système par ajout de molécules d'eau et d'ions.

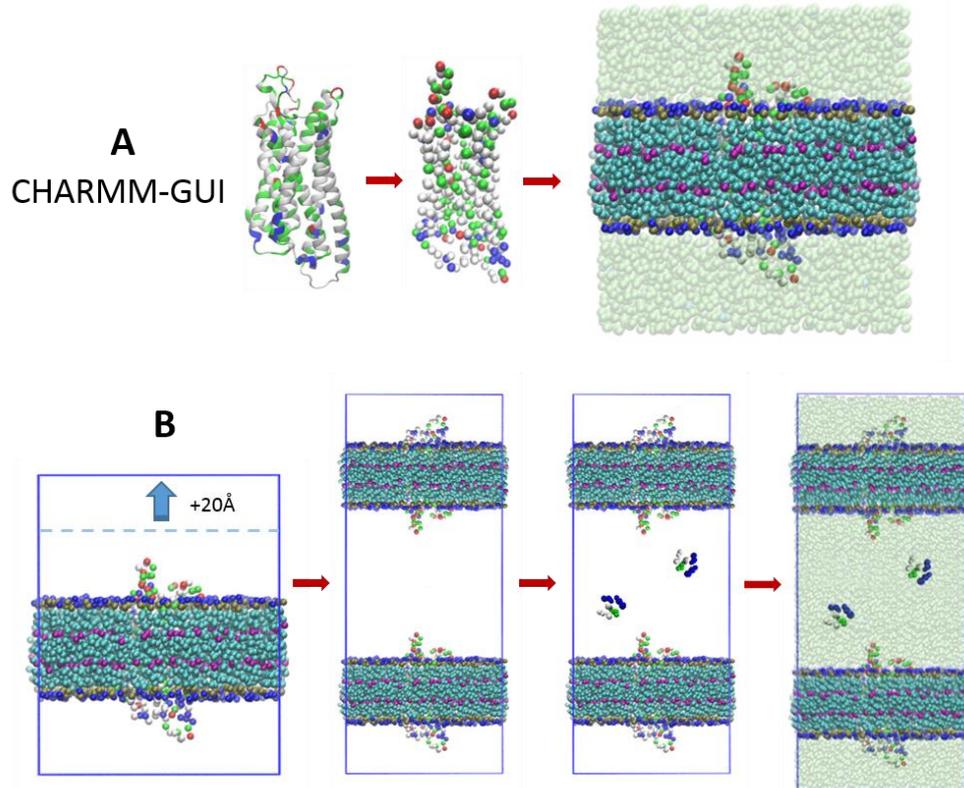


Figure 37 : Schématisation du processus de création des « systèmes doubles » ; A) Via le serveur CHARMM-GUI ([charmm-gui.org](http://charmm-gui.org)) le récepteur est converti d'un modèle tout-atomes à un modèle gros-grains avant d'être placé dans une membrane POPC entourée d'eau et d'ions ; B) L'eau et les ions sont ensuite retirés, la boîte de simulation agrandie avant de dupliquer le système, de rajouter le ou les peptides dans le compartiment extra-cellulaire, l'eau et les ions sont à nouveau ajoutés

Le système ainsi obtenu est soumis à des étapes de minimisation d'énergie et équilibré à toutes les températures requises pour les simulations REMD (**Figure 38**).

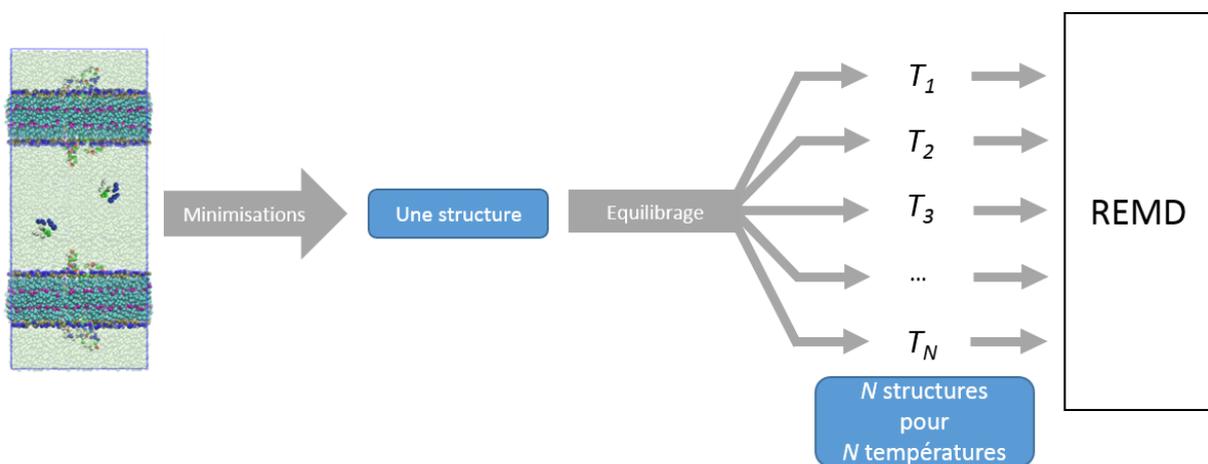


Figure 38 : Schématisation du processus de préparation des systèmes pour les simulations de dynamique moléculaire de « Replica-Exchange ».

### 2.5 Moyens informatiques utilisés

#### 2.5.1 Clusters de calcul

Durant ma thèse, j'ai travaillé sur une station de travail Linux équipée de 2 processeurs Intel Xeon Gold 5118 (24 cœurs) cadencés à 2,30 GHz et 64 Go de mémoire vive et d'une carte GPU NVIDIA GTX1080Ti. Sur cette station j'ai réalisé la préparation des différents systèmes moléculaires, des calculs tests de dynamique moléculaire, mais aussi la plupart des analyses statistiques et figures présentées dans ce manuscrit.

Au laboratoire, nous disposons d'un cluster de calcul qui nous a permis de réaliser notamment les phases d'équilibrage et des calculs tests de REMD sur les différents systèmes. Ce cluster est composé d'un nœud maître de 32 cœurs (processeurs Intel Xeon CPU E5-2650 cadencés à 2,00 GHz) et 16 Go de mémoire vive, de sept nœuds de calcul 32 cœurs Intel Xeon CPU E5-2660 cadencés à 2,20 GHz et 32 Go de mémoire vive et de cinq nœuds 40 cœurs Intel Xeon CPU E5-2660 v2 cadencés à 2,20 GHz et 32 Go de mémoire vive.

La plupart des calculs CG-REMD ont été déportés sur des machines externes au laboratoire, *via* des demandes de support au niveau régional, national et européen. Au niveau régional nous avons eu accès à la machine MUSE de MESO@LR (anciennement HPC@LR). Il s'agit d'une plateforme de calcul de l'université de Montpellier et hébergée au CINES. Le cluster MUSE est composé de nœuds de 28 ou de 80 cœurs, avec 128 ou 1 To de mémoire vive. Sur ce cluster, un total de 3 millions d'heures CPU ont permis de réaliser les premiers tests et la validation du protocole décrit dans cette thèse.

Nos calculs portant sur l'étude de la sélectivité de peptides dérivés de la Neurotensine ont été réalisés sur le cluster EOS du centre de calcul CALMIP localisé à Toulouse grâce à l'attribution de 1,4 million d'heures de calcul. Ce cluster est composé de nœuds Intel IVYBRIDGE 10-cœurs cadencés à 2,80 GHz avec 64Go de mémoire vive. Ces simulations ont été poursuivies et terminées grâce à l'obtention d'heures au CINES sur le cluster OCCIGEN2 *via* les appels nationaux pilotés par GENCI pour un total de 8,6 millions d'heures. Le cluster OCCIGEN2 est composé de nœuds Intel Xeon E5-2690 v3 12 cœurs cadencés à 2,60 GHz et de nœuds Intel Xeon E5-2690 v4 14-cœurs cadencés à 2,60 GHz.

Enfin, nous avons récemment obtenu un accès au centre de calcul du TGCC *via* un appel à projet PRACE au niveau européen. Grâce à ce projet, 5 millions d'heures de calcul nous ont

été attribuées sur la machine Curie nous permettant notamment de simuler la dynamique de dimères de RCPGs et 20 millions d'heures sur la machine IRENE (2018-2019) qui nous ont déjà permis de réaliser le travail sur le récepteur Ghréline, mais aussi sur l'étude de la liaison des peptides C-terminaux des protéines-G. Cette dernière machine est composée de nœuds Intel Skylake 8168 48-cœurs cadencés à 2,7 GHz avec 192 Go de mémoire vive.

### 2.5.2 Logiciels utilisés

Afin de préparer les systèmes pour les dynamiques moléculaires, les structures issues de la PDB possédaient des parties non résolues. De plus certains récepteurs n'ont pas encore de structure et ont été construits par homologie. Dans les deux cas, le logiciel MODELLER [Sali et Blundell 1993] a été utilisé. Les éventuelles mutations ont été introduites avec l'outil mutagenèse de PyMOL [Schrödinger]. Les modèles ont ensuite été préparés avec le champ de forces gros-grains de MARTINI [Marrink et al. 2007] en utilisant le serveur CHARMM-GUI [Jo et al. 2008] pour les étapes initiales de construction, d'insertion du récepteur dans une membrane, d'ajout de l'eau et des ions.

Toutes nos simulations ont ensuite été réalisées avec le logiciel GROMACS [Berendsen, van der Spoel, et van Drunen 1995] d'abord dans sa version 4.6.7 [Páll et al. 2014] pour les études de validation de la méthode, de prédiction de la sélectivité et l'application du protocole sur la Mambaquarétine, puis dans sa version 5.1.2 ou 5.1.4 [Abraham et al. 2015] pour les études des peptides se liant au récepteur de la Ghréline, des peptides C-terminaux des protéines-G et du dimère DRD2:GHSR-1a.

La visualisation et analyses des trajectoires obtenues (calculs de RMSD, mesures de distances, calcul de pourcentages de liaison, etc...) ont été effectuées avec VMD [Humphrey, Dalke, et Schulten 1996]. Le logiciel PyMOL a également été utilisé pour la visualisation de nos modèles (structures initiales et conformations issues de nos calculs). Enfin le logiciel R [R Development Core Team (2008)] a été utilisé pour réaliser un certain nombre d'analyses statistiques et la génération de graphiques.

# 2.6 Synthèse et test biologiques de peptides inhibiteurs des interactions protéine-G:protéine-G

## 2.6.1 Synthèse

Pour la synthèse des peptides, une stratégie sur support solide (« Solid-Phase Peptide Synthesis » ou SPPS) a été choisie. Dans cette méthodologie, la synthèse est effectuée avec le peptide accroché par une liaison covalente à une bille de résine par son extrémité C-terminale. En effet, la synthèse peptidique va de l'extrémité C-terminale à l'extrémité N-terminale du peptide afin d'éviter des problèmes de racémisation. De plus, les chaînes latérales sont protégées par des groupements résistants aux conditions de couplage (protection orthogonale). L'accrochage à une bille de taille importante permet d'éliminer facilement les phases aqueuses des réactions et des lavages par simple filtration. On peut ainsi travailler avec un excès de réactifs et gagner du temps sur les différentes étapes. Une fois les acides aminés ajoutés un par un, on décroche simplement le peptide de la résine qui est éliminée par filtration.

### 2.6.1.a Préparation des résines

La première étape de la synthèse des différents peptides potentiellement inhibiteurs du dimère DRD2:GHSR-1a est l'accrochage du premier acide aminé sur la résine (synthèse peptidique du C-terminal vers le N-terminal). Il s'agit d'une résine 2-chloro chlorotrytyle (100-200mesh) (Clt). Un équivalent de résine est gonflé dans du N,N-diméthylformamide (DMF) avant d'ajouter 1,2 équivalents d'acide aminé (AA) avec l'amine protégé par un groupement fluorénylméthoxycarbonyle (Fmoc), et 4 équivalents de N,N-diisopropyléthylamine (DIEA) en solution dans du DMF. Le mélange est laissé sous agitation pendant 2H à température ambiante (TA).

La résine est lavée successivement par 1x un mélange dichlorométhane (DCM)/méthanol (MeOH)/DIEA 17/2/1, 2x DMF, 2x DCM et 2x ether diéthylique (Et<sub>2</sub>O) avant d'être placée sous vide. La charge de la résine est mesurée en prenant  $m = 10\text{mg}$  de résine, traitée par 1mL de pipéridine (Pip):20%/DMF sous agitation à TA pendant 30min pour déprotéger les fonctions amines. Un volume de 100 $\mu\text{L}$  de cette solution est diluée  $D = 100$  fois dans du DMF. Avec un spectrophotomètre Genesys 10 (Thermo Electron Corporation) avec une cuve en quartz de  $l = 1\text{cm}$ , l'absorbance  $A$  est mesurée, et dont la valeur est directement reliée à la charge de la résine

par l'**Equation 13** où  $\varepsilon$  est le coefficient d'extinction molaire. La mesure est réalisée à  $\lambda = 289,8\text{nm}$  avec  $\varepsilon = 6089\text{L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$ .

$$A = \varepsilon * c * l \leftrightarrow \text{Charge}(\text{mmol}\cdot\text{g}^{-1}) = \frac{A * D}{\varepsilon * l * m} * 1000 = \frac{A}{m} * 16421 \quad (13)$$

### 2.6.1.b Robot synthétiseur

La suite de la séquence est synthétisée avec le synthétiseur micro-ondes automatisé Liberty Blue (CEM). Tous les peptides ont été synthétisés selon la même méthode avec l'hexafluorophosphate de (diméthylamino)-N,N-diméthyl(3H-[1,2,3]triazolo[4,5-b]pyridin-3-yloxy)méthaniminium (HATU) comme activateur, DIEA en tant que base et DMF comme solvant. Les réactions sont réalisées dans le réacteur de l'appareil, chauffées par micro-ondes, sous une pression de 20psi et sous azote.

Tout d'abord la résine est gonflée par 10mL de solvant. Ensuite chaque acide aminé est ajouté selon le même protocole : tout d'abord deux étapes de déprotection du groupement Fmoc de 5 et 10min respectivement par un mélange Pip:20%/DMF, la résine est ensuite lavée trois fois par le solvant, et enfin deux étapes de couplage de 25min en présence de 4 équivalents de Fmoc-AA, 4 équivalents de HATU et 8 équivalents de DIEA. Finalement, le dernier acide aminé de la séquence subit une déprotection finale selon le même protocole de déprotection que décrit plus haut. Enfin la résine est sortie du robot et placée sous vide avant d'être lavée au DCM, puis à l'Et<sub>2</sub>O.

### 2.6.1.c Décrochage et déprotection des peptides linéaires

Pour les peptides linéaires, la dernière étape est le décrochage du peptide de la résine et déprotection des chaînes latérales qui ont lieu lors de la même étape. La résine est transvasée dans un carrousel avec du DCM. Un volume de 15mL d'un mélange acide trifluoroacétique (TFA)/triisopropylamine (TIS):3%/H<sub>2</sub>O:3%, ou TFA/TIS:6% dans le cas du peptide comportant une Méthionine afin de ne pas l'oxyder, est ajouté. Le mélange est laissé sous agitation à 45°C pendant 45min.

Le mélange est passé sur verre fritté pour retirer la résine, puis lavé à l'Et<sub>2</sub>O. Les solvants, sont évaporés, puis le liquide restant est transvasé dans un tube qui est rempli à l'Et<sub>2</sub>O à froid. Le mélange est alors agité et régulièrement dégazé, avant d'être laissé à TA pendant 2H. Le mélange est passé à la centrifugeuse pendant 10min avec une agitation de 500G. Le surnageant

est retiré avant de laver le solide restant à l'Et<sub>2</sub>O à froid. L'étape de la centrifugeuse est réitérée encore deux fois. Le solide est solubilisé dans un minimum d'acétonitrile (AcN) et d'H<sub>2</sub>O avant d'être congelé à l'azote liquide et lyophilisé.

### 2.6.1.d Cyclisation, décrochage et déprotection du peptide cyclique

Pour le peptide à cycliser, il faut décrocher le peptide de la résine sans retirer les protections latérales de celui-ci. Pour se faire, la résine est mise en présence de 10mL de DCM/TFA:1%/TIS:3% et laissée sous agitation à TA pendant 5min. La phase liquide est mélangée à 50mL de DCM/DIEA:1%, et cette étape est répétée encore trois fois. La phase liquide est ensuite évaporée et le solide solubilisé dans AcN puis congelé à l'azote liquide avant d'être lyophilisé.

Le peptide va pouvoir être cyclisé par une liaison amine intra-moléculaire. Le peptide est solubilisé dans un minimum de DMF et laissé sous agitation à TA. 3 équivalents de DIEA sont ajoutés. Après 2min, 1 équivalent de HATU solubilisé dans un minimum de DMF est ajouté. La réaction est laissée sous agitation à TA pendant 4H. Les solvants sont ensuite évaporés, le produit est solubilisé dans AcN et lyophilisé.

Enfin, le peptide cyclisé est placé dans 20mL d'un mélange TFA/TIS:2,5%/H<sub>2</sub>O:2,5% afin de retirer les groupements protecteurs des chaînes latérales. Le mélange est laissé sous agitation à TA pendant 1H. Le brut réactionnel est alors évaporé et le peptide précipité par Et<sub>2</sub>O à froid. Le mélange est passé à la centrifugeuse pendant 10min avec une agitation de 500G. Le surnageant est retiré avant de laver le solide restant à l'Et<sub>2</sub>O à froid. L'étape de la centrifugeuse est réitérée encore deux fois. Le solide est solubilisé dans un minimum d'AcN et d'H<sub>2</sub>O avant d'être congelé là encore sous azote liquide puis lyophilisé.

### 2.6.1.e Purification et analyses

Une fois les peptides synthétisés, ils sont purifiés par RP-HPLC sur une PLC 2050 (Gilson) équipé d'une colonne DeltaPak C18 (Waters) (15µm, 300Å, 25x100mm, débit de 45mL.min<sup>-1</sup>). En LC-MS, on identifie au moment de l'élution du composé le gradient AcN/H<sub>2</sub>O nécessaire à sa sortie et qui sera utilisé pour la purification par HPLC. La méthode de purification suit ainsi le gradient AcN/H<sub>2</sub>O présenté dans la **Figure 39**. Les fractions comportant le produit sont ensuite rassemblées, évaporées et lyophilisées afin d'obtenir le peptide purifié.

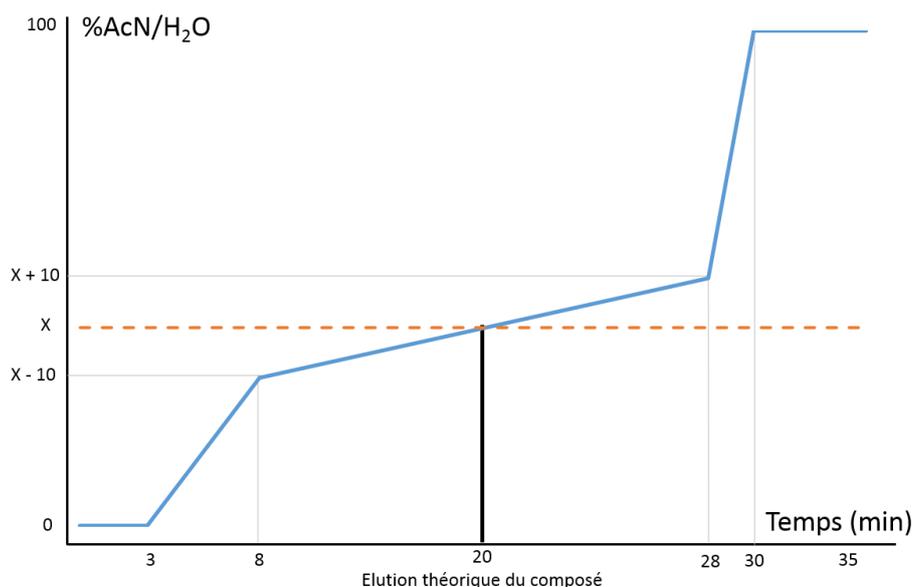


Figure 39 : Généralisation du gradient d'éluant utilisé pour la purification par HPLC d'un composé avec une élution en LC-MS à un gradient d'élution X.

Le long de la synthèse, les suivis réactionnels, l'identification des fractions en purification et les analyses finales ont été réalisés par LC-MS. Le système LC-MS comprend une HPLC Alliance 2695 (Waters) couplée à un spectromètre Micromass ZQ (Waters) (ionisation par électrospray, ESI+). La colonne utilisée est une Chromolith HighResolution RP-18e (Merk) ( $15\mu\text{m}$ ,  $15\text{nm}$ - $1,15\mu\text{m}$ ,  $4,6\times 25\text{mm}$ , débit de  $3,0\text{mL}\cdot\text{min}^{-1}$ ). Pour la méthode de séparation, un gradient de 100%  $\text{H}_2\text{O}$ /Acide méthanoïque ( $\text{HCO}_2\text{H}$ ):1% à 100%  $\text{AcN}/\text{HCO}_2\text{H}$ :1% avec un débit de  $3\text{mL}\cdot\text{min}^{-1}$  pour un temps d'élution de 2,5min ont été utilisés. La détection UV a été réalisée à 214nm. Le spectre de masse par ESI a été réalisé avec un débit de  $200\mu\text{L}\cdot\text{min}^{-1}$  de solvant. L'azote a été utilisé comme gaz nébuliseur et gaz séchant.

## 2.6.2 Tests biologiques

### 2.6.2.a Préparation des systèmes

Pour mesurer la capacité des peptides à inhiber la formation du complexe  $\text{DRD2}/\text{G}_i:\text{GHSR-1a}/\text{G}_q$ , un test biologique se basant sur la mesure de l'activité de  $\text{G}_{\alpha i}$  a été mis en place. Cette activité est mesurée par la vitesse de liaison du GTP à  $\text{G}_{\alpha i}$ . En effet, cette liaison entraîne un changement conformationnel de la sous unité  $\text{G}_\alpha$  qui peut être suivi par la variation de l'émission de fluorescence des tryptophanes de cette sous-unité. Or la vitesse de liaison est augmentée dans le complexe par la présence de  $\text{G}_q$  [Damian et al. 2018]. Pour éviter des

## 2 Matériels et Méthodes

interférences liées à la présence de  $G_{\alpha q}$ ,  $G_{\alpha i}$  est marqué par un 5-Hydroxytryptophane (5-HW) qui peut être excité sélectivement [Mesnier et Banères 2004]. La schématisation de l'effet de  $G_q$  et d'un peptide inhibiteur sur l'activité de  $G_i$  est présentée dans la **Figure 40**.

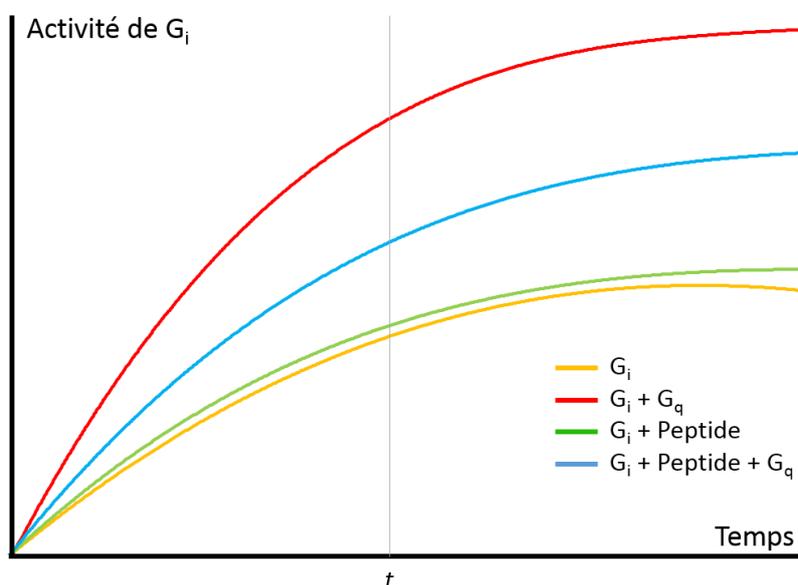


Figure 40 : Schématisation de l'activité de  $G_i$ , à un temps  $t$  après l'ajout du GTP, seule, en présence de  $G_q$ , d'un peptide inhibiteur (ici partiellement) de la formation du dimère, ou des deux.

Le test est réalisé avec le  $GTP\gamma S$ , un GTP dans lequel le dernier phosphate est remplacé par un soufre afin d'empêcher son hydrolyse par  $G_{\alpha i}$  et ainsi permettre la stabilisation du signal de fluorescence [Harrison et Traynor 2003]. Le dimère de récepteur DRD2:GHSR-1a est purifié en solubilisant des vésicules lipidiques contenant les récepteurs avec du copolymère acide maléique/styrène (SMA) [Logez et al. 2016; Dörr et al. 2016]. On obtient alors des disques membranaires (SMAPLs) de 20nm de diamètre où les dimères sont encapsulés. Enfin, on sépare les hétéro-dimères désirés des homo-dimères par chromatographie en tandem (**Figure 41**).

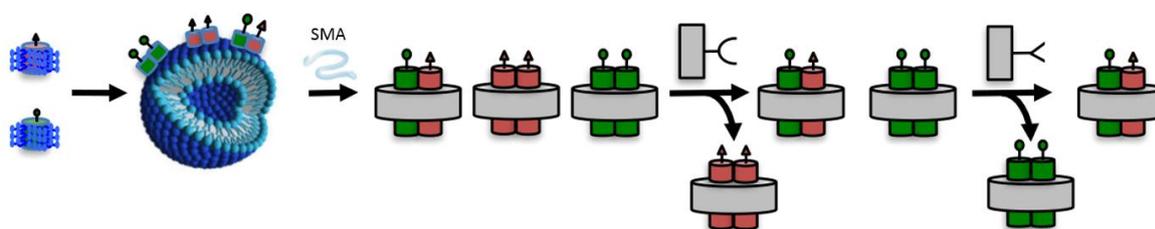


Figure 41 : Schématisation du processus de production et purification de l'hétéro-dimère DRD2:GHSR-1a dans un nanodisque membranaire ; d'après [Damian et al. 2018].

Le récepteur GHSR-1a est exprimé par *Escherichia coli* [Damian et al. 2012]. Le récepteur DRD2 est exprimé par *Pichia pastoris* [André et al. 2006]. Les sous-unités  $G_{\alpha q}$  et  $G_{\beta\gamma}$  sont exprimées par *sf9* [Kozasa 2004]. La sous-unité  $G_{\alpha i}$  est exprimée par *Escherichia coli* [Eps et al. 2011] et 5-HW est introduit durant l'expression bactérienne [Mesnier et Banères 2004]. La Dopamine et le GTP $\gamma$ S sont des produits commerciaux.

### 2.6.2.b Mise en place du test

Pour les tests, les peptides sont dilués à 1mM dans l'eau. Dans un eppendorf sont ajoutés les récepteurs (R), la Dopamine (Dopa), la protéine  $G_i$ . En fonction de l'échantillon, la protéine  $G_q$  et/ou le peptide sont ajoutés. L'eppendorf est complété à un volume de 49 $\mu$ L par une solution tampon Trishydroxyméthylaminométhane (TRIS):50nM/Chlorure de sodium (NaCl):150nM pour un pH de 8. Les différents types d'échantillons, de volumes d'injection et de concentrations sont présentés dans le **Tableau 6**.

Echantillon	R ( $\mu$ L)	Dopa ( $\mu$ L)	$G_i$ ( $\mu$ L)	$G_q$ ( $\mu$ L)	Peptide ( $\mu$ L)	Tampon ( $\mu$ L)
$G_i$	10	1	5	0	0	33
$G_i$ +Peptide	10	1	5	0	5	28
$G_i$ + $G_q$	10	1	5	5	0	28
$G_i$ + $G_q$ +Peptide	10	1	5	5	5	23
Concentration	20nM	10 $\mu$ M	100nM	100nM	100 $\mu$ M	/

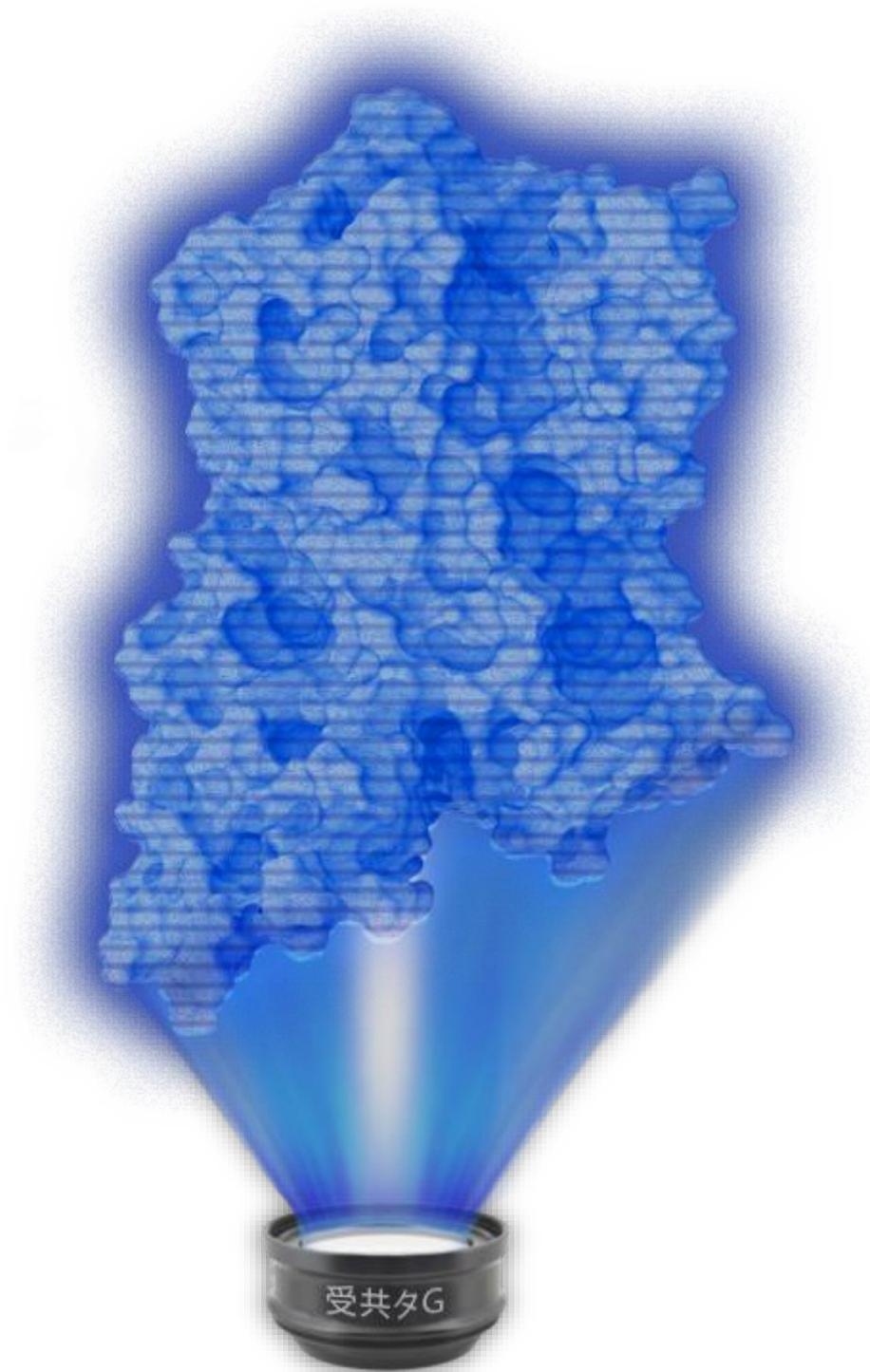
Tableau 6 : Composition des différents échantillons et concentration de chaque composé (volume final 50 $\mu$ L).

Les eppendorfs sont placés dans un bain-marie à 20°C. L'ajout du GTP $\gamma$ S (1 $\mu$ L) initie alors les changements conformationnels et de fluorescence de  $G_i$ . La réaction est stoppée à 10min en plongeant le tube dans la glace. La fluorescence est ensuite mesurée sur un spectrofluorimètre Cary Eclipse (Varian) à des longueurs d'ondes d'émission et d'absorption de 315nm et 340nm, respectivement.



# Chapitre 3

## Résultats





## 3 Résultats

### 3.1 Prédiction de la liaison de peptide aux RCPGs

#### 3.1.1 Validation de la méthode

##### 3.1.1.a Contexte

La prédiction de la liaison d'un ligand à un Récepteur Couplé aux Protéines-G est un défi assez difficile. La cristallographie aux rayons-X et plus récemment la Cryo-EM ont permis de résoudre les structures de nombreux RCPGs complexés avec des ligands agonistes, antagonistes ou encore allostériques, donnant ainsi de précieuses informations sur les contacts établis par ces ligands avec leurs récepteurs respectifs, dans la poche de liaison mais aussi possiblement dans des positions intermédiaires. Mais si ces structures fournissent des données importantes, elles ne peuvent pas retransmettre toute la dynamique de liaison du ligand au récepteur. Pour étudier ce cheminement, essentiel pour comprendre tous les mécanismes mis en jeu, la dynamique moléculaire (MD) a été déjà employée par le passé. Toutefois, l'échelle de temps nécessaire à l'observation de l'entrée d'un ligand dans son récepteur est telle qu'elle nécessite l'utilisation soit de méthodes biaisées de MD, dont les résultats sont souvent particulièrement difficiles à interpréter car extrêmement dépendants de la (les) coordonnée(s) réactionnelle(s) utilisée(s), soit de ressources de calcul conséquentes pour simuler cette même liaison sans contrainte. Ceci est d'autant plus vrai pour des peptides présentant un nombre de degrés de libertés important.

L'idée développée dans ce manuscrit est d'utiliser les champs de forces gros-grains (CG) afin de simplifier ces simulations, les rendant ainsi accessibles à un bien plus grand nombre de chercheurs. Comme présenté dans l'introduction de ce manuscrit, les modèles CG ont déjà été utilisés pour simuler le comportement des RCPGs, notamment pour étudier leur dimérisation [Periole et al. 2007; Ghosh, Sonavane, et Joshi 2014; Prasanna, Sengupta, et Chattopadhyay 2016; Pluhackova et al. 2016; Prada-Gracia et al. 2017]. Toutefois, avec la diminution de la résolution de ces modèles, il est à prouver que l'on conserve suffisamment d'informations structurales pour décrire correctement la liaison d'un peptide à son récepteur. Le champ de forces gros-grains MARTINI [Marrink, de Vries, et Mark 2004] a été développé et validé initialement pour les lipides et permet aussi de simuler des peptides, protéines, ainsi que bien

évidemment l'eau et les ions. Ceci en faisait un candidat parfait pour étudier nos systèmes membranaires. Dans notre travail, nous avons utilisé MARTINI avec ElNeDyn [Periole et al. 2009] permettant de maintenir les structures secondaires et tertiaires des protéines. En effet, des tests réalisés avec MARTINI sans ElNeDyn montrent des distorsions des structures des RCPGs, notamment des mouvements inter-hélices peu réalistes. Pour plus de facilité et pour nous éviter le paramétrage de petits ligands organiques avec MARTINI, nous avons utilisé les deux seuls complexes cristallographiques connus à l'époque présentant un RCPG co-cristallisé avec un ligand peptidique : le récepteur de la Neurotensine-1 du rat (NTR1r) (code PDB : 4BUO) avec son agoniste naturel la Neurotensine [8-13] (NT[8-13]) [Egloff et al. 2014] et le récepteur de Chimiochine C-X-C de type-4 (CXCR4) (code PDB : 3OE0) lié à un antagoniste : le peptide CVX15 [B. Wu et al. 2010]. La nature complètement différente de ces deux peptides, agoniste, court (6 résidus) et flexible pour NT[8-13], et antagoniste, un peu plus long (16 résidus) et rigide pour CVX15, en font des cas d'étude idéaux pour une preuve de concept.

Nos premières simulations effectuées sur NTR1r ne produisirent pas tout de suite le résultat escompté. En effet, parmi les dizaines de simulations lancées (à l'échelle chacune de la dizaine de  $\mu$ s), la liaison de la Neurotensine à la poche orthostérique de son récepteur ne fut observée qu'une seule fois. En revanche, cette simulation démontrait déjà la capacité du modèle gros-grains utilisé à bien prédire la dynamique de liaison de ce peptide à son récepteur, avec une position prédite à la fin de la trajectoire très proche de celle connue en cristallographie de rayons-X (RMSD inférieur à 3Å). L'analyse des trajectoires ne présentant pas de liaison a permis de montrer que le principal problème était la capture de positions intermédiaires du peptide présentant un nombre conséquent de contacts favorables (minima locaux), le plus souvent entre les boucles ECL2 et ECL3 et à l'entrée de la poche du récepteur du côté extra-cellulaire, mais aussi du côté intra-cellulaire à cause des conditions périodiques appliquées lors de nos simulations. Pour remédier à ces problèmes d'échantillonnage et augmenter la probabilité d'observer la liaison du peptide à son récepteur, le champ de forces MARTINI a été combiné à la méthode de dynamique moléculaire « replica-exchange » (REMD) [Sugita et Okamoto 1999], méthode facilitant l'échantillonnage en apportant de l'énergie au système (voir partie 2.2.3). Avec ce nouveau protocole, nous avons rapidement conclu à une bien meilleure reproductibilité de la liaison du peptide à son récepteur, et ce à l'échelle de la  $\mu$ s. Toutefois, le peptide se retrouvait encore bien souvent au contact de la face intra-cellulaire du récepteur. Afin de supprimer ce problème, nous avons opté pour la construction et la simulation de « systèmes

doubles » où deux systèmes complets (récepteur, membrane, peptide) se retrouvent en vis-à-vis afin de créer un « compartiment extra-cellulaire » fini (voir partie 2.4.9).

La première partie de mon travail a été de faire la preuve de concept de ce protocole. Pour cela, j'ai effectué des simulations CG-REMD sur les deux complexes NTR1r:NT[8-13] et CXCR4:CVX15. Dans nos simulations, deux copies du même peptide ont été placées dans le compartiment extra-cellulaire, suffisamment loin des deux récepteurs pour ne pas introduire de biais initial. Pour chaque complexe, nous avons produits 2 $\mu$ s de simulations à 58 températures différentes séparées d'un pas de 3K (de 300 à 471K) pour un total de temps de simulation de 116 $\mu$ s sur un système comportant deux récepteurs et deux peptides. L'analyse des trajectoires montre une liaison presque systématique des peptides à leurs récepteurs lors des 58 répliques ; le clustering des trajectoires montre que les conformations les plus représentées sont celles où le peptide est dans la poche de liaison, conformations présentant un RMSD bas autour de 3Å lorsque comparées aux structures cristallographiques de référence. Une fois converties en modèles tout-atomes, ces conformations montrent une bonne stabilité des contacts récepteur:peptide prédits avec le modèle gros-grains, avec environ 60% des contacts natifs conservés tout au long des différentes trajectoires de dynamique moléculaire obtenues.

Grâce à cette méthode, nous démontrons ici que nous sommes capables de prédire la position de peptides dans leurs récepteurs, mais aussi d'apporter des hypothèses quant au chemin emprunté par ces peptides avant d'atteindre leurs positions finales. Un autre avantage de notre approche est qu'elle peut être facilement appliquée à d'autres couples peptide:récepteur (ou autres protéines d'ailleurs) car elle demande des moyens de calcul assez raisonnables (cluster de laboratoire).

### 3.1.1.b Article

« Reprinted with the permission from [Delort, Bartholomé, Pedro Renault, Landry Charlier, Florent Raussin, Jean Martinez, et Nicolas Floquet. 2017. *Journal of Chemical Information and Modeling* 57 (3): 562-71.]. Copyright 2017 American Chemical Society. »

# Coarse-Grained Prediction of Peptide Binding to G-Protein Coupled Receptors

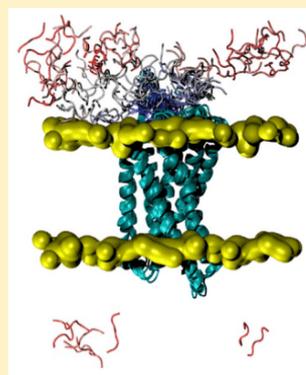
Bartholomé Delort,<sup>†</sup> Pedro Renault,<sup>†,‡</sup> Landry Charlier,<sup>†</sup> Florent Raussin,<sup>†</sup> Jean Martinez,<sup>†</sup>  
 and Nicolas Floquet<sup>\*,†</sup>

<sup>†</sup>Institut des Biomolécules Max Mousseron (IBMM), UMR5247, CNRS, Université de Montpellier, Ecole Nationale Supérieure de Chimie de Montpellier, Faculté de Pharmacie, 15 Avenue Charles Flahault, BP 14491, 34093 Montpellier, Cedex 05, France

<sup>‡</sup>Centre de Biochimie Structurale (CBS), UMR5048, CNRS, Université de Montpellier, INSERM U1054, 29 rue de Navacelles, 34090 Montpellier, France

 Supporting Information

**ABSTRACT:** In this study, we used the Martini Coarse-Grained model with no applied restraints to predict the binding mode of some peptides to G-Protein Coupled Receptors (GPCRs). Both the Neurotensin-1 and the chemokine CXCR4 receptors were used as test cases. Their ligands, NTS<sub>8-13</sub> and CVX15 peptides, respectively, were initially positioned in the surrounding water box. Using a protocol based on Replica Exchange Molecular Dynamics (REMD), both opening of the receptors and entry of the peptides into their dedicated pockets were observed on the  $\mu$ s time-scale. After clustering, the most statistically representative orientations were closely related to the X-ray structures of reference, sharing both RMSD lower than 3 Å and most of the native contacts. These results demonstrate that such a model, that does not require access to tremendous computational facilities, can be helpful in predicting peptide binding to GPCRs as well as some of the receptor's conformational changes required for this key step. We also discuss how such an approach can now help to predict, *de novo*, the interactions of GPCRs with other intra- or extracellular peptide/protein partners.



## ■ INTRODUCTION

G-Protein Coupled Receptors (GPCRs) form one of the largest families of human proteins (more than 800 members) and display a wide range of biological functions both at the physiological and pathological levels.<sup>1</sup> These receptors are characterized by seven hydrophobic, helical trans-membrane segments that form a well-defined binding crevice. This orthosteric binding pocket can accommodate ligands of highly diverse nature and size, including small, rigid organic molecules, as well as peptides and proteins.<sup>2</sup> Since 2007, many X-ray structures of GPCRs have been solved, describing these receptors in interaction not only with different ligands but also with their favorite intracellular partners, e.g. heterotrimeric G-proteins<sup>3</sup> or beta-arrestin.<sup>4</sup>

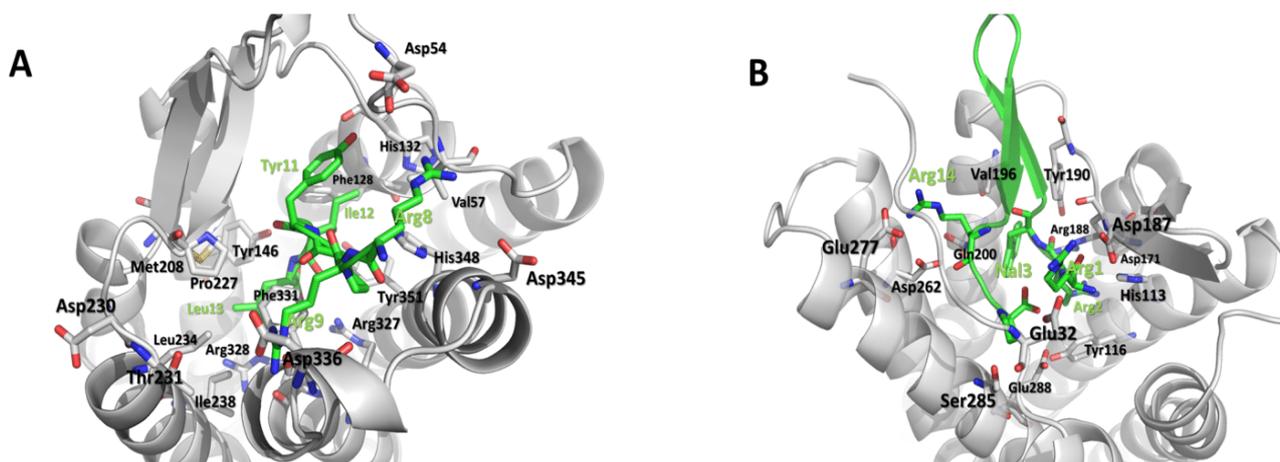
To date, the structures of nearly 30 different “rhodopsin-like” class-A GPCRs have been solved. Nevertheless, only two of these structures describe a GPCR in complex with an extracellular peptide: (i) the Neurotensin receptor bound to the linear agonist NTS<sub>8-13</sub> peptide<sup>5,6</sup> and (ii) the CXCR4 chemokine receptor bound to the cyclic antagonist CVX15 peptide.<sup>7</sup> Despite these structural data, it is still unclear how peptides bind to their receptors from a dynamic perspective at the molecular level. It can nevertheless be hypothesized that such binding should involve quite large amplitude motions, including the opening/closing of the whole extracellular moiety of the receptor. Such conformational changes would not be required for the interaction of GPCRs with smaller, rigid molecules. This binding is thought to appear within the

microsecond time-scale. However, experimental studies are still unable to provide a time-resolved description of such an event. Nevertheless, theoretical calculations including Molecular Dynamics (MD) simulations can help to shed light on these dynamic aspects.

Accordingly, recent papers have shown that large computational facilities combined with highly efficient algorithms can achieve *de novo* prediction of ligand binding to their respective targets.<sup>8</sup> Using the beta-2 adrenergic receptor and some small, rigid ligands as an example, it has been shown that semiempirical force fields including that of CHARMM<sup>9</sup> could also predict interactions between a drug and its target, both at the molecular level and at the microsecond time-scale.<sup>10</sup> However, such simulations performed on large molecular systems (most of the time over 100000 atoms) require access to tremendous computational facilities that remain inaccessible for most laboratories in the world. To reduce the required computational time to hundreds of nanoseconds, the accelerated molecular dynamics method<sup>11</sup> was also applied to the beta-2 adrenergic receptor, allowing the sampling of large amplitude motions<sup>12</sup> that normally occur throughout larger time-scales.<sup>13</sup> The same powerful method was as successful in predicting the binding mode of an M3 ligand to its receptor at a similar time-scale.<sup>14</sup> Despite an increased sampling efficiency, this approach still requires significant computational time,

**Received:** August 25, 2016

**Published:** February 23, 2017



**Figure 1.** Orientation of the two NT-8:13 (A) and CVX15 (B) peptides in the rat Neurotensin-1 and CXCR4 receptors, respectively.

especially for large systems such as GPCRs with an explicit representation of the membrane environment. Moreover, despite the acceleration factor inherent to this method, several trajectories are still required for statistical robustness, and fine-tuning of the boost parameters may also be necessary.

When it comes to the simulation of peptide binding to GPCRs, adequate sampling becomes even harder to achieve, because of the flexibility of the peptide that must be taken into account, as well as the presumably larger structural rearrangements of the receptor. Therefore, while the approaches mentioned above work well to simulate the binding of small molecules, more efficient computational methods are necessary to study peptide binding. Coarse-Grained (CG) molecular dynamics simulations are a convenient alternative to tackle various problems in biomolecular dynamics.<sup>15</sup>

Through a significant decrease of the number of particles to be treated, Coarse-Grained models, including the Martini force field for both lipids and proteins,<sup>16,17</sup> now allow easy access to  $\mu$ s range MD simulations of very large molecular systems<sup>18,19</sup> on a typical cluster machine. The portability of MD codes on graphical processing units (GPUs) now even enables such calculations to be performed on a desktop workstation. Recently, we used the Martini Coarse-Grained model to predict both the insertion and orientation of a viral protein at the membrane surface, as well as its ability to capture specific lipids.<sup>20</sup> In the present study we have used the same CG model in a Replica Exchange Molecular Dynamics scheme, to predict the dynamic binding of a peptide to its receptor, with no applied restraints. Our main purpose was to demonstrate that a relatively inexpensive simulation protocol is able to reach reasonable binding poses, that can be refined by further computational studies, or used to suggest site-directed mutagenesis experiments.

To validate our approach, we used two specific test cases: the rat Neurotensin-1 (NTSR1) and the Chemokine (CXCR4) receptors. These receptors have been respectively cocrystallized with the endogenous agonist Neurotensin 8–13 hexapeptide ( $\text{NH}_3^+$ -Arg-Arg-Pro-Tyr-Ile-Leu- $\text{CO}_2^-$ ) and with the antagonist CVX15 cyclic peptide ( $\text{NH}_3^+$ -Arg-Arg-Nal-Cys-Tyr-Gln-Lys-DPro-Pro-Tyr-Arg-Cit-Cys-Arg-Gly-DPro- $\text{CO}_2^-$ ), two ligands that exhibit nanomolar affinity for their targets.<sup>21,22</sup>

The X-ray binding modes of these two peptides within their receptors are reported in Figure 1. The NTS<sub>8–13</sub> peptide binds deeply into its receptor, its C-terminus lying in a pocket formed by residues Arg327, Arg328, Tyr146, and Tyr351. The two

hydrophobic side chains of Leu13 and Ile12 are oriented on opposite sides of the main chain. The Leu13 residue establishes contacts with Met208, Pro227, Ile238, Leu234, Phe331, whereas Ile12 contacts residues His348, Phe128, Val57, and His132. In the bound conformation, the Tyr11 residue of the peptide orients toward another subpocket mainly formed by the N-terminal part of the receptor but without establishing any strong interactions. Similarly, the N-terminal Arg8 and Arg9 residues of the peptide, that are localized at the surface of the receptor, are not involved in ionic interactions. The high B-factor values of these two residues nevertheless suggest they could interact with residues Asp54, Asp230, Thr231, Asp336, and/or Asp345. Interestingly, a quick comparison of the bound and unbound<sup>23</sup> conformations of the NTS<sub>8–13</sub> peptide suggests that no significant conformational change is required for binding to its receptor (data not shown), the peptide adopting in both cases an unordered, mostly extended conformation. Like NTS<sub>8–13</sub>, the CVX15 peptide shares positively charged residues (Arg1, Arg2, Lys7, Arg11, Arg14) that clearly contribute to stabilize its position inside its receptor binding pocket. Among them, Arg1 interacts directly with residues Glu32 and Asp187, whereas Arg2 projects deep inside the receptor in a subpocket formed by residues His113, Tyr116, Asp171, and Arg188. The X-ray structure also shows a direct interaction of Arg14 with Glu277 and possibly with Asp262. The Naphthalene3 (Nal3) residue of the peptide is located between residues Arg188, Tyr190, Val196, and Gln200. Finally, the peptide's C-terminal Pro16 contacts Ser285 and Glu288 in the receptor. Contrary to NTS<sub>8–13</sub>, CVX15 is highly restrained by the disulfide bridge between Cys4 and Cys13 residues. Therefore, NTSR1 and CXCR4 receptors constitute two quite different but complementary examples in this study.

## METHODS

**Systems Setup.** As described in the original references,<sup>6,7</sup> the models of both receptors corresponded to thermostable mutants. We have not modified their structures or sequences, since our objective was to reproduce the binding poses observed in the crystal structures.

Simulations of the Neurotensin-1 rat receptor (NTSR1) were carried out using the X-ray structure of the receptor bound to its peptide agonist (PDB ID: 4BUO)<sup>6</sup> that was determined at 2.75 Å resolution. To simulate ligand binding, the NTS<sub>8–13</sub> peptide ( $\text{NH}_3^+$ -Arg-Arg-Pro-Tyr-Ile-Leu- $\text{CO}_2^-$ ) was first removed from its receptor. To start with a complete

structure of the receptor, the intracellular loop 1 (ICL1) was built with Modeler,<sup>24</sup> whereas residues missing in the intracellular loop 3 (ICL3) were replaced by six alanine residues to make the connection between TM5 and TM6. The disulfide bridge between Cys91 and Cys174 was included in our simulations. The receptor was converted into a Coarse-Grained (CG) Martini<sup>16</sup> representation using the Charmm-Gui server<sup>25,26</sup> and oriented according to its *z*-axis across a 100% palmitoyl-oleoyl-phosphatidyl-choline (POPC) membrane. The system was then duplicated and rotated by 180 deg along the *z*- and *x*-axes, respectively, to prevent exploration of the intracellular compartment by the peptide, due to periodic boundary conditions. Water and chloride anions were added at this stage to complete and neutralize the global charge of the system.

The scripts provided by Charmm-Gui were then used for a first equilibration and production step, using the Martini 2.0 force field<sup>17</sup> together with the GROMACS software (version 4.6.7).<sup>27</sup> Elastic network (EINeDyn 22) was turned on, setting the elastic bond force constant to 500 kJ mol<sup>-1</sup>.nm<sup>-2</sup> and the lower and upper elastic bond cutoffs to 0.5 and 0.9 nm, respectively. Next, water and ions were removed so that two copies of the NTS<sub>8-13</sub> peptide could be added in between two facing receptors, in the “extracellular” compartment (see supplemental Figure S1). The system building was completed by reading water and ions, resulting in a total of 18000 CG beads.

Simulations of the CXCR4 chemokine receptor were carried out using the same protocol as the one employed for NTSR1, starting with the X-ray structure of the human receptor (PDB ID: 3OE0)<sup>7</sup> bound to the CVX15 peptide (NH<sub>3</sub><sup>+</sup>-Arg-Arg-Nal-Cys-Tyr-Gln-Lys-DPro-Pro-Tyr-Arg-Cit-Cys-Arg-Gly-DPro-CO<sub>2</sub>-), and determined at a 2.9 Å resolution. Again, the peptide was initially removed from the receptor binding pocket. The T4 lysozyme inserted into the intracellular loop 3 (ICL3) was also deleted, ICL1 and ICL3 loops being completed/built with Modeler,<sup>24</sup> as mentioned above for NTSR1. The two disulfide bridges of the receptor between Cys81:Cys158 and Cys4:Cys244 were included.

As for NTSR1, the system was then duplicated along the *z*-axis and rotated by 180 deg along the *x*-axis. An initial step of equilibration was performed in the presence of water and ions, which were later removed to enable introduction of two copies of the CVX15 peptide ligand in the “extracellular” compartment. Water and ions were once again included before the production runs, to finally reach 17500 beads. The peptide disulfide bridge between Cys4 and Cys13 was included. Because of lacking CG parameters, four residue modifications were applied to the CVX15 peptide as compared to the cocrystallized ligand. Because proline residues are represented by only two beads in Martini, the D-Pro to L-Pro conversion was thought to have no influence on the final results. So, both DPro8 and DPro16 residues were replaced by natural Lpro; also, Citruline12 and Naphthalene3 residues were respectively replaced by Arginine and Tryptophan, as described in evolutionary related derived peptides.<sup>22,28</sup>

**Replica Exchange MD Simulations.** All REMD simulations were performed with the GROMACS software package (version 4.6.7).<sup>27</sup> For each system, 58 replicas were launched simultaneously using a temperature range from 300 to 471 K. For such systems, the used temperature interval of 3 K corresponded to an exchange probability of 0.2 between two adjacent replicas (see supplemental Figure S2). The metropolis

criterion required to exchange coordinates was tested every 1000 MD steps (20 ps). Each replica was first heated and equilibrated at the desired temperature through NVT/NPT molecular dynamics. Integration was performed with the leapfrog algorithm, using an integration step of 20 fs. Long-range nonbonded interactions were truncated at 12 Å. Temperature was coupled to the system with the velocity rescale thermostat, and the Berendsen barostat was used for pressure control in the NPT simulations. The production phase was stopped after 2 μs, leading to 58 \* 2 = 116 μs of simulation for each peptide:receptor complex. Simulations were repeated several times to verify the reproducibility and the convergence of the obtained results.

**Clustering Analysis.** The clustering procedure was the same for NTSR1 and CXCR4 receptors. In each case, after concatenation of the trajectories of all replicas, rotations and translations were removed by least-squares fitting to the backbone beads of the CG converted crystal structure. Clustering of the fitted trajectory was based on peptide beads and on receptor beads closer than 15 Å to the bound peptide. It was performed with the *g\_cluster* tool of the GROMACS package, using the algorithm described by Daura et al.<sup>29</sup> and a 2.5 Å cutoff.

**Principal Component Analysis.** Principal Component Analysis (PCA) of the concatenated and fitted trajectories consisted in the diagonalization of the covariance matrix

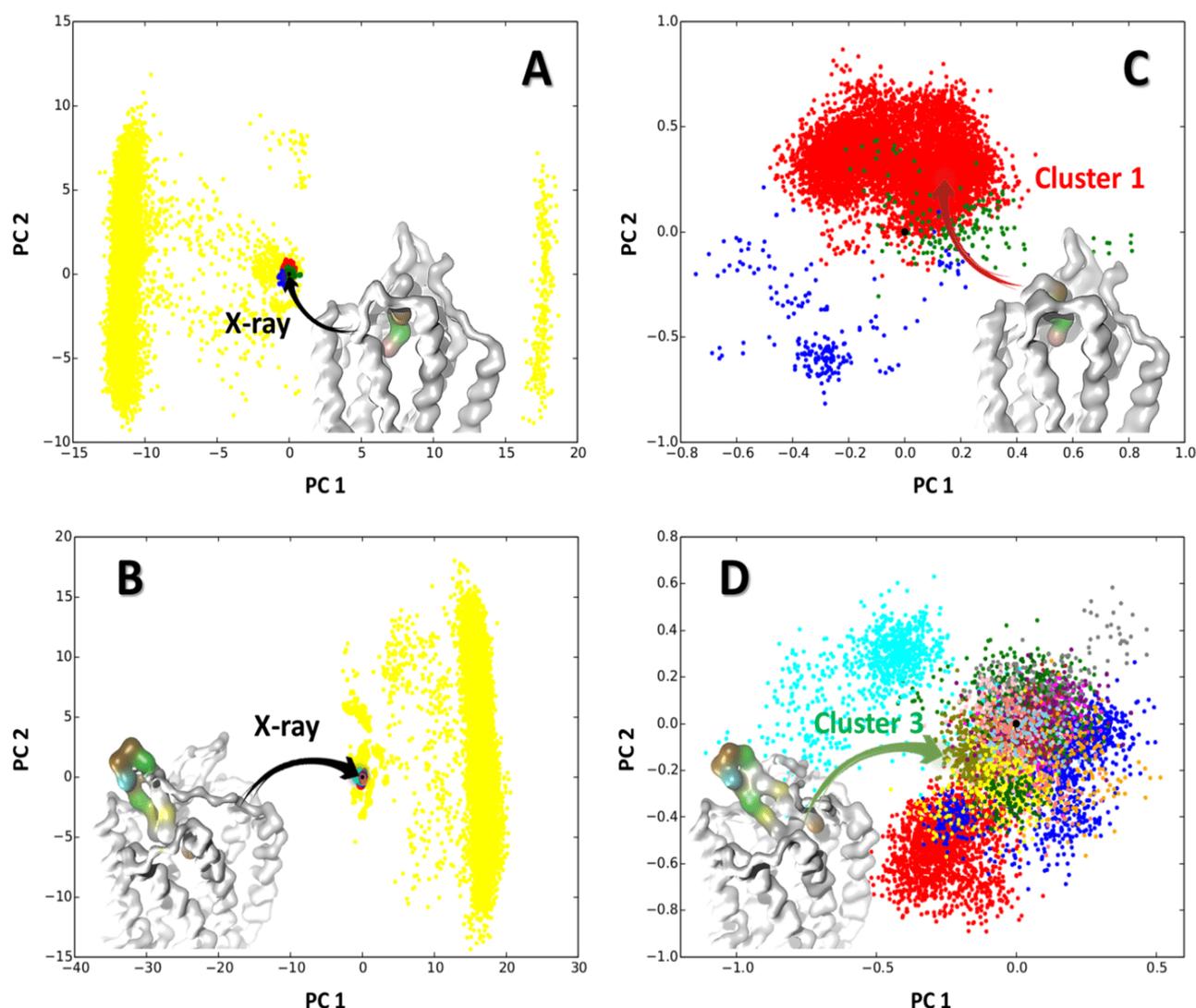
$$C_{ij} = \langle (r_i - \langle r_i \rangle) \cdot (r_j - \langle r_j \rangle) \rangle \quad (1)$$

where *i* and *j* denote all pairs of the 3N Cartesian coordinates of the N backbone atoms of the receptor; *r<sub>i</sub>* indicates the instantaneous value of coordinate *i*, and  $\langle r_i \rangle$  is the average value of this coordinate in the ensemble of conformations. The analysis and the projection of the trajectories in the space of the respective first two PC's were performed with the ProDy software.<sup>30</sup>

**Remapping to All-Atom Models for MD Simulations.** For both our peptide:receptor complexes, the center of each of the three most populated clusters was used for all-atom MD simulations. These centers were then used for remapping to an all-atom model, using the procedure described in Martini Maker<sup>26</sup> and related scripts,<sup>29</sup> as described in Charmm-Gui. The remapped all-atom models were then used as starting configurations for conventional molecular dynamics simulations. For these calculations, the CHARMM (version 36) force field was used. A step of 2 fs was used for integration of the Newton's equations of motion, fixing in length all bonds involving hydrogen atoms. PME (Particle Mesh Ewald) was used for the treatment of long-range interactions, by applying a switch function between 10 and 12 Å.

Using Charmm-Gui, the complex was first inserted in a POPC bilayer, and the system including the protein and the membrane was solvated by a box of TIP3P water molecules. Ions were added to neutralize the global charge of the system. After energy minimization, the system was equilibrated through successive steps of MD simulations performed in the NVT and in the NPT (300 K, 1 bar) ensembles, as described,<sup>31</sup> and by progressively decreasing the restraints applied to the membrane and the protein atoms. In the last stage of equilibration (200 ps), as well as during the production stage, no residual restraint was applied to the system.

**Identification of Peptide:Receptor Contacts and Binding Affinity Calculations.** The identification of contacts between each peptide and its receptor was made with the PRODIGY



**Figure 2.** Results from clustering analyses and projections of the conformational ensembles obtained for both the NTS-1 (A, C) and CXCR4 (B, D) receptors. The X-ray structure (reported in boxes A and B) was materialized by a black point on all graphs. C and D boxes correspond to a zoom-in on the most populated clusters, the three most populated clusters being reported in red, blue, and green, respectively. The models in boxes C and D correspond to the clusters that were the closest to the X-ray structure (cluster 1 and cluster 3, respectively).

software.<sup>32</sup> Two residues were considered in contact when a pair of atoms belonging to these residues was closer than 5.5 Å. For each system, contacts were identified in the X-ray crystal structure of the complex and also in the all-atom MD simulation of the cluster center closest to the crystal structure (see the subsection “Remapping to All-Atom Models for MD Simulations” above); 600 snapshots were extracted every 20 ps from the last stage of equilibration and from the first 1 ns of the production stage. The probability of occurrence of each X-ray (native) contact was then computed among these MD snapshots using in-house scripts.

We also used the PRODIGY software to estimate binding affinities between each peptide and its receptor. For details about the method of estimation, the reader is referred to the original publication.<sup>33</sup> Briefly, the method uses a simple but robust descriptor of binding affinity based on a combination of the number of contacts at the interface of a protein–protein complex and on properties of the noninteracting surface. In terms of accuracy, this method showed a Pearson’s Correlation coefficient of 0.73 between the predicted and measured binding affinities on the benchmark described by Kastriitis et al.<sup>34</sup>

For both our peptide:receptor complexes, the estimations were based on all-atom MD simulations of the three most populated clusters (see the subsection “Remapping to All-Atom Models for MD Simulations” above). The predicted values of binding affinities were computed on snapshots extracted every 10 ps from the first nanosecond of the production phase.

**TMD.** Targeted molecular dynamics simulations of the Neurotensin-1 receptor were performed using the CHARMM36 force field and the NAMD software,<sup>35</sup> using the same parameters and procedure as that described above in the “Remapping to All-Atom Models for MD Simulations” subsection. In TMD, the bias  $U_{\text{TMD}}$  applied to the chosen set of  $N$  atoms is given by  $U_{\text{TMD}} = 1/2 * k/N * [\text{RMSD}_{\text{obs}} - \text{RMSD}_{\text{targ}}]^2$ , where  $\text{RMSD}_{\text{obs}}$  denotes the current value of RMSD at time  $t$  and  $\text{RMSD}_{\text{targ}}$  is the expected value of RMSD at this time. The expected values at any time were determined from a linear interpolation of the initial and final (0 Å) values of RMSD. The same starting structure adopted in Coarse-Grained simulations was used (PDB: 4BUO<sup>6</sup>). After insertion of the receptor in a POPC membrane, the peptide was moved outside the binding pocket and put at four different positions in the

surrounding water box, leading to RMSD in the 30–60 Å range as compared to its X-ray position of reference. After NVT equilibration, the production phase consisted in 30 ns of NPT simulation (300 K, 1 bar): the first 10 ns were performed without any bias; TMD was performed in the last 20 ns, by applying a force constant of  $k = 20 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{Å}^{-2}$  to a set of atoms that included the whole peptide and the three Tyr95, Arg251, and Arg252 residues of the receptor. Eight different trajectories were produced, two for each starting position of the peptide, after modification of the initial velocities. The position of the peptide in the X-ray structure was used as a target.

## RESULTS

For both peptide:receptor systems a REMD simulation was performed, simulating the same system at 58 different temperatures, for a total time of 116  $\mu\text{s}$ . The trajectories at different temperatures were concatenated for analysis, resulting in two trajectories, one for each peptide:receptor system.

**The Most Populated Clusters Contain Conformations That Resemble the X-ray Structure.** For both systems, clustering analysis was performed on the obtained conformations using the beads of both the peptide and the receptor's pocket (see the [Methods](#) section). Using a cutoff value of 2.5 Å, a total of 1627 clusters were found for the NTSR1 receptor, but only 3 of them contained more than 100 conformations. In the vast majority of clusters, the peptide was far from its receptor. Similar observations were made for the CXCR4 receptor, with a total of 5996 clusters, only 17 of which contained more than 100 conformations.

Principal component analysis was performed on both trajectories. For each one, the coordinates of the receptor and of the peptide, as well as the corresponding X-ray crystal structure of the complex, were projected onto the space of the respective first two principal components. The obtained projections are shown in [Figures 2A](#) and [2B](#) for the NTSR1 and CXCR4 receptors, respectively. Clusters were sorted and numbered according to their respective populations, the cluster 1 being the most populated. Conformations of low-populated clusters, including all configurations in which the peptide was free in solution, were reported in yellow. The other colors represent clusters comprising more than 100 conformations, the red, green, and blue clusters corresponding to the three most populated ones, respectively. A closer view of these highly populated clusters is shown in [Figures 2C](#) and [2D](#). In both cases, the most populated clusters contain conformations that resemble that of the crystal structure. For the NTSR1 receptor, the crystal structure was part of the first cluster (reported in red in [Figure 2C](#)), that contained 9061 of the 20000 explored conformations. For the CXCR4 receptor, the crystal structure was 2.7 Å away from the center of the third most populated cluster (1147 out of the total 20000 conformations, reported in green in [Figure 2D](#)).

[Table 1](#) summarizes the RMSD between the centers of the three most populated clusters of each system and their respective X-ray structure of reference.

These results demonstrate that orientations resembling that of X-ray structures of reference could be predicted, *de novo*, for both peptide:receptor complexes, the NTSR1 cluster 1 and the CXCR4 cluster 3 being the closest to the X-ray structures. As reported in [Table 1](#), the minimum RMSDs computed on the peptide reached 4.5 Å for the NTS<sub>8–13</sub> peptide in cluster 1 and 4.1 Å for the CVX15 peptide in cluster 3. For the NTSR1 receptor, this RMSD was notably explained by a bad

**Table 1. RMSD Values Computed between the Centers of the Three Most Populated Clusters of Each System and the CG Model of the Respective X-ray Structure of Reference<sup>a</sup>**

	NTSR1			CXCR4		
	cluster 1	cluster 2	cluster 3	cluster 1	cluster 2	cluster 3
complex	1.6	1.9	2.0	4.3	3.6	2.2
pocket + peptide	1.7	1.8	2.5	5.7	5.0	2.7
pocket	1.4	1.5	1.7	2.7	1.7	2.6
peptide	4.5	4.8	8.0	14.6	13.6	4.1
peptide (BB)	2.7	3.8	6.5	14.0	13.3	3.8

<sup>a</sup>These values have been computed on different parts of the models.

orientation of the Tyr11 residue of the NTS<sub>8–13</sub> peptide when compared to the X-ray structure, whereas the main chain was well oriented (as shown by a RMSD of 2.7 computed on backbone beads of the peptide). For the CVX15 peptide, this value of 4.1 Å was due to a slight translation of the whole peptide in the receptor pocket, leading to a closely related value when computed on backbone beads only.

The other clusters were all sharing higher RMSDs, corresponding, in each case, to significantly different orientations of the peptides in their respective pockets.

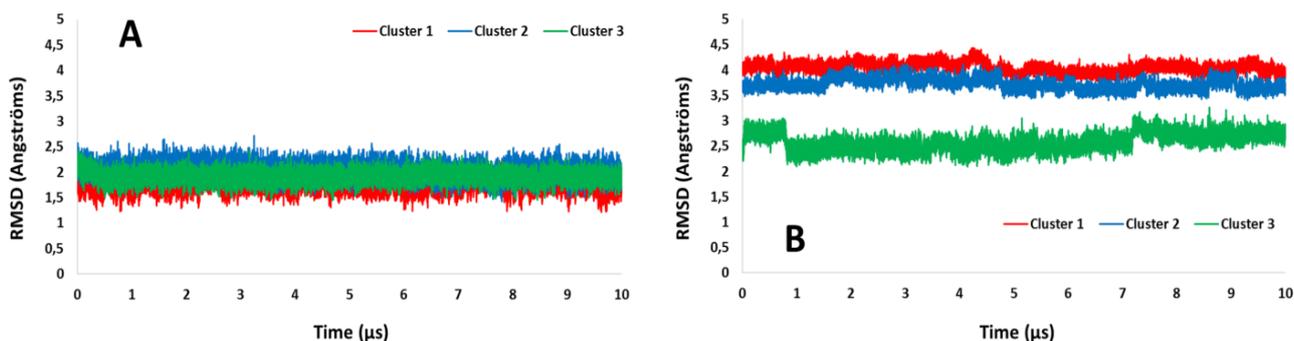
**The Obtained Clusters Are Stable and Belong to the Conformational Ensemble of the X-ray Structure.** To assess their stability, an additional 10  $\mu\text{s}$  CG MD simulation was performed on each of the three most populated clusters of each peptide:receptor complex. In each case, the center of each cluster was used as a starting configuration. RMSD values reported in [Figure 3](#) were computed on all beads of both the peptide and the receptor and show that all the predicted orientations were remarkably stable and remained relatively similar to the X-ray structures of reference. Indeed, values lower than 2 Å were observed for the three most populated clusters of the NTSR1 receptor ([Figure 3A](#)). For CXCR4, cluster 3 gave the most interesting results with RMSDs  $\sim 2.5$  Å ([Figure 3B](#)).

To confirm that these values were related to the model limitation, we used the X-ray crystal structure of each peptide:receptor complex as a starting configuration to perform five additional CG MD simulations of 10  $\mu\text{s}$ . Interestingly, and as shown in [Figure 4A](#), we explored different conformations of the NTSR1:NTS<sub>8–13</sub> complex, some of which sharing a low peptide RMSD value of 2.0 Å to the center of cluster 1. This confirmed that the same alternate orientation of the Tyr11 residue was also observed starting from the X-ray crystal structure. This observation was of general interest as we also observed such a rotation of this residue using an all-atom model, demonstrating its role in NTS1/NTS2 human receptor selectivity (unpublished results).

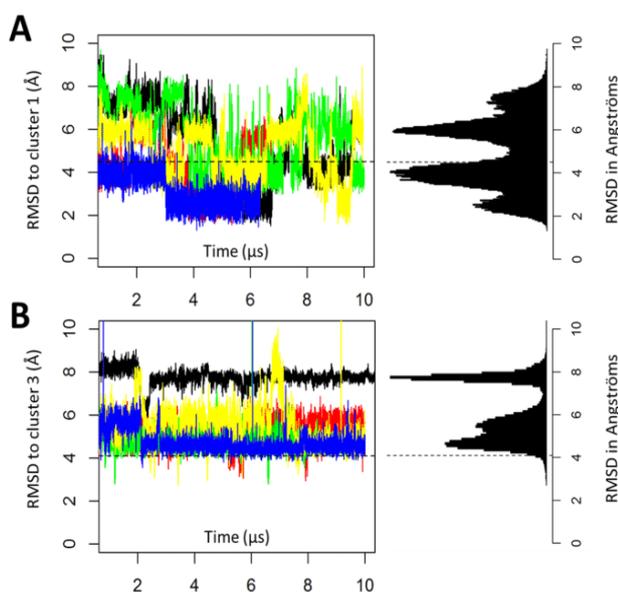
Starting from the X-ray crystal structure of the CVX15: CXCR4 complex, different conformations were also explored but rarely sharing a peptide RMSD lower than 4 Å to the center of cluster 3 (see [Figure 4B](#)).

Considered together, these results showed that cluster 1 of NTSR1 and cluster 3 of CXCR4 constitute reasonably good models to start with other simulations, because they belong to the same conformational ensemble as that of the X-ray structure.

Interestingly, we also observed the free binding of the NTS<sub>8–13</sub> peptide during a conventional (non REMD) Coarse-Grained MD trajectory, without any applied restraint, at 300 K.



**Figure 3.** RMSD evolution along classical coarse-grained MD trajectories at 300 K of the three most populated clusters obtained for both the NTSR1 (A) and CXCR4 (B) receptors. RMSDs were computed against the CG mapped X-ray structure using all the CG beads of the full peptide:receptor complex. Cluster 1 (NTSR1) and cluster 3 (CXCR4) were the closest to the X-ray structure.

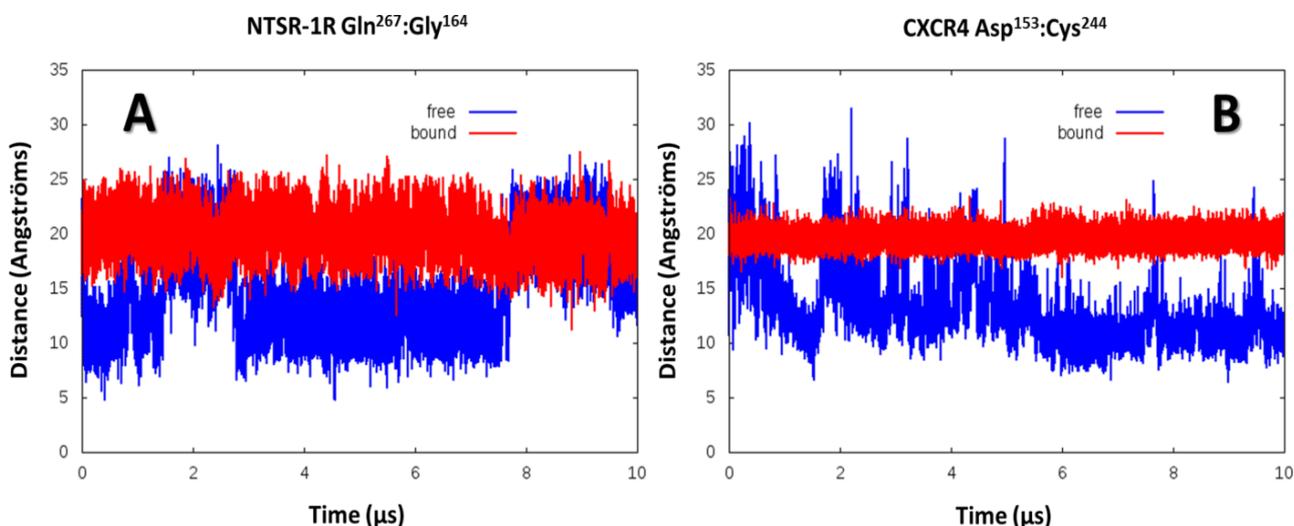


**Figure 4.** Peptide RMSDs to cluster 1 or cluster 3 computed along 5 different MD simulations, starting from the X-ray crystal structure of NTSR1 (A) and CXCR4 (B). Initial RMSDs of the peptide alone computed to cluster 1 (NTSR1) and cluster 3 (CXCR4) were 4.5 and 4.1 Å, respectively, and were materialized by dashed lines.

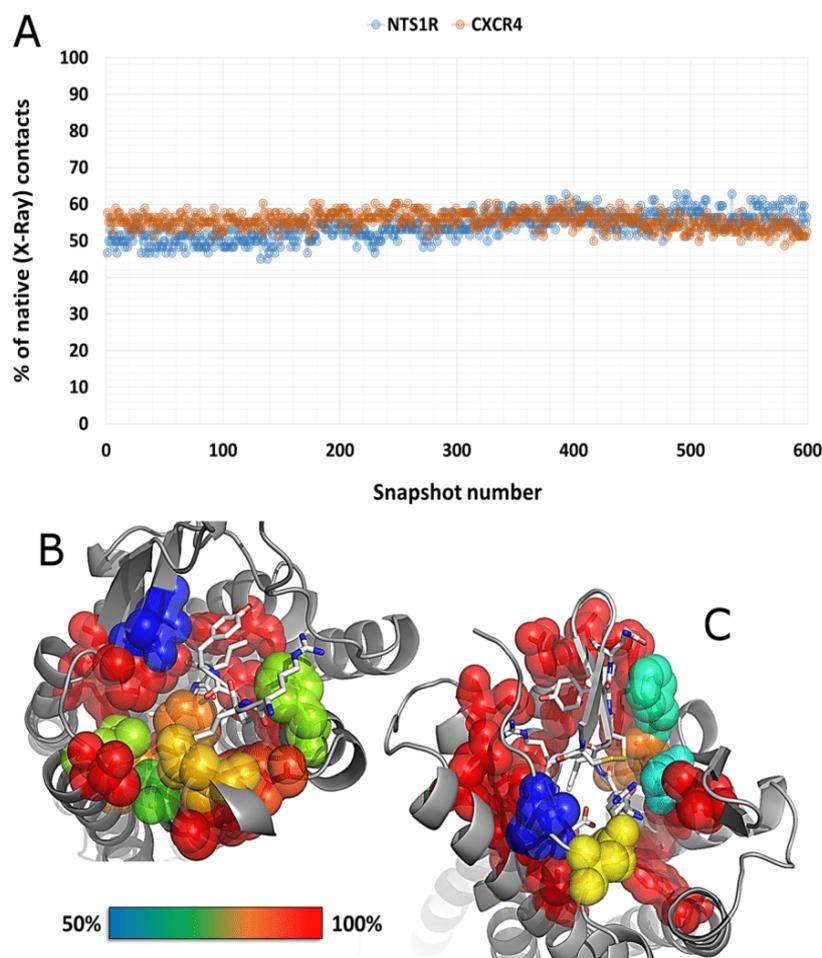
However, this rare event was observed only once among tens of trajectories, indicative of the absolute requirement to use enhanced sampling methods such as REMD. A movie is provided in the [Supporting Information](#) (supplemental movie S1) showing this binding event. In this simulation, the final RMSD of the peptide computed against the X-ray position was only 2.0 Angströms.

**Opening of the Receptor Is Required for Peptide Binding.** Two distances between Gly164 and Gln267 for the NTSR1 receptor, and between Asp153 and Cys244 for the CXCR4 receptor, were chosen for analysis of the open/closed states of each receptor along our MD simulations (see [Figure 5](#)). In both receptors, one of the chosen residues was located in the Extra-Cellular Loop 2 (ECL2), whereas the other was located in the ECL3.

To compare the behavior of both bound and unbound receptors, an additional 10  $\mu$ s CG MD simulation was performed on each receptor, starting again from the X-ray crystal structure, but this time after removing the bound peptide. The distances reported in [Figure 5](#) clearly show that the two unbound receptors fluctuate between a closed and an open conformation, the latter being stabilized in the presence of the peptide. This result is of particular interest as it shows that the used elastic network allows such an opening/closing motion. More interestingly, we observed that this opening motion corresponded to a collective motion involving nearly all



**Figure 5.** Distance measurements in both NTSR1 (A) and CXCR4 (B) receptors showing the opening/closing of the extracellular part of the free- (blue) or peptide-bound (red) receptors.



**Figure 6.** (A) Percentage of native contacts observed along the all-atom MD trajectories computed for both the NTS1R and CXCR4 receptors. Contacts retrieved more than 50% of the whole trajectory were also reported for the NTS1R (B) and CXCR4 (C) receptors. The color code was the same for both systems: lighter colors indicated a lower probability, whereas red indicated a probability of 1, *i.e.*, a contact that was present in all the analyzed snapshots.

the extracellular residues of the receptor. Principal component analyses revealed this pattern for both receptors, indicating that this opening/closing motion of the binding pocket was coupled to the most prevalent collective motions along the obtained trajectories. For the NTS1R receptor, the spreading of residues Gln267 and Gly164 was captured by the second PC (see Figure S3). For the CXCR4 receptor, the first two PCs described a collective motion that affected the distance between Asp153 and Cys244 residues (see Figure S3).

To verify if such a motion could also be observed with an all-atom representation, we performed targeted molecular dynamics (TMD) simulations on the NTS1R receptor with the CHARMM force field. Eight different trajectories of 30 ns were computed, starting from different positions of the peptide outside the receptor. The simulations were then biased, so that the peptide could reach a position resembling that of the X-ray crystal structure at the end of each trajectory. In half the simulations the peptide could not bind, because of the flexible N-terminal part of the receptor that interfered with binding at the surface of the receptor. In the other simulations, the peptide bound successfully to its receptor, to finally adopt a position reminiscent of the X-ray crystal structure (RMSD of the peptide  $<2$  Å, see Figure S4 as a supplement). Interestingly, analysis of these four successful trajectories showed the same opening motion of the receptor as that observed in our CG

simulations, with the same motion of both ECL2 and ECL3 loops (see Figure S4).

**After Conversion to an All-Atom Model, the Predicted Poses Capture Most of the Peptide:Receptor Native Contacts.** For both peptide:receptor systems, the center of the cluster closest to the crystal structure was remapped to an all-atom model, inserted into a POPC membrane, and subjected to a short conventional MD simulation (equilibration +1 ns of production).

For the NTS1R:NTS<sub>8-13</sub> complex, the analysis of contacts was made for the trajectory initiated from the center of cluster 1; for the CXCR4:CVX15 complex, we used the center of cluster 3. Our analyses focused on native contacts, *e.g.* on contacts that appear in the X-ray structure. With PRODIGY, analysis of the 4BUO and 3OE0 structures led to 62 and 78 contacts, respectively.

Figure 6A shows that, for both systems,  $\sim 60\%$  of these native contacts were retrieved along our simulations. For a detailed analysis of each individual native contact, please refer to Figure S5. Interestingly, and as shown in Figures 6B and 6C, many of these contacts were maintained throughout the obtained trajectories ( $\sim 100\%$  occurrence), thus permitting (1) a good delimitation of the binding pocket and (2) a fine definition of the binding mode of each peptide to its receptor.

Contacts analyses performed on the other obtained clusters (see Figure S5) confirmed the RMSD analyses reported in

**Table 1.** Indeed, the clusters 2 and 3 obtained for the NTSR1:NTS<sub>8-13</sub> complex captured only 45% and 33% of the X-ray contacts, respectively. For the CXCR4:CVX15 complex, the clusters 1 and 2 captured very few of the expected X-ray contacts (less than 10%).

## DISCUSSION

G-Protein Coupled Receptors (GPCRs) constitute a great pool of targets for the development of new drugs. Therefore, a better knowledge of the binding mechanism of drugs to these key receptors is of general pharmacological interest. Molecular simulations constitute an important tool in this context, as they allow for investigating the role of dynamics in the interaction between the receptor and its different ligands at the atomic level. Ideally, approaches based on molecular dynamics simulations could overcome some of the limitations of static methods, including molecular docking, by taking into account the role of the solvent and of the flexibility of both the receptor and the ligand. However, the sampling problems associated with all-atom MD simulations practically impair their use as a tool to predict binding mechanisms or poses in systems as those studied here. This is particularly true when the ligands are peptides, instead of small, relatively rigid molecules. In such cases, an increased sampling is required to capture the flexibility of both the peptide and the receptor properly, the latter undergoing larger conformational changes.

In this study, to overcome these sampling problems, we used a Coarse-Grained model that was coupled to a Replica-Exchange MD protocol. Despite the obvious loss of details in the description of the systems and limitations on the representation of the dynamics, we were able to observe, for the apo form of both receptors, collective motions of relatively large amplitude, enabling to explore both open and closed conformations. This allowed each peptide studied to bind to its target within the  $\mu\text{s}$  time-scale. After this binding, we observed the stabilization of an open conformation of each receptor that was also sampled by the unbound protein, thus suggesting a phenomenon of conformational selection. These findings are in agreement with the well-known structural plasticity of GPCRs. It has also been observed that, given the diversity in GPCR ligands, conformational selection or induced fit mechanisms could occur for these receptors.<sup>36</sup> In particular, previous computational studies have shown conformational selection of the  $\beta_2$  adrenergic receptor by small ligands,<sup>37,38</sup> more recently, conformational selection in the activation of the adenosine A<sub>2A</sub> receptor was observed experimentally.<sup>39</sup> In our study, the stabilization of open conformations is evidence for conformational selection of the receptors by their respective peptide ligands.

Although we have only demonstrated the validity of our protocol in two specific cases, its sampling efficiency clearly encourages new tests. Other calculations are still in progress to verify if this approach can be applied to other peptide:GPCR complexes and decipher their binding pathway/mechanism.

In a predictive approach in which the structure of the peptide:receptor complex is unknown, the main difficulty of this protocol will be to choose the best orientation(s) of the peptide among the obtained clusters, for further refinement. We have already shown that all-atom simulations of the most populated clusters of the CG REMD simulation can lead to models containing most of the native peptide:receptor contacts. Therefore, without any other additional calculations, these data would have enabled us to successfully propose peptide

modifications or site-directed mutagenesis, to perturb the binding properties of each peptide to its receptor. Nevertheless, here we tested the ability of the PRODIGY method to predict  $\Delta G_{\text{binding}}$  differences between clusters 1, 2, and 3 for both the NTS<sub>8-13</sub> and CVX15 peptides. The PRODIGY method was then employed to compute the  $\Delta G_{\text{binding}}$  on snapshots extracted every 20 ps from the first nanosecond of each production trajectory. The values predicted for the three clusters pointed to cluster 1 of NTSR1 and to cluster 3 of CXCR4 as sharing the best mean values of  $\Delta G_{\text{binding}}$  (see Figure S6). Interestingly, these two clusters displayed the greatest similarity to the respective X-ray structures of reference.

In this work, we started our simulations from receptor structures that were cocrystallized with the peptides under study. In agreement, one can ask if similar results would have been obtained starting from a different X-ray structure. The CXCR4 receptor has been cocrystallized with other ligands than the CVX15 peptide (PDB codes 3ODU, 3OE6, 3OE8, 3OE9, 4RWS). However, a cross RMSD analysis performed on the binding pocket of these structures indicated a maximum discrepancy of 1.6 Å (see Figure S7B). For the NTSR1 receptor, this value was only 1.1 Å (PDB codes 3ZEV, 4BV0, 4BWB, 4GRV, 4XEE, 4XES, 5T04).

To confirm that our simulations explored conformations of the binding pocket enclosing all those captured experimentally, we proceeded to a PCA analysis of a conformational set containing all the X-ray structures and the conformations extracted from the equilibration step (at 58 different temperatures). The resulting projections of these conformations onto the space defined by the first two PCs are reported in Figure S7C and S7D for the NTSR1 and the CXCR4 receptors, respectively. For both systems, this confirmed that the set of pocket conformations explored during our simulations, and before the peptides bind to their receptors, was much larger than that described by available X-ray structures, suggesting that similar results could have been obtained, whatever the structure used as a starting point.

However, because receptor/ligand recognition can also be affected by small, local conformational changes, it cannot be excluded that even small differences as observed here among the different X-ray structures could affect the final results. In agreement, further calculations are still required to definitively verify this hypothesis.

The protocol we describe here is not very expensive from a computational standpoint and could be easily used to simulate the binding of multiple peptides to multiple receptors. Calculations are still in progress to demonstrate whether such calculations could effectively help to predict peptide selectivity. Other calculations are in progress to simulate binding of other partners to GPCRs, including extra- and intracellular protein partners.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.6b00503.

Descriptions of the figures and movie (PDF)

Figure S1, representation of the typical system studied in this work (PDF)

Figure S2, temperature dependent overlaps of the total energy during REMD simulations (PDF)

Figure S3, analysis of the opening of the two receptors (PDF)

Figure S4, results from targeted MD simulations (PDF)

Figure S5, analyses of X-ray contacts for the three most populated clusters of each system (PDF)

Figure S6, analyses of deltaG predictions for the three most populated clusters of each system (PDF)

Figure S7, analysis of pocket conformational diversity of the simulations versus X-ray structures (PDF)

Movie of the binding of the NTS<sub>8–13</sub> peptide to its receptor (MPG)

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: nicolas.floquet@umontpellier.fr.

### ORCID

Jean Martinez: 0000-0002-4551-4254

Nicolas Floquet: 0000-0002-3883-2852

### Author Contributions

B.D. and P.R. contributed equally to this work.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank the University of Montpellier for financial support of Bartholomé Delort. We thank Agence Nationale de la Recherche (ANR) for financial support of Pedro Renault via the BIP-BIP “Investissement d’Avenir” project. We thank ANR “Ghredyn” for financial support. We thank DARI and CINES for attributing computational resources (project no. 2016077530). This study was also performed with the support of HPC@LR, a Centre of competence in High-Performance Computing from the “Région Languedoc Roussillon”. This work was also performed using HPC resources from CALMIP (Grant 2017-P17011). We thank Dr. May C. Morris for proofreading the manuscript.

## REFERENCES

- (1) Lagerström, M. C.; Schiöth, H. B. Structural Diversity of G Protein-Coupled Receptors and Significance for Drug Discovery. *Nat. Rev. Drug Discovery* **2008**, *7*, 339–357.
- (2) Venkatakrisnan, A. J.; Deupi, X.; Lebon, G.; Tate, C. G.; Schertler, G. F.; Babu, M. M. Molecular Signatures of G-Protein-Coupled Receptors. *Nature* **2013**, *494*, 185–194.
- (3) Rasmussen, S. G. F.; DeVree, B. T.; Zou, Y.; Kruse, A. C.; Chung, K. Y.; Kobilka, T. S.; Thian, F. S.; Chae, P. S.; Pardon, E.; Calinski, D.; Mathiesen, J. M.; Shah, S. T. A.; Lyons, J. A.; Caffrey, M.; Gellman, S. H.; Steyaert, J.; Skiniotis, G.; Weis, W. I.; Sunahara, R. K.; Kobilka, B. K. Crystal Structure of the  $\beta$ 2 Adrenergic Receptor-Gs Protein Complex. *Nature* **2011**, *477*, 549–555.
- (4) Zhou, X. E.; Gao, X.; Barty, A.; Kang, Y.; He, Y.; Liu, W.; Ishchenko, A.; White, T. A.; Yefanov, O.; Han, G. W.; Xu, Q.; de Waal, P. W.; Suino-Powell, K. M.; Boutet, S.; Williams, G. J.; Wang, M.; Li, D.; Caffrey, M.; Chapman, H. N.; Spence, J. C. H.; Fromme, P.; Weierstall, U.; Stevens, R. C.; Cherezov, V.; Melcher, K.; Xu, H. E. X-Ray Laser Diffraction for Structure Determination of the Rhodopsin-Arrestin Complex. *Sci. Data* **2016**, *3*, 160021.
- (5) White, J. F.; Noinaj, N.; Shibata, Y.; Love, J.; Kloss, B.; Xu, F.; Gvozdenovic-Jeremic, J.; Shah, P.; Shiloach, J.; Tate, C. G.; Grishammer, R. Structure of the Agonist-Bound Neurotensin Receptor. *Nature* **2012**, *490*, 508–513.
- (6) Egloff, P.; Hillenbrand, M.; Klenk, C.; Batyuk, A.; Heine, P.; Balada, S.; Schlinkmann, K. M.; Scott, D. J.; Schütz, M.; Plückthun, A. Structure of Signaling-Competent Neurotensin Receptor 1 Obtained by Directed Evolution in Escherichia Coli. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, E655–E662.
- (7) Wu, B.; Chien, E. Y. T.; Mol, C. D.; Fenalti, G.; Liu, W.; Katritch, V.; Abagyan, R.; Brooun, A.; Wells, P.; Bi, F. C.; Hamel, D. J.; Kuhn, P.; Handel, T. M.; Cherezov, V.; Stevens, R. C. Structures of the CXCR4 Chemokine Receptor in Complex with Small Molecule and Cyclic Peptide Antagonists. *Science* **2010**, *330*, 1066–1071.
- (8) Shan, Y.; Kim, E. T.; Eastwood, M. P.; Dror, R. O.; Seeliger, M. A.; Shaw, D. E. How Does a Drug Molecule Find Its Target Binding Site? *J. Am. Chem. Soc.* **2011**, *133*, 9181–9183.
- (9) Brooks, B. R.; Brooks, C. L., 3rd; Mackerell, A. D., Jr; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoseck, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545–1614.
- (10) Dror, R. O.; Pan, A. C.; Arlow, D. H.; Borhani, D. W.; Maragakis, P.; Shan, Y.; Xu, H.; Shaw, D. E. Pathway and Mechanism of Drug Binding to G-Protein-Coupled Receptors. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 13118–13123.
- (11) Hamelberg, D.; Mongan, J.; McCammon, J. A. Accelerated Molecular Dynamics: A Promising and Efficient Simulation Method for Biomolecules. *J. Chem. Phys.* **2004**, *120*, 11919–11929.
- (12) Tikhonova, I. G.; Selvam, B.; Ivetac, A.; Wereszczynski, J.; McCammon, J. A. Simulations of Biased Agonists in the  $\beta$ (2) Adrenergic Receptor with Accelerated Molecular Dynamics. *Biochemistry* **2013**, *52*, 5593–5603.
- (13) Grossfield, A.; Feller, S. E.; Pitman, M. C. Convergence of Molecular Dynamics Simulations of Membrane Proteins. *Proteins: Struct., Funct., Genet.* **2007**, *67*, 31–40.
- (14) Kappel, K.; Miao, Y.; McCammon, J. A. Accelerated Molecular Dynamics Simulations of Ligand Binding to a Muscarinic G-Protein-Coupled Receptor. *Q. Rev. Biophys.* **2015**, *48*, 479–487.
- (15) Hyeon, C.; Thirumalai, D. Capturing the Essence of Folding and Functions of Biomolecules Using Coarse-Grained Models. *Nat. Commun.* **2011**, *2*, 487.
- (16) Marrink, S. J.; de Vries, A. H.; Mark, A. E. Coarse Grained Model for Semiquantitative Lipid Simulations. *J. Phys. Chem. B* **2004**, *108*, 750–760.
- (17) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111*, 7812–7824.
- (18) Arnarez, C.; Marrink, S. J.; Periole, X. Molecular Mechanism of Cardiolipin-Mediated Assembly of Respiratory Chain Supercomplexes. *Chem. Sci.* **2016**, *7*, 4435–4443.
- (19) Periole, X.; Huber, T.; Marrink, S.-J.; Sakmar, T. P. G Protein-Coupled Receptors Self-Assemble in Dynamics Simulations of Model Bilayers. *J. Am. Chem. Soc.* **2007**, *129*, 10126–10132.
- (20) Charlier, L.; Louet, M.; Chaloin, L.; Fuchs, P.; Martinez, J.; Muriaux, D.; Favard, C.; Floquet, N. Coarse-Grained Simulations of the HIV-1 Matrix Protein Anchoring: Revisiting Its Assembly on Membrane Domains. *Biophys. J.* **2014**, *106*, 577–585.
- (21) Fanelli, R.; Besserer-Offroy, É.; René, A.; Côté, J.; Tétréault, P.; Collette-Tremblay, J.; Longpré, J.-M.; Leduc, R.; Martinez, J.; Sarret, P.; Cavelier, F. Synthesis and Characterization in Vitro and in Vivo of (L)-(Trimethylsilyl)alanine Containing Neurotensin Analogues. *J. Med. Chem.* **2015**, *58*, 7785–7795.
- (22) DeMarco, S. J.; Henze, H.; Lederer, A.; Moehle, K.; Mukherjee, R.; Romagnoli, B.; Robinson, J. A.; Brianza, F.; Gombert, F. O.; Locuro, S.; Ludin, C.; Vrijbloed, J. W.; Zumbunn, J.; Obrecht, J.-P.; Obrecht, D.; Brondani, V.; Hamy, F.; Klimkait, T. Discovery of Novel, Highly Potent and Selective Beta-Hairpin Mimetic CXCR4 Inhibitors with Excellent Anti-HIV Activity and Pharmacokinetic Profiles. *Bioorg. Med. Chem.* **2006**, *14*, 8396–8404.
- (23) Coutant, J.; Curmi, P. A.; Toma, F.; Monti, J.-P. NMR Solution Structure of Neurotensin in Membrane-Mimetic Environments:

Molecular Basis for Neurotensin Receptor Recognition. *Biochemistry* **2007**, *46*, 5656–5663.

(24) Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinf.* **2014**, *47*, 5.6.1–5.6.32.

(25) Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y.; Jo, S.; Pande, V. S.; Case, D. A.; Brooks, C. L.; MacKerell, A. D.; Klauda, J. B.; Im, W. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **2016**, *12*, 405–413.

(26) Qi, Y.; Ingólfsson, H. I.; Cheng, X.; Lee, J.; Marrink, S. J.; Im, W. CHARMM-GUI Martini Maker for Coarse-Grained Simulations with the Martini Force Field. *J. Chem. Theory Comput.* **2015**, *11*, 4486–4494.

(27) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

(28) Tamamura, H.; Hori, A.; Kanzaki, N.; Hiramatsu, K.; Mizumoto, M.; Nakashima, H.; Yamamoto, N.; Otaka, A.; Fujii, N. T140 Analogs as CXCR4 Antagonists Identified as Anti-Metastatic Agents in the Treatment of Breast Cancer. *FEBS Lett.* **2003**, *550*, 79–83.

(29) Daura, X.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. Reversible Peptide Folding in Solution by Molecular Dynamics Simulation. *J. Mol. Biol.* **1998**, *280*, 925–932.

(30) Bakan, A.; Meireles, L. M.; Bahar, I. ProDy: Protein Dynamics Inferred from Theory and Experiments. *Bioinformatics* **2011**, *27*, 1575–1577.

(31) Wu, E. L.; Cheng, X.; Jo, S.; Rui, H.; Song, K. C.; Dávila-Contreras, E. M.; Qi, Y.; Lee, J.; Monje-Galvan, V.; Venable, R. M.; Klauda, J. B.; Im, W. CHARMM-GUI Membrane Builder Toward Realistic Biological Membrane Simulations. *J. Comput. Chem.* **2014**, *35*, 1997–2004.

(32) Xue, L. C.; Rodrigues, J. P.; Kastiris, P. L.; Bonvin, A. M.; Vangone, A. PRODIGY: A Web Server for Predicting the Binding Affinity of Protein-Protein Complexes. *Bioinformatics* **2016**, *32*, 3676–3678.

(33) Vangone, A.; Bonvin, A. M. J. J. Contacts-Based Prediction of Binding Affinity in Protein-Protein Complexes. *eLife* **2015**, *4*, e07454.

(34) Kastiris, P. L.; Moal, I. H.; Hwang, H.; Weng, Z.; Bates, P. A.; Bonvin, A. M. J. J.; Janin, J. A Structure-Based Benchmark for Protein-Protein Binding Affinity. *Protein Sci. Publ. Protein Soc.* **2011**, *20*, 482–491.

(35) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

(36) Deupi, X.; Kobilka, B. K. Energy Landscapes as a Tool to Integrate GPCR Structure, Dynamics, and Function. *Physiology* **2010**, *25*, 293–303.

(37) Bhattacharya, S.; Vaidehi, N. Computational Mapping of the Conformational Transitions in Agonist Selective Pathways of a G-Protein Coupled Receptor. *J. Am. Chem. Soc.* **2010**, *132*, 5205–5214.

(38) Niesen, M. J. M.; Bhattacharya, S.; Vaidehi, N. The Role of Conformational Ensembles in Ligand Recognition in G-Protein Coupled Receptors. *J. Am. Chem. Soc.* **2011**, *133*, 13197–13204.

(39) Ye, L.; Van Eps, N.; Zimmer, M.; Ernst, O. P.; Prosser, R. S. Activation of the A2A Adenosine G-Protein-Coupled Receptor by Conformational Selection. *Nature* **2016**, *533*, 265–268.

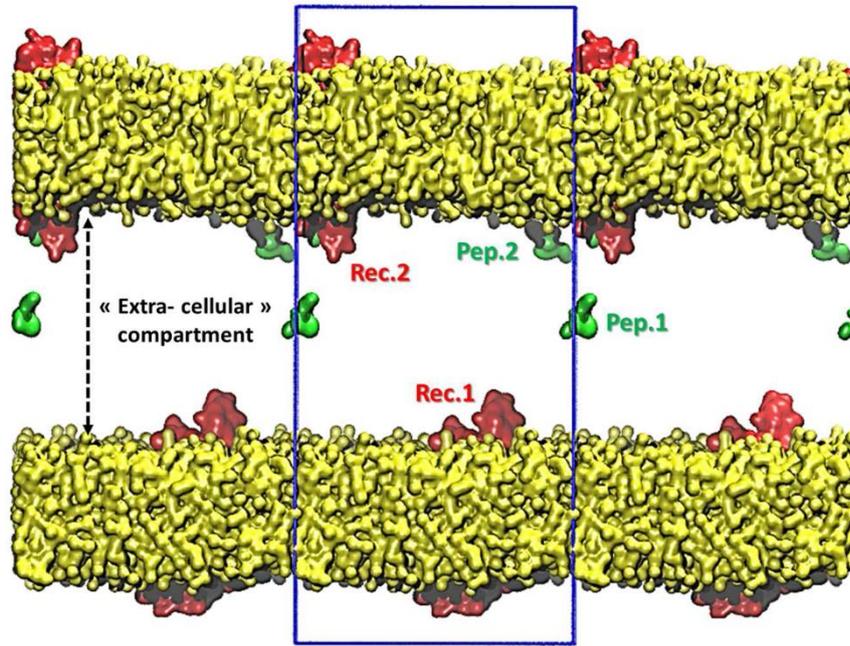


Figure S1: The typical system used in this study (NTSR1 receptor).

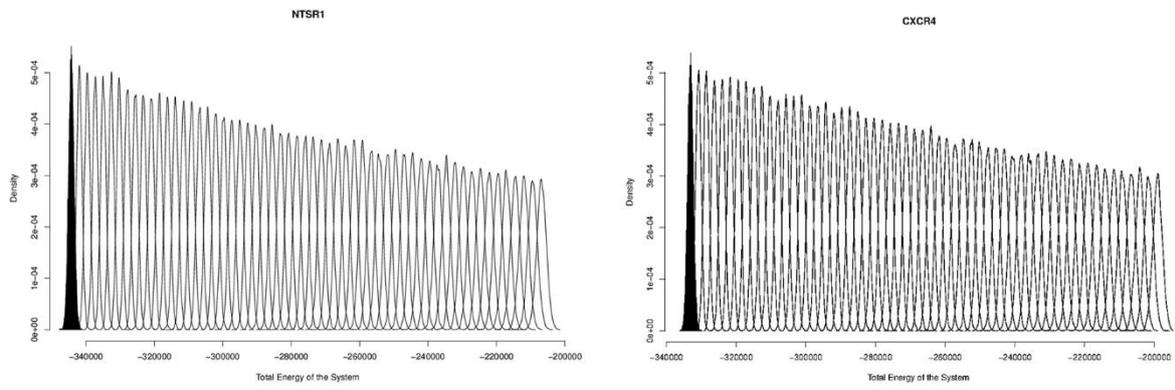


Figure S2: Energy distribution of each system according to the 58 different temperature.

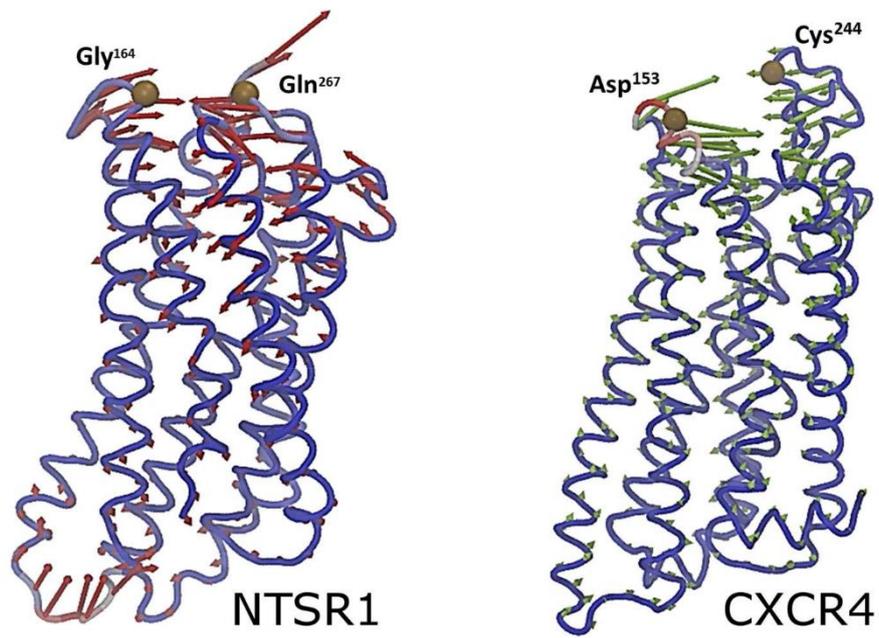


Figure S3: Direction of the motion along the first principal component of each receptor.

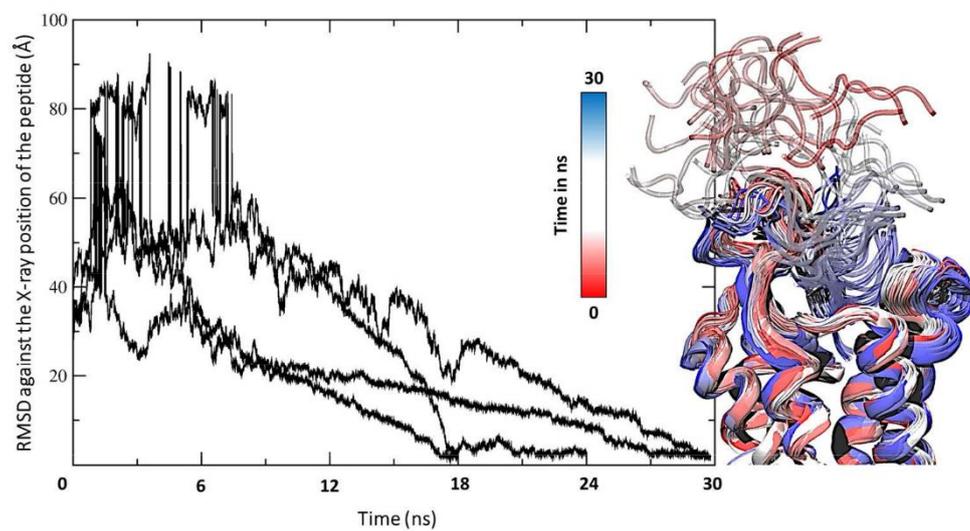


Figure S4: Results from Targeted MD simulations performed on the  $\text{NTS}_{8-13}$ :NTSR1 complex.

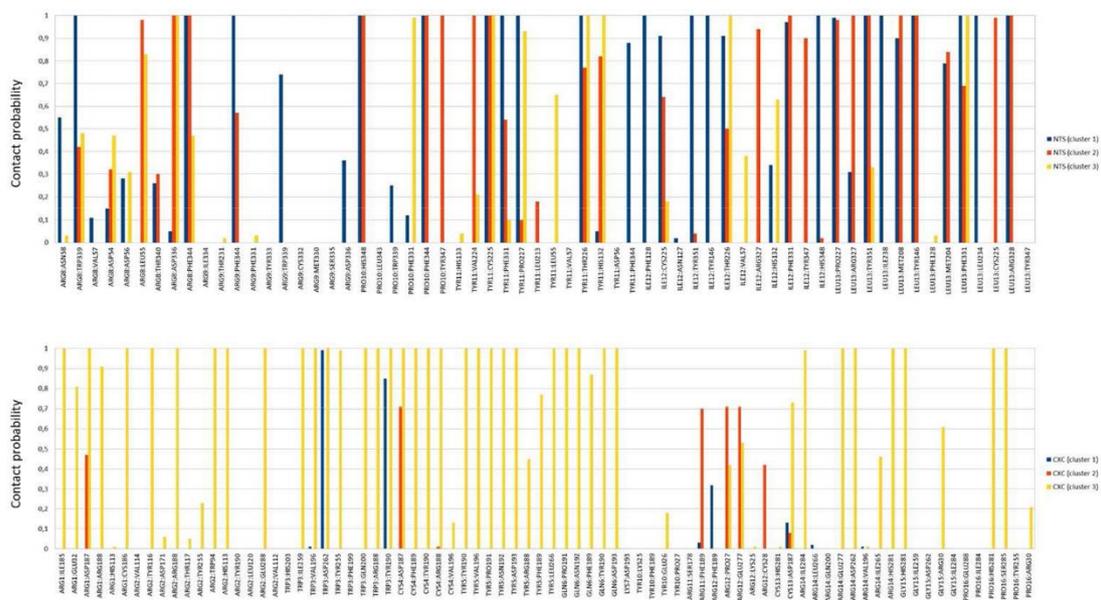


Figure S5: Percentage of each native X-ray contact along all-atom MD simulations of the NTSR1 (upper graph) and CXCR4 (lower graph) receptors. The max value was 100% in each case.

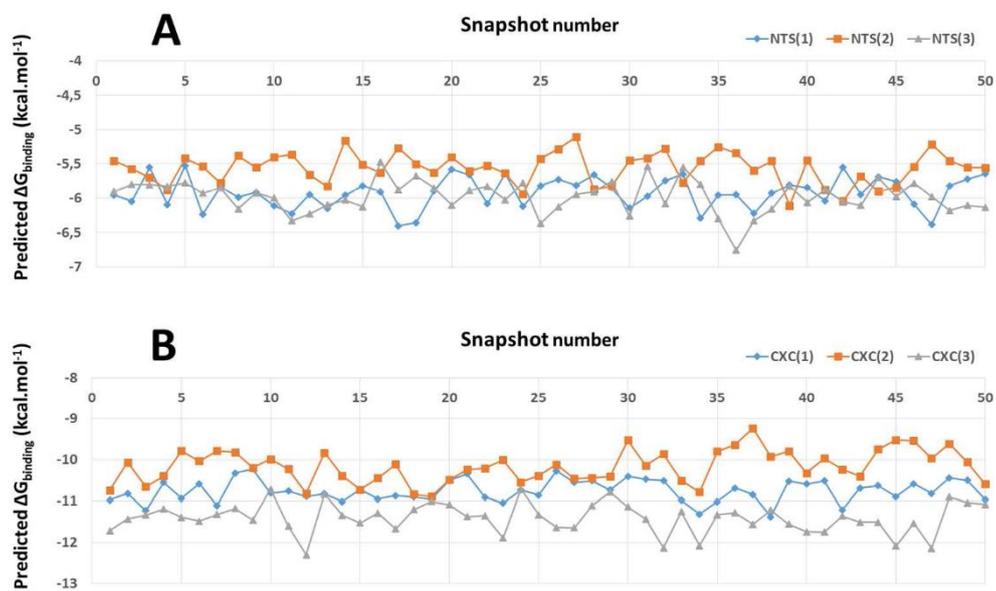


Figure S6:  $\Delta G_{\text{binding}}$  values predicted for the three most populated clusters of NTSR1 (A) and CXCR4 (B) receptors after all-atoms MD simulations.

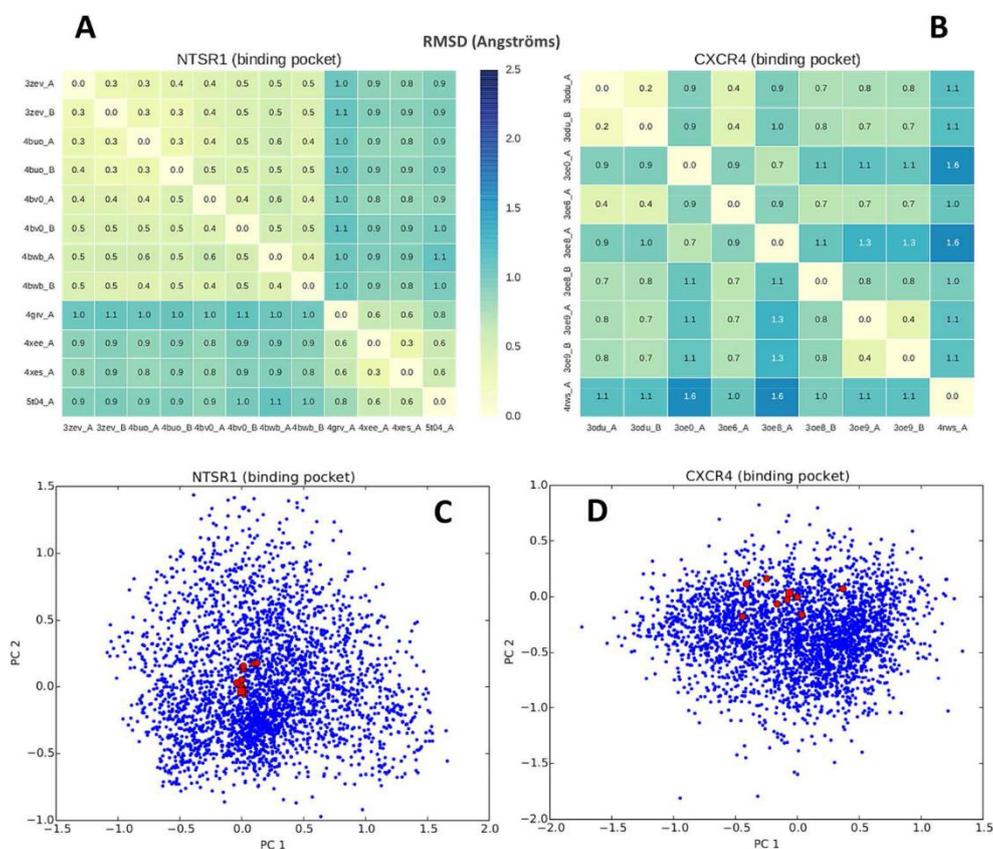


Figure S7: Cross-RMSD analyses of all X-ray structures available for the NTSR1 (A) and CXCR4 (B) receptors. These RMSD values were computed on pocket CG beads. (C,D) After PCA of a conformational set composed of these structures and the conformation sampled during the equilibration phase, all these configurations were projected onto the space of the first two PCs. The X-ray structures used as starting points of our simulations are reported as red squares. The other X-ray structures are represented by red spheres. Conformations from equilibration trajectories are reported in blue.

### 3.1.1.c Commentaires

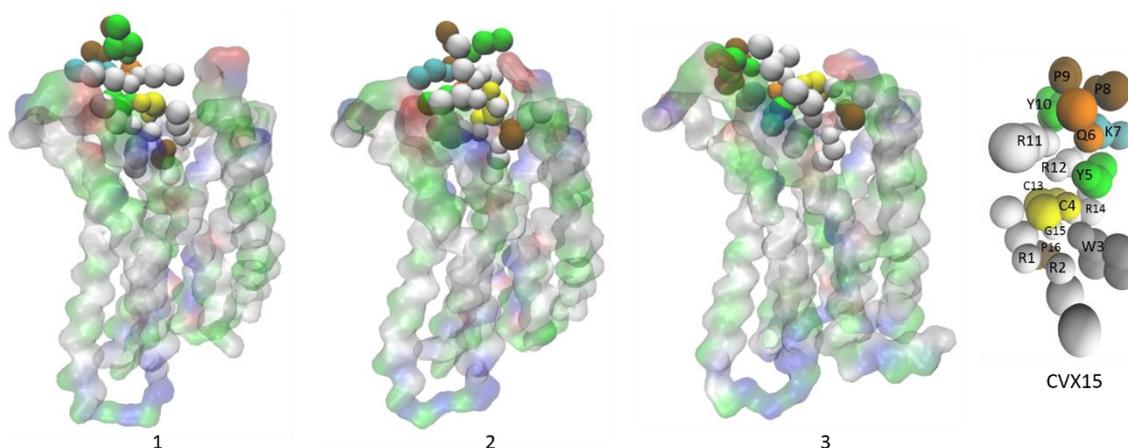
Cet article a été soumis le 25 Août 2016 au journal « *Journal of Chemical Information and Modeling* » pour être finalement publié le 23 Février 2017. Ce délai important s'explique par deux phases successives de révisions majeures. Parmi les différentes remarques, les deux plus importantes concernèrent le choix du modèle de départ.

Premièrement, puisque nous utilisons un modèle de récepteur déjà lié au ligand, un des évaluateurs nous a demandé de prouver que nous aurions obtenu le même résultat en partant d'un autre modèle du récepteur, par exemple présentant un ligand différent, et donc une conformation de la poche différente. En effet, comme c'est le cas par exemple lors d'expériences de docking dans lesquelles la structure du récepteur est rigide, il est bien plus facile de prédire la position d'un ligand à son récepteur lorsque la structure initiale contient déjà ce ligand, ou un ligand de la même famille. Même si dans notre approche dynamique, les degrés de liberté du peptide et du récepteur sont explorés, une des réserves émises concernait l'utilisation du réseau élastique qui aurait permis au récepteur de conserver la conformation de la poche, prête à accueillir le peptide. En réalisant des calculs de dynamique CG des deux récepteurs NTR1r et CXCR4, cette fois sans ligand, nous avons pu montrer un équilibre dynamique entre des conformations ouvertes et fermées de la poche (**Figure 5** et **S3** de l'article). Ce dernier ne reste donc pas bloqué dans la conformation de départ. Pour définitivement répondre à ce point, nous avons néanmoins comparé toutes les conformations de CXCR4 et NTR1r disponibles dans la PDB, (**Figure S7** de l'article) à celles obtenues lors de la phase d'équilibrage de nos systèmes aux différentes températures. On montre ainsi pour les deux récepteurs un large échantillonnage de la conformation de la poche de liaison, bien plus large que celui déduit des structures disponibles, démontrant par la même occasion que le choix d'une autre structure de départ aurait eu peu d'influence sur nos résultats.

La seconde interrogation des évaluateurs portait sur la robustesse de la méthode et sa capacité à prédire la liaison d'un peptide à un modèle de RCPG cette fois construit par homologie. Dans cet article, nous ne présentons aucun résultat impliquant un récepteur construit par homologie, l'objectif étant ici avant tout de faire la preuve de concept de la méthode. Néanmoins, pour répondre à cette question, nous avons construit un modèle de CXCR4 basé sur la structure du récepteur  $\kappa$ -Opioïdes humain (code PDB : 4DJH) [H. Wu et al. 2012] avec lequel il possède 28% d'identité. En appliquant le même protocole, nous avons pu montrer qu'il était encore possible d'obtenir la position cristallographique parmi les clusters les plus peuplés

### 3 Résultats

comme montré dans l'étude originale. Plus précisément, la conformation la plus proche de la structure cristallographique correspondait initialement au 3<sup>ème</sup> cluster le plus peuplé. Avec le modèle construit par homologie, ce cluster apparaît même en première position (cluster le plus peuplé) (**Figure 42**). Même si cela pourrait être tentant de parler de « meilleurs résultats » obtenus avec le modèle par homologie, nous nous contenterons de dire ici que l'ordre des premiers clusters reste difficile à reproduire et peut être dû à une limite de l'échantillonnage ; par contre, nous montrons bien ici la capacité de notre protocole à prédire la liaison de peptides à des RCPGs pour lesquels aucune structure cristallographique n'est encore disponible, ce qui représente la très grande majorité de ces récepteurs. Il est d'ailleurs important de noter ici que dans la plupart des autres applications décrites dans cette thèse nous sommes partis de modèles construits par homologie.



*Figure 42 : Comparaison de la structure cristallographique et des conformations obtenues lors des simulations en gros-grains effectuées sur le récepteur CXCR4, le modèle 1 correspond à la structure de référence (code PDB : 3OE0), le modèle 2 représente le centre du cluster 3 (3<sup>ème</sup> cluster le plus peuplé) issu de la simulation CG-REMD utilisant cette structure comme point de départ, le modèle 3 représente le centre du cluster 1 (le plus peuplé) issu de la simulation CG-REMD utilisant un récepteur CXCR4 construit par homologie.*

Il est à noter enfin, que nous avons observé une différence importante entre la position connue en cristallographie de la Neurotensine et celle prédite par nos calculs. En effet, si la chaîne principale du peptide est placée exactement à la même position, le résidu Tyrosine en position 11 présente deux orientations possibles de sa chaîne latérale : soit vers des histidines situées sous la boucle ECL1 comme montré en cristallographie (orientation « C »), soit vers la boucle ECL3 (orientation « S ») (**Figure 43A**). Lors de nos simulations l'orientation « C » n'est que très rarement observée. Pour essayer de comprendre ce phénomène, nous avons réalisé des dynamiques CG en partant des deux conformations de la NT[8-13] déjà liée. En partant de

l'orientation « C », on observe très rapidement une réorientation de la chaîne latérale de la Tyrosine, celle-ci se relevant d'abord en direction de la boucle ECL2 (**Figure 43B**) puis se plaçant dans la conformation « S ». Sur une échelle de temps de l'ordre de la  $\mu$ s, on observe un équilibre entre ces deux conformations « S » et « C », très nettement en faveur de l'orientation « S ». Cet équilibre est retrouvé à l'identique lors de simulations où cette fois l'orientation « S » a été choisie comme modèle de départ. Il est difficile de discuter exactement de l'origine de ce possible réarrangement : simple limitation du modèle ou bien capture par la cristallographie d'un seul des deux états possibles de la NT liée ? Toutefois, une donnée importante en ressort : on observe une possible interaction du résidu Y11 du peptide avec la boucle ECL2 du récepteur. C'est une information qui sera importante pour l'étude présentée dans le chapitre sur le design de peptides sélectifs par dynamique moléculaire (voir partie 3.2.1).

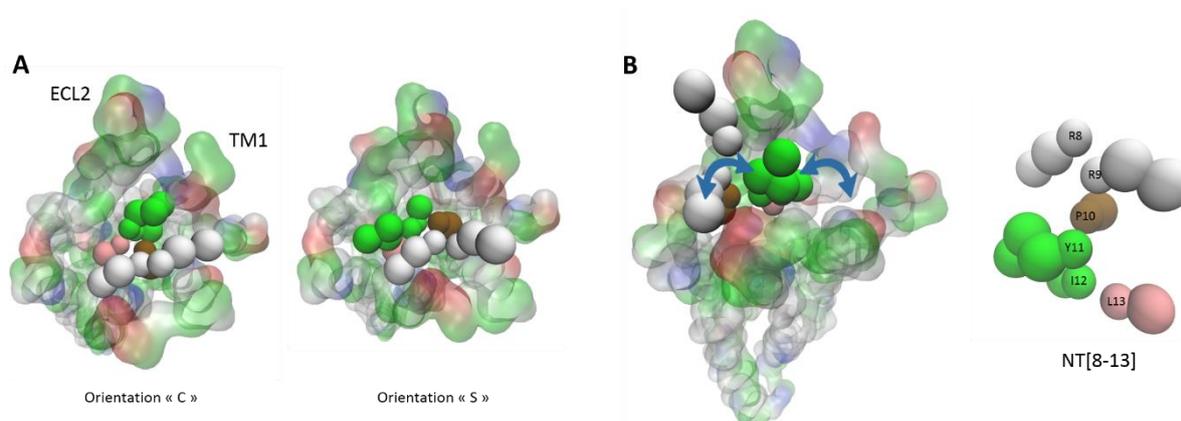


Figure 43 : A) Conformations adoptées par la Tyrosine du peptide NT[8-13] lié au récepteur de la Neurotensine; B) Conformation intermédiaire de la Tyrosine capable de basculer vers les orientations « C » et « S » et qui montre une interaction directe avec ECL2.

### 3.1.1.d Le cas des peptides « scrambled »

Après avoir validé notre approche sur deux complexes connus en cristallographie et montré que la structure de départ du récepteur avait peu d'influence sur nos résultats, nous nous sommes posé la question suivante : en CG-REMD, le récepteur peut-il faire la différence entre plusieurs peptides ? Autrement dit, est-ce que le modèle utilisé pourrait permettre de discriminer la liaison de plusieurs peptides à un même récepteur ? Pour répondre à cette question, j'ai construit trois peptides dits « scrambled » de NT[8-13], c'est-à-dire présentant la même composition en acides-aminés que le peptide « natif » mais dans un ordre aléatoire, et donc (en théorie) incapables de se lier au récepteur (**Tableau 7**). Le protocole de CG-REMD fut ensuite appliqué à chacun de ces peptides en utilisant le même modèle du récepteur NTR1r.

Peptide	Séquence
NT[8-13]	$\text{NH}_3^+\text{-Arg-Arg-Pro-Tyr-Ile-Leu-CO}_2^-$
S1	$\text{NH}_3^+\text{-Ile-Pro-Arg-Leu-Arg-Tyr-CO}_2^-$
S2	$\text{NH}_3^+\text{-Leu-Tyr-Ile-Arg-Pro-Arg-CO}_2^-$
S3	$\text{NH}_3^+\text{-Tyr-Ile-Arg-Arg-Leu-Pro-CO}_2^-$

Tableau 7 : Séquences du peptide Neurotensine et de ses peptides « scrambled » testés dans nos simulations.

Nos premières analyses ont montré que les peptides « scrambled » étaient capables de se lier au récepteur et ce malgré leur séquence aléatoire. A 300K, nous avons observé une probabilité de présence de ces peptides dans la poche du récepteur de l'ordre de 80%. En comparaison, le peptide natif NT[8-13] présente une liaison à plus de 95% à la même température. Objectivement, cette différence entre le peptide natif et les peptides « scrambled » serait difficile à discuter en termes de prédiction, et il paraît ici impossible de parler de peptides capables ou non de se lier à leur récepteur.

Pour discuter de ces aspects et de la stabilité de ces peptides dans la poche de liaison de leur récepteur, nous nous sommes donc intéressé au pourcentage de liaison de chacun des peptides à chacune des températures simulées, sans faire de différence ici entre les différentes conformations liées du peptide (**Figure 44**). En procédant ainsi, la différence entre le peptide « natif » et les peptides « scrambled » est alors bien plus marquée. On observe en effet que le peptide natif présente une courbe stable et ce, malgré l'augmentation de la température, indiquant qu'il se décroche du récepteur bien plus difficilement que les peptides « scrambled ».

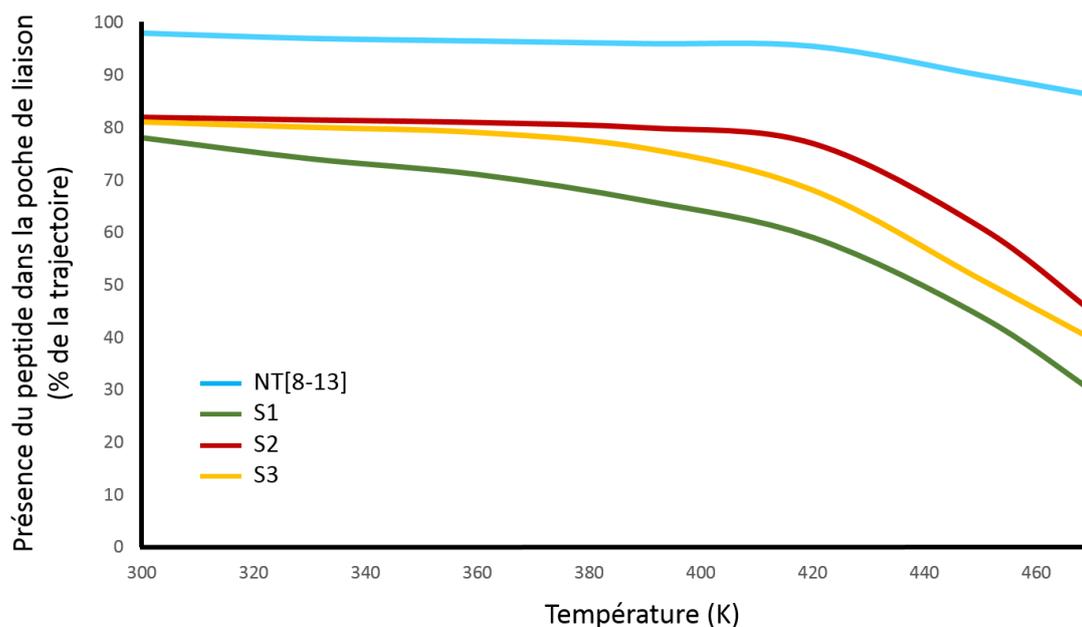


Figure 44 : Evolution du pourcentage de la liaison du peptide natif et des peptides « scrambled » de la Neurotensine en fonction de la température lors de nos simulations CG-REMD.

Ces résultats encourageants nous ont amené par la suite à étudier avec le même protocole la liaison de plusieurs peptides dérivés de la Neurotensine aux deux récepteurs humains NTR1 et NTR2 avec pour objectif de concevoir des peptides sélectifs du récepteur NTR2 et surtout de comprendre les mécanismes mis en jeu dans la sélectivité NTR2/NTR1 (voir partie 3.2.2).

## 3.2 Prédiction de la sélectivité de peptides à différents récepteurs

Après l'étape de validation de notre protocole et la mise en avant de sa capacité à prédire du point de vue dynamique la liaison d'un peptide à un RCPG, nous avons cherché à étendre son champ d'application à la prédiction de la sélectivité de différents peptides pour différents récepteurs. Aussi, nous avons utilisé nos données dynamiques obtenues sur le récepteur de la Neurotensine pour concevoir des peptides présentant une sélectivité NTR2/NTR1 et comprendre cette sélectivité. Nous nous sommes pour cela basés sur notre observation par le calcul de deux orientations possibles de la chaîne latérale de la Tyrosine 11 de la Neurotensine liée. En effet, des études expérimentales ont montré par le passé le rôle de ce résidu du peptide dans la sélectivité de la Neurotensine [Boules et al. 2010, 2011; Einsiedel et al. 2011; Hapău et al. 2016].

### 3.2.1 Design de peptides sélectifs de NTR2 par dynamique moléculaire

#### 3.2.1.a Contexte

La Neurotensine et son récepteur sont des peptide et protéine ayant fait l'objet de très nombreuses études au sein du laboratoire [Belmeguenai et al. 2000; Sarret et al. 2005; Bredeloux et al. 2008; Fanelli et al. 2015]. La NT peut, entre autres récepteurs, se lier avec une affinité équivalente aux deux récepteurs de la Neurotensine NTR1 et NTR2. Ces récepteurs sont impliqués dans une pléiade de rôles physiologiques. Le développement de molécules spécifiques du NTR2 trouve un intérêt pharmacologique dans le traitement des douleurs neuropathiques, car évite les effets indésirables liés à la liaison au récepteur NTR1 et qui incluent l'hypothermie et l'hypotension [Dobner 2006; Osadchii 2015].

Ainsi, des simulations de dynamique moléculaire tout-atomes avaient été lancées à mon arrivée au laboratoire, visant à étudier la stabilité du peptide de la Neurotensine dans la poche orthostérique des deux récepteurs humains NTR1 et NTR2, et ce en partant de la structure cristallographique du récepteur NTR1 de rat (code PDB : 4GRV) [White et al. 2012]. Dans ce travail, j'ai construit les deux récepteurs humains de la Neurotensine par homologie en utilisant comme structure de référence la structure du récepteur de rat, mais surtout en conservant le

peptide lié dans la poche tout au long du processus de construction. Pour ces deux nouveaux complexes, des simulations de 180ns de dynamique moléculaire tout-atomes ont été réalisées. De façon très intéressante, les trajectoires obtenues ont montré un même réarrangement du peptide et la rotation de la chaîne latérale de son résidu Y11 comme observé lors de nos simulations gros-grains (voir partie 3.1.1). Ces calculs confirment ainsi que ce résidu peut établir un contact polaire avec un résidu du récepteur situé dans la boucle ECL2. Or ce résidu est très différent entre les récepteurs NTR1 et NTR2 qui présentent respectivement à cette position une Arginine (R212) et un Glutamate (E179). Nous avons par la suite exploité cette hypothèse d'un contact du résidu Y11 essentiel pour la sélectivité du peptide [Einsiedel et al. 2011] avec deux résidus de charges opposées au niveau du récepteur.

Notre hypothèse de travail était alors de remplacer le résidu Y11 par un résidu chargé positivement ou négativement, pour obtenir des peptides sélectifs des deux récepteurs NTR1 et NTR2. Plusieurs peptides ont été synthétisés au laboratoire (collaboration avec Roberto FANELLI et Florine CAVELIER, Institut des Biomolécules Max Mousseron, Montpellier) avant d'être testés sur les deux récepteurs (collaboration avec Philippe SARRET, Université de Sherbrooke, Canada). Comme prédit par nos modèles, l'introduction d'une charge positive en position 11 (notamment d'une lysine) augmente la sélectivité des peptides vis-à-vis du récepteur NTR2. L'implication du résidu E179/R212 du récepteur dans cette sélectivité a pu également être confirmée *in vitro* par la production d'un récepteur NTR1 muté (R212E), mimant le récepteur NTR2, et présentant une affinité pour le peptide NT:Y11K (11Lys) proche de celle observée dans le récepteur NTR2 natif. Ces résultats très intéressants confirment donc un contact entre le résidu 11 du peptide et la boucle ECL2 des deux récepteurs, à l'origine de la sélectivité.

### 3.2.1.b Article

« Reprinted with the permission from [Fanelli, Roberto, Nicolas Floquet, Élie Besserer-Offroy, Bartholomé Delort, Mélanie Vivancos, Jean-Michel Longpré, Pedro Renault, Jean Martinez, Philippe Sarret, et Florine Cavelier. 2017. *Journal of Medicinal Chemistry* 60 (8): 3303-13]. Copyright 2017 American Chemical Society. »

## Use of Molecular Modeling to Design Selective NTS2 Neurotensin Analogues

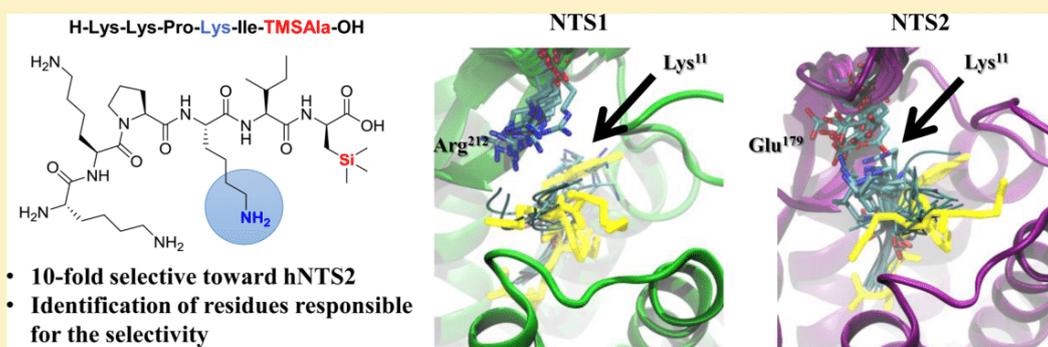
Roberto Fanelli,<sup>†</sup> Nicolas Floquet,<sup>†,‡</sup> Élie Besserer-Offroy,<sup>‡</sup> Bartholomé Delort,<sup>†</sup> Mélanie Vivancos,<sup>‡</sup> Jean-Michel Longpré,<sup>‡</sup> Pedro Renault,<sup>†,§</sup> Jean Martinez,<sup>†,‡</sup> Philippe Sarret,<sup>‡,||</sup> and Florine Cavalier<sup>\*,†,||,‡</sup>

<sup>†</sup>Institut des Biomolécules Max Mousseron, IBMM, UMR-5247, CNRS, Université Montpellier, ENSCM, Place Eugène Bataillon, 34095 Montpellier cedex 5, France

<sup>‡</sup>Department of Pharmacology and Physiology, Institut de Pharmacologie de Sherbrooke, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Sherbrooke J1H 5N4, Québec, Canada

<sup>§</sup>Centre de Biochimie Structurale (CBS), UMR-5048, CNRS, Université de Montpellier, INSERM U1054, 29 rue de Navacelles, 34090 Montpellier, France

### Supporting Information



**ABSTRACT:** Neurotensin exerts potent analgesia by acting at both NTS1 and NTS2 receptors, whereas NTS1 activation also results in other physiological effects such as hypotension and hypothermia. Here, we used molecular modeling approach to design highly selective NTS2 ligands by investigating the docking of novel NT[8-13] compounds at both NTS1 and NTS2 sites. Molecular dynamics simulations revealed an interaction of the Tyr<sup>11</sup> residue of NT[8-13] with an acidic residue (Glu<sup>179</sup>) located in the ECL2 of hNTS2 or with a basic residue (Arg<sup>212</sup>) at the same position in hNTS1. The importance of the residue at position 11 for NTS1/NTS2 selectivity was further demonstrated by the design of new NT analogues bearing basic (Lys, Orn) or acid (Asp or Glu) function. As predicted by the molecular dynamics simulations, binding of NT[8-13] analogues harboring a Lys<sup>11</sup> exhibited higher affinity toward the hNTS1-R212E mutant receptor, in which Arg<sup>212</sup> was substituted by the negatively charged Glu residue.

### INTRODUCTION

Chronic pain is the most common reason for medical consultations, and conventional opioid analgesics remain the pharmacological cornerstone of pain therapy. However, despite their proven analgesic efficacy, the management of the opioid adverse effects still represents a major clinical challenge.<sup>1</sup> There is therefore an urgent need to develop new lead compounds for improving the pain pharmacopeia.

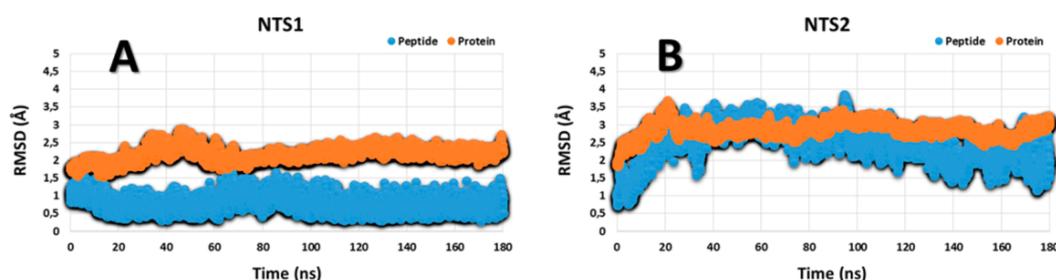
Neurotensin (NT) is a tridecapeptide first isolated in 1973 by Carraway and Leeman from bovine hypothalamus extract.<sup>2</sup> Following central administration, NT exerts a wide range of biological functions including hypothermia,<sup>3</sup> analgesia,<sup>4</sup> and antipsychotic-like properties.<sup>5</sup> Importantly, the analgesic activity of NT, mediated by its interaction with two G protein-coupled receptors, namely NTS1 and NTS2, was found to be as potent as that of opioids, such as morphine.<sup>6</sup> Additionally, it was notably reported that NTS1 activation

results in significant antinociception but also causes marked hypotension and hypothermia.<sup>3</sup> In sharp contrast, the neurotensin receptor subtype 2 (NTS2) has emerged as an important pain target because NTS2-selective analogues exhibit potent analgesic activity in both acute and chronic pain conditions in dose-dependent analgesic effects without inducing drop in blood pressure or body temperature.<sup>7</sup> Thus, the development of new NTS2-selective agonists acting as nonopioid analgesics could be useful for relieving chronic pain.

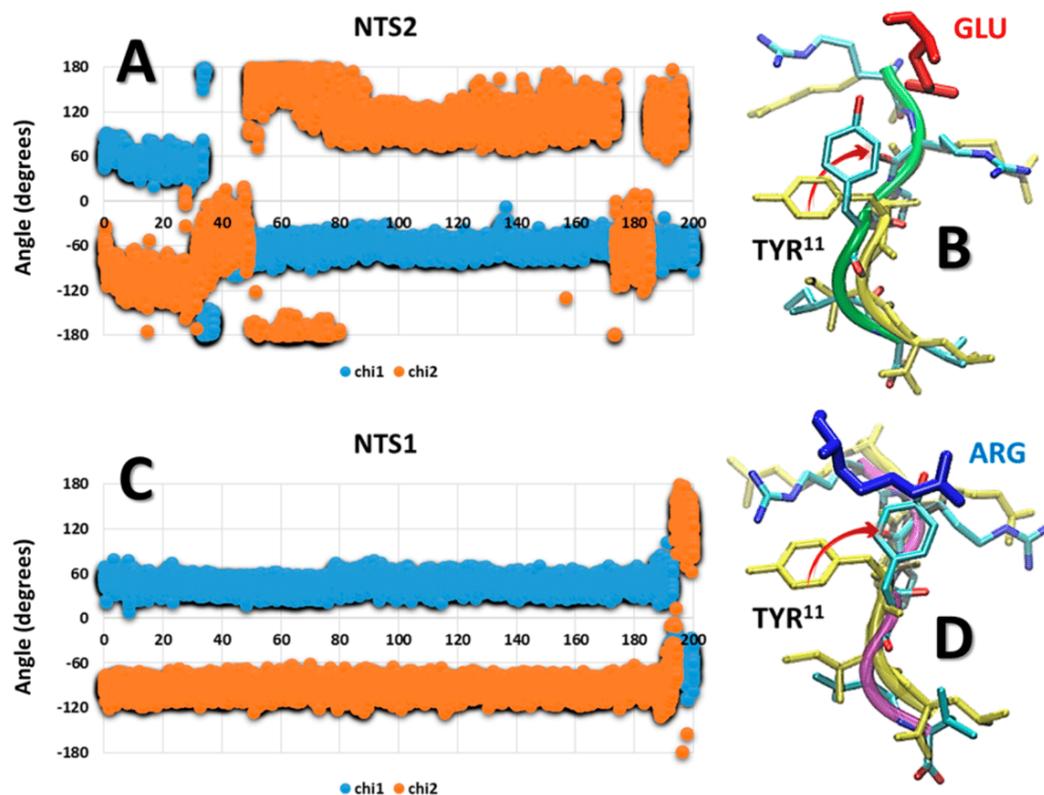
To develop new potent and selective NTS2 receptor ligands as possible pain treatment, there is a crucial need to provide valuable information on the structure–activity relationship (SAR) of the ligand–receptor interactions, which serve as a tool to the drug design process. To date, the design of

Received: December 20, 2016

Published: April 3, 2017



**Figure 1.** RMSD values computed on the backbone atoms of either receptor (orange) or peptide (blue) along the two trajectories obtained for hNTS1 (A) and hNTS2 (B) receptors, respectively. In each case, the receptor was bound to the native NT[8-13] peptide.



**Figure 2.** Chi1 (in blue) and chi2 (in orange) values of the Tyr<sup>11</sup> residue in the NT [8-13] peptide along the two trajectories computed for NTS2 (A) and NTS1 (C) receptors (180 ns of classical MD + 20 ns of accelerated MD). (B,D) Orientation of the Tyr<sup>11</sup> residue after its rotation in both receptors. In yellow is reported the reference position of the peptide in the rat NTS1 X-ray structure with an orientation of Tyr<sup>11</sup> corresponding to initial chi1 and chi2 values of 53 and -93 degrees, respectively.

minimum active fragments of the biologically active NT tridecapeptide has led to the identification of the six carboxy-terminal residues (Arg<sup>8</sup>-Arg<sup>9</sup>-Pro<sup>10</sup>-Tyr<sup>11</sup>-Ile<sup>12</sup>-Leu<sup>13</sup>, NT[8-13]),<sup>8,9</sup> therefore constituting a good starting point to provide insight into the molecular determinants of the ligand–receptor interactions. Interestingly, we recently demonstrated that the replacement of the carboxy-terminal residue of NT[8-13] by a (L)-(trimethylsilyl)alanine (TMSAla) enhanced its binding affinity for both NTS1 and NTS2 by improving its interaction with the hydrophobic ligand binding pocket of both receptors.<sup>10</sup> Additionally, position 11 has been found to be critical for the selectivity toward NTS2 but no clear SAR has emerged from these studies.<sup>11–14</sup> Recently, the X-ray crystallography structure of the rat NTS1 receptor was resolved in an agonist-bound conformation.<sup>15</sup> This crystal structure of rNTS1 bound to its agonist peptide NT[8-13] has allowed

deciphering of the key features governing receptor–ligand interactions and now gives the structural basis for investigating the effects of bound ligands on receptor dynamics.

In the present study, starting from the high-resolution crystal data of NTS1, we built structural models for both hNTS1 and hNTS2 receptors bound to the NT[8-13] peptide. Our results revealed the existence of some divergence at the ligand–receptor interface, especially the NT receptor residues facing the tyrosine residue in position 11 of the NT[8-13] peptide, namely residue 212 in NTS1 and residue 179 in NTS2. On the basis of these observations, we synthesized a series of NT[8-13] analogues modified at position 11 to evaluate their selectivity toward NTS1 and NTS2 receptors. As predicted by the molecular dynamics simulations, our binding experiments were able to demonstrate that the residue at position 11 in NT[8-13] was crucial for the NTS1/NTS2 selectivity. Finally, to further

support these findings, we found that a single amino acid mutation at position 212 in the second extracellular loop of the NTS1 receptor (hNTS1-R212E) was able to restore affinity to NTS2-selective analogues. Altogether, these results will help to improve the design of the next-generation of NT[8-13] analogues acting at the NTS2 receptor site.

## RESULTS AND DISCUSSION

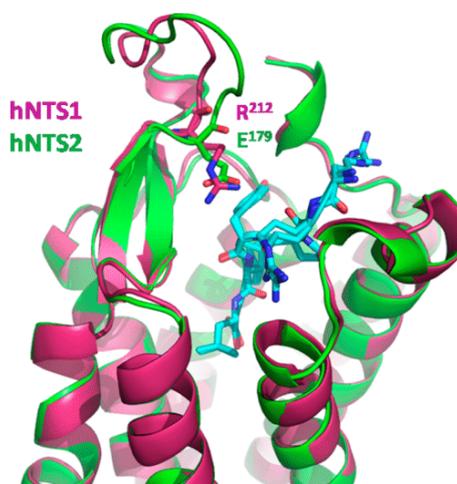
**Molecular Modeling.** Molecular dynamics simulations were performed on both human NTS1 and NTS2 receptors, each trajectory reaching 180 ns. In agreement with the known NTS1 X-ray structure used as an initial template for homology modeling (PDB 4BUO), both hNTS1 receptor conformation and NT[8-13] ligand orientation inside the binding pocket were remarkably stable all along the trajectory, with mean root mean square deviations (RMSD) of 2.4 and 0.8 Å, respectively (Figure 1A). For the hNTS2 model, we observed an increased flexibility of both receptor (mean RMSD of 2.9 Å) and peptide, with a RMSD oscillating between 2.0 and 3.0 Å (Figure 1B).

Initially, this slight increase of the NTS2 receptor flexibility was attributed to its reduced rate of sequence identity with the rat NTS1 receptor (44% versus 88% for the hNTS1/rNTS1 couple). However, after inspection of both peptide and receptor conformations, it appeared that this RMSD increase was mainly due to a concerted motion of both the side-chain of the Tyr<sup>11</sup> residue in NT[8-13] and the extracellular loop 2 (ECL2) of the receptor. This rotation was clearly established, showing a concerted shift of the two  $\chi_1$  and  $\chi_2$  angles of Tyr<sup>11</sup> after 40 ns (Figure 2A), leading to the interaction of Tyr<sup>11</sup> with Glu<sup>179</sup> of NTS2 (KHE<sup>179</sup>LE) (Figure 2B). Such rotation of the Tyr<sup>11</sup> residue was not observed during the first 180 ns of the trajectory obtained for the hNTS1 receptor. Therefore, to improve the sampling of the conformations of the hNTS1 and hNTS2 peptide/receptor complexes, the two trajectories were extended to ~200 ns, performing 20 ns of additional accelerated dynamics (aMD).

Accelerated molecular dynamics (aMD) is an enhanced-sampling method that improves the conformational space sampling by reducing energy barriers separating different states of a same system. In this method, the potential energy landscape is altered by adding a bias potential to the true potential such that the escape rates from potential wells are enhanced. This method has been recently applied successfully to GPCRs, in particular to muscarinic receptors, to study the dynamics of activation and the ligand binding.<sup>16</sup> Using such a protocol, the same rotation of the Tyr<sup>11</sup> residue was also observed in the hNTS1 receptor, orienting Tyr<sup>11</sup> toward Arg<sup>212</sup> (QNR<sup>212</sup>SA), as depicted in Figures 2C,D.

Interestingly, these two positively charged Arg<sup>212</sup> and negatively charged Glu<sup>179</sup> residues were aligned on the initial sequence alignment used for homology modeling step and were therefore located at the same position of ECL2 in both receptors at the entry of the binding site (Figure 3).

At this step of the work, our simulations suggested that the Tyr<sup>11</sup> residue of the native peptide could adopt a different orientation than that observed in the available X-ray structure of reference. This rotation was observed in both NTS1 and NTS2 receptors. However, the residue promoting this rotation was sharing an opposite charge according to the receptor subtype: a glutamate in NTS2 and an arginine in NTS1. Therefore, considering that this orientation could be another binding mode of the peptide in its receptor, or at least a transient conformation along its binding pathway, we



**Figure 3.** Localization of R<sup>212</sup> and E<sup>179</sup> residues in the extracellular loop 2 of hNTS1 (in magenta) and hNTS2 (in green) receptors, respectively. These models were both built by homology modeling using the structure of the rat neurotensin receptor type 1 (PDB 4BUO) as a template.

hypothesized that this residue could significantly contribute to the selectivity of peptides sharing the complementary charge in position 11.

## CHEMISTRY

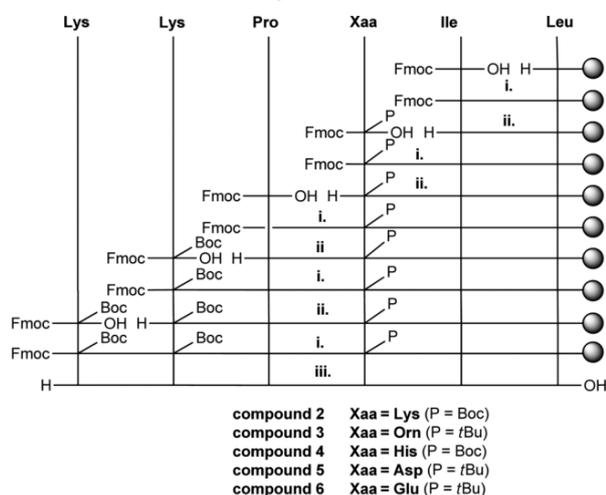
All NT analogues prepared in this study were initially modified to replace the two Arg residues of NT[8-13] (i.e., in positions 8 and 9) by Lys. Detailed SAR studies have revealed that the presence of basic residues with positively charged side chains at positions 8 and 9 is of critical importance for high-affinity receptor binding and peptide activity.<sup>17</sup> However, Lys is often used in position 8, 9, or both because protection of amine side chains is easier to handle in peptide synthesis. Importantly, this modification does not affect the biological activities of the NT analogues, as previously reported for compound 1 (JMV438).<sup>9,18</sup> Therefore, the replacement of the two Arg residues was done for all the newly generated analogues.

The molecular modeling results prompted us to synthesize a series of NT[8-13] peptide analogues in which the aromatic residue at position 11 was substituted with amino acids able to provide ionic interactions with the corresponding residue on the receptor at the entry of the binding pocket. The Tyr<sup>11</sup> residue was therefore changed for amino acids bearing either a basic or an acidic side chain. Tyr<sup>11</sup> was successively replaced by lysine to obtain compound 2 (H-Lys-Lys-Pro-Lys-Ile-Leu-OH, JMV 5836), by ornithine in compound 3 (H-Lys-Lys-Pro-Orn-Ile-Leu-OH, JMV 5837), and by histidine in compound 4 (H-Lys-Lys-Pro-His-Ile-Leu-OH, JMV 5838), which presents a charged side chain conserving the aromatic ring. Replacement of Tyr<sup>11</sup> by acid amino acids, such as aspartic and glutamic acid residues, gave rise to compounds 5 (H-Lys-Lys-Pro-Asp-Ile-Leu-OH, JMV 5839) and 6 (H-Lys-Lys-Pro-Glu-Ile-Leu-OH, JMV 5963), respectively.

Compounds 2–6 were prepared by solid phase peptide synthesis (SPPS) at a 0.25 mmol scale, using Fmoc-Leu preloaded Wang resin, on a Liberty CEM microwave-assisted peptide synthesizer. All Fmoc-amino acids were coupled by using HATU as coupling reagent in the presence of DIPEA. Deprotection cycles were carried out using 20% piperidine in

DMF. Final cleavage from the resin and removal of protecting groups were performed by using a TFA/TIS mixture, leading to the fully deprotected peptides, which were purified by preparative HPLC, affording compounds 2–6 (Scheme 1).

Scheme 1. Synthetic Strategy for Compounds 2–6



i) HATU, DIPEA, DMF; (ii) Piperidine, DMF; (iii) TFA/TIS.

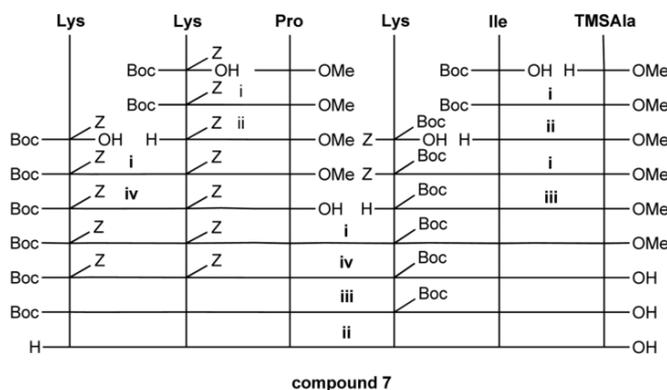
Very recently, we also demonstrated that incorporation of the silylated amino acid (*L*)-(trimethylsilyl)-alanine (TMSAla) replacing the natural occurring leucine residue at the C-terminal end of NT[8-13] led to a new analogue H-Lys-Lys-Pro-Tyr-Ile-TMSAla-OH (JMV 2007),<sup>10</sup> exhibiting high affinity but no selectivity toward NT receptors. To exploit the gain in affinity provided by this end terminal substitution, compound 7 (H-Lys-Lys-Pro-Lys-Ile-TMSAla-OH, JMV 5965) was also synthesized in which Tyr<sup>11</sup> and Leu<sup>13</sup> were respectively replaced by Lys and TMSAla residues.

Compound 7 was prepared in solution exploiting a 3 + 3 fragment condensation between the TMSAla-containing tripeptide and the Boc-Lys(Z)-Lys(Z)-Pro-OH (Scheme 2). The methyl ester of H-TMSAla-OH<sup>19</sup> was prepared in quantitative yield by using thionyl chloride in methanol and

was directly coupled with Boc-Ile-OH in the presence of BOP and DIPEA, affording the dipeptide Boc-Ile-TMSAla-OMe in very good yield. Removal of the Boc protecting group with TFA afforded the free amine, which was directly coupled with Z-Lys(Boc)-OH under the same experimental conditions (as described above), affording the tripeptide Z-Lys(Boc)-Ile-TMSAla-OMe in good yield. Cleavage of the Z protecting group was achieved by catalytic hydrogenation in methanol in the presence of one molar equivalent of 1 M hydrochloric acid in order to avoid possible side reaction due to the nucleophilicity of the primary amine. The tripeptide was then coupled with Boc-Lys(Z)-Lys(Z)-Pro-OH, affording the fully protected hexapeptide Boc-Lys(Z)-Lys(Z)-Pro-Lys(Boc)-Ile-TMSAla-OMe in moderate yield. Total deprotection was achieved in three steps: first, saponification of the methyl ester moiety afforded the free carboxylic acid, followed by removal of the Z protecting group of the lysine residues at positions 8 and 9 by catalytic hydrogenation, and finally removal of Boc in the presence of TFA and triisopropylsilane as scavenger. Purification using preparative HPLC afforded compound 7.

**Binding Properties.** The ability of this series of NT[8-13] derivatives to inhibit the binding of <sup>125</sup>I-[Tyr<sup>3</sup>]-NT on membranes prepared from cells stably expressing either hNTS1 or hNTS2 receptors was first evaluated and the results are summarized in Table 1. Insertion of Lys in position 11 (compound 2) resulted in an important loss of affinity on hNTS1 of more than 5000-fold compared to NT[8-13] ( $K_i$  of  $5700 \pm 750$  nM for compound 2 versus  $1.07 \pm 0.05$  nM for NT[8-13]). This compound also showed a decreased affinity for the hNTS2 receptor with a  $K_i$  value of  $261.8 \pm 72.8$  nM when compared to NT[8-13]. Nevertheless, compound 2 showed more than 20-fold selectivity toward hNTS2, which was consistent with the molecular modeling results described above. When position 11 was modified for an ornithine residue, a lysine analogue with a shorter side chain (compound 3), the binding affinity was found to be decreased for both receptors although some selectivity was conserved for hNTS2 ( $K_i$  of  $>10000$  nM toward hNTS1 and  $619.5 \pm 198$  nM toward hNTS2). Therefore, these results suggest that the length of the side chain exerts only a minor impact on the ionic interaction stability with Glu<sup>179</sup> facing residue in the NTS2 receptor. We

Scheme 2. Synthetic Strategy for Compound 7

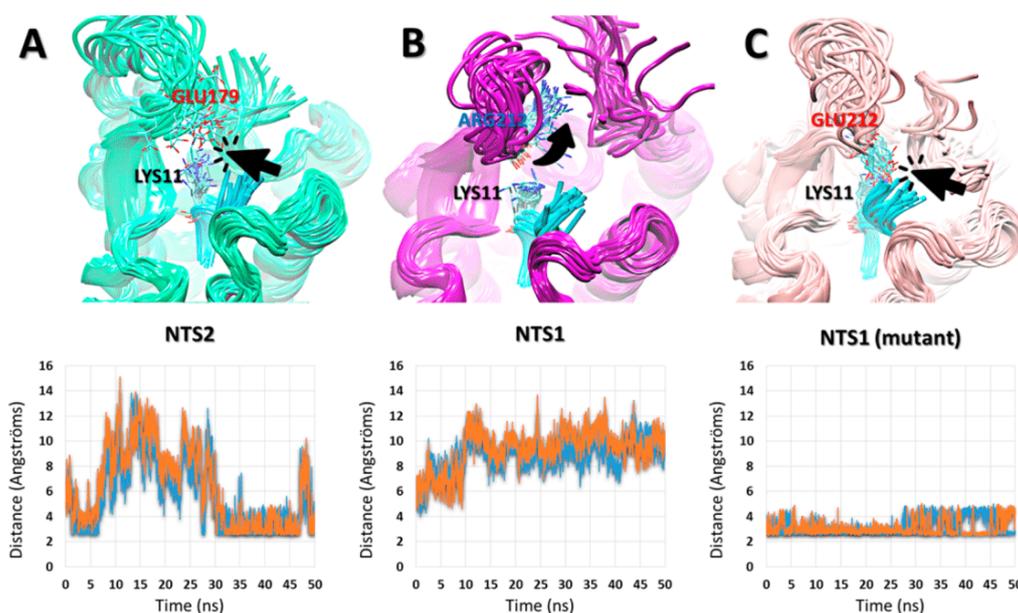


(i) BOP, DIPEA, DMF; (ii) TFA/TIS; (iii) H<sub>2</sub>, Pd/C; (iv) KOH (4N), MeOH.

**Table 1.** Binding Affinities of the Reference Compound NT[8-13] and NT Analogues toward the hNTS1 and hNTS2 Receptors

compd	sequence	$K_i$ binding (nM) <sup>a</sup>		
		hNTS1	hNTS2	selectivity hNTS2/hNTS1
NT[8-13]	H-Arg-Arg-Pro-Tyr-Ile-Leu-OH	1.07 ± 0.05	6.57 ± 2.18	0.16
1	H-Lys-Lys-Pro-Tyr-Ile-Leu-OH	0.33 ± 0.002	0.95 ± 0.003	0.35 <sup>c</sup>
2	H-Lys-Lys-Pro-Lys-Ile-Leu-OH	5700 ± 750	261.8 ± 72.8	21.8
3	H-Lys-Lys-Pro-Orn-Ile-Leu-OH	>10000	619.5 ± 198	nd <sup>b</sup>
4	H-Lys-Lys-Pro-His-Ile-Leu-OH	455.6 ± 78.4	474.1 ± 82.0	0.96
5	H-Lys-Lys-Pro-Asp-Ile-Leu-OH	>10000	4200 ± 740	nd <sup>b</sup>
6	H-Lys-Lys-Pro-Glu-Ile-Leu-OH	>10000	1600 ± 370	nd <sup>b</sup>
7	H-Lys-Lys-Pro-Lys-Ile-TMSAla-OH	662.7 ± 80.1	67.0 ± 26.0	9.90

<sup>a</sup> $K_i$  refers to the equilibrium dissociation constant of a ligand determined by competitive ligand binding assay using <sup>125</sup>I-[Tyr<sup>3</sup>]-NT as reference probe. <sup>b</sup>Not determinable. <sup>c</sup>IC<sub>50</sub> values taken from ref 18.



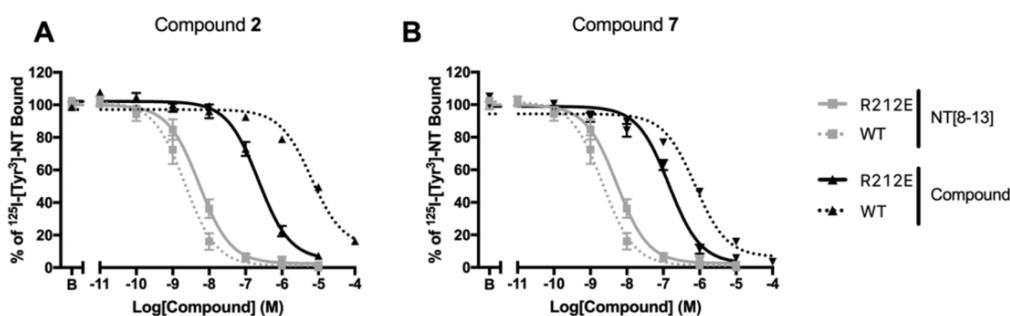
**Figure 4.** Snapshots extracted from MD simulations performed on compound **2** in complex with hNTS2 (A), hNTS1 (B), or hNTS1-R212E (C). In each case, two distances were also computed between the nitrogen atom of Lys<sup>11</sup> and the two oxygen (Glu) or nitrogen (Arg) atoms of the facing residue in each receptor. This highlights the interaction of the Lys<sup>11</sup> residue with Glu<sup>179</sup> (A) and Glu<sup>212</sup> (C) and the absence of this interaction in the hNTS1 receptor.

next evaluated the influence of both positive charge and aromaticity with compound **4**, where Tyr<sup>11</sup> was replaced by a histidine. Binding results indicated a complete loss of selectivity due to an increased affinity toward hNTS1 compared to compounds **2** and **3** ( $K_i$  of 455.6 ± 78.4 and 474.1 ± 82.0 nM toward hNTS1 and hNTS2, respectively). The aromaticity of the imidazole ring seems therefore to be responsible of the restored affinity toward hNTS1. Furthermore, these results reinforce the importance of the complementary positively charged residue in position 11 for the binding toward hNTS2.

Replacement of Tyr<sup>11</sup> by negatively charged amino acids, i.e., aspartic (compound **5**) or glutamic (compound **6**), resulted in an important loss of affinity for hNTS2, as predicted by the presence of the negatively charged Glu<sup>179</sup> in the ECL2 of hNTS2 ( $K_i$  of 4200 ± 740 and 1600 ± 370 nM for compounds **5** and **6**, respectively). It was, however, not expected that this mutation also induced a complete loss in binding affinity for hNTS1 ( $K_i$  of >10000 nM for both analogues). On the basis of the binding data with compound **4** carrying an aromatic residue in position 11, we can hypothesize that this loss in binding affinity of compounds **5** and **6** for NTS1 may be attributed in

part to the absence of aromaticity. We can also speculate that compounds **5** and **6** did not bind efficiently to the hNTS1 receptor due to an intramolecular interaction between the negatively charged incorporated residue and one of the positively charged residues of the hexapeptide, thus leading to a drastic structural change on the active NT peptide conformation. Finally, we can suppose that both hNTS1 and hNTS2 receptors share several negatively charged residues in their extracellular surface, impacting on the binding of compounds **5** and **6** to both hNTS1 and hNTS2 receptors.

Finally, insertion of a TMSAla at position 13, resulting in compound **7**, induced a gain of binding affinity on both receptor subtypes compared to compound **2** ( $K_i$  for hNTS1 of 662.7 ± 80.1 versus 67.0 ± 26.0 nM for hNTS2). However, this compound still exhibited a 10-fold selectivity toward the hNTS2 receptor. Altogether, as predicted by the molecular dynamics simulations, these binding experiments reinforce that Tyr<sup>11</sup> is a key residue for maintaining receptor binding and that the residue at position 11 in NT[8-13] is crucial for NTS1/NTS2 selectivity.

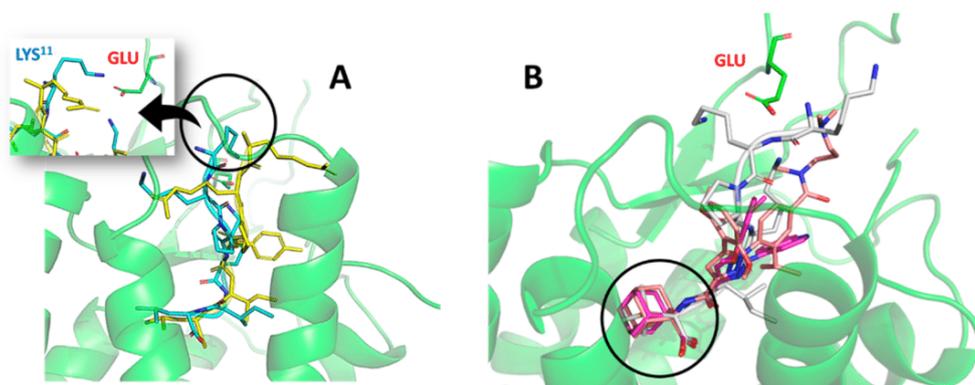


**Figure 5.** Displacement curves of <sup>125</sup>I-[Tyr<sup>3</sup>]-NT with NT[8-13], (A) compound 2, and (B) compound 7 on the wild-type (WT) form of the hNTS1 receptor (dotted line) and on the hNTS1-R212E mutant (solid line).

**Table 2.** Binding Affinities of NT Analogues toward the Wild-Type hNTS1 and Mutated hNTS1-R212E Receptors

compd	sequence	$K_i$ binding (nM) <sup>a</sup>		affinity gain R212E/WT
		hNTS1-WT	hNTS1-R212E	
NT[8-13]	H-Arg-Arg-Pro-Tyr-Ile-Leu-OH	1.07 ± 0.05	4.69 ± 0.75	-4.4
2	H-Lys-Lys-Pro-Lys-Ile-Leu-OH	5700 ± 750	199.2 ± 31.6	28.6
5	H-Lys-Lys-Pro-Asp-Ile-Leu-OH	>10000	>10000	nd <sup>b</sup>
6	H-Lys-Lys-Pro-Glu-Ile-Leu-OH	>10000	>10000	nd <sup>b</sup>
7	H-Lys-Lys-Pro-Lys-Ile-TMSAla-OH	662.7 ± 80.1	126.9 ± 17.1	5.22

<sup>a</sup> $K_i$  refers to the equilibrium dissociation constant of a ligand determined by competitive ligand binding assay using <sup>125</sup>I-[Tyr<sup>3</sup>]-NT as reference probe. <sup>b</sup>Not determinable.



**Figure 6.** Docking results of compounds 2 (A), and 7 (B) in the hNTS2 receptor, showing (A) the interaction of the Lys<sup>11</sup> residue of compound 2 with Glu<sup>179</sup> of hNTS2, and (B) the position taken by the TMSAla group of compound 7 in the bottom of the pocket (in white), compared to the adamantane group of SR142948A (in rust) and SR48692 (in magenta). The native NT[8-13] peptide is reported in yellow in (A).

As observed by molecular modeling, the selectivity of NT ligands toward hNTS2 seemed to be driven by the interaction between the residues in position 11 of NT[8-13] and Glu<sup>179</sup> of hNTS2. To confirm this, we performed molecular dynamics simulations of compound 2 bound either to hNTS1 or hNTS2. For each receptor, we started from the conformation obtained at the end of the trajectories described in Figures 1 and 2, e.g., after rotation of the Tyr<sup>11</sup> residue of the native NT[8-13] peptide. The later was then replaced by compound 2 by replacing the two N-terminal Arg<sup>8</sup> and Arg<sup>9</sup> as well as the Tyr<sup>11</sup> residue into lysine. For each peptide:receptor complex, a new MD trajectory of 50 ns was then computed. As shown in Figure 4A,B, these trajectories confirmed that the Lys<sup>11</sup> residue of compound 2 was able to stably interact with Glu<sup>179</sup> of hNTS2, whereas no residue was able to play this role in hNTS1. On the contrary, the charge:charge repulsion between Lys<sup>11</sup> and Arg<sup>212</sup> in the hNTS1 was leading to a rotation of this last residue.

To confirm these findings, a mutant of hNTS1 was expressed in which the positively charged Arg<sup>212</sup> residue was replaced by the negatively charged Glu residue (hNTS1-R212E).

As shown in Figure 4C, MD simulations reinforced our hypothesis suggesting that compound 2 could bind to the hNTS1-R212E mutant and then interacts with the receptor's Glu residue, as shown for hNTS2. Affinities of compounds 2, 5, 6, and 7 were evaluated on the hNTS1-R212E mutant by competition binding experiments using <sup>125</sup>I-[Tyr<sup>3</sup>]-NT. As shown in Figure 5 and Table 2, the binding affinity increased for compounds 2 and 7 when tested on the hNTS1-R212E mutant receptor. Compound 2 showed a more important gain in binding affinity with a 28-fold increase, whereas compound 7 displayed a moderate affinity gain of 5-fold. As expected, compounds 5 and 6 did not display any gain in binding affinity on the R212E mutant, as compared to the wild type NTS1 receptor. These results indicated that the ionic interaction

directly influences the binding affinity and plays a significant role for the selectivity of NT ligands toward NTS1.

Docking of the different peptides was then performed in the most representative conformations of the hNTS2 (20 conformations) binding pocket. First, the docking method was validated on the NT[8-13] peptide, confirming the ability of the PLANTS docking software in properly predicting the binding mode of the native peptide into the native structure. Our docking results confirmed that the Lys<sup>11</sup> residue of compounds 2 or 7 was able to form the expected salt bridge with Glu<sup>179</sup> in the hNTS2 receptor, while the conformation of the peptide was very close to that of the native NT[8-13] peptide (reported in yellow) (Figure 6A).

Our calculations also showed how the TMSAla moiety in compound 7 fits quite nicely at the bottom of the binding pocket. This was also confirmed by comparison with the adamantane groups of SR142948A and SR48692, two non-peptide compounds that showed binding affinity toward NTS2<sup>19–21</sup> when docked using the same experimental protocol. Again, in this model we also observed the Lys<sup>11</sup>:Glu<sup>179</sup> interaction (Figure 6B).

## CONCLUSION

In conclusion, we report here that the Tyr<sup>11</sup> residue of the NT peptide plays a crucial role for both affinity and selectivity toward the hNTS1 and hNTS2 receptor subtypes. These results are in accordance with previous studies supporting the importance of aromatic residues in the structure–activity relationship for the affinity toward NT receptors.<sup>11–14</sup> The extension of the aromaticity and changes in side chain orientation seem to favor hNTS2 selectivity. Indeed, substitution of the L-Tyr in position 11 by unnatural D-amino acids such as D- $\alpha$ -naphthylalanine is well-tolerated by NTS2 and leads to substantial loss in NTS1 binding.<sup>22–24</sup> Furthermore, high NTS2 selectivity has also been observed with a peptide–peptoid hybrid resulting from replacement of Tyr with N-homotyramine<sup>12,25</sup> and more recently with tetrahydrofuran amino acids.<sup>26</sup>

Finally, hydrophobicity of the C-terminal end is essential for receptor binding. Indeed, as demonstrated by alanine scan strategy, which sequentially replaced each amino acid of NT[8-13] with alanine, the isopropyl group of leucine at position 13 has been found to be particularly important for the binding and biological activity of the peptide.<sup>27</sup> Accordingly, replacement of Ile or/and Leu with more hydrophobic non-natural amino acids, such as trimethylsilylalanine, results in improved affinity.<sup>10,14</sup> However, changing the L-configuration in position 13 and introducing  $\beta$ -amino acid to prevent exoprotease cleavages appear to induce an important loss in receptor binding.<sup>17</sup> Altogether, these molecular modeling and binding results will help to improve the design of the next-generation of NTS2-selective pain-reliever drugs.

## EXPERIMENTAL SECTION

**Computational Methods. Molecular Dynamics Simulations.** Both the human NTS1 and NTS2 receptors were built from their Swissprot sequences (sequences IDs P30989 and O9S665) using the MODELER software<sup>28</sup> and the 4BUO PDB structure as a template.<sup>29</sup> This structure describes the rat NTS1 receptor complexed to the NT[8-13] agonist peptide (H-Lys-Lys-Pro-Tyr-Ile-Leu-OH) and was chosen among of others because of a better resolution and a higher level of completeness. The bound hexapeptide was included during the whole homology-modeling step. The human NTS1 and NTS2

receptors shared 88% and 44% sequence identity rate with the initial template, respectively. Twenty models were built for each receptor, the one sharing the best score being selected for further calculations. The CHARMM-gui interface<sup>30</sup> was then used to insert the NTS1 and NTS2 human receptors in a 100 nm<sup>2</sup> membrane bilayer composed of 100% POPC (1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine). Water molecules and ions were then added to bring this model to completion and neutralize the global charge of the system. Molecular dynamics simulations were all performed with NAMD<sup>31</sup> and the CHARMM forces field.<sup>32</sup> After energy minimization, equilibration was completed in six several successive steps, as provided by CHARMM-gui in NAMD scripts, using NVT dynamics for the two first equilibration steps and then shifting to NPT for the four following ones. Harmonic restraints applied to heavy atoms of the protein or used to hold the position of lipid head groups were progressively decreased for these six successive steps.<sup>33</sup> For the production phase, an integration step of 2 fs was used. Particle Mesh Ewald (PME) was employed for the treatment of long-range interactions. For all calculations, a classical cutoff scheme was used, electrostatic interactions being truncated at 12 Å, applying a switching function between 10 and 12 Å. After a total of 675 ps for equilibration, a production trajectory of 180 ns was obtained for each receptor.

To improve sampling of the possible conformations of each peptide:receptor complex, 20 ns of accelerated molecular dynamics (aMD)<sup>34</sup> were also performed. In this method, whenever the potential decreased below a threshold value, a boost potential was added, increasing the rate of conformational transitions. The modified potential was  $V^*(r) = V(r) + \Delta V$ ,

$$\Delta V = \begin{cases} 0 & V(r) \geq E \\ \frac{(E - V(r))^2}{\alpha + E - V(r)} & V(r) < E \end{cases}$$

$E$  is the threshold, and  $\alpha$  is the acceleration parameter that shaped the modified potential.

We used a dual-boost scheme, with the application of boost potentials to dihedral angles and to the total potential energy of the entire system. Boost parameters were obtained according to the protocol described by Miao et al., which is adequate for protein–membrane systems:<sup>16</sup>  $E_{\text{dihed}} = V_{\text{dihed\_avg}} + 0.3 \times V_{\text{dihed\_avg}}$ ,  $\alpha_{\text{dihed}} = 0.3 \times V_{\text{dihed\_avg}}$ ,  $E_{\text{total}} = V_{\text{total\_avg}} + 0.2 \times \text{total number of atoms}$ ,  $\alpha_{\text{total}} = 0.2 \times \text{total number of atoms}$ . The average values  $V_{\text{dihed\_avg}}$  and  $V_{\text{total\_avg}}$  were obtained from the previous conventional MD runs (180 ns) and restarting from the last structure of the respective conventional MD simulation.

**Peptide:Receptor Docking.** Representative conformations of NTS1 and NTS2 receptor binding sites were obtained by using a clustering procedure. In each case, after concatenation of the trajectories (classical and accelerated dynamics), rotations and translations were removed by least-squares fitting to the backbone atoms of the receptor. Clustering of the fitted trajectory was based on receptor atoms closer than 15 Å to the bound peptide in the reference crystal structure. It was performed with the g\_cluster tool of the Gromacs package<sup>35</sup> using the algorithm of Daura et al.<sup>36</sup> and a 2.5 Å cutoff. The different peptides were built with the Schrödinger Maestro software (<http://www.schrodinger.com/>) and converted to the mol2 file format for further docking calculations. Docking of the different peptides was then performed in the two sets of representative conformations with PLANTS<sup>37</sup> using the ChemPLP scoring function. A sphere of 25 Å around the residues Arg<sup>212</sup> (NTS1) or Glu<sup>179</sup> (NTS2) was used for docking site definition, encompassing both the whole binding site of the receptor and its whole extracellular surface. The 10 best poses obtained for each peptide in each receptor conformation were then conserved for further analyses.

**Chemistry.** All solvents and reagents for the synthesis of compounds 1–4 were purchased from Sigma-Aldrich, Fluka, and Alfa Aesar in gradient grade or reagent quality. All reactions involving air-sensitive reagents were performed under nitrogen or argon. Purifications were performed with column chromatography using

silica gel (Merck 60, 230–400 mesh) or with a Biotage instrument Isolera 4 using SNAP KP-SIL flash cartridges. Nuclear magnetic resonance NMR spectra were recorded on a Bruker spectrometer Avance 300 at 600 MHz. Chemical shifts ( $\delta$ ) are reported from tetramethylsilane with the solvent resonance as internal standard. LC/MS system consisted of a Waters Alliance 2690 HPLC, coupled to a ZQ spectrometer (Manchester, UK), fitted with an electrospray source operated in the positive ionization mode (ESI<sup>+</sup>). All the analyses were carried out using a C18 Chromolith Flash 25 mm  $\times$  4.6 mm column operated at a flow rate of 3 mL/min. A gradient of 0% or 0.1% aqueous TFA (solvent A) to 100% of acetonitrile containing 0.1% TFA (solvent B) was developed over 3 min. Positive-ion electrospray mass spectra were acquired at a solvent flow rate of 100–200  $\mu$ L/min. Nitrogen was used for both the nebulizing and drying gas. The data were obtained in a scan mode ranging from 200 to 1700  $m/z$  in 0.1 s intervals. A total of 10 scans was summed up to get the final spectrum. All the fully unprotected peptides were purified using a gradient composed of water/acetonitrile with 0.1% TFA at 50 mL/min flow rate performed on a PLC2020 Gilson using a 75 mm  $\times$  21.2 mm Phenomenex Luna 5 $\mu$  C18(2) column. Purity was determined by RP-Analytic HPLC performed on a Agilent 1220 using a 50 mm  $\times$  4.6 mm Chromolith high resolution column. Compounds were separated using a linear gradient system (0–100% solvent B in 10 min) using a constant flow rate of 3 mL $\cdot$ min<sup>-1</sup>. All peptides were obtained with purity  $\geq$ 95%. High-resolution mass spectra (HRMS) were performed by the "Laboratoire de Mesures Physiques" of Montpellier University on a Micromass Q-ToF spectrometer equipped with electrospray source ionization (ESI), using phosphoric acid as internal standard.

**General Procedures. General Procedure for Coupling Reaction in Solution (Procedure 1).** To a solution of the NH<sub>2</sub>-free amino ester or peptide in DMF, was added the suitable N-protected amino acid or peptide followed by addition of BOP (1 equiv) and *N,N*-diisopropylethylamine (3 equiv). The mixture was stirred 24 h at room temperature. The solvent was evaporated in vacuo and the residue diluted with EtOAc and consecutively extracted with 1 M KHSO<sub>4</sub> (2 $\times$ ), aqueous NaHCO<sub>3</sub> (2 $\times$ ), and brine (2 $\times$ ), dried over Na<sub>2</sub>SO<sub>4</sub>, and the solvent evaporated under reduced pressure to afford the crude product, which was then purified by chromatography on a silica gel column and characterized by LC-MS.

**General Procedure for Z Deprotection of Peptides in Solution (Procedure 2).** The protected peptide was dissolved in methanol (5 mL/mmol) in the presence of 1 M HCl (1 equiv) and 10% Pd/C under hydrogen atmosphere. The mixture was stirred for 3 h at room temperature. The mixture was filtered on a pad of Celite, which was washed with methanol. The solvent was evaporated under reduced pressure to afford the free amine derivative.

**General Procedure for Boc Deprotection of Peptides in Solution (Procedure 3).** The protected peptide was dissolved in TFA (4 mL/mmol). The mixture was stirred for 2.5 h at room temperature. Volatiles were removed under reduced pressure. To precipitate the compound, diethyl ether was added and evaporated to afford the corresponding TFA salt.

**General Procedure for Saponification of Methyl Esters (Procedure 4).** The peptide methyl ester was dissolved in methanol (10 mL/mmol), and a solution of 4N KOH (5 equiv) was added. The reaction was stirred until disappearance of the starting material. The solvent was removed under reduced pressure and the residue dissolved in water (10 mL/mmol) and acidified with citric acid 15% to pH 4–5. The solution was extracted with ethyl acetate (4 $\times$ ), and the combined organic phases were dried over Mg<sub>2</sub>SO<sub>4</sub> and volatiles removed under reduced pressure to afford the desired carboxylic acid derivative, which was used without further purification.

**General Procedure for Solid Phase Peptide Synthesis (Procedure 5).** All peptides were prepared using a Liberty CEM microwave-assisted peptide synthesizer at a 0.25 mmol scale on a Wang resin preloaded with Fmoc-leucine (loading 0.8 mmol/g). All Fmoc-amino acids (4 equiv) were coupled by using a solution of HATU in DMF (4 equiv) as coupling reagent and a solution of DIPEA (10 equiv) in DMF under microwave irradiation at 40 W for 300 s at 70 °C. The reactant and solvent were then filtered and the resin washed with

DMF (3  $\times$  15 mL). Two deprotection cycles were carried out using 10 mL of a 20% piperidine solution in DMF under microwave irradiation at 40 W for 30 s at 75 °C. After each deprotection cycle, the solvent was removed by filtration and the resin washed with DMF. For the cleavage from the resin and final deprotection, the resin was suspended in a mixture of TFA/TIS (99/1) and stirred for 3 h at room temperature. The solution was filtered and the peptide precipitated with ice-cold diethyl ether. After centrifugation and elimination of the supernatant, the crude was purified by preparative HPLC.

**H-Lys-Lys-Pro-Lys-Ile-Leu-OH (2).** H-Lys-Lys-Pro-Lys-Ile-Leu-OH was prepared on SPPS following procedure 5. ESI-MS: 726.7 [M + H]<sup>+</sup>, 363.8 [M + 2H]<sup>2+</sup>. RP-LC:  $t_R$  = 2.60 min. HRMS: calcd For C35H68N9O7 [M + H]<sup>+</sup>, 726.5244; found, 726.5242.

**H-Lys-Lys-Pro-Orn-Ile-Leu-OH (3).** H-Lys-Lys-Pro-Orn-Ile-Leu-OH was prepared on SPPS following procedure 5. ESI-MS: 712.5 [M + H]<sup>+</sup>, 356.7 [M + 2H]<sup>2+</sup>. RP-LC:  $t_R$  = 2.57 min. HRMS: calcd For C34H66N9O7 [M + H]<sup>+</sup>, 712.5085; found, 712.5081.

**H-Lys-Lys-Pro-His-Ile-Leu-OH (4).** H-Lys-Lys-Pro-His-Ile-Leu-OH was prepared on SPPS following procedure 5. ESI-MS: 735.4 [M + H]<sup>+</sup>, 368.2 [M + 2H]<sup>2+</sup>. RP-LC:  $t_R$  = 2.98 min. HRMS: calcd For C35H63N10O7 [M + H]<sup>+</sup>, 735.4881; found, 735.4875.

**H-Lys-Lys-Pro-Asp-Ile-Leu-OH (5).** H-Lys-Lys-Pro-Asp-Ile-Leu-OH was prepared on SPPS following procedure 5. ESI-MS: 713.4 [M + H]<sup>+</sup>. RP-LC:  $t_R$  = 2.74 min. HRMS: calcd For C33H61N8O9 [M + H]<sup>+</sup>, 713.4568; found, 713.4562.

**H-Lys-Lys-Pro-Glu-Ile-Leu-OH (6).** H-Lys-Lys-Pro-Glu-Ile-Leu-OH was prepared on SPPS following procedure 5. ESI-MS: 727.3 [M + H]<sup>+</sup>. RP-LC:  $t_R$  = 2.71 min. HRMS: calcd For C34H63N8O9 [M + H]<sup>+</sup>, 727.4722; found, 727.4726.

**H-Lys-Lys-Pro-Lys-Ile-TMSAla-OH (7).** Boc-Ile-TMSAla-OMe. TFA-H-TMSAla-OH (223 mg, 0.67 mmol) was dissolved in methanol (6 mL), thionyl chloride was added (245  $\mu$ L, 5 equiv), and the mixture stirred at reflux 4 h 30 min. The solvent was removed under reduced pressure to afford H-TMSAla-OMe-HCl, which was used without further purification for the coupling with Boc-Ile-OH-1/2H<sub>2</sub>O (160 mg, 1 equiv) following the general procedure 1. The dipeptide Boc-Ile-TMSAla-OMe was obtained as a white foam (220 mg, 83%).  $R_f$ : 0.4 (cyclohexane/ethyl acetate 7:3). ESI-MS: 389.4 [M + H]<sup>+</sup>. RP-LC:  $t_R$  = 1.92 min.

**Z-Lys(Boc)-Ile-TMSAla-OMe.** Boc-Ile-TMSAla-OMe (350 mg, 0.9 mmol) was deprotected following general procedure 3. The TFA salt was directly coupled with Z-Lys(Boc)-OH following the general procedure 1. The tripeptide Z-Lys(Boc)-Ile-TMSAla-OMe was obtained as a white foam (375 mg, 64%).  $R_f$ : 0.2 (cyclohexane/ethyl acetate 6:4). ESI-MS: 651.4 [M + H]<sup>+</sup>. RP-LC:  $t_R$  = 1.98 min.

**Boc-Lys(Z)-Lys(Z)-Pro-Lys(Boc)-Ile-TMSAla-OMe.** Z-Lys(Boc)-Ile-TMSAla-OMe (170 mg, 0.26 mmol) was deprotected following general procedure 2. Without further purification it was directly coupled with Boc-Lys(Z)-Lys(Z)-Pro-OH (1 equiv) following the general procedure 1. The hexapeptide was obtained as a white foam (56 mg, 17%).  $R_f$ : 0.47 (dichloromethane/methanol 9:1). ESI-MS: 1238.8 [M + H]<sup>+</sup>. RP-LC:  $t_R$  = 2.20 min.

**Boc-Lys(Z)-Lys(Z)-Pro-Lys(Boc)-Ile-TMSAla-OH.** Starting from Boc-Lys(Z)-Lys(Z)-Pro-Lys(Boc)-Ile-TMSAla-OMe (56 mg, 0.045 mmol), Boc-Lys(Z)-Lys(Z)-Pro-Lys(Boc)-Ile-TMSAla-OH (55 mg, 100%) was obtained from following general procedure 4 as a white foam. ESI-MS: 1224.9 [M + H]<sup>+</sup>. RP-LC:  $t_R$  = 2.07 min.

**Boc-Lys-Lys-Pro-Lys(Boc)-Ile-TMSAla-OH.** Starting from Boc-Lys(Z)-Lys(Z)-Pro-Lys(Boc)-Ile-TMSAla-OH (55 mg, 0.045 mmol), Boc-Lys-Lys-Pro-Lys(Boc)-Ile-TMSAla-OH (43 mg, 99%) was obtained from following general procedure 2. ESI-MS: 956.6 [M + H]<sup>+</sup>. RP-LC:  $t_R$  = 1.18 min.

**H-Lys-Lys-Pro-Lys-Ile-TMSAla-OH (7).** Compound 4 was obtained starting from Boc-Lys-Lys-Pro-Lys(Boc)-Ile-TMSAla-OH (43 mg, 0.044 mmol) by following the general procedure 3 and purified by preparative HPLC. ESI-MS: 756.6 [M + H]<sup>+</sup>. RP-LC:  $t_R$  = 2.87 min. HRMS: calcd For C35H70N9O7 [M + H]<sup>+</sup>, 756.5158; found, 756.5167.

**In Vitro Procedures. Generation of HA-hNTS1-R212E by Site-Directed Mutagenesis.** The cDNA clone of the human HA-tagged

NTS1 receptor was purchased from cdna.org (St-Louis, MO). NTS1 mutant R212E was generated using the Q5 Site-directed mutagenesis kit from New England Biolabs (Ipswich, MA) following manufacturer's recommendations. Primers used to introduce the mutation were synthesized by Integrated DNA Technologies (Coralville, IA) and are described in Table 3 (changed bases in primer's sequence are in lower case). Construct was verified by DNA sequencing.

**Table 3**

primer	sequence
forward	5'-C GAG CAG AAC gaa AGC GCC GAC G-3'
reverse	5'-C CCA TGG TGA ACA GCA TAG-3'

**Cell Culture and Transfections.** CHO-K1 cells stably expressing hNTS1 and 1321N1 cells stably expressing hNTS2 were cultured respectively in Ham's F12 and DMEM. Culture media were supplemented with 10% FBS, 100 U/mL penicillin, 100 µg/mL streptomycin, 20 mM HEPES, and 0.4 mg/mL G418 at 37 °C in a humidified chamber at 5% CO<sub>2</sub>.

HEK293 cells were cultured in DMEM supplemented with 10% FBS, 100 U/mL penicillin, 100 µg/mL streptomycin, and 20 mM HEPES. HEK293 were seeded at  $2.2 \times 10^6$  cells in a 10 cm Petri dish and transfected 24 h after plating with 10 µg of DNA coding for the human HA-tagged NTS1-R212E receptor using PEI as a transfection agent. Membrane preparation was conducted 48 h after transfection, as described below.

**Competitive Radioligand Binding Assay on WT Neurotensin Receptors.** Binding experiments were carried out on freshly prepared membranes, as previously described.<sup>10</sup> Briefly, competitive radioligand binding experiments were performed by incubating 15 µg of cell membranes expressing the hNTS1 receptor with 45 pM of [<sup>125</sup>I]-[Tyr<sup>3</sup>]-NT (2200 Ci/mmol) or 50 µg of cell membranes expressing the hNTS2 receptor with 255 pM of [<sup>125</sup>I]-[Tyr<sup>3</sup>]-NT (2200 Ci/mmol) in binding buffer (50 mM Tris-HCl, pH7.5, 0.2% BSA) in the presence of increasing concentrations of the analogues ranging from 10<sup>-11</sup> to 10<sup>-5</sup> M for 30 min at 25 °C. After incubation, the binding reaction mixture was transferred in polyethylenimine-coated 96-well filter plates (glass fiber filters GF/B, Millipore, Billerica, MA). Reaction was terminated by filtration, and plates were washed three times with 200 µL of ice-cold binding buffer. Glass filters were then counted in a γ-counter (2470 Wizard2, PerkinElmer, Mississauga, Ontario, Canada). Non-specific binding was measured in the presence of 10<sup>-5</sup> M unlabeled NT [8-13] and represented less than 5% of total binding. IC<sub>50</sub> values were determined from the competition curves as the unlabeled ligand concentration inhibiting half of the [<sup>125</sup>I]-[Tyr<sup>3</sup>]-NT-specific binding. Data were plotted using GraphPad Prism 6 using the One-site-Fit Log(IC<sub>50</sub>) and represented the mean ± SEM of at least three separate experiments.

**Competitive Radioligand Binding Assay on the hNTS1-R212E Receptor.** Binding experiments were carried out on freshly prepared membranes 48 h after transfection. Competitive radioligand binding experiments were performed by incubating 50 µg of cell membranes expressing the hNTS1-R212E receptor with 130 pM of [<sup>125</sup>I]-[Tyr<sup>3</sup>]-NT (2200 Ci/mmol) in binding buffer (50 mM Tris-HCl, pH7.5, 0.2% BSA). All other steps in binding experiments were unchanged comparing to the one using the WT receptors.

**Data Analysis.** Competitive radioligand binding data were plotted using Prism 7 (GraphPad, La Jolla, CA) using the One-site-Fit Log(IC<sub>50</sub>) and represented the mean ± SEM of at least three separate experiments.

IC<sub>50</sub> calculated from the competitive radioligand binding assays were then transformed into K<sub>i</sub> using the Cheng-Prusoff equation:<sup>38</sup>

$$K_i = \frac{IC_{50}}{\left(1 + \frac{[L]}{K_d}\right)}$$

where *L* refers to the concentration of radiolabeled tracer (<sup>125</sup>I-[Tyr<sup>3</sup>]-NT) and K<sub>d</sub> refers to the equilibrium dissociation constant of the radioligand.

## ■ ASSOCIATED CONTENT

### 📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jmedchem.6b01848.

HPLC analyses, HRMS spectra for compounds 2–7, <sup>1</sup>H NMR and <sup>13</sup>C NMR spectra of compounds 2 and 7, and saturation binding curves of hNTS1-WT and hNTS1-R212E receptors (PDF)

Molecular formula strings (CSV)

NTS1 (PDB)

NTS2 (PDB)

## ■ AUTHOR INFORMATION

### Corresponding Author

\*Phone: +33 467 143 765. Fax: +33 467 144 866. E-mail: florine.cavelier@umontpellier.fr.

### ORCID

Nicolas Floquet: 0000-0002-3883-2852

Jean Martinez: 0000-0002-4551-4254

Florine Cavelier: 0000-0001-5308-6416

### Author Contributions

<sup>||</sup>These authors contributed equally to directing this study.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work was granted access to the HPC resources of CINES under the allocation no. c2016077530. This work was also realized with the support of HPC@LR, a Center of Competence in High-Performance Computing from the Languedoc-Roussillon region, funded by the Languedoc-Roussillon region, the Europe and the Université de Montpellier. The HPC@LR Center is equipped with an IBM hybrid supercomputer. P.S. is the recipient of the Canada Research Chair in Neurophysiopharmacology of Chronic Pain. This work was supported by research funding from Montpellier University (F.C. for postdoctoral fellowship to R.F., N.F. for Ph.D. fellowship to B.D.) and the Canadian Institutes of Health Research (P.S.). The ANR BIP-BIP is acknowledged for financial support of P.R.

## ■ ABBREVIATIONS USED

BOP, (benzotriazol-1-yloxy)tris(dimethylamino)phosphonium hexafluorophosphate; DMF, *N,N*-dimethylformamide; TMSAla, (L)-(trimethylsilyl)alanine; NT, neurotensin; NTS1, neurotensin receptor subtype 1; NTS2, neurotensin receptor subtype 2; TFA, trifluoroacetic acid; DIPEA, *N,N*-diisopropylethylamine; Z, benzyloxycarbonyl; Boc, *tert*-butyloxycarbonyl; Fmoc, 9-fluorenyl-methoxycarbonyl; HATU, 1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxide hexafluorophosphate

## ■ REFERENCES

(1) Trang, T.; Al-Hasani, R.; Salvemini, D.; Salter, M. W.; Gutstein, H.; Cahill, C. M. Pain and Poppies: The good, the bad, and the ugly of opioid analgesics. *J. Neurosci.* **2015**, *35*, 13879–13888.

- (2) Carraway, R.; Leeman, S. E. The isolation of a new hypotensive peptide, neurotensin, from bovine hypothalamus. *J. Biol. Chem.* **1973**, *248*, 6854–6861.
- (3) Dubuc, I.; Costentin, J.; Doulut, S.; Rodriguez, M.; Martinez, J.; Kitabgi, P. JMV 449: a pseudopeptide analogue of neurotensin-(8-13) with highly potent and long-lasting hypothermic and analgesic effects in the mouse. *Eur. J. Pharmacol.* **1992**, *219*, 327–9.
- (4) Nemeroff, C. B.; Osbahr, A. J., 3rd; Manberg, P. J.; Ervin, G. N.; Prange, A. J., Jr. Alterations in nociception and body temperature after intracisternal administration of neurotensin, beta-endorphin, other endogenous peptides, and morphine. *Proc. Natl. Acad. Sci. U. S. A.* **1979**, *76*, 5368–5371.
- (5) Skoog, K. M.; Cain, S. T.; Nemeroff, C. B. Centrally administered neurotensin suppresses locomotor hyperactivity induced by d-amphetamine but not by scopolamine or caffeine. *Neuropharmacology* **1986**, *25*, 777–782.
- (6) Kleczkowska, P.; Lipkowski, A. W. Neurotensin and neurotensin receptors: characteristic, structure-activity relationship and pain modulation—a review. *Eur. J. Pharmacol.* **2013**, *716*, 54–60.
- (7) Tetreault, P.; Beaudet, N.; Perron, A.; Belleville, K.; Rene, A.; Cavelier, F.; Martinez, J.; Stroth, T.; Jacobi, A. M.; Rose, S. D.; Behlke, M. A.; Sarret, P. Spinal NTS2 receptor activation reverses signs of neuropathic pain. *FASEB J.* **2013**, *27*, 3741–3752.
- (8) Coutant, J.; Curmi, P. A.; Toma, F.; Monti, J. P. NMR solution structure of neurotensin in membrane-mimetic environments: molecular basis for neurotensin receptor recognition. *Biochemistry* **2007**, *46*, 5656–5663.
- (9) Granier, C.; Van Rietschoten, J.; Kitabgi, P.; Poustis, C.; Freychet, P. Synthesis and characterization of neurotensin analogues for structure/activity relationship studies. *Eur. J. Biochem.* **1982**, *124*, 117–125.
- (10) Fanelli, R.; Besserer-Offroy, É.; René, A.; Côté, J.; Tetreault, P.; Colletette-Tremblay, J.; Longpré, J.-M.; Leduc, R.; Martinez, J.; Sarret, P.; Cavelier, F. Synthesis and characterization in vitro and in vivo of (L)-(trimethylsilyl)alanine containing neurotensin analogues. *J. Med. Chem.* **2015**, *58*, 7785–7795.
- (11) Hapău, D.; Rémond, E.; Fanelli, R.; Vivancos, M.; René, A.; Côté, J.; Besserer-Offroy, É.; Longpré, J.-M.; Martinez, J.; Zaharia, V.; Sarret, P.; Cavelier, F. Stereoselective synthesis of  $\beta$ -(5-arylthiazolyl)  $\alpha$ -amino acids and use in neurotensin analogues. *Eur. J. Org. Chem.* **2016**, *2016*, 1017–1024.
- (12) Einsiedel, J.; Held, C.; Hervet, M.; Plomer, M.; Tschammer, N.; Hübner, H.; Gmeiner, P. Discovery of highly potent and neurotensin receptor 2 selective neurotensin mimetics. *J. Med. Chem.* **2011**, *54*, 2915–2923.
- (13) Boules, M.; Johnston, H.; Tozy, J.; Smith, K.; Li, Z.; Richelson, E. Analgesic synergy of neurotensin receptor subtype 2 agonist NT79 and morphine. *Behav. Pharmacol.* **2011**, *22*, 573–581.
- (14) Boules, M.; Shaw, A.; Liang, Y.; Barbut, D.; Richelson, E. NT69L, a novel analgesic, shows synergy with morphine. *Brain Res.* **2009**, *1294*, 22–38.
- (15) White, J. F.; Noinaj, N.; Shibata, Y.; Love, J.; Kloss, B.; Xu, F.; Gvozdenovic-Jeremic, J.; Shah, P.; Shiloach, J.; Tate, C. G.; Grishammer, R. Structure of the agonist-bound neurotensin receptor. *Nature* **2012**, *490*, 508–513.
- (16) Miao, Y.; Nichols, S. E.; Gasper, P. M.; Metzger, V. T.; McCammon, J. A. Activation and dynamic network of the M2 muscarinic receptor. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 10982–10987.
- (17) Kitabgi, P.; Poustis, C.; Granier, C.; Van Rietschoten, J.; Rivier, J.; Morgat, J. L.; Freychet, P. Neurotensin binding to extraneural and neural receptors: comparison with biological activity and structure–activity relationships. *Mol. Pharmacol.* **1980**, *18*, 11–19.
- (18) Rémond, E.; Martin, C.; Martinez, J.; Cavelier, F. Silicon-containing amino acids: synthetic aspects, conformational studies, and applications to bioactive peptides. *Chem. Rev.* **2016**, *116*, 11654–11684.
- (19) Rene, A.; Vanthuyne, N.; Martinez, J.; Cavelier, F. (L)-(Trimethylsilyl)alanine synthesis exploiting hydroxypinane-induced diastereoselective alkylation. *Amino Acids* **2013**, *45*, 301–317.
- (20) Gully, D.; Canton, M.; Boigegrain, R.; Jeanjean, F.; Molimard, J. C.; Poncelet, M.; Gueudet, C.; Heaulme, M.; Leyris, R.; Brouard, A. Biochemical and pharmacological profile of a potent and selective nonpeptide antagonist of the neurotensin receptor. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 65–69.
- (21) Schaeffer, P.; Laplace, M. C.; Bernat, A.; Prabonnaud, V.; Gully, D.; Lespy, L.; Herbert, J. M. SR142948A is a potent antagonist of the cardiovascular effects of neurotensin. *J. Cardiovasc. Pharmacol.* **1998**, *31*, 545–550.
- (22) Richelson, E.; McCormick, D. J.; Pang, Y. P.; Phillips, K. S. Peptide analogs that are potent and selective for human neurotensin receptor subtype 2. U.S. Pat. Appl. Publ. WO 2008137720, November 13, 2008.
- (23) Tyler, B. M.; Douglas, C. L.; Fauq, A.; Pang, Y.-P.; Stewart, J. A.; Cusack, B.; McCormick, D. J.; Richelson, E. In vitro binding and CNS effects of novel neurotensin agonists that cross the blood–brain barrier. *Neuropharmacology* **1999**, *38*, 1027–1034.
- (24) Dubuc, I.; Sarret, P.; Labbe-Julie, C.; Botto, J. M.; Honore, E.; Bourdel, E.; Martinez, J.; Costentin, J.; Vincent, J. P.; Kitabgi, P.; Mazella, J. Identification of the receptor subtype involved in the analgesic effect of neurotensin. *J. Neurosci.* **1999**, *19*, 503–510.
- (25) Held, C.; Plomer, M.; Hübner, H.; Meltretter, J.; Pischetsrieder, M.; Gmeiner, P. Development of a metabolically stable neurotensin receptor 2 (NTS2) ligand. *ChemMedChem* **2013**, *8*, 75–81.
- (26) Simeth, N. A.; Bause, M.; Dobmeier, M.; Kling, R. C.; Lachmann, D.; Hübner, H.; Einsiedel, J.; Gmeiner, P.; König, B. NTS2-selective neurotensin mimetics with tetrahydrofuran amino acids. *Bioorg. Med. Chem.* **2017**, *25*, 350–359.
- (27) Heyl, D. L.; Seffler, A. M.; He, J. X.; Sawyer, T. K.; Wustrow, D. J.; Akunne, H. C.; Davis, M. D.; Pugsley, T. A.; Heffner, T. G.; Corbin, A. E.; et al. Structure-activity and conformational studies of a series of modified C-terminal hexapeptide neurotensin analogues. *Int. J. Pept. Protein Res.* **1994**, *44*, 233–238.
- (28) Webb, B.; Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinf.* **2014**, *47*, 5.6.1–5.6.32.
- (29) Egloff, P.; Hillenbrand, M.; Klenk, C.; Batyuk, A.; Heine, P.; Balada, S.; Schlinkmann, K. M.; Scott, D. J.; Schütz, M.; Plücker, A. Structure of signaling-competent neurotensin receptor 1 obtained by directed evolution in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, E655–E662.
- (30) Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y.; Jo, S.; Pande, V. S.; Case, D. A.; Brooks, C. L.; MacKerell, A. D.; Klauda, J. B.; Im, W. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* **2016**, *12*, 405–413.
- (31) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (32) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caffisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **2009**, *30*, 1545–1614.
- (33) Wu, E. L.; Cheng, X.; Jo, S.; Rui, H.; Song, K. C.; Dávila-Contreras, E. M.; Qi, Y.; Lee, J.; Monje-Galvan, V.; Venable, R. M.; Klauda, J. B.; Im, W. CHARMM-GUI membrane builder toward realistic biological membrane simulations. *J. Comput. Chem.* **2014**, *35*, 1997–2004.

(34) Hamelberg, D.; Mongan, J.; McCammon, J. A. Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. *J. Chem. Phys.* **2004**, *120*, 11919–11929.

(35) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

(36) Daura, X.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. Reversible peptide folding in solution by molecular dynamics simulation. *J. Mol. Biol.* **1998**, *280*, 925–932.

(37) Korb, O.; Stützle, T.; Exner, T. E. Empirical scoring functions for advanced protein–ligand docking with PLANTS. *J. Chem. Inf. Model.* **2009**, *49*, 84–96.

(38) Cheng, Y.; Prusoff, W. H. Relationship between the inhibition constant ( $K_I$ ) and the concentration of inhibitor which causes 50% inhibition ( $I_{50}$ ) of an enzymatic reaction. *Biochem. Pharmacol.* **1973**, *22*, 3099–3108.

### 3.2.1.c Commentaires

Cet article a été soumis le 20 Décembre 2016 au journal « *Journal of Medicinal Chemistry* » pour être finalement publié le 23 Février 2017. L'article a connu une phase de révision majeure avec un grand nombre d'éléments à préciser, tous domaines confondus. Concernant la modélisation moléculaire, on nous a demandé de décrire plus en détail les résultats de nos simulations et de produire des simulations supplémentaires validant notre hypothèse dans le mutant (simulations du peptide modifié « Lysine » dans les trois récepteurs NTR1, NTR2 et NTR1 muté). Une remarque surtout est importante à préciser pour ce manuscrit. Pour toutes les simulations, les Arginines 8 et 9 de la séquence de NT[8-13] ( $\text{NH}_3^+$ -**Arg-Arg-Pro-Tyr-Ile-Leu-CO<sub>2</sub><sup>-</sup>**) ont été remplacées par des lysines afin de correspondre aux données expérimentales. En effet, des études de structure-activité effectuées au laboratoire ont montré que ces deux arginines strictement nécessaires pour l'affinité pouvaient être remplacées par des lysines permettant une synthèse peptidique plus simple au niveau de la protection des fonctions des chaînes latérales, sans pour autant modifier l'activité biologique des peptides résultants [Rémond et al. 2016].

Si à ce point du travail il était déjà très encourageant d'avoir conçu, sur la base de nos simulations, des peptides sélectifs du récepteur NTR2 en introduisant une charge positive en position 11, les résultats expérimentaux obtenus ont en revanche invalidé l'hypothèse inverse. En effet, nous pensions que l'introduction d'une charge négative à la même position du peptide contribuerait cette fois à produire des peptides sélectifs du récepteur NTR1. Or, les peptides présentant une charge négative en position 11 ne se lient ni au NTR1, ni au NTR2. Il y a donc derrière la sélectivité observée expérimentalement un phénomène certainement plus complexe à expliquer que par un simple contact récepteur:peptide.

## 3.2.2 Étude de la sélectivité de peptides dérivés de la Neurotensine

### 3.2.2.a Contexte

Lors de la validation de notre protocole pour la prédiction de la liaison de peptides à leurs récepteurs (voir partie 3.1.1), nous nous sommes intéressés à des peptides « scrambled » de la Neurotensine. Lors de nos simulations CG-REMD nous avons ainsi pu observer une liaison significative à 300K de peptides n'ayant pourtant (*a priori*) pas d'affinité pour le récepteur NTR1. Toutefois, nos résultats suggèrent qu'il est possible de discriminer assez facilement le peptide natif des peptides « scrambled » grâce à l'analyse de cette liaison en fonction de la température. Dans ce chapitre nous montrons la capacité de notre protocole à prédire la sélectivité de différents peptides pour un même récepteur.

Pour cette partie du travail, nous nous sommes naturellement intéressés aux peptides dérivés de la Neurotensine et décrit dans le paragraphe 2.4.2. Ces peptides présentent une sélectivité certaine pour les deux récepteurs NTR1 et NTR2 humains. De plus, ils se lient au récepteur de la Neurotensine sur lequel notre méthode a déjà été validée dans notre premier article. Aussi, nous avons réalisé des calculs CG-REMD sur plusieurs peptides incluant : le peptide « natif » (11Tyr) présentant une tyrosine en position 11 et utilisé comme référence car non sélectif et les peptides dérivés présentant une charge en position 11: 11Asp et 11Glu qui ne se lient ni au récepteur NTR1 ni au récepteur NTR2, et surtout 11Lys qui présente une sélectivité d'environ 20 fois supérieur pour le récepteur NTR2 que pour le récepteur NTR1. En plus des deux récepteurs sauvages, construits par homologie à partir de la structure cristallographique du récepteur NTR1 de rat, nous avons également construit le récepteur NTR1 mutant R212E, permettant au peptide 11Lys de retrouver une affinité équivalente au récepteur NTR2 natif. Pour comprendre le rôle joué par cette position clé de la boucle ECL2, nous avons également simulé la liaison des différents peptides à un récepteur NTR2 présentant également une inversion de charge à cette même position (E179R). Des études expérimentales sont en cours pour caractériser la liaison de ces peptides dans ce mutant « reverse » (collaboration avec Philippe SARRET, Université de Sherbrooke, Canada).

Pour chacun des 16 complexes (4 peptides *versus* 4 récepteurs), j'ai réalisé 2 $\mu$ s de simulation à 58 températures différentes séparées d'un pas de 3K (de 300K à 471K) pour un total de temps de simulation de 116 $\mu$ s sur un système comportant deux récepteurs et deux

peptides. L'analyse de la liaison des peptides dans la poche de chacun des récepteurs en fonction de la température montre des résultats similaires à ceux obtenus avec les peptides « scrambled ». En effet, ces analyses permettent une séparation des différents couples peptide:récepteur en accord avec l'affinité observée expérimentalement. Ainsi, nos calculs prédisent bien le peptide natif comme celui se liant le mieux aux différents récepteurs, mais aussi la liaison bien plus faible des peptides chargés négativement. Le résultat le plus probant vient de la prédiction de la sélectivité du peptide 11Lys pour le récepteur NTR2, et de l'annulation de cette sélectivité dans le récepteur NTR1 muté.

En outre, les courbes en température montrent un comportement très différent du récepteur NTR1 par rapport aux autres récepteurs (**Figure 1** de l'article), avec un pourcentage de liaison plus faible pour chacun des peptides, et ce à chacune des températures. De manière intéressante, la seule mutation de NTR1 suffit à redonner à ce récepteur un comportement du type NTR2. En revanche, la mutation introduite dans NTR2 ne produit pas l'effet inverse, et contribue seulement à diminuer un peu la liaison de chacun des peptides, avec un effet par contre bien plus prononcé pour le peptide 11Lys. Cette observation nous a amené à montrer le rôle joué par ce résidu de la boucle ECL2 dans l'ouverture/fermeture de ces récepteurs. Nos modèles montrent que l'inversion de charge à cette position joue un rôle clé dans la dynamique intrinsèque de ces récepteurs car elle modifie le réseau d'interactions ioniques à leurs surfaces. Plus précisément, dans le cas du récepteur NTR1, le résidu R212 de la boucle ECL2 fait face à deux résidus chargés négativement dans la boucle ECL3 ; dans le cas du récepteur NTR2, les deux boucles ECL2 et ECL3 présentent chacune deux résidus chargés négativement. Dans l'article qui suit, nous avons étudié en détails ce réseau d'interaction et montré son rôle dans l'ouverture/fermeture de ces récepteurs, notamment en validant notre hypothèse par la simulation d'un récepteur NTR2 double muté (E179R et E181R) montrant un comportement proche du récepteur NTR1 sauvage.

Nous montrons en revanche que cette capacité à s'ouvrir ou se fermer, même si elle peut évidemment jouer un rôle, n'est pas suffisante à expliquer la sélectivité des différents peptides. Enfin, un des points assez remarquables de cette étude est de montrer que le protocole et les méthodes utilisés, malgré leurs approximations, sont assez sensibles pour détecter des variations entre peptides et récepteurs ne présentant qu'un résidu de différence.

### 3.2.2.b Article

## **Dynamical study of peptide binding to GPCRs: the route towards selectivity?**

Bartholomé DELORT <sup>1</sup>, Pedro RENAULT <sup>1</sup>, Roberto FANELLI <sup>1</sup>, Élie BESSERER-OFFROY <sup>2</sup>, Philippe SARRET <sup>2</sup>, Florine CAVELIER <sup>1</sup>, Maxime LOUET <sup>1</sup> and Nicolas FLOQUET <sup>1</sup>

<sup>1</sup> Institut des Biomolécules Max Mousseron (IBMM), CNRS UMR5247, Université de Montpellier, ENSCM, France.

<sup>2</sup> Département de Pharmacologie-Physiologie/Institut de Pharmacologie de Sherbrooke, Université de Sherbrooke, Québec, Canada.

### **Abstract**

Predicting the dynamical binding of a peptide to its target without any bias is a hard task that requires access to huge computational facilities. In this study, we show how a simple but robust modelling protocol combining Replica Exchange Molecular Dynamics and the MARTINI Coarse-Grained model can be useful in qualitatively predicting, without any applied bias, the binding selectivity of several neurotensin (NT) derived peptides against human NTS-1 and NTS-2 native/mutant receptors. In agreement with the experiment, we show how such simulations can capture the effect(s) of a single residue difference among these peptides or receptors. Our results also suggest that the selectivity of these peptides is explained by neither their final orientation in the orthosteric binding site, nor the route they take to reach this position, but rather by their respective capabilities to take this route. Because our protocol does not require extensive computational resources, we demonstrate here that it can be easily applied to series of peptides in order to characterize key binding features/differences for the development of new selective drugs targeting GPCRs.

## INTRODUCTION

G-Protein Coupled Receptors (GPCRs) constitute a large class of druggable membrane proteins as they are involved in nearly all physiological processes<sup>1</sup>. Despite many structural data now available in the Protein Data Bank for these key receptor targets, the dynamical process of peptide binding to GPCR still remains poorly defined. Indeed, state of the art experiments cannot yet bring the required information to describe this binding process in details at both atomic level and  $\mu\text{s}$  time-scale. However, as suggested by previous studies<sup>2</sup>, a better knowledge of intermediate positions of peptides at the receptor interface could probably help in designing new ligands, possibly with a more controlled effect.

In some recent studies, it has been shown that Molecular Dynamics (MD) simulations can help in this task<sup>3</sup>, bringing some dynamical information on ligand:receptor interactions that could not or difficultly be captured by experimental methods<sup>2,4</sup>. However, because ligand binding occurs at the  $\mu\text{s}$  and larger time-scales, the computational cost associated with the use of classical semi-empirical force fields strongly limits such studies to laboratories having access to huge computational facilities.

Due to their inherent structural simplifications, coarse-grained simulations allow to reach more easily biological relevant time-scales and have been used to predict the dynamical binding of ligands to their targets<sup>5-7</sup>.

In a recent study<sup>8</sup>, we have shown that the combination of a Coarse-Grained (CG) model (the MARTINI force field<sup>9</sup> and a Replica Exchange Molecular Dynamics protocol (REMD) was successful in predicting, without any applied intermolecular restraint or bias, the binding poses of peptides in their cognate receptor. Applied to two different examples: the rat neurotensin NTS-1<sup>10</sup> and the human chemokine CXCR4<sup>11</sup> receptors, our protocol was able to predict ~60% of the peptide:receptor native (X-ray) contacts, after backward mapping of the best (most populated) clusters into a classical “all-atoms” model. As compared to other approaches employing classical models, our CG/REMD presents the advantage to be much less computationally expensive and can therefore be easily applied to several peptide receptor complexes, *e.g.* at the medium throughput scale.

In the present study, we applied this modelling protocol to the prediction of the binding properties of four different neurotensin (NT) peptides against the two human NTS-1 and NTS-2 receptors. NTS receptors play various roles both at the physiological and pathological levels<sup>12</sup>. Among them, the NTS-1 receptor is reported to be involved in the regulation of the cardiovascular system<sup>13</sup>, whereas the NTS-2 receptor was found to be associated with the analgesic responses of centrally administered NT<sup>14</sup>. Other roles of NTS receptors have been highlighted, including body weight disorders<sup>15</sup>, hypothermia<sup>16</sup>, and oncogenic role in various types of cancer<sup>17</sup>. Therefore, the design of selective peptides targeting either the NTS-1 or the NTS-2 receptors find many pharmacological applications. Such a design is nevertheless difficult as the binding pockets of these two receptors are almost 100% conserved.

Recently, we designed the peptides reported in **Table 1** on the basis of classical MD simulations suggesting that, in the bound form, the Tyr<sup>11</sup> residue of the unselective neurotensin agonist NT:8-13 hexa-peptide ( $\text{NH}_3^+\text{-Arg}^8\text{-Arg}^9\text{-Pro}^{10}\text{-Tyr}^{11}\text{-Ile}^{12}\text{-Leu}^{13}\text{-CO}_2^-$ ) could indifferently interact with Arg<sup>212</sup> (NTS-1) or Glu<sup>179</sup> (NTS-2) located at the same position of the extra-cellular loop 2 (ECL2) of these two receptors<sup>18</sup>. In agreement, the replacement of Tyr<sup>11</sup> by a positively charged Lys residue (further named peptide 2 in this study) validated our hypothesis, this peptide sharing a low affinity for the NTS-1 receptor ( $>5\mu\text{M}$ ) and a significant (~20 fold higher than for NTS-1) selectivity for the NTS-2 receptor (see **Table 1**). Of interest, the production of a R<sup>212</sup>E NTS-1 mutant confirmed the quasi-exclusive role of this residue in this selectivity, the peptide 2 recovering some binding properties for the R<sup>212</sup>E NTS-1 mutant that were closely related to those observed in the native NTS-2 receptor. Unexpectedly however,

the Asp and Glu-modified peptides (peptides 3 and 4 in **Table 1**) were not selective of the NTS-1 receptor, these two peptides being unable to bind to any of the three receptors, without any clear structural explanation.

Peptide	SEQUENCE	Ki (nM)		
		hNTS-1	hNTS-2	hNTS-1 (R212E)
NT[8-13]	H-Arg-Arg-Pro-Tyr-Ile-Leu-OH	<b>1.07 ± 0.075</b>	<b>6.57 ± 2.18</b>	<b>45.08 ± 0.81</b>
<b>1</b>	H-Lys-Lys-Pro-Tyr-Ile-Leu-OH	<b>7.18 ± 0.28</b>	<b>0.95 ± 0.003</b>	<b>1.17 ± 0.33</b>
<b>2</b>	H-Lys-Lys-Pro-Lys-Ile-Leu-OH	5700 ± 750	<b>261.8 ± 72.8</b>	<b>199.2 ± 31.6</b>
<b>3</b>	H-Lys-Lys-Pro-Asp-Ile-Leu-OH	>10,000	4200 ± 740	>10,000
<b>4</b>	H-Lys-Lys-Pro-Glu-Ile-Leu-OH	>10,000	1600 ± 370	>10,000

Table 1: binding properties of the four peptides used in this study. To get more details on the experimental procedures please report to <sup>18</sup>.

In this study, we applied our CG/REMD protocol to each of these 12 complexes (4 peptides, 3 receptors) and show its ability to qualitatively reproduce these experimental observations. After this validation step, we focused on the peptide receptor interactions and simulated other complexes to explain, at the molecular level, the selectivity of the different peptides for the different receptors.

## METHODS

### Building of initial models of the receptors / peptides

Models of the human NTS-1 and NTS-2 WT/mutant receptors were built from their swissprot sequences, using the agonist bound X-ray structure of the rat NTS-1 receptor <sup>19</sup> as a template. Missing residues in the intra-cellular loop 3 (ICL3) were replaced by six alanine residues to make the connection between the TM5 and TM6 trans-membrane segments. Sequence identity with the template was reaching 83 and 41 %, for the NTS-1 and NTS-2 receptors, respectively. For homology modeling, we used the MODELLER software <sup>20</sup>, generating 100 models and keeping the one showing the best objective function value for further calculations. The selected model was then converted to the Coarse-Grained MARTINI force field <sup>9</sup> version 2.2 + Elnedyn, by using tools available in CHARMM-GUI <sup>21,22</sup>, setting the elastic bond force constant to 500 kJ.mol<sup>-1</sup>.nm<sup>-2</sup> and the lower and upper elastic bond cutoffs to 0.5 and 0.9 nm, respectively. Each receptor was then inserted in a 100% POPC bilayer. The protein+membrane system was then duplicated/rotated along the z-axis to prevent exploration by the peptide of the intra-cellular side of the receptor, as we did previously <sup>8</sup>. All the peptides were built in their all-extended conformation with the Schrödinger's software MAESTRO (<https://www.schrodinger.com/maestro>) and then converted to the same Coarse-Grained model. Two copies of each peptide were then inserted in the "extra-cellular" compartment between the two facing receptors and at a distance from both receptors and from each other higher than the cut-off to prevent any initial bias. The system was then completed by water and ion beads to neutralize the system. The same procedure was used for each of the four NTS-1(WT), NTS-1(R<sup>212</sup>E), NTS-2(WT) and NTS-2(E<sup>179</sup>R) receptors and for each of the four NT[8-13] derived peptides (Tyr, Lys, Asp or Glu residue in position 11), leading to 16 different peptide receptor complexes. The same procedure was also applied to the NTS-2 (E<sup>179</sup>R+E<sup>181</sup>R) double mutant.

## **Binding simulations**

The binding of each peptide to each receptor was achieved by using the same Replica Exchange Molecular Dynamics (REMD) protocol we described recently<sup>8</sup>. In brief, each system was simulated in the temperature range 300:474 K using a 3K interval between adjacent replicas, leading to 58 different trajectories of 2  $\mu$ s and to a total simulation time of 116  $\mu$ s. Equilibration and production steps were achieved in the NVT/NPT ensembles, respectively, using classical parameters described for coarse-grained simulations with MARTINI, and including an integration time-step of 20 fs and a cut-off value of 12 Angströms for the computation of non-bonded interactions. The energy difference between models from adjacent replicas was tested every 20 ps. Using such parameters on related molecular systems, we have previously shown that the exchange probability between coordinates of adjacent replicas was around 0.2. All simulations were performed with the Gromacs software version 5.1.2<sup>23</sup>.

## **Conformational clustering of the trajectories**

The clustering procedure was the same for all studied complexes. Each obtained demuxed trajectory (*e.g.* continuous with respect to simulation time) including two receptors and two peptides was first decomposed into four different sub-complexes (two copies of the peptide *versus* two copies of the receptor) to simplify analyses and scripting. All the resulting files were then concatenated and filtered, to only keep the snapshots having at least one peptide receptor contact (distance  $\leq 7.0$  Angströms). All the resulting conformations were fit onto the backbone beads of the receptor. The clustering was then performed with the algorithm by Daura et al. implemented in Gromacs<sup>24</sup>, using a RMSD cutoff of 3 Angströms computed on all beads of the peptide.

## **Computation of the binding rate and contacts for each peptide receptor complex**

For each peptide receptor complex, we checked the occupancy of the binding site by the peptide at each simulated temperature; for that purpose, each system including two receptors and two peptides was again decomposed into the four different sub-complexes. Three residues of the binding pocket were chosen for each receptor, *e.g.* Phe<sup>127</sup>, Thr<sup>225</sup> and Phe<sup>326</sup> (NTS-1 numbering) or their equivalent for other receptors. The peptide was considered in its binding pocket when at least one of its beads was at a maximum distance of 10 Å of all three residues of the pocket at the same time. For each temperature, we finally computed a percentage of conformations in which the peptide was effectively bound to its receptor along the trajectories. As a final result, temperature dependent binding curves were obtained for each peptide receptor complex. For contact analysis, we used the same set of trajectories as those used for the conformational clustering (distance peptide-receptor  $\leq 7$  Angströms). The same method was also applied to the native NT[8-13] peptide in the rat NTS-1 X-ray structure<sup>19</sup>. With these criteria, the “native” orientation was defined by a set of ~50 peptide receptor contacts. This reference was used to compute a percentage of “native” contacts along each obtained trajectory.

## RESULTS

### The selectivity predicted for all the peptides is in qualitative agreement with the experiment

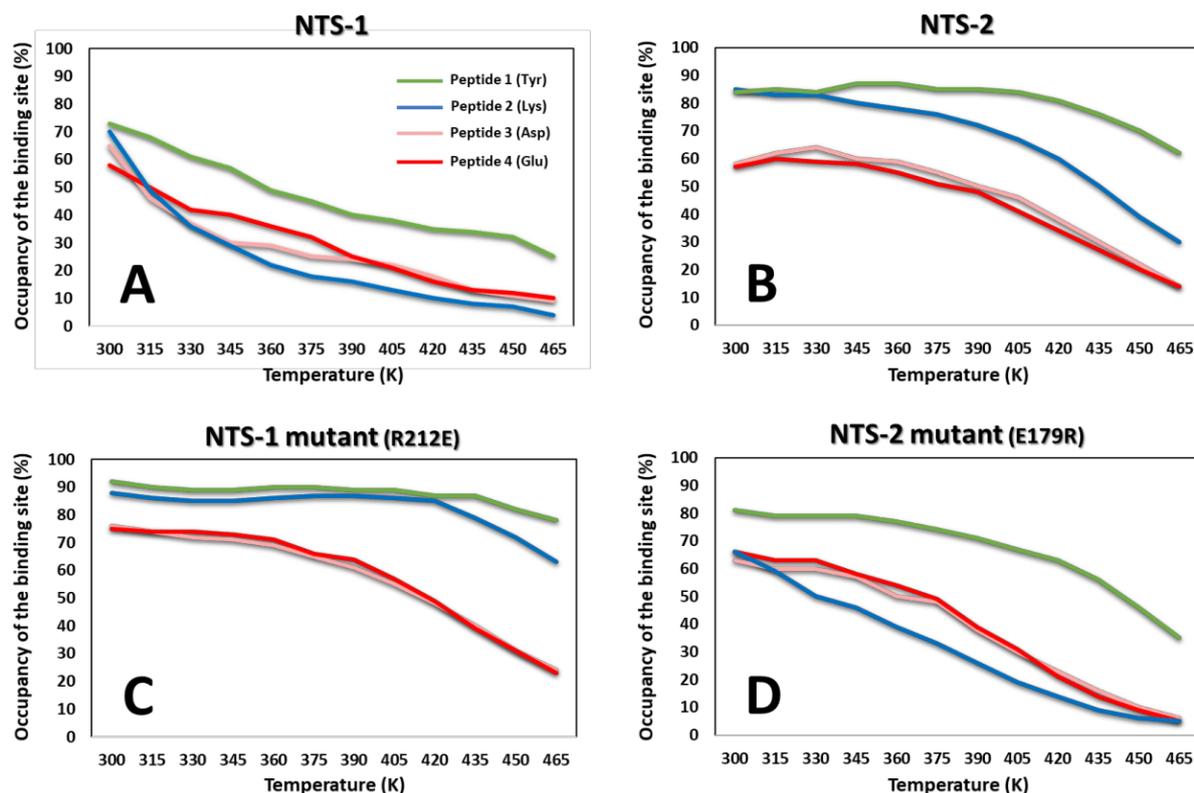
In a first step, our CG/REMD binding protocol was applied to each of the 4 peptides (peptides 1 to 4 in **Table 1**) in each of the three NTS-1 wild-type, NTS-2 wild-type and NTS-1 R<sup>212</sup>E mutant receptors. For each complex, a total simulation time of 116  $\mu$ s was obtained.

At each studied temperature, we checked the percentage of presence of each peptide inside each receptor's binding pocket. The graphs reported in **Figure 1** (boxes A, B and C) show that all peptides significantly bound to any of the three receptors at 300 K, with a minimum of 60%. As expected, this value was decreasing as a function of the temperature, but with a sensitivity directly depending on the peptide receptor couple. Of interest, the respective sensitivities to temperature were in perfect qualitative agreement with our previously described experimental data (reported above in **Table 1**). Indeed, the "native" peptide 1 displaying a Tyrosine residue in position 11 was always giving the most potent binding results, both in the two native and in the designed R<sup>212</sup>E mutant receptors (boxes A, B and C in **Figure 1**). Our last experiments showed that the  $K_i$  of this peptide ( $K_i \cong 1$ nM) was considerably decreased ( $K_i > \mu$ M) when this Tyrosine was replaced by a negatively charged Glutamate or Aspartate residue. Indeed, we observed a higher sensitivity to the temperature for the peptides 3 and 4 in all three receptors.

The most interesting result came from the Lys modified peptide (peptide 2). From the experiments, we concluded that this peptide shows a significant selectivity for the NTS-2 receptor with a  $K_i \sim 200$  nM *versus* a  $K_i > 5 \mu$ M in the NTS-1 receptor. Again, results from our simulations were in good agreement with these observations, showing a reasonably good binding level of this peptide 2 to the NTS-2 receptor (**Figure 1B**) whereas it behaved very similarly to the negatively charged peptides in the NTS-1 receptor (**Figure 1A**). More surprisingly, our simulations were also able to predict the increased binding rate of this Lys-modified peptide 2 for the NTS-1 R<sup>212</sup>E mutant (**Figure 1C**), as observed experimentally. It was particularly noteworthy here because our simulations were able to capture differences among peptides and receptors that only differed by a single residue, despite the use of a Coarse-Grained model.

To confirm the role of this R<sup>212</sup> residue in the binding of the different peptides and especially of peptide 2, we also performed additional binding simulations of the four peptides in a designed NTS-2 E<sup>179</sup>R "reverse" mutant, the E<sup>179</sup> residue of NTS-2 corresponding to the R<sup>212</sup> residue of NTS-1<sup>18</sup>. As shown in **Figure 1D**, the binding curves obtained *in silico* for this NTS-2 mutant were highly related to those obtained in the native NTS-2 receptor, exception made for peptide 2 that reached a level even lower than those of peptides 3 and 4.

Our protocol was successful in capturing, at least qualitatively, the key molecular differences leading to the experimentally observed binding discrepancies. These results also confirmed a direct role for the R<sup>212</sup>/E<sup>179</sup> residues in the binding of the different peptides, and especially in the observed selectivity of the Lys-modified peptide 2.



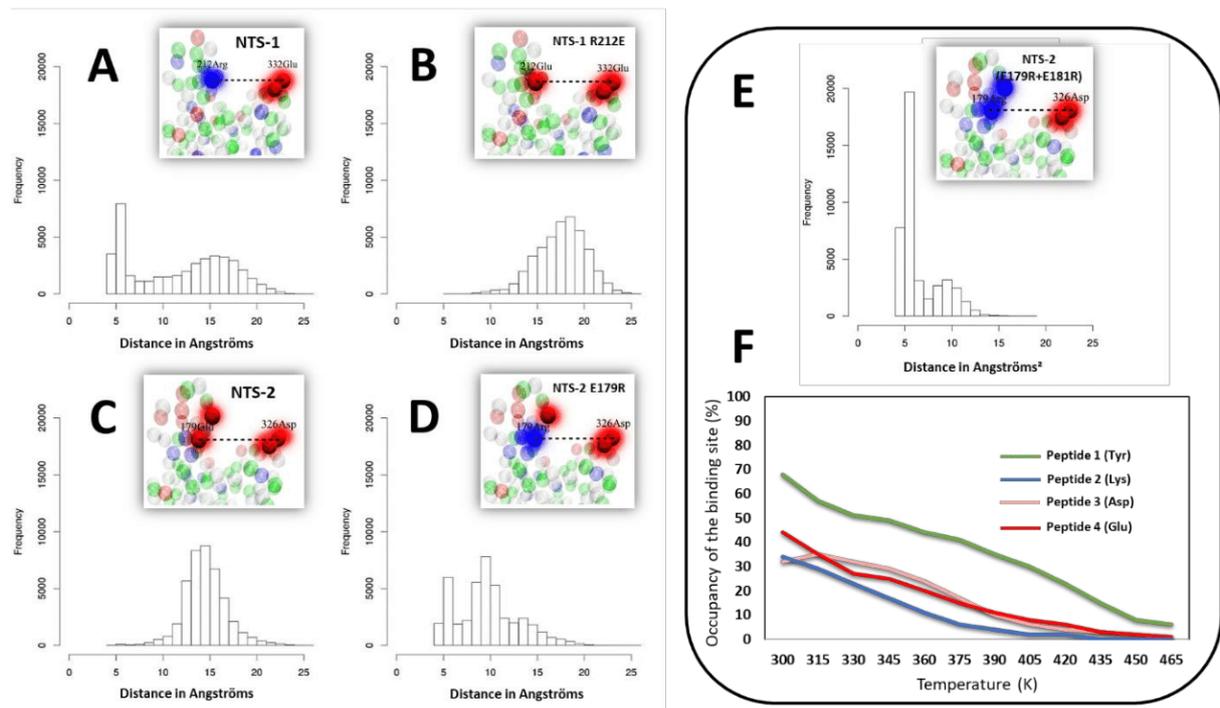
**Figure 1:** Temperature dependent occupancy of the binding site (expressed in % of snapshots) computed for peptides 1 to 4 in the four NTS-1, NTS-2, NTS-1(R<sup>212</sup>E) and NTS-2(E<sup>179</sup>R) Wild-Type/mutant receptors.

## The introduced mutations affect the opening/closing properties of the receptors

A key observation concerned the binding curves obtained for all the peptides in the NTS-1 receptor. Indeed, the temperature-dependent binding rates revealed an increased sensitivity of the wild-type NTS-1 receptor compared to those obtained for the three other receptors (see **Figure 1A**). While wild-type and mutant NTS-1 receptors differed only by one residue, the single point R<sup>212</sup>E mutation was nevertheless sufficient to recover some binding curves close to those observed for the NTS-2 receptor as shown in **Figure 1C**. This observation suggested a role of this mutation in the global dynamics of the NTS-1 receptor. To precise this role, we performed an additional Coarse-Grained (non-REMD) MD simulation of each isolated receptor. During these simulations, a closer look at the residue R<sup>212</sup> of NTS-1 showed a direct interaction with the facing, oppositely charged E<sup>322</sup> residue in ECL3 (see **Figure 2A**). A plot of the related distance distribution permitted to conclude that the receptor fluctuates between an opened and a closed conformation, corresponding to mean distances of 15 and 5 Angströms, respectively (see **Figure 2A**). We already observed such behavior in our previous work performed on the NTS-1 rat receptor<sup>18</sup>. The same analysis performed on the R<sup>212</sup>E NTS-1 mutant receptor showed that the single point mutation disrupted this interaction and opened the receptor (**Figure 2B**). A measure of the equivalent E<sup>179</sup>:D<sup>326</sup> distance in the NTS-2 receptor showed that this receptor was mainly open (**Figure 2C**). As expected, the E<sup>179</sup>R mutation introduced in the NTS-2 receptor contributed to close this receptor (**Figure 2D**), as shown by the appearance of conformations with mean distances of 5 Angströms. Other conformations were however observed in the NTS-2 E<sup>179</sup>R mutant corresponding to a mean distance around 10 Angströms. A possible explanation was the presence of a second negatively charged residue nearby, *e.g.* the E<sup>181</sup> residue hindering the R<sup>179</sup>:D<sup>326</sup> interaction (**Figure 2D**). In agreement, simulations

performed on the NTS-2 E<sup>179</sup>R+E<sup>181</sup>R double mutant confirmed this hypothesis and showed the prevalence of a closed conformation (**Figure 2E**). Accordingly, the resulting binding curves for the NTS-2 E<sup>179</sup>R+E<sup>181</sup>R double mutant exhibited a similar profile to those obtained with the wild-type NTS-1 receptor (**Figure 2F**).

Together, these results show that the inserted R<sup>212</sup>E or E<sup>179</sup>R mutations affect the global behavior of NTS-1 and NTS-2 receptors, the decreased binding/occupancy of all the peptides in the NTS-1 receptor being explained by a more important closure of this receptor. However, it is important to note that this analysis was not sufficient to explain the selectivity of the different peptides for the different receptors.



**Figure 2:** Effect of the introduced R<sup>212</sup>E and E<sup>179</sup>R mutations on the opening/closure of NTS-1 and NTS-2 mutant receptors.

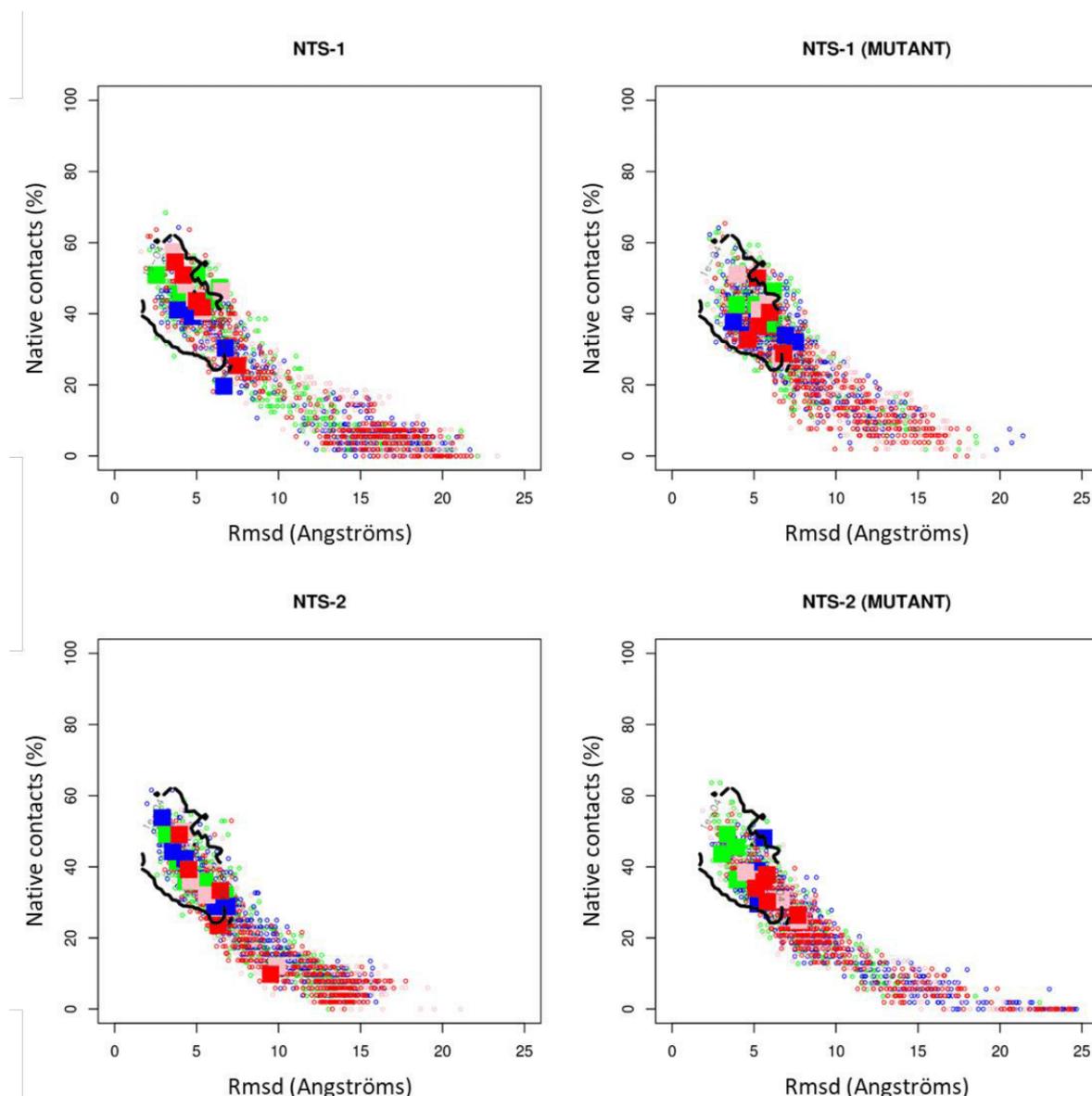
### A same binding route is observed for all the complexes

For each complex, all the resulting billions of conformations were clustered according to Root Mean Square Deviations (RMSD) computed on the peptide's backbone beads after alignment onto the receptors. Representative conformations (centroids of clusters) were extracted and analyzed. For simplification these representative conformations were further named "clusters". For each obtained cluster, we then computed both a RMSD against the peptide's reference position and a percentage of native contacts (see Methods).

The plots reported in **Figure 3** confirmed that a near-native orientation of the peptide was found in each of the 16 studied complexes with many clusters showing RMSD values lower than 5 Angströms and a percentage of native contacts greater than 40%. As reported in **Table S1**, such kind of values were notably retrieved among the 5 most populated clusters (reported as colored squares on each graph of **Figure 3**).

Interestingly, these values were in the range of those extracted from a standard CG simulation performed on the X-ray structure of reference, as delimited by the black thick line in **Figure 3**.

This figure showed that an index of 40% of the native contacts could be considered as a reasonably good orientation of the peptide in its receptor. This result was explained by the flexibility of the bound peptide, and in particular of its two N-terminus Arginines, all native contacts being usually not retrieved at the same time.



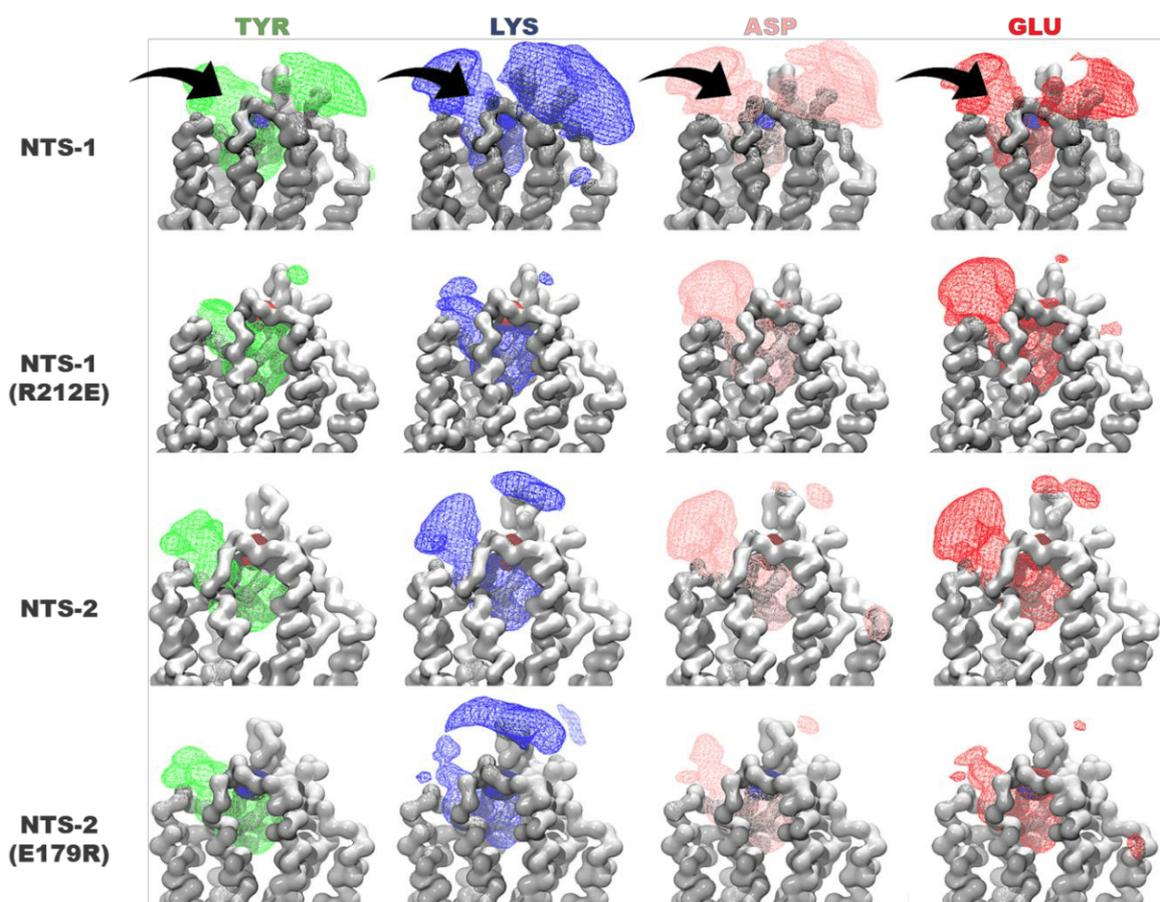
**Figure 3:** Relation between the percentage of native contacts and the RMSD of the peptide with respect to the reference structure, as computed for the 500 first clusters of each of the 16 studied complexes. The colored squares delineate the 5 most populated clusters. The same color scheme as for other figures has been employed. The thick black line indicates the region of the graph explored by a standard (non REMD) CG simulation starting from the reference structure. The squares represent the 5 most populated clusters retrieved in each complex.

	NTS-1				NTS-1 (R212E)				NTS-2				NTS-2 (E179R)			
	Y	K	D	E	Y	K	D	E	Y	K	D	E	Y	K	D	E
<b>Rmsd (Å)</b>	3,25	3,45	3,55	3,60	4,20	3,60	3,75	4,65	3,06	3,09	3,96	3,83	2,83	4,65	3,92	5,43
<b>Contacts (%)</b>	50,88	41,07	57,41	54,55	46,30	37,74	50,98	50,00	49,06	53,85	50,00	49,02	49,09	48,15	38,46	37,74

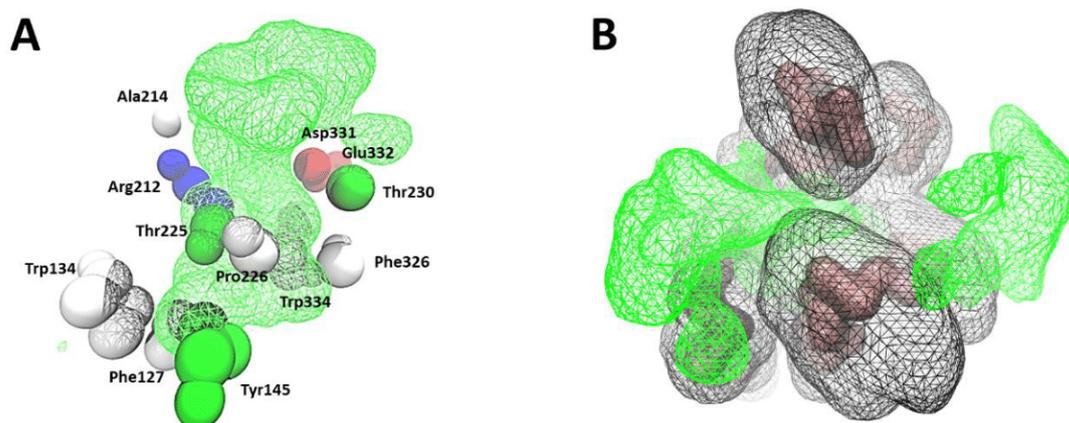
**Table S1:** Minimum RMSD of the peptide with respect to the reference structure and maximum % of native contacts retrieved among the 5 most populated clusters of each of the 16 studied complexes.

Positions of peptides in respect to the receptors were checked along all trajectories. Thus volumetric maps were computed for each complex, depicting the statistical distribution of each peptide at the surface of each receptor (see **Figure 4**). To simplify this figure, only the C-terminal Leu<sup>6</sup> residue of each peptide was considered, this residue lying at the bottom of the pocket in the bound conformation. These maps were also computed on a reduced set of snapshots extracted from the demuxed trajectories and sharing at least one peptide receptor contact. Together, these maps show that all the peptides followed a same route to bind to any of the four receptors. This route passed between the extra-cellular loops 2 and 3, starting at the top/end of TM5/ECL2, respectively. This route was interestingly lined by the residue that was mutated in our experiments (R<sup>212</sup> or E<sup>179</sup>). Other important receptor's residues along this route included Asp<sup>265</sup>, Glu<sup>266</sup>, Thr<sup>180</sup>, Phe<sup>260</sup>, Trp<sup>268</sup>, Pro<sup>176</sup>, Thr<sup>175</sup>, Ala<sup>164</sup>, Tyr<sup>95</sup>, Phe<sup>77</sup> and Trp<sup>84</sup> (NTS-1 numbering) as reported in **Figure S2A**.

These maps also permitted to conclude that an opening of the receptor was strictly required for the peptide binding. This was shown by the overlapping volumetric maps obtained for both the peptide and the receptor, as represented for the NTS-1:peptide 1 complex in **Figure S2B**.



**Figure 4:** Volumetric maps computed for each of the 16 complexes and showing the distribution of each peptide at the surface of each receptor. The same statistical threshold of  $p=0.0001$  was chosen in each case. The receptor's residue that was mutated was reported as a blue/red filled surface.



**Figure S2:** (A) receptor's residues (NTS-1 numbering) delineating the conserved binding route of the different peptides in the different receptors. (B) volumetric maps obtained at the same statistical threshold for both the NTS-1 receptor (backbone beads, black meshes) and the peptide 1 (C-terminus residue Leu<sup>6</sup>, green meshes) showing that an opening of the receptor is strictly required for peptide binding.

## DISCUSSION

Predicting the dynamical binding of a peptide to its target is a hard task and requires a proper sampling of all the conformations of both the isolated and the complexed partners. Here, we addressed such a task for a series of disordered peptides ~~sharing~~ displaying a selectivity for human NTS-1 and NTS-2 wild-type/mutant receptors. Our results show that the selectivity of these peptides for these different receptors could have been predicted, at least at the qualitative level, and without applying any intermolecular bias or restraint.

This was particularly noteworthy as we started here from homology models of the human NTS-1 and NTS-2 receptors, whereas until now we only applied our protocol to available X-ray structures. Of course, the studied receptors were here closely related to their template as the global identity rates with the NTS-1 rat receptor were 83% and 41% for NTS-1 and NTS-2 human receptors, respectively, and reaching almost 100% in the region of the binding site. Further studies will therefore be required to confirm that this protocol can also be applied to less related homology models of GPCRs. In other simulations still in progress, we already have demonstrated that this protocol can also be useful to predict the interactions of GPCRs with more structured peptides/small proteins (data not shown).

As compared to standard CG MD simulations, the used REMD based protocol allows to increase the sampling and the overcoming of energy barriers which otherwise trap the system in local energy minima, as it can be guessed in other studies using the same force field on related systems<sup>25-27</sup>. Through this study, we also show that the MARTINI force field is successful in capturing the main differences existing among a set of peptide receptor complexes that differ only by one single residue.

The use of a dynamical protocol not only helps to predict the final position of these peptides in the binding site, but also some intermediate ones that also play a key role in the observed selectivity. In the present study, both the final position and the binding route were conserved among all the studied complexes. Therefore, the selectivity of each peptide for each receptor could only be explained by its ability (easiness or hardness) to take this route.

We have shown that this route can be more or less ~~open~~ accessible according to the receptor, the later encountering some important conformational changes to allow the peptide entry.

From these calculations, we also concluded how hard it is to delineate the role of a particular peptide/receptor residue in the binding process, as most peptide receptor contacts are cross-correlated. This is particularly true because the percentage of each peptide receptor contact can vary significantly among trajectories, even in the present case where the binding pathway is conserved. Such analyses are therefore strongly conditioned by a good convergence of the data; in our case such a good convergence can be guessed by the obtained temperature dependent binding curves. Because of these convergence requirements, it would be very difficult at this time (if not impossible) to obtain the same kind of analyses at the medium throughput scale using a classical “all-atoms” model.

### Acknowledgements

We thank CALMIP, HPC@LR, CINES and PRACE for calculation resources. This work has been funded by ANR (Agence nationale de la recherche) via the projects BIP:BIP and ALLOSIG; We thank the CHEMISYST labex for financial support for PR and the University of Montpellier for BD; this work is also supported by the Canadian Institute of Health Research awarded to PS [grant number FDN-148413]. PS holds a Canada Research Chair in Neurophysiopharmacology of Chronic Pain.

### REFERENCES

1. Rosenbaum, D. M., Rasmussen, S. G. F. & Kobilka, B. K. The structure and function of G-protein-coupled receptors. *Nature* **459**, 356–363 (2009).
2. Dror, R. O. *et al.* Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 13118–13123 (2011).
3. Borhani, D. W. & Shaw, D. E. The future of molecular dynamics simulations in drug discovery. *J. Comput. Aided Mol. Des.* **26**, 15–26 (2012).
4. Kappel, K., Miao, Y. & McCammon, J. A. Accelerated molecular dynamics simulations of ligand binding to a muscarinic G-protein-coupled receptor. *Q. Rev. Biophys.* **48**, 479–487 (2015).
5. Charlier, L. *et al.* Coarse-Grained Simulations of the HIV-1 Matrix Protein Anchoring: Revisiting Its Assembly on Membrane Domains. *Biophys. J.* **106**, 577–585 (2014).
6. Chang, C.-E. A., Trylska, J., Tozzini, V. & McCammon, J. A. Binding pathways of ligands to HIV-1 protease: coarse-grained and atomistic simulations. *Chem. Biol. Drug Des.* **69**, 5–13 (2007).

7. Negami, T., Shimizu, K. & Terada, T. Coarse-grained molecular dynamics simulations of protein–ligand binding. *J. Comput. Chem.* **35**, 1835–1845 (2014).
8. Delort, B. *et al.* Coarse-Grained Prediction of Peptide Binding to G-Protein Coupled Receptors. *J. Chem. Inf. Model.* (2017). doi:10.1021/acs.jcim.6b00503
9. Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P. & de Vries, A. H. The MARTINI force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **111**, 7812–7824 (2007).
10. White, J. F. *et al.* Structure of the agonist-bound neurotensin receptor. *Nature* **490**, 508–513 (2012).
11. Wu, B. *et al.* Structures of the CXCR4 chemokine receptor in complex with small molecule and cyclic peptide antagonists. *Science* **330**, 1066–1071 (2010).
12. Mustain, W. C., Rychahou, P. G. & Evers, B. M. The role of neurotensin in physiologic and pathologic processes. *Curr. Opin. Endocrinol. Diabetes Obes.* **18**, 75–82 (2011).
13. Osadchii, O. E. Emerging role of neurotensin in regulation of the cardiovascular system. *Eur. J. Pharmacol.* **762**, 184–192 (2015).
14. Dobner, P. R. Neurotensin and pain modulation. *Peptides* **27**, 2405–2414 (2006).
15. Schroeder, L. E. & Leininger, G. M. Role of central neurotensin in regulating feeding: Implications for the development and treatment of body weight disorders. *Biochim. Biophys. Acta* **1864**, 900–916 (2018).
16. Gu, X. *et al.* Pharmacologically induced hypothermia attenuates traumatic brain injury in neonatal rats. *Exp. Neurol.* **267**, 135–142 (2015).
17. Ouyang, Q. *et al.* Oncogenic role of neurotensin and neurotensin receptors in various cancers. *Clin. Exp. Pharmacol. Physiol.* **44**, 841–846 (2017).

18. Fanelli, R. *et al.* Use of Molecular Modeling to Design Selective NTS2 Neurotensin Analogues. *J. Med. Chem.* (2017). doi:10.1021/acs.jmedchem.6b01848
19. Egloff, P. *et al.* Structure of signaling-competent neurotensin receptor 1 obtained by directed evolution in *Escherichia coli*. *Proc. Natl. Acad. Sci.* **111**, E655–E662 (2014).
20. Sali, A., Potterton, L., Yuan, F., van Vlijmen, H. & Karplus, M. Evaluation of comparative protein modeling by MODELLER. *Proteins* **23**, 318–326 (1995).
21. Lee, J. *et al.* CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **12**, 405–413 (2016).
22. Wu, E. L. *et al.* CHARMM-GUI Membrane Builder Toward Realistic Biological Membrane Simulations. *J. Comput. Chem.* **35**, 1997–2004 (2014).
23. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
24. Daura, X., Jaun, B., Seebach, D., van Gunsteren, W. F. & Mark, A. E. Reversible peptide folding in solution by molecular dynamics simulation. *J. Mol. Biol.* **280**, 925–932 (1998).
25. Provasi, D., Boz, M. B., Johnston, J. M. & Filizola, M. Preferred supramolecular organization and dimer interfaces of opioid receptors from simulated self-association. *PLoS Comput. Biol.* **11**, e1004148 (2015).
26. Prasanna, X., Chattopadhyay, A. & Sengupta, D. Cholesterol modulates the dimer interface of the  $\beta_2$ -adrenergic receptor via cholesterol occupancy sites. *Biophys. J.* **106**, 1290–1300 (2014).
27. Gahbauer, S. & Böckmann, R. A. Membrane-Mediated Oligomerization of G Protein Coupled Receptors and Its Implications for GPCR Function. *Front. Physiol.* **7**, (2016).

### 3.2.2.c Commentaires

Cet article sera prochainement soumis au journal « *Journal of Chemical Information and Modeling* », après avoir été refusé par le journal « *E-Life* ». Les principales remarques émises par l'éditeur d'*E-Life* concernaient le choix d'un modèle gros-grains ainsi que l'utilisation de récepteurs construits par homologie. Pourtant, notre premier article a déjà largement contribué à répondre à ces questions, comme déjà discuté plus haut dans ce manuscrit (voir partie 3.1.1). De plus, nous montrons une bonne concordance entre résultats théoriques et expérimentaux. La principale réserve émise ici et relayée par d'autres collègues est le manque de corrélation directe et quantitative entre les résultats obtenus en simulation et ceux obtenus par l'expérience. Au moment de l'écriture de ce manuscrit, nous réfléchissons à comment utiliser des méthodes d'analyses quantitatives, notamment la méthode WHAM [Kumar et al. 1992] adaptée à l'analyse de données issues de simulations REMD, pour évaluer l'énergie libre de liaison de chacun de nos peptides dans chacun de leurs récepteurs, et confronter ces résultats aux valeurs de  $K_i$  mesurées.

Cet article a mis beaucoup de temps à être finalisé, malgré l'obtention assez rapide des résultats bruts, notamment grâce à la rapidité de notre approche et à l'accès à des centres nationaux de calcul.

En effet, nous avons cherché pendant longtemps à expliquer la sélectivité au niveau moléculaire *via* notamment l'analyse par ACP de contacts par paires de résidus récepteur:peptide, ou bien encore l'analyse de ces contacts au cours du temps. Malheureusement, ces analyses n'ont pas permis de proposer un mécanisme simple. En effet, nous avons pu constater que la simple mutation d'un résidu de la boucle ECL2 du récepteur était capable d'entraîner un changement global du comportement de sa partie extra-cellulaire. D'autres analyses ont montré que le réseau d'interaction récepteur:peptide s'en retrouve complètement modifié et ce malgré un chemin d'entrée commun à tous les couples récepteur:peptide. Ce chemin qui passe entre les boucles ECL2 et ECL3 du côté des TM5 et TM6 oblige le peptide à franchir un pont ionique qui fait intervenir le résidu identifié dans nos études comme la clé de la sélectivité. Ce pont est également le moteur de la dynamique d'ouverture/fermeture du récepteur. La nature du résidu en position 11 du peptide va donc possiblement lui permettre d'emprunter plus ou moins facilement ce même chemin. Dans ce contexte, la liaison diminuée des peptides chargés négativement pourrait être expliquée par un problème de pré-positionnement à la surface du (des) récepteur(s). En effet, les boucles ECL2

et ECL3 étant très chargées négativement, ces peptides auraient alors du mal à se placer correctement à l'entrée de la poche.

Au final, cette étude montre tout l'intérêt de l'approche CG-REMD développée durant ma thèse. En des temps de simulations très réduits, on est capable d'observer la liaison d'un peptide à un RCPG sans la moindre contrainte inter-moléculaire appliquée. Le peptide va simplement diffuser dans l'eau avant d'interagir avec le récepteur et se placer dans la poche de liaison dans une conformation quasiment identique aux données expérimentales existantes. Malgré les simplifications liées à l'utilisation d'un champ de forces gros-grains, la méthode reste assez sensible pour pouvoir comparer (au moins sur le plan qualitatif) la liaison de différents peptides à leur(s) récepteur(s). Ces résultats très encourageants nous ont amenés à appliquer cette même méthode à d'autres cas d'études notamment pour prédire des interactions protéine-protéine impliquant les RCPGs. Ces différentes applications seront discutées dans la dernière partie de ce manuscrit (voir partie 3.4).

## 3.3 Prédiction des interactions RCPG:RCPG

### 3.3.1 Dimère de RCPGs

#### 3.3.1.a Contexte

Un modèle de dimère comprenant deux RCPGs et deux protéines-G a été proposé il y a quelques années dans l'équipe. Ce modèle se base sur trois structures cristallographiques différentes (**Figure 45**). Tout d'abord le complexe  $\beta 2AR:G_s$  (code PDB : 3SN6) résolu à 3,2Å [Rasmussen et al. 2011] a été utilisé pour modéliser l'interface RCPG:pG. La structure décrivant un dimère du récepteur CXCR4 (code PDB : 3ODU) résolu à 2,5Å [B. Wu et al. 2010] a été ensuite utilisée pour créer l'interface entre deux complexes  $\beta 2AR:G_s$ . En procédant ainsi, l'ouverture en deux de la sous-unité  $\alpha$  de la protéine-G telle que décrite dans la structure du complexe  $\beta 2AR:G_s$  conduit à des clashes stériques importants entre les deux sous-unités  $\alpha$ . En revanche, lorsqu'on utilise une structure inactive de la protéine-G hétéro-trimérique telle que  $G_i$  (code PDB : 1GP2), résolue à 2,3Å [Wall et al. 1995], le mode de dimérisation connu pour CXCR4 permet une imbrication parfaite des deux protéines-G. Ces assemblages faisant intervenir plusieurs protéines-G ont par ailleurs été proposés dans la littérature [Navarro et al. 2016; Cordoní et al. 2015]. Toutefois, aucun de ces modèles ne montre d'interactions protéine-G:protéine-G aussi complémentaires que dans notre modèle.

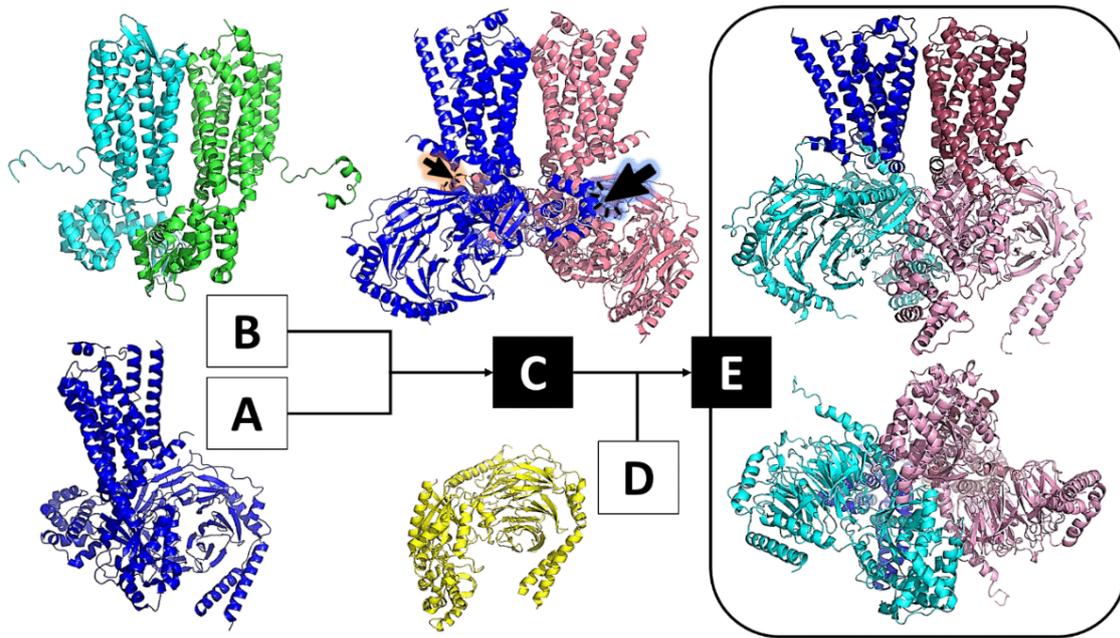


Figure 45 : Schématisation du processus de construction du dimère ; A) Complexe RCPG:pG donné par  $\beta 2AR:G_s$  (code PDB : 3SN6) [Rasmussen et al. 2011]; B) Complexe RCPG:RCPG donné par le dimère de CXCR4 (code PDB : 3ODU)[B. Wu et al. 2010]; C) Observation de clashes stériques entre les sous-unités  $G_\alpha$ ; D) Remplacement de  $G_s$  par une  $G_i$  inactive (code PDB : 1GP2) [Wall et al. 1995]; E) Dimère obtenu permettant une imbrication parfaite des deux Protéines-G.

Récemment, notre modèle de dimère a été validé dans le cadre d'une ANR multidisciplinaire ayant pour objectif d'étudier le récepteur de la Ghréline (GHSR-1a). Dans le travail présenté ci-après, nous avons étudié la capacité du récepteur GHSR-1a à se dimériser avec le récepteur de la Dopamine D2 (DRD2). Ce dimère est connu pour moduler les voies de signalisation dépendantes de la Dopamine [Kern et al. 2012]. L'originalité de ce travail a été de montrer que dans ce dimère DRD2:GHSR-1a, l'activation de  $G_i$  couplée au récepteur DRD2 est augmentée en présence de la protéine  $G_q$ , couplée quant à elle au récepteur GHSR-1a. Cette interdépendance suggère des interactions fortes entre les deux protéines-G, comme proposé dans notre modèle.

Le rôle de la modélisation dans cet article a été de proposer un modèle du dimère DRD2/ $G_i$ :GHSR-1a/ $G_q$  afin de discuter des contacts à l'interface de celui-ci. Pour construire ce dimère nous sommes partis du modèle proposé quelques années plus tôt. Puisque ce modèle était basé sur un dimère  $\beta 2AR:G_s$ , un important travail de modélisation par homologie a été réalisé avec le logiciel MODELLER. Le récepteur DRD2 a été construit par homologie en utilisant sa séquence issue de la base UniProt (P14416) qui présente un taux d'identité de 34%

avec  $\beta 2AR$ . Il est à noter que lorsque ce travail a été soumis pour publication, la structure du récepteur DRD2 n'avait pas encore été résolue [S. Wang et al. 2018]. De la même façon, le récepteur GHSR-1a a été construit par homologie (UniProt Q92847), en utilisant le récepteur  $\beta 2AR$  comme référence. Les deux récepteurs ne présentent que 23% d'identité de séquence (contre environ 35% entre GHSR-1a et le récepteur de la Neurotensine-1, utilisé plus loin dans ce manuscrit pour d'autres calculs). Aussi ce choix du récepteur  $\beta 2AR$  comme référence a été fait ici pour rester au plus près du modèle initialement construit, pour ne pas modifier l'interface entre les deux récepteurs. Dans notre modèle, la protéine  $G_i$  a été construite directement à partir de sa structure (code PDB : 1GP2), après alignement sur la protéine  $G_s$  co-cristallisée avec le récepteur  $\beta 2AR$ . Il en est de même pour  $G_q$  dont la structure cristallographique est également disponible (code PDB : 3AH8) [Nishimura et al. 2010] (résolution de 2,9Å). Pour valider expérimentalement ce modèle présentant deux récepteurs et deux protéines-G, des distances croisées entre les récepteurs et les protéines-G ont été mesurées par LRET (Transfert d'Énergie de Luminescence par Résonance). Les sondes utilisées en LRET ont été introduites à des endroits bien choisis de la (les) séquence(s) en se basant sur les résultats d'une simulation de dynamique moléculaire tout-atomes de 100ns du dimère ainsi construit. Ce choix s'est orienté vers les régions non structurées les moins flexibles du dimère. Des distances croisées  $G_{\alpha i}$ - $G_{\alpha q}$ ,  $G_{\alpha i}$ -GHSR-1a et DRD2-GHSR-1a ont ainsi été proposées puis validées par les mesures LRET effectuées au sein du laboratoire (**Tableau 8**). Pour la distance DRD2-GHSR-1a, il a été impossible d'exprimer le récepteur DRD2 marqué, expliquant l'absence de valeur expérimentale.

Distance	Point A	Point B	Distance simulation (Å)	Distance LRET (Å)
1	$G_{\alpha q}$ -Met1	$G_{\alpha i}$ -Glu289	23	20
2	GHSR-1a-THR156	$G_{\alpha i}$ -Lys314	38	35
3	GHSR-1a-THR156	DRD2-Arg150	30	/

Tableau 8 : Distances mesurées lors de simulations AA du dimère et validées par des mesures de LRET.

Cette étude très complète qui combine la modélisation moléculaire et la biophysique à des tests pharmacologiques a permis de proposer un mécanisme d'assemblage de ce dimère et surtout de montrer l'importance des contacts aux interfaces protéine-G:protéine-G dans sa fonctionnalité.

### 3.3.1.b Article



# GHSR-D2R heteromerization modulates dopamine signaling through an effect on G protein conformation

Marjorie Damian<sup>a</sup>, Véronique Pons<sup>b</sup>, Pedro Renault<sup>a</sup>, Céline M'Kadmi<sup>a</sup>, Bartholomé Delort<sup>a</sup>, Lucie Hartmann<sup>c</sup>, Ali I. Kaya<sup>d</sup>, Maxime Louet<sup>a</sup>, Didier Gagne<sup>a</sup>, Khoubaib Ben Haj Salah<sup>a</sup>, Séverine Denoyelle<sup>a</sup>, Gilles Ferry<sup>e</sup>, Jean A. Boutin<sup>e</sup>, Renaud Wagner<sup>c</sup>, Jean-Alain Fehrentz<sup>a</sup>, Jean Martinez<sup>a</sup>, Jacky Marie<sup>a</sup>, Nicolas Floquet<sup>a</sup>, Céline Galès<sup>b</sup>, Sophie Mary<sup>a</sup>, Heidi E. Hamm<sup>d</sup>, and Jean-Louis Banères<sup>a,1</sup>

<sup>a</sup>Institut des Biomolécules Max Mousseron, CNRS, Ecole Nationale Supérieure de Chimie de Montpellier, Université Montpellier, 34093 Montpellier, France; <sup>b</sup>I2MC, INSERM U1048, Université de Toulouse, F-31432 Toulouse, France; <sup>c</sup>Integral Membrane Proteins Research and Services (IMPREsS), UMR 7242 CNRS, Ecole Supérieure de Biotechnologie de Strasbourg, Université de Strasbourg, F-67412 Illkirch, France; <sup>d</sup>Department of Pharmacology, Vanderbilt University Medical Center, Nashville, TN 37232-6600; and <sup>e</sup>Pole d'expertise Biotechnologie, Chimie, Biologie, Institut de Recherches Servier, F-78290 Croissy-sur-Seine, France

Edited by Robert J. Lefkowitz, Howard Hughes Medical Institute and Duke University Medical Center, Durham, NC, and approved February 21, 2018 (received for review July 22, 2017)

**The growth hormone secretagogue receptor (GHSR) and dopamine receptor (D2R) have been shown to oligomerize in hypothalamic neurons with a significant effect on dopamine signaling, but the molecular processes underlying this effect are still obscure. We used here the purified GHSR and D2R to establish that these two receptors assemble in a lipid environment as a tetrameric complex composed of two each of the receptors. This complex further recruits G proteins to give rise to an assembly with only two G protein trimers bound to a receptor tetramer. We further demonstrate that receptor heteromerization directly impacts on dopamine-mediated Gi protein activation by modulating the conformation of its  $\alpha$ -subunit. Indeed, association to the purified GHSR:D2R heteromer triggers a different active conformation of G $\alpha$ i that is linked to a higher rate of GTP binding and a faster dissociation from the heteromeric receptor. This is an additional mechanism to expand the repertoire of GPCR signaling modulation that could have implications for the control of dopamine signaling in normal and physiopathological conditions.**

GPCR | G protein | heteromer | conformational dynamics | signaling

**G**protein-coupled receptors (GPCRs) are membrane proteins involved in regulating virtually every aspect of physiology. Most membrane receptors assemble as oligomeric complexes to fulfill their function (1). GPCRs do not escape the rule, as homo- and hetero-oligomers have been described for many different receptors (2). Although monomers can activate G proteins and recruit arrestins (3–5), association of different GPCRs within the same assembly can give rise to complexes with distinct and unique properties (6). This heteromer-directed signaling process has been proposed to be pivotal in modulating some of the physiological function of GPCRs as well as in disease-associated deregulations of signaling (6).

Ghrelin is a peptide hormone that has emerged as a major gut–brain signal controlling, among others, growth hormone secretion, food intake, and reward-seeking behaviors (7). These effects result from its interaction with the growth hormone secretagogue receptor (GHSR), a typical peptide-activated class A GPCR (7). Multiple data indicate that GHSR dimerizes with a wide array of GPCRs and that heteromerization could have an impact on its signaling properties and on those of the associated receptors (8). In this context, dimerization of GHSR with the dopamine D2 receptor is of particular interest. Indeed, assembly of D2R with apo-GHSR, which has been demonstrated in heterologous systems and hypothalamic neuron cultures, modulates dopamine-dependent signaling through the  $\beta\gamma$  subunits of D2R-associated Gi (9). However, so far, the molecular processes responsible for the impact of GHSR on D2R pharmacology remain to be illuminated.

To disentangle the mechanism through which interaction of apo-GHSR with D2R could affect dopamine signaling, we assembled

the purified ghrelin and dopamine receptors into lipid vesicles. By doing so, we demonstrate that, in liposomes, these two receptors interact to form a tetrameric complex comprising two each of GHSR and D2R. We further bring evidence that, in this model system, heteromerization alters dopamine-dependent Gi activation by modulating the conformational features of the G $\alpha$ <sub>i1</sub> subunit.

## Results

### Hetero-Oligomeric Arrangement of the Ghrelin and Dopamine Receptors in Lipid Vesicles.

We first investigated the propensity of GHSR to oligomerize when inserted in 200-nm PC/PE/PS/cholesterol large unilamellar vesicles. Large unilamellar vesicles were selected to limit segregation effects associated with membrane curvature (10). Receptor oligomerization was investigated using three-color FRET where fluorescein (Fluo), Cy3, and Cy5 were used as the donor, the transmitter, and the acceptor (11). To this end, isolated GHSR was labeled with either of the fluorophores on a unique reactive cysteine in the extracellular part of TM7 (C<sup>304</sup>) (12), mixed in equimolar amounts, and inserted into the liposomes. After proteoliposome assembly, a significant signal was observed at the emission wavelength of both Cy3 and Cy5 upon

## Significance

**G protein-coupled receptors (GPCRs) are one of the largest cell surface receptor family that transmit their signal through coupling to intracellular partners, such as G proteins. Receptor oligomerization has been shown to be pivotal in this signaling process. To address how oligomerization can impact on signaling in a major physiological process, dopamine signaling, we used a purified GPCR heteromer composed of the ghrelin and dopamine receptors to which we applied a variety of state-of-the-art biochemical and biophysical approaches. By doing so, we provide a direct experimental evidence for a mechanism where receptor heteromerization affects the conformation of the associated G protein. This sheds light on the way a GPCR oligomer can affect G protein activation to modulate signaling.**

Author contributions: M.D., V.P., P.R., J. Marie, N.F., C.G., and J.-L.B. designed research; M.D., V.P., P.R., C.M., B.D., L.H., A.I.K., M.L., D.G., R.W., N.F., and C.G. performed research; K.B.H.S., S.D., G.F., J.A.B., and J.-A.F. contributed new reagents/analytic tools; M.D., V.P., A.I.K., D.G., J.A.B., R.W., J.-A.F., J. Martinez, J. Marie, N.F., C.G., S.M., H.E.H., and J.-L.B. analyzed data; and A.I.K., J.A.B., R.W., J.-A.F., J. Martinez, J. Marie, N.F., S.M., H.E.H., and J.-L.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

<sup>1</sup>To whom correspondence should be addressed. Email: jean-louis.baneres@umontpellier.fr.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1712725115/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1712725115/-DCSupplemental).

Fluo excitation (Fig. 1A). This points to an oligomeric assembly comprising at least three GHSR protomers.

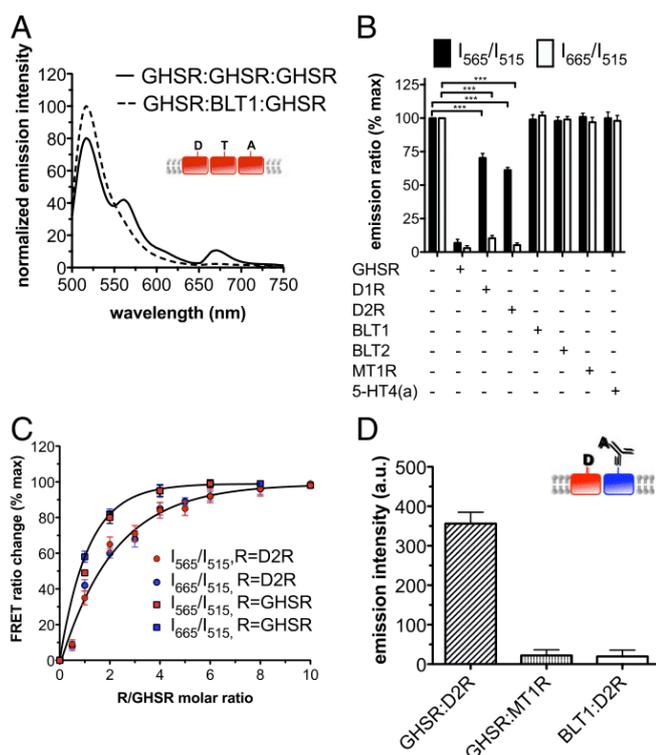
We then assembled into the liposomes labeled GHSR with unlabeled potential partners and monitored through Fluo-Cy3 (two-color) and Fluo-Cy3-Cy5 (three-color) FRET whether the additional receptor competed with GHSR homo-oligomerization. Because GHSR is essentially monomeric before insertion into the liposomes (*SI Appendix*, Fig. S1), the assay reports on the propensity of isolated receptors to form oligomers upon detergent-to-lipid exchange. To validate the method, Fluo-, Cy3-, and Cy5-labeled GHSRs were first mixed with a 10-fold molar excess of unlabeled GHSR before assembly into liposomes. Under such conditions, a decrease of both two-color and three-color FRET signals was observed (Fig. 1B), indicating that the unlabeled receptor competed with the assembly of labeled-GHSR oligomers. Consistently, saturation-like plots were obtained when varying the molar ratio of unlabeled-to-labeled GHSR (Fig. 1C), with a percent of fluorescence change at a 1:1 molar ratio in the same range than that measured in a cellular system (9). We then investigated whether other purified GPCRs could compete with GHSR oligomer formation. These included two receptors that had been shown to dimerize with GHSR (D1R, D2R) and unrelated GPCRs [BLT1, BLT2, MT1R, 5-HT4(a)]. Either of these receptors was mixed with GHSR labeled with Fluo, Cy3, and

Cy5 at a 10-fold GPCR-to-GHSR molar ratio and assembled into liposomes. A significantly lower FRET signal was observed in the presence of D1R and D2R only (Fig. 1B). In both cases, whereas a two-color FRET signal persisted, no more detectable three-color FRET could be observed. Of importance, most of the D2R and of the negative control receptors (BLT1, MT1R) inserted into the lipid vesicle in the same inside-in orientation as GHSR (*SI Appendix*, Fig. S2). A saturation-like profile was also obtained when the FRET signal was measured at increasing D2R-to-GHSR molar ratios (Fig. 1C). The ratio at which half of the change in the signal occurred was similar for two- and three-color FRET, suggesting that both changes result from the same molecular event. Overall, these data indicate that D2R disrupts GHSR oligomers while maintaining some homodimeric interactions. GHSR:D2R proximity in the lipid vesicles was further confirmed by measuring FRET between GHSR labeled with a Tb-cryptate donor on Cys<sup>304</sup> and D2R labeled with an XL255-labeled antibody reacting with its N-terminal Flag-tag (Fig. 1D).

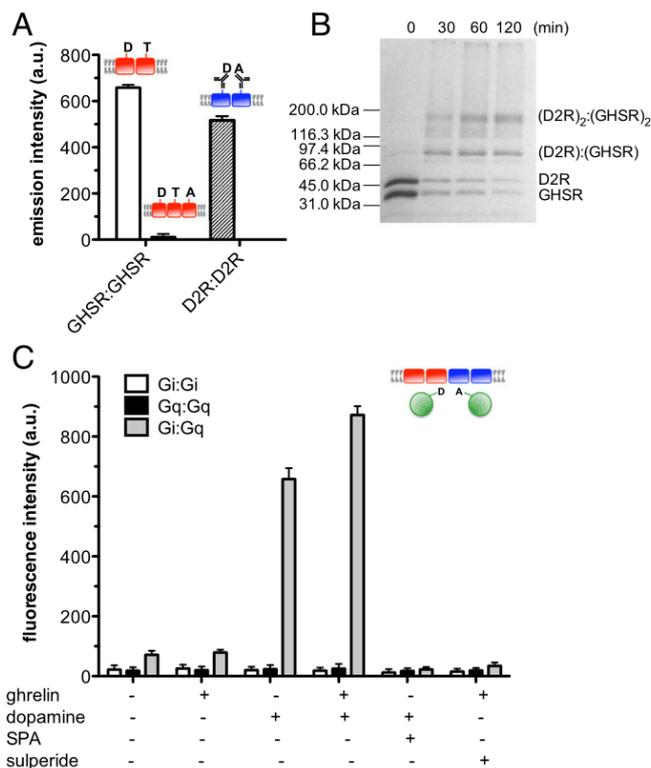
**The GHSR:D2R Heteromer.** We then purified the GHSR:D2R complex from the proteoliposomes. To this end, lipid vesicles containing GHSR and D2R were solubilized with the styrene maleic acid (SMA) copolymer (13, 14) to get membrane discs (SMALPs) encapsulating the receptors. SMA solubilization of purified vesicles is a fast process (15) so that the complexes initially present in the proteoliposomes are likely to be trapped as such into the SMALPs. Particles in the 20-nm range were obtained after solubilization of the proteoliposomes with SMA at low polymer-to-lipid ratios and purification by tandem chromatography (*SI Appendix*, Fig. S3). The complex obtained under such conditions was functional with regard to ligand-binding, G protein activation, and arrestin recruitment (*SI Appendix*, Fig. S4), with both D2R and GHSR fully ligand-competent (*SI Appendix*, Fig. S5).

A significant two-color signal could be measured when the SMALPs were prepared from liposomes containing D2R and Fluo-, Cy3-, and Cy5-labeled GHSR, whereas no three-color FRET could be detected (Fig. 2A). This suggests the occurrence of two but no more GHSR protomers in the SMALPs. These SMALPs likely included more than a single receptor, at least in the case of GHSR, as suggested by a GHSR dimer purification assay (*SI Appendix*, Fig. S6). A significant homogenous time-resolved fluorescence (HTRF) signal could also be measured upon labeling with Tb- and XL255-labeled antibodies recognizing the Flag-tag of D2R (Fig. 2A), indicating the presence of at least two D2R protomers. Finally, the highest molecular-weight limit species after chemical cross-linking (16) had an electrophoretic mobility compatible with that of a tetramer composed of two D2R and two GHSR (Fig. 2B and *SI Appendix*, Fig. S7). Taken together, these data suggest that the major complex trapped from the proteoliposomes could be a tetramer composed of two each of GHSR and D2R.

We then considered the organization of the complex formed between the purified D2R:GHSR heteromer in SMALPs and G proteins. To this end, we labeled  $G\alpha_q$  and  $G\alpha_{q1}$  at their N terminus with either a fluorescence donor (AF-350) or an acceptor (AF-488) (17) for G protein:G protein intermolecular FRET.  $G\alpha_{q1}$  was used because GHSR does not couple to this G protein subtype *in vitro* (5) and in HEK cells (18). We first assembled the purified heteromer in SMALPs with AF350- and AF488-labeled  $G\alpha_q\beta\gamma$  in equimolar amounts. In the absence or presence of dopamine, no FRET signal could be observed (Fig. 2C). This was not a limiting effect of the receptor-to-G protein ratio used, as increasing this ratio did not induce any FRET signal. This suggests that the tetrameric complex recruits only one  $G\alpha_q\beta\gamma$  trimer upon D2R activation. In the same way, no  $G\alpha_q:G\alpha_{q1}$  FRET could be detected in the absence or in the presence of ghrelin when incubating AF350- and AF488-labeled  $G\alpha_q\beta\gamma$  with the D2R:GHSR heteromer (Fig. 2C). In contrast, a FRET signal was observed in the presence of dopamine with AF350-labeled  $G\alpha_q$  and AF488-labeled  $G\alpha_{q1}$ , whether ghrelin was present or not (Fig. 2C). A possibility would be to consider a model where both G



**Fig. 1.** GHSR:D2R oligomerization in proteoliposomes. (A) Fluorescence emission spectra ( $\lambda_{exc}$ : 495 nm) from proteoliposomes containing Fluo-GHSR, Cy3-GHSR, and Cy5-GHSR (plain lines) or Fluo-GHSR, Cy3-BLT1, and Cy5-GHSR (dotted lines). (B) Changes in Fluo-Cy3 (I<sub>565</sub>/I<sub>515</sub>) and Fluo-Cy3-Cy5 (I<sub>665</sub>/I<sub>515</sub>) FRET in proteoliposomes containing Fluo-, Cy3-, and Cy5-GHSR and a 10-fold higher molar ratio of unlabeled receptors. (C) Changes in normalized emission intensity of Cy3 (I<sub>565</sub>/I<sub>515</sub>) and Cy5 (I<sub>665</sub>/I<sub>515</sub>) as a function of the GHSR:GHSR (boxes) or GHSR:D2R (circles) molar ratio. Data are presented as the percent of maximal change in FRET signal observed at the 8:1 (GHSR:GHSR) or 10:1 (D2R:GHSR) ratio. (D) XL255 emission intensity after Tb-cryptate excitation of proteoliposomes containing Tb-cryptate-labeled GHSR and XL255-labeled D2R, Tb-cryptate GHSR, and XL255 MT1R, or Tb-cryptate BLT1 and XL255 D2R. The different species are schematically represented (red: GHSR, blue: D2R; D: Fluo donor, T: Cy3 transmitter, A: Cy5 acceptor). Data in B and D are the mean  $\pm$  SD of three experiments (\*\**P* < 0.001).



**Fig. 2.** Stoichiometry of the GHSR:D2R:G protein complex in SMALPs. (A) Two- and three-color GHSR:GHSR and Tb-cryptate:XL255 D2R:D2R FRET for heteromer-containing SMALPs. (B) SDS/PAGE monitored DMS cross-linking kinetics (0, 30, 60, and 120 min) of the heteromer in SMALPs. (C) FRET between G protein trimers where the  $G\alpha$  subunit was labeled with AF-350 and AF-488 in the presence of the GHSR:D2R heteromer in the absence or in the presence of ligands. The different species are schematically presented (red: GHSR; blue: D2R; green: G protein; A: acceptor; D: donor; T: transmitter). Data in A and C are the mean  $\pm$  SD of three different experiments.

proteins would be associated to the receptor heterotetramer, with Gq preassembled with GHSR (17) and  $G\alpha_i$  recruited to D2R upon dopamine-mediated activation (19). Accordingly, no FRET signal could be observed in the presence of the GHSR inverse agonist SPA that dissociates the preassembled GHSR:Gq complex (17) (Fig. 2C). Finally, a significant  $G\alpha_i$ : $G\alpha_q$  FRET signal was observed upon stimulation with both dopamine and ghrelin. This signal was slightly different from that in the absence of ghrelin (Fig. 2C), possibly because of a different arrangement of the preassembled and active GHSR:Gq complexes (17).

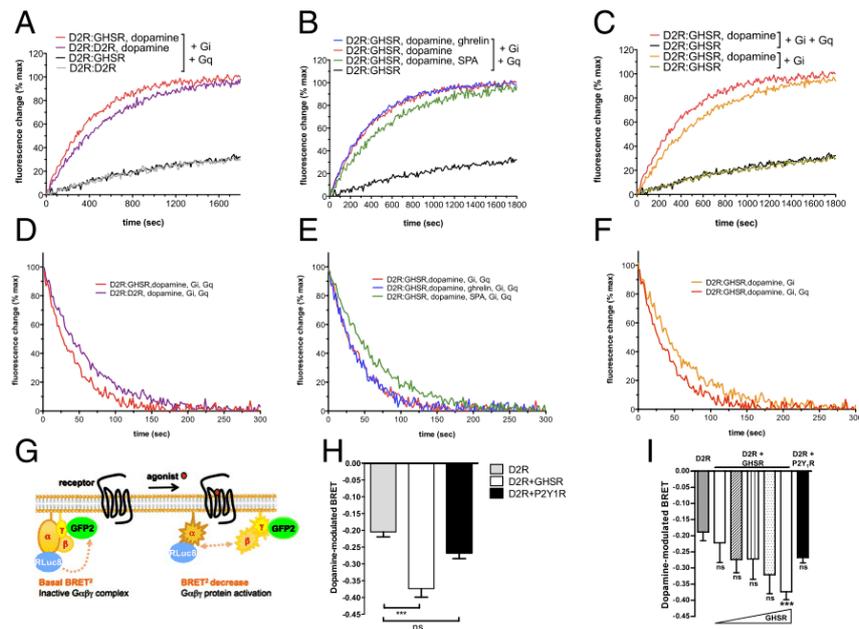
**Impact of GHSR:D2R Association on G Protein Activation.** We then analyzed whether heteromerization affected G protein activation. In this case, the model system included either the D2R homomer or the D2R:GHSR heteromer in SMALPs,  $G\alpha_{i1}\beta\gamma$  and  $G\alpha_q\beta\gamma$ . The D2R homomer was obtained after solubilization of D2R-only proteoliposomes (SI Appendix, Fig. S8). We first investigated  $G_i$  activation by measuring the rate of association of GTP $\gamma$ S to  $G\alpha_{i1}\beta\gamma$  in the preformed ternary complex. This rate was monitored through the changes in Trp emission that accompany GTP $\gamma$ S binding to  $G\alpha_i$  (SI Appendix, Fig. S9). To avoid any interference from  $G\alpha_q$ ,  $G\alpha_{i1}$  was labeled with 5-hydroxytryptophan (5HW) to selectively excite it without affecting other Trp residues (20). The receptors in SMALPs were incubated in the absence or presence of dopamine with both G proteins, and the rate of GTP $\gamma$ S association to  $G\alpha_i$  was measured. This rate was faster for the heteromer than for the homomer (Fig. 3A and SI Appendix, Table S1). The difference between the homomer and the heteromer was

not affected by ghrelin but was no more observed in the presence of SPA or in the absence of Gq (Fig. 3B and C).

We then investigated how G protein dissociated from the receptor after its activation. To this end, proteoliposomes were reconstituted in the presence of Fluo-labeled PE. SMALPs containing the fluorescent lipid and the different receptor protomers were then purified and assembled in the presence of dopamine with  $G\alpha_q\beta_1\gamma_2$  and  $G\alpha_i\beta_1\gamma_2$ , with the  $\beta\gamma$  subunits of  $G_i$  only labeled with AF-350 (21). These subunits were selected for monitoring dissociation as, in cell-based assays, modulation of D2R signaling by GHSR involved  $\beta\gamma$  from  $G_i$  (9). To avoid anchoring of the G proteins to the lipid bilayer and the associated nonspecific FRET, we used the soluble unlipidated mutants of all G protein subunits (22). Accordingly, no major FRET signal was observed with empty SMALPs or with SMALPs containing only GHSR (SI Appendix, Fig. S10). We then monitored the changes in FRET between the G protein and the fluorescein-labeled SMALPs as a function of time after addition of GTP $\gamma$ S. This time-course was different, depending whether the D2R homomer or the D2R:GHSR heteromer were considered (Fig. 3D). While the difference was essentially not affected by ghrelin, it was abolished by SPA or in the absence of Gq (Fig. 3E and F).

To then assess whether the effects we observed with the purified heteromer could impact on D2R signaling in a cellular environment, we monitored dopamine-mediated  $G\alpha_{i1}$  activation in HEK293T/17 cells using the bioluminescence resonance energy transfer (BRET) signal between RLuc8- $G\alpha_{i1}$  and GFP2- $G\gamma_2$  (Fig. 3G). In this assay, dopamine promotes D2R activation and subsequent dissociation between  $G\alpha$  and  $G\beta\gamma$ , thereby increasing the distance between the  $G\alpha_{i1}$  energy donor and the  $G\gamma_2$  energy acceptor and leading to a decrease in the BRET signal compared with the basal state (23–25). Hence, any effect of GHSR on D2R-dependent  $G_i$  activation should result in a change in the dopamine-modulated BRET signal. As shown in Fig. 3H, coexpressing GHSR with D2R decreased this BRET signal, indicative of an increase in  $G_i$  activation. This effect was dependent on the amount of GHSR and was not observed with an unrelated Gq-coupled GPCR, the purinergic receptor P2Y1 (Fig. 3H and I). Of importance, cotransfecting the different receptors did not result in a significant variation in the level of  $G\alpha_{i1}$  expression that could be responsible for the changes in BRET signal (SI Appendix, Fig. S11). These data suggest that apo-GHSR can increase dopamine-dependent  $G\alpha_{i1}$  activation in a cellular context as well.

We finally investigated the conformational features of  $G\alpha_i$  bound to homomer- and heteromer-containing SMALPs using intramolecular time-resolved luminescence resonance energy transfer (LRET) (26) with a lanthanide donor and a fluorescein acceptor both attached to  $G\alpha_i$ . To this end, we used  $G\alpha_i$  mutants with two unique reactive cysteines at different positions (i.e., 90 and 238, 141 and 333, or 171 and 276) (27). LRET measurements were carried out by monitoring the decay in the acceptor-sensitized emission, so that only the donor and acceptor engaged in LRET were detected (26) for the doubly labeled  $G\alpha_i$  of the  $G\alpha_i\beta\gamma$  trimer in its free- and receptor-bound states (Fig. 4). The free-state corresponded to  $G_i$  in the presence of Gq and empty SMALPs, and the receptor-bound state to  $G_i$  in the presence of Gq, SMALPs containing either the D2R homomer or the D2R:GHSR heteromer, and saturating concentrations in dopamine. In all cases, measurements were carried out with the empty state of the G protein (no GTP added, 30-min apyrase treatment) and dopamine was preincubated with the receptor. All decay plots were best fitted by two exponentials with a short and a long component (SI Appendix, Table S2). The short component was always predominant in the absence of receptor, whereas the long one was the major component in the presence of the receptor. This suggests that  $G\alpha_i$  is conformationally dynamic with at least two states whose relative abundance is governed by its coupling to the receptor. Whereas the short component was essentially unchanged, the long one was different depending whether the G protein was associated to the dopamine-activated homomer or heteromer. The major difference



**Fig. 3.** Heteromer modulation of G protein activation. (A) GTP $\gamma$ S binding to G $\alpha_{i1}$  catalyzed by the D2R homomer or the GHSR:D2R heteromer in the absence or the presence of dopamine and in the presence of Gq. (B) GTP $\gamma$ S binding to G $\alpha_{i1}$  catalyzed by the GHSR:D2R heteromer in the presence of Gq and in the absence of dopamine, in the presence of dopamine only or in the presence of dopamine and 10  $\mu$ M ghrelin or SPA, or (C) in the absence of dopamine and in the absence or presence of Gq. (D) FRET-monitored release of the  $\beta\gamma$  subunits of Gi from homomer- or heteromer-containing SMALPs in the presence of dopamine and Gq. (E) Release of G $\beta\gamma$  from heteromer-containing SMALPs in the presence of Gq and of dopamine or of dopamine and GHSR ligands, or (F) in the presence of dopamine and in the presence or absence of Gq. (G) Schematic representation of the BRET-based assay for G $\alpha_{i1}$  activation. (H) BRET signal in HEK293T/17 cells coexpressing RLuc8-G $\alpha_{i1}$ , GFP2-G $\gamma_2$ , and G $\beta_1$  in the presence of D2R alone or together with GHSR or P2Y $_1$ R after stimulation for 1 min with 10  $\mu$ M dopamine. (I) BRET signal between RLuc8-G $\alpha_{i1}$  and GFP2-G $\gamma_2$  in the presence of D2R alone or together with increasing amounts of GHSR or with P2Y $_1$ R after stimulation for 1 min with dopamine (10  $\mu$ M). In H and I, results are expressed as the difference in the BRET signal measured in the presence and the absence of dopamine (basal BRET signal measured in the presence of D2R alone taken as reference) and data represent the mean  $\pm$  SEM of at least three independent experiments. Statistical significance was assessed using one-way ANOVA followed by Dunnett's test (\*\*\*)  $P < 0.001$ ; ns, not statistically significant).

was observed for the 90–238 and 171–276 pairs (Fig. 4 B and C). No significant difference was observed for the 141–333 pair (Fig. 4D), possibly because the distance between the probes does not fall within a region of high sensitivity in LRET. Again, the LRET profiles of the heteromer were unaffected by ghrelin, whereas the difference was abolished by SPA (*SI Appendix, Fig. S12*) or in the absence of Gq (Fig. 4).

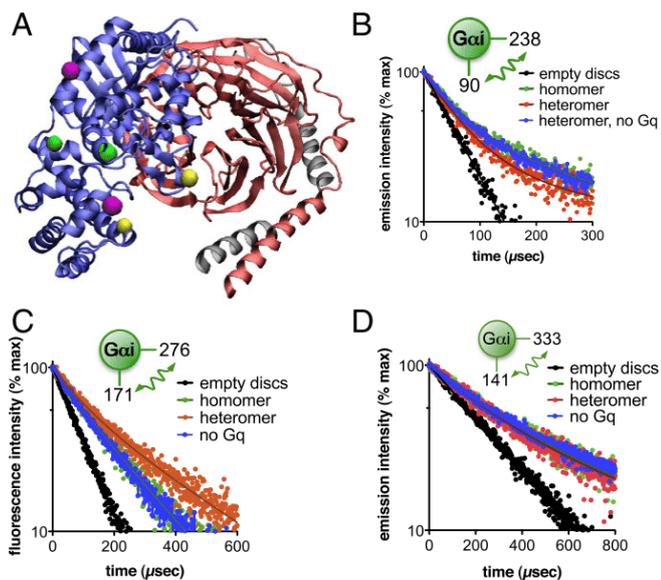
**Possible G Protein:G Protein Contacts in the Complex with the Heteromer.** The effect of heteromerization on Gi was not observed in the absence of Gq, suggesting that it could be related to Gq:Gi contacts. We used molecular dynamics to tentatively assess whether such contacts could possibly occur. Because of the lack of structural information on higher-order GPCR oligomers, we restricted the model to the putative central GHSR:D2R protomers. The model was built by homology modeling using the  $\beta_2$ -adrenergic receptor:Gs crystal structure for the GPCR:G protein complex (28) (Fig. 5 and *SI Appendix, Figs. S13–S15*). In this orientation, the main interface between the G proteins essentially involved the two  $\alpha$ -subunit ras-like domains (*SI Appendix, Figs. S13E and S14*), with the  $\alpha$ -helical one still mobile (*SI Appendix, Fig. S15*), in agreement with LRET. Although alternative organizations can be considered (29), this model nevertheless indicates possible allosteric contacts between the G proteins. Of importance, the second GHSR or D2R protomer added to account for a tetrameric arrangement couldn't accommodate a G protein (*SI Appendix, Fig. S16*), in agreement with the experimental stoichiometry.

We then carried out a preliminary series of measurements where LRET was used as an intermolecular ruler to assess whether some of the features of the model could be compatible with the arrangement of the purified assembly. When LRET was measured between G $\alpha_i$  labeled at its N terminus with fluorescein

and G $\alpha_i$  labeled with Lumi4-Tb on a unique reactive cysteine at position 289 (Fig. 5B), the acceptor-sensitized emission decay profile displayed a major component corresponding to a donor-acceptor distance in the 20-Å range (Fig. 5C). This distance was compatible with that measured from the model (in the 23-Å range) (Fig. 5B and E). When LRET was measured between Fluo bound at position 156 in the i2 loop of GHSR (17) and Lumi4-Tb bound to position 314 of G $\alpha_i$  Hexa I, the decay curve displayed two lifetime components corresponding to donor-acceptor distances in the 35- and 60-Å ranges. Only the shorter distance was compatible with the model (38 Å) (Fig. 5B and E). The longer one could result from LRET between G $\alpha_i$  and the second GHSR protomer, as both protomers were labeled within the heterotetramer. However, an exhaustive search procedure in the molecular dynamics experiments did not allow identifying an unambiguous interaction interface that would be compatible with the 60-Å distance (*SI Appendix, Fig. S16*), suggesting a higher dynamics of this protomer (*SI Appendix*). At the present stage, the model is thus considered a tentative one, the experimental validation of which will have to be strengthened further.

## Discussion

It had been shown that GHSR and D2R form a complex in living cells that modulates dopamine-mediated signaling (9). Our results confirm the specific interaction between these two receptors even in a simplified lipid system, suggesting this is an intrinsic feature of the receptors. As for the  $\beta_2$ -adrenergic receptor or rhodopsin (30), oligomerization occurred only upon reconstitution in liposomes, indicating that the interaction is dependent upon lipids. Additionally, our data show that the major heteromeric species formed in liposomes is likely a tetramer composed of two each of GHSR and D2R. However, the cellular environment could further modulate this stoichiometry, as it could also affect the relative



**Fig. 4.** Heteromer modulation of G protein conformation. (A) Ribbon model of  $G\alpha_i\beta\gamma$  (1GP2) showing the position of the labeled cysteines (yellow: C<sup>90</sup>, C<sup>238</sup>, magenta: C<sup>141</sup>, C<sup>333</sup>, green: C<sup>276</sup>, C<sup>171</sup>). (B–D) Sensitized-emission decays from  $G\alpha_{i1}$  labeled with Tb and fluorescein in the presence of empty SMALPs or of dopamine-activated D2R homomer or GHSR:D2R heteromer, in the absence or in the presence of Gq. The labeled positions are schematically presented in each case.

stability of the complex. Indeed, our data on receptor protomer exchange dynamics in the D2R:GHSR heteromer differ from what was reported in cellular systems for the D2R homomer. Indeed, the protomers appear to be in slow exchange in liposomes, whereas D2R was shown to form tetramers (31) with very dynamic (32) or even totally transient interactions (33) in cells. The plasma membrane is a complex environment with multiple proteins, mixtures of different lipids and cholesterol, asymmetry between the outer and inner layers, distinct domains, and no chemical equilibrium (34). This complexity can hardly be reproduced in vitro. Because both the lipid bilayer and additional proteins have been shown to modulate the association dynamics of membrane proteins through specific and nonspecific effects (35, 36), it is tempting to speculate that the differences in oligomerization dynamics result from a further modulation of the process we observed in simple lipid models by the different components of the cellular environment. These would include membrane composition and physical properties (lateral pressure, curvature elastic stress, thickness) (37), microdomains (38), interfacial lipids (39), as well as ancillary proteins (35, 40).

Association of GHSR and D2R within the same complex has been shown to affect dopamine-mediated signaling in a cellular system (9). Our data suggest that this effect on signaling could be related to an impact of GHSR:D2R association on the active conformation of  $G\alpha_i$ . In particular, the larger distance between LRET probes could translate a more open structure with an increased displacement of the helical vs. the nucleotide domains (27) when  $G\alpha_i$  is bound to the heteromer. Alternatively, the differences in LRET efficiency could result from distinct conformational dynamics between open and closed conformations of  $G\alpha_i$ . This increased amplitude and dynamics could be responsible for the faster nucleotide binding rates in the in vitro model system and the related increase in dopamine-mediated  $G_i$  activation we observed in HEK cells.

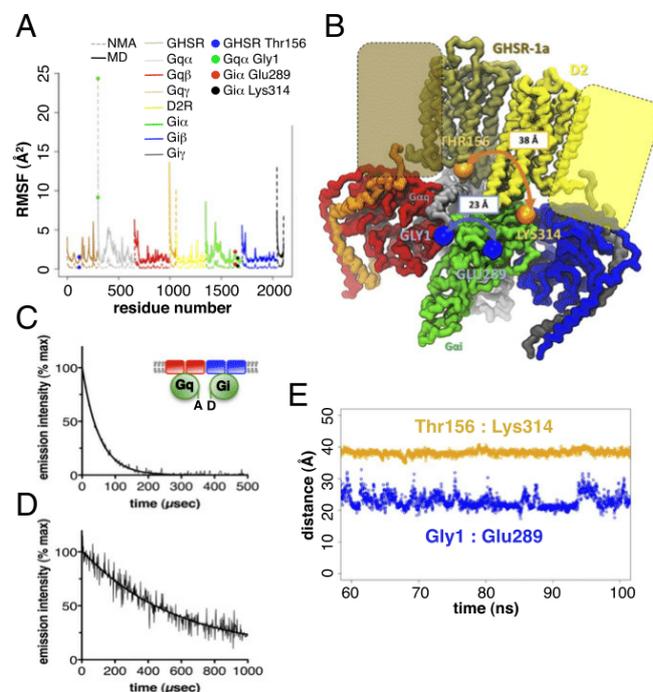
These differences in  $G\alpha_i$  conformation could be the consequence of specific protein:protein contacts within the (D2R)<sub>2</sub>:(GHSR)<sub>2</sub>: $G_i$ :Gq assembly that would allosterically modulate the conformational landscapes of both the receptors and the G proteins. A possibility would be that apo-GHSR in the heteromer

exerts an allosteric effect on the adjacent receptors that modifies D2R active conformation and its ability to trigger G protein activation. A second, nonexclusive mechanism would be to consider an influence of Gq on  $G_i$  activation, as in the GPCR priming model (41). Accordingly, the effects we observed were abolished in the absence of Gq or in the presence of the GHSR inverse agonist substance P analog that was shown to dissociate the preassembled GHSR:Gq complex. To be noted, an inverse agonist derived from substance P also abolished GHSR-dependent modulation of D2R signaling in living cells (9). However, caution needs to be exerted as, in contrast to our observations, Gq-directed siRNAs did not significantly reduce the effect of apo-GHSR on D2R signaling in cellular systems. A possibility would be that, in a cell, other G proteins that couple GHSR ( $G_{11}$ ,  $G_{13}$ ) (18) could take over Gq. Additional functional experiments in cellular systems and in vivo will thus be required to determine the broader physiological significance of our findings. At the present stage, our data nevertheless provide some evidence for a process where heteromerization could modulate G protein activation by promoting distinct changes in its conformation and in its conformational dynamics. This is reminiscent of what has been proposed to explain differences in G protein-activation efficacy by biased ligands of the calcitonin receptor (42), and suggests that receptor-dependent allosteric modulation of G protein conformation could be a way to extend the repertoire of mechanisms for modulating GPCR-mediated responses.

## Materials and Methods

All procedures are detailed in *SI Appendix*.

**Protein Preparation.** GHSR, BLT1, BLT2, and 5-HT4(a) were produced in *Escherichia coli* (5, 43). D1R, D2R, and MT1R were expressed in *Pichia pastoris* (44). The soluble unlipidated mutants of  $G\alpha_q$  and  $G\beta\gamma$  were produced in *sf9*



**Fig. 5.** The receptor:G protein model. (A) Root mean-squared fluctuations for each residue of the GHSR:D2R:Gq:Gi complex. Solid line: fluctuations from a 100 ns MD simulation; dotted lines: fluctuations obtained by normal mode analysis. (B) Model of the GHSR:D2R:Gi:Gq complex showing the positions of the LRET probes. (C and D) Sensitized-emission decays from (C) Tb-labeled  $G_i$  and Fluo-labeled Gq or (D) Tb-labeled  $G_i$  and Fluo-labeled GHSR. The different species are schematically presented (red: GHSR; blue: D2R; green: G proteins). (E) Time evolution of the corresponding distances along the last 50 ns of the molecular dynamics simulation.

(22).  $G\alpha_i$  was expressed in *E. coli* (27) and 5HW incorporated during bacterial expression (20).

**Protein Labeling.** GHSRC<sup>304</sup> or BLT1C<sup>266</sup> were labeled with fluorescein (ThermoFisher), Cy3-, or Cy5-maleimide (GE Healthcare) by incubating them with the dye at 4 °C for 16 h. For LRET, GHSR was labeled with Click-IT Alexa Fluor 488 DIBO Alkyne (Life Technologies) (18) and  $G\alpha_i$  Hexa I by adding the fluorophores in a stepwise manner.  $G\alpha_q$ ,  $G\alpha_{11}$ , and  $G\beta\gamma$  were labeled on their N terminus using the NHS derivative of the fluorophore (17, 21).

**Proteoliposome Assembly.** Receptors were inserted into  $\beta$ -DDM solubilized 200 nm DOPC, DOPE, DOPS (40/40/20) (Avanti lipids) liposomes containing cholesterol (0.2 cholesterol-to-lipid molar ratio) (13) and further purified on a 5–30% linear sucrose gradient.

**FRET, HTRF, LRET, and BRET Measurements.** Fluorescence emission spectra were recorded on a Cary Eclipse spectrofluorimeter (Varian) with an excitation at 495 nm. For HTRF-monitored dimerization assays, the signal was optimized by varying concentrations in the acceptor-labeled antibody at fixed donor concentration (45). In LRET assays, the donor lifetimes in the presence of the acceptor were measured through the acceptor-sensitized emission at 515 nm ( $\lambda_{exc}$ : 337 nm). G protein-activation experiments in HEK293T/17 cells were performed as previously described (23–25).

- Klingenberg M (1981) Membrane protein oligomeric structure and transport function. *Nature* 290:449–454.
- Franco R, Martínez-Pinilla E, Lanciego JL, Navarro G (2016) Basic pharmacological and structural evidence for class A G-protein-coupled receptor heteromerization. *Front Pharmacol* 7:76.
- Whorton MR, et al. (2007) A monomeric G protein-coupled receptor isolated in a high-density lipoprotein particle efficiently activates its G protein. *Proc Natl Acad Sci USA* 104:7682–7687.
- Bayburt TH, et al. (2011) Monomeric rhodopsin is sufficient for normal rhodopsin kinase (GRK1) phosphorylation and arrestin-1 binding. *J Biol Chem* 286:1420–1428.
- Damian M, et al. (2012) High constitutive activity is an intrinsic feature of ghrelin receptor protein: a study with a functional monomeric GHS-R1a receptor reconstituted in lipid discs. *J Biol Chem* 287:3630–3641.
- Ferré S, et al. (2014) G protein-coupled receptor oligomerization revisited: functional and pharmacological perspectives. *Pharmacol Rev* 66:413–434.
- Müller TD, et al. (2015) Ghrelin. *Mol Metab* 4:437–460.
- Wellman M, Abizaid A (2015) Growth hormone secretagogue receptor dimers: A new pharmacological target. *eNeuro* 2:ENEURO.0053-14.2015.10.1523/ENEURO.0053-14.2015.
- Kern A, Albarran-Zeckler R, Walsh HE, Smith RG (2012) Apo-ghrelin receptor forms heteromers with DRD2 in hypothalamic neurons and is essential for anorexigenic effects of DRD2 agonism. *Neuron* 73:317–332.
- Rosholm KR, et al. (2017) Membrane curvature regulates ligand-specific membrane sorting of GPCRs in living cells. *Nat Chem Biol* 13:724–729.
- Olejko L, Cywiński PJ, Bald I (2016) An ion-controlled four-color fluorescent telomeric switch on DNA origami structures. *Nanoscale* 8:10339–10347.
- Mary S, et al. (2012) Ligands and signaling proteins govern the conformational landscape explored by a G protein-coupled receptor. *Proc Natl Acad Sci USA* 109:8304–8309.
- Logez C, et al. (2016) Detergent-free isolation of functional G protein-coupled receptors into nanometric lipid particles. *Biochemistry* 55:38–48.
- Dörr JM, et al. (2016) The styrene-maleic acid copolymer: a versatile tool in membrane research. *Eur Biophys J* 45:3–21.
- Oluwole AO, et al. (2017) Solubilization of membrane proteins into functional lipid-bilayer nanodiscs using a diisobutylene/maleic acid copolymer. *Angew Chem Int Ed Engl* 56:1919–1924.
- Arcemibéhère L, et al. (2010) Leukotriene BLT2 receptor monomers activate the  $G(i)2$  GTP-binding protein more efficiently than dimers. *J Biol Chem* 285:6337–6347.
- Damian M, et al. (2015) Ghrelin receptor conformational dynamics regulate the transition from a preassembled to an active receptor:Gq complex. *Proc Natl Acad Sci USA* 112:1601–1606.
- M'Kadmi C, et al. (2015) Agonism, antagonism, and inverse agonism bias at the ghrelin receptor signaling. *J Biol Chem* 290:27021–27039.
- Bondar A, Lazar J (2017) The G protein  $G_{11}$  exhibits basal coupling but not preassembly with G protein-coupled receptors. *J Biol Chem* 292:9690–9698.
- Mesnier B, Banères JL (2004) Cooperative conformational changes in a G-protein-coupled receptor dimer, the leukotriene B(4) receptor BLT1. *J Biol Chem* 279:49664–49670.
- Zurawski Z, et al. (2017)  $G\beta\gamma$  directly modulates vesicle fusion by competing with synaptotagmin for binding to neuronal SNARE proteins embedded in membranes. *J Biol Chem* 292:12165–12177.

**Liposome Solubilization and SMALP Preparation.** The proteoliposomes in 50 mM Tris-HCl, 200 mM NaCl pH 8 were incubated at 25 °C with the SMA copolymer at a lipid-to-polymer molar ratio of 0.10. The S-tag of GHSR and the Flag-tag of D2R were used to purify heteromer-containing SMALPs (20).

**Chemical Cross-Linking.** Proteins in SMALPs were cross-linked in the presence of DMS to a final concentration of 1 mg/mL. The reaction was stopped by addition of glycine to a concentration of 50 mM and the cross-linked species submitted to SDS/PAGE after extraction by  $CH_3OH/CHCl_3/H_2O$  (15).

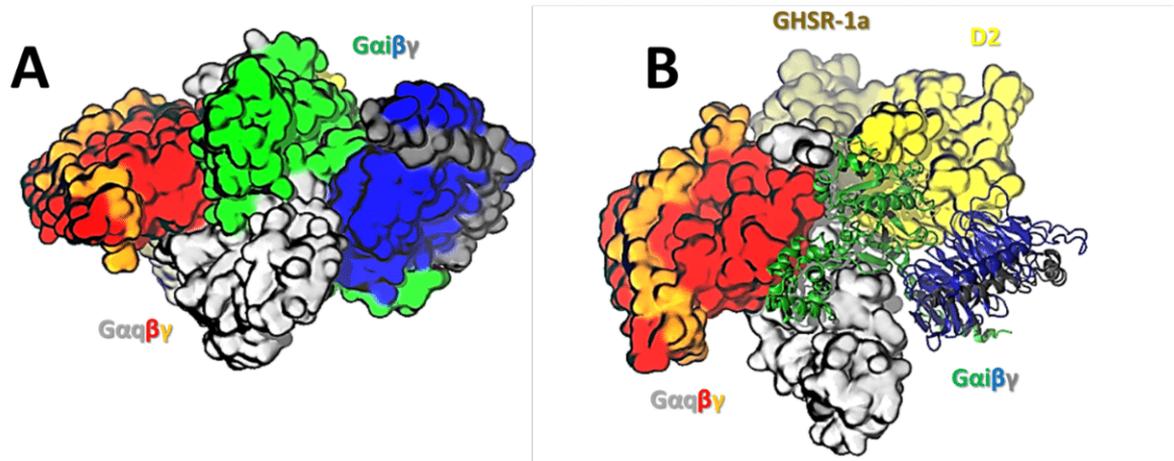
**GTP $\gamma$ S Binding and G Protein Dissociation Assays.** GTP $\gamma$ S binding experiments were carried out using Bodipy-FL GTP $\gamma$ S (5). The rate of GTP $\gamma$ S binding to  $G_i$  was determined by monitoring the relative increase in the intrinsic 5HW fluorescence ( $\lambda_{exc}$ : 315 nm;  $\lambda_{em}$ : 350 nm) as a function of time after the addition of GTP $\gamma$ S.

**ACKNOWLEDGMENTS.** We thank J.-P. Pin (Institut de Génétique Fonctionnelle) for a critical reading of the manuscript; R. K. Sunahara (University of California, San Diego) for his comments on the G protein luminescence resonance energy transfer data; and E. Trinquet and T. Roux (CisBio) for their advice on homogenous time-resolved fluorescence. This work was supported by CNRS, Université Montpellier, Agence Nationale de Recherches (ANR-13-BSV8-0006-01), the Chemist Labex (P.R.), and the Labex EpiGenMed, an “Investissements d’avenir” Program ANR-10-LABX-12-01 (to K.B.H.S.). Demande d’Attribution de Ressources Informatiques and Centre Informatique National de l’Enseignement Supérieur provided computational resources (Project 2016077530). This study was performed with the support of HPC@LR.

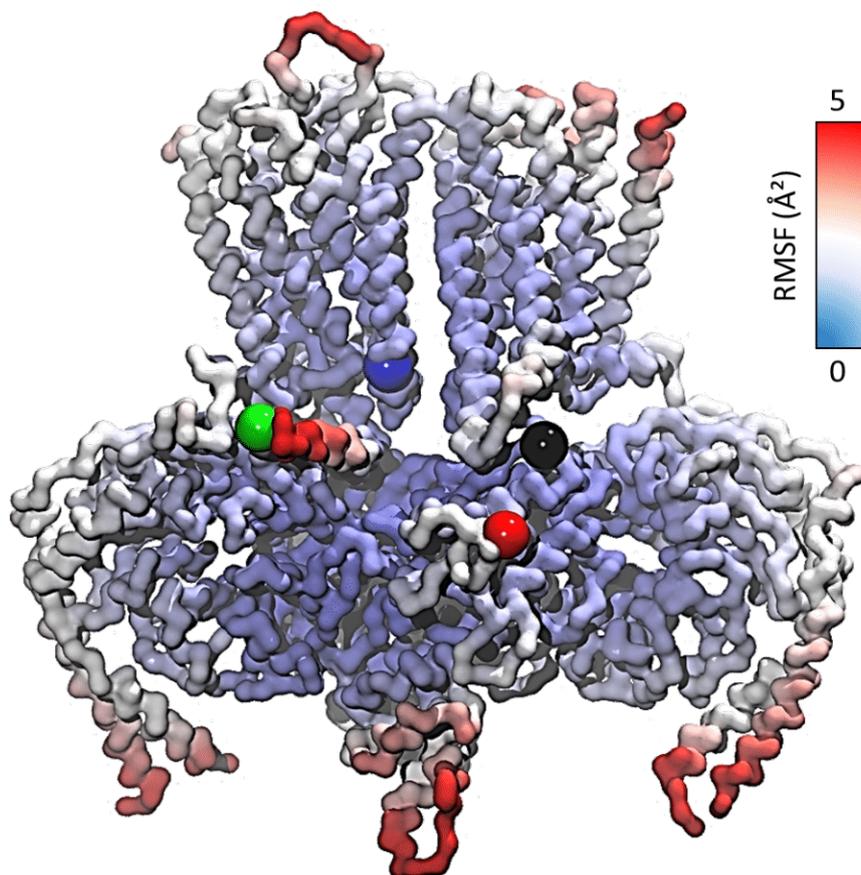
- Kozasa T (2004) Purification of G protein subunits from Sf9 insect cells using hexahistidine-tagged alpha and beta gamma subunits. *Methods Mol Biol* 237:21–38.
- Galandrin S, et al. (2016) Cardioprotective angiotensin-(1-7) peptide acts as a natural-biased ligand at the angiotensin II type 1 receptor. *Hypertension* 68:1365–1374.
- Bellot M, et al. (2015) Dual agonist occupancy of AT1-R- $\alpha$ 2C-AR heterodimers results in atypical Gs-PKA signaling. *Nat Chem Biol* 11:271–279.
- Saulière A, et al. (2012) Deciphering biased-agonism complexity reveals a new active AT1 receptor entity. *Nat Chem Biol* 8:622–630.
- Selvin PR (2000) The renaissance of fluorescence resonance energy transfer. *Nat Struct Biol* 7:730–734.
- Van Eps N, et al. (2011) Interaction of a G protein with an activated receptor opens the interdomain interface in the alpha subunit. *Proc Natl Acad Sci USA* 108:9420–9424.
- Rasmussen SG, et al. (2011) Crystal structure of the  $\beta_2$  adrenergic receptor-Gs protein complex. *Nature* 477:549–555.
- Navarro G, et al. (2016) Quaternary structure of a G-protein-coupled receptor heterotetramer in complex with  $G_i$  and Gs. *BMC Biol* 14:26.
- Goren MA, et al. (2014) Constitutive phospholipid scramblase activity of a G protein-coupled receptor. *Nat Commun* 5:5115.
- Guo W, et al. (2008) Dopamine D2 receptors form higher order oligomers at physiological expression levels. *EMBO J* 27:2293–2304.
- Tabor A, et al. (2016) Visualization and ligand-induced modulation of dopamine receptor dimerization at the single molecule level. *Sci Rep* 6:33233.
- Fonseca JM, Lambert NA (2009) Instability of a class A G protein-coupled receptor oligomer interface. *Mol Pharmacol* 75:1296–1299.
- Sezgin E, Levental I, Mayor S, Eggeling C (2017) The mystery of membrane organization: composition, regulation and roles of lipid rafts. *Nat Rev Mol Cell Biol* 18:361–374.
- Hong H, Bowie JU (2011) Dramatic destabilization of transmembrane helix interactions by features of natural membrane environments. *J Am Chem Soc* 133:11389–11398.
- Pawar AB, et al. (2015) Thermodynamic and kinetic characterization of transmembrane helix association. *Phys Chem Chem Phys* 17:1390–1398.
- Soubias O, Teague WE, Jr, Hines KG, Gawrisch K (2015) Rhodopsin/lipid hydrophobic matching-rhodopsin oligomerization and function. *Biophys J* 108:1125–1132.
- Pluhackova K, Gahbauer S, Kranz F, Wassenaar TA, Böckmann RA (2016) Dynamic cholesterol-conditioned dimerization of the G protein coupled chemokine receptor type 4. *PLOS Comput Biol* 12:e1005169.
- Gupta K, et al. (2017) The role of interfacial lipids in stabilizing membrane protein oligomers. *Nature* 541:421–424.
- Navarro G, et al. (2014) Intracellular calcium levels determine differential modulation of allosteric interactions within G protein-coupled receptor heteromers. *Chem Biol* 21:1546–1556.
- Gupte TM, Malik RU, Sommese RF, Ritt M, Sivaramakrishnan S (2017) Priming GPCR signaling through the synergistic effect of two G proteins. *Proc Natl Acad Sci USA* 114:3756–3761.
- Furness SG, et al. (2016) Ligand-dependent modulation of G protein conformation alters drug efficacy. *Cell* 167:739–749.e11.
- Dahmane T, Damian M, Mary S, Popot JL, Banères JL (2009) Amphipol-assisted in vitro folding of G protein-coupled receptors. *Biochemistry* 48:6516–6521.
- André N, et al. (2006) Enhancing functional production of G protein-coupled receptors in *Pichia pastoris* to levels required for structural studies via a single expression screen. *Protein Sci* 15:1115–1126.
- Albizu L, et al. (2010) Time-resolved FRET between GPCR ligands reveals oligomers in native tissues. *Nat Chem Biol* 6:587–594.

**Modeling.** As a first step towards the building of the full GHSR:D2R:Gq:Gi model, we built a dimer of the  $\beta_2$ -adrenergic receptor (B2AR):Gs complex. The structure of this complex (PDB CODE 3SN6) (46) was duplicated and the receptors were first positioned according to the orientation of the CXCR4 dimer (PDB CODE 3ODU) (47), with each B2AR fit to one of the CXCR4 monomers. However, when the B2ARs adopt such an orientation, steric clashes between the G proteins were observed. Therefore, a full model of the hetero-trimeric G protein was built by homology modeling using the MODELLER software (48) and the structure of the GDP bound hetero-trimeric Gi protein (PDB CODE 1GP2) (49) as a template. Lacking parts in the  $\alpha$  subunit were reconstructed from available structures in the PDB. Lacking parts of B2AR were built by homology using the swissprot sequence as a query. At this step, residues allowing anchoring of the receptors (palmytoylated Cys<sup>326</sup>) and the G proteins (myristoylated Gly and palmytoyl Cys at positions 1 and 2 of the  $\alpha$  subunit and geranyl Cys at position 68 of the  $\gamma$  subunit) to the lipid membrane were added using the Charmm-gui server (50). The dimer was oriented towards the Z-axis and inserted in a 100% POPC membrane model. Water molecules and ions were then added prior to equilibration of the full system through different steps of restrained/unrestrained molecular dynamics, as previously described (51). After equilibration, a 100 ns NPT (1 bar, 300K) MD simulation of the system further refined the model. The NAMD software (52) was used together with the CHARMM 36 forcefield (53) for proteins and lipids and the TIP3P model for water molecules (54). The integration step was fixed to 2 fs, using a cut-off value of 12Å for the computation of non-bonded terms and the Particle Mesh Ewald summation for long-range electrostatic interactions (55). The so obtained refined model was further used as a reference to build the GHSR:D2R:Gq:Gi model. Again, using MODELLER for homology modeling, the first B2AR:Gs complex was replaced by a GHSR:Gq pair of proteins using the GHSR sequence (uniprot Q92847) and the structure of the Gq protein (56) as building materials. The other complex was modified into a dopamine D2R:Gi couple of proteins, the Gi protein being built directly from its X-ray structure whereas the D2R receptor was built from uniprot sequence P14416. This full GHSR:D2R:Gq:Gi system was then refined by MD simulation using the same protocol as that described for the B2AR:Gs dimeric complex.

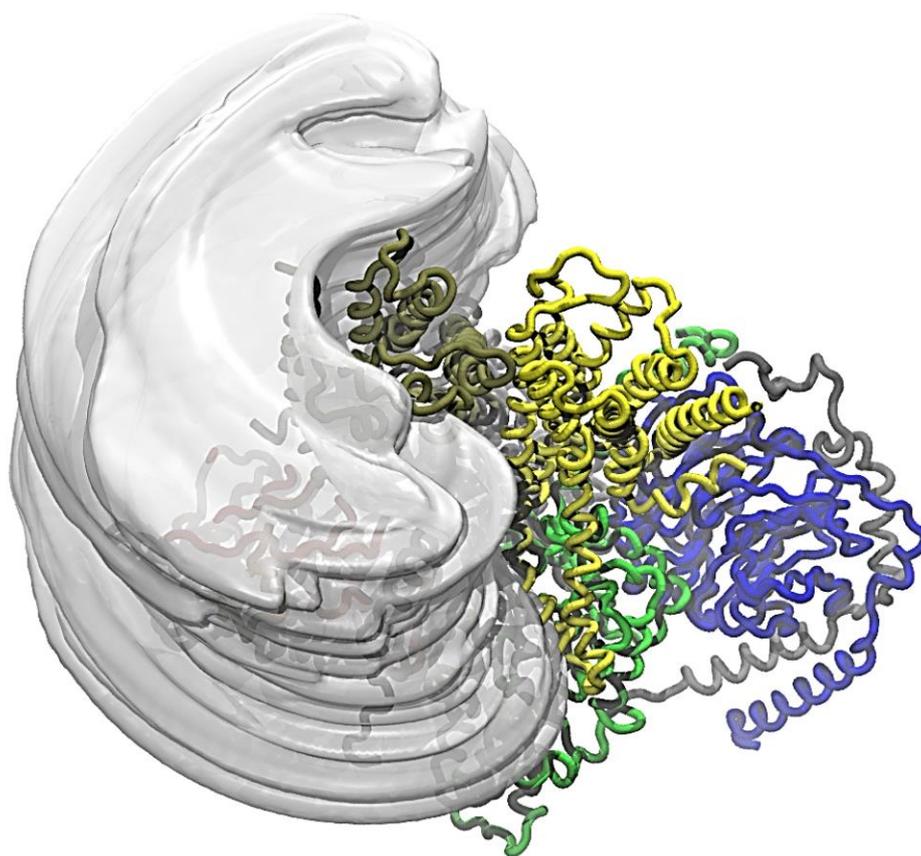
**Normal Mode Analysis.** Normal mode analysis was based on an elastic network model where only the  $\alpha$ -carbons were represented. The force constants of the springs connecting interacting atoms were defined according to the parameters developed by Hinsen et al. (57), which basically distinguish between nearest-neighbors and the other pairs of atoms. These calculations were performed with the Bio3d version 2.0 (58,59).



**Figure S14.** Two different bottom views of the GHSR:D2R:Gq:Gi model showing imbrication of the two G proteins and the interactions at the interface  $G\alpha:G\alpha$  and  $G\alpha:G\beta$ .



**Figure S15.** Model of the GHSR:D2R:Gq:Gi complex colored according to the root mean squared fluctuations (RMSF), showing more flexible regions in red. These include the  $\alpha$ -helical domain of  $G\alpha$ , in full agreement with their dynamics observed in the LRET measurements, with fluctuation amplitudes during the MD trajectories compatible with the LRET data. The positions of the fluorescent probes used to estimate the distances in LRET experiments are indicated by spheres.



**Figure S16.** Volumetric map resulting from an exhaustive search procedure showing the possible locations of the second GHSR protomer that are compatible with the experimentally measured distance of 60Å between C<sup>314</sup> of G $\alpha_i$  and T<sup>156</sup> of GHSR. This receptor must be in close contact with the first GHSR protomer, but several positions/orientations around the GHSR:D2R:Gq:Gi complex are consistent with the experimentally measured distance. By symmetry, similar considerations apply to the second D2R protomer.

#### 3.3.1.c Commentaires

Cet article a été soumis le 22 Juillet 2017 au journal « *Proceedings of the National Academy of Sciences* » pour être finalement publié le 21 Février 2018. L'article a connu deux phases de révisions majeures qui concernaient principalement l'obtention de contrôles supplémentaires. Les révisions ont également porté sur la possible extrapolation de ces résultats à un modèle cellulaire. Concernant plus précisément la modélisation, la principale remarque portait sur l'aspect « spéculatif » de la construction du modèle initial du dimère et le choix, *a priori*, d'un mode de dimérisation parmi d'autres possibles. En réponse à cette question, nous avons pu démontrer que les autres modes de dimérisation connus des RCPGs ne permettaient pas de valider les distances mesurées expérimentalement.

Cet article permet de mieux comprendre les mécanismes de formation du dimère entre le récepteur de la Ghréline et le récepteur de la Dopamine, mais surtout apporte une explication structurale à la modulation par le récepteur GHSR-1a de la signalisation liée à la Dopamine. La modélisation ouvre ici des perspectives intéressantes car notre modèle nous permet de discuter des interactions aux interfaces protéine-G:protéine-G.

Une autre information importante de cet article est la présence non pas de deux mais des quatre récepteurs dans cet assemblage complexe. La présence de ces deux récepteurs supplémentaires a été déduite lors des mesures de FRET (Transfert d'Énergie de Fluorescence par Résonance) à trois couleurs montrant clairement l'implication d'au moins trois récepteurs. Des analyses supplémentaires ont permis d'identifier que les deux récepteurs GHSR-1a et DRD2 supplémentaires interagissaient avec leurs homologues impliqués dans l'hétéro-dimère central (hétéro-dimère d'homo-dimères). Cet assemblage a ensuite été confirmé lors des mesures des distances croisées en LRET. En effet, la mesure de la distance entre  $G_{oi}$  et GHSR-1a a montré un second signal à une distance minimum de  $60\text{\AA}$ , permettant de délimiter une zone possible de présence du second récepteur GHSR-1a (**Figure S16 de l'article**).

Ces observations ouvrent un certain nombre de perspectives pour l'étude de la dynamique des RCPGs par modélisation. En effet, en se basant sur les interfaces protéine-G:protéine-G décrites par notre modèle, nous avons synthétisé des peptides, mimant ces interfaces, et potentiellement capables d'inhiber la formation du dimère de protéines-G. Dans le même temps, nous avons utilisé notre protocole CG-REMD pour prédire les interactions protéine:protéine au sein de ce gros complexe, notamment la position des deux récepteurs supplémentaires pour proposer un modèle complet DRD2/DRD2/ $G_i$ :GHSR-1a/GHSR-1a/ $G_q$ .

### 3.3.2 Synthèse et test de peptides inhibiteurs des interactions protéine-G:protéine-G

#### 3.3.2.a Choix des peptides

Afin d'apporter une validation supplémentaire de notre modèle DRD2/G<sub>i</sub>:GHSR-1a/G<sub>q</sub>, j'ai réalisé la synthèse de peptides ayant pour objectif d'inhiber la formation de ce complexe. Nous avons ciblé pour cela l'interface entre les deux protéines-G, montrant une forte complémentarité et une large surface d'interaction. L'hypothèse de départ réside dans la capacité de ces peptides à se placer à la surface des deux protéines G<sub>i</sub> et G<sub>q</sub>, dans les régions impliquées dans le dimère, afin d'empêcher leur interaction par compétition.

Notre choix s'est porté sur la synthèse de fragments de G<sub>i</sub>, en espérant qu'ils se fixent à la surface de G<sub>q</sub>. Le choix de mimer G<sub>i</sub> plutôt que son homologue vient de l'observation expérimentale d'un pré-couplage de G<sub>q</sub> avec son récepteur GHSR-1a. La visualisation de nos modèles a permis d'identifier quatre fragments continus de G<sub>i</sub>, présentant une structuration particulière et une bonne complémentarité de charge avec G<sub>q</sub>. Des analyses des trajectoires de dynamique moléculaire ont par ailleurs montré que ces segments sont en contact permanent avec G<sub>q</sub>. Les peptides que nous avons sélectionnés pour se fixer à la surface de G<sub>q</sub> sont représentés sur la **Figure 46**.

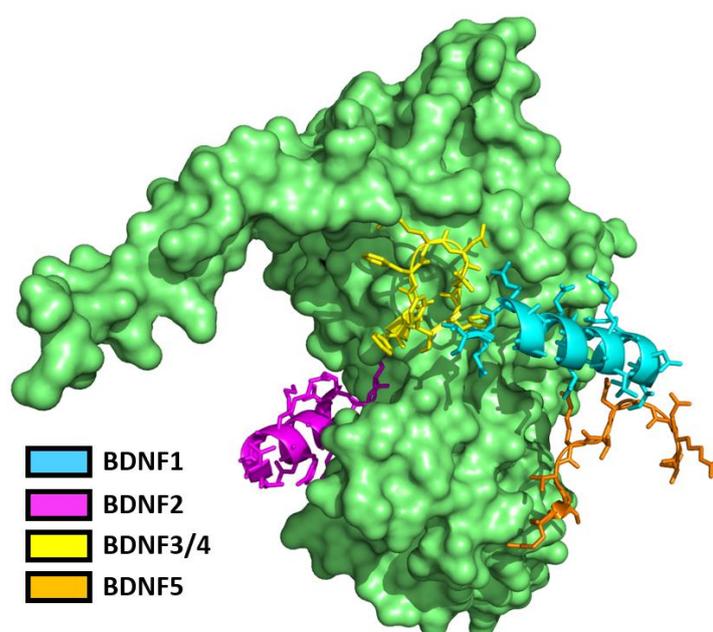


Figure 46 : Structures des segments extraits de G<sub>i</sub> à l'interface G<sub>i</sub>:G<sub>q</sub>, G<sub>αq</sub> est représenté en vert, les séquences des peptides sont données dans le **Tableau 9**.

### 3 Résultats

Deux de ces peptides sont structurés en hélice ; le troisième décrit un coude ; le quatrième présente une structure désordonnée. Pour celui présentant un coude, nous avons décidé de le synthétiser sous sa forme linéaire, mais également sous une forme cyclique pour tenter de conserver son repliement. Les séquences de ces peptides sont données dans le **Tableau 9**.

Peptide	Sous-unité d'origine	Numéros des résidus	Séquence	Masse moléculaire (g.mol <sup>-1</sup> )
BDNF1	G <sub>αi</sub>	17-33	H-MIDRNLRLEDGEKAAREV-OH	2002
BDNF2	G <sub>αi</sub>	60-72	H-YSEEECKQYKAVV-OH	1575
BDNF3	G <sub>αi</sub>	188-198	H-FTFKDLHF-OH	1053
BDNF4	G <sub>αi</sub>	188-198	<u>FTFKDLHF</u>	1035
BDNF5	G <sub>β</sub>	128-137	H-TREGNVRVSR-OH	1173

Tableau 9 : Séquences des peptides synthétisés et leur numérotation d'origine dans la protéine-Gi.

#### 3.3.2.b Synthèse des peptides

La synthèse des peptides a été réalisée avec une stratégie sur support solide. La première étape consistait en l'accrochage du premier acide aminé sur la résine (résidu C-terminal en synthèse peptidique). La suite de la séquence a été incorporée par un robot synthétiseur automatisé. Les peptides BDNF3 et BDNF4 sont issus de la même synthèse sur le robot, avant d'être séparés pour donner le peptide linéaire BDNF3 et le peptide cyclique BDNF4. Une fois récupérés, les peptides linéaires BDNF1, BDNF2, BDNF3, et BDNF5 ont été décrochés des résines dans des conditions permettant de supprimer les protections latérales. Pour le peptide cyclique BDNF5, il a été dans l'ordre décroché de la résine, cyclisé par une liaison amide entre ses extrémités et enfin déprotégé. Les peptides ont été ensuite purifiés par HPLC préparatoire. Les protocoles des réactions sont décrits dans la partie Matériels et Méthodes (voir partie 2.6.1).



Figure 47 : Schéma réactionnel de l'accrochage du premier acide aminé à la résine.

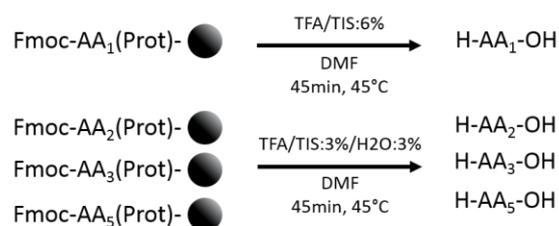
Pour le peptide BDNF5 nous disposons au laboratoire de résine de Wang avec une arginine greffée Fmoc-R(Pbf)-Wang avec une charge de 0,64mmol.g<sup>-1</sup>. Il n'a donc pas été nécessaire de réaliser cette première étape. Pour les autres peptides, le premier acide aminé, Valine pour BDNF1 et BDNF2 et Phénylalanine pour BDNF3/4, a été accroché sur une résine 2 chloro-Chlorotrytyl (Clt). Pour la synthèse de Fmoc-V-Clt, 1g de résine avec une charge à

1,6mmol.g<sup>-1</sup> a réagi avec 651,6mg de Fmoc-V et 816mg de DIEA. 1,370g de résine a été obtenue avec une charge mesurée à 0,91mmol.g<sup>-1</sup> pour un rendement de 84% (**Tableau 10**). Pour la synthèse de Fmoc-F-Clt, 1g de résine avec une charge à 1,6mmol.g<sup>-1</sup> a réagi avec 743,8mg de Fmoc-F et 816mg de DIEA. 1,357g de résine a été obtenue avec une charge mesurée à 0,94mmol.g<sup>-1</sup> pour un rendement de 92% (**Tableau 10**).

Résine	m (mg)	Absorbance	Charge (mmol.g <sup>-1</sup> )	Charge moyenne (mmol.g <sup>-1</sup> )	Charge attendue (mmol.g <sup>-1</sup> )
Val-Ctl	13,3	0,663	0,81	0,91	1,08
	11,5	0,742	1,05		
	10,0	0,532	0,87		
Phe-Ctl	10,4	0,714	1,12	0,94	1,02
	12,6	0,625	0,81		
	10,1	0,598	0,90		

Tableau 10 : Mesure de la charge des résines avec le premier acide aminé greffé, selon : charge = (A/m)\*16423 (voir partie 2.6.1.a).

Pour la complétion des séquences, une quantité de 0,25mmol des résines préparées a été utilisée pour tous les peptides, ce qui correspond à 275mg de Fmoc-V-Ctl pour les peptides BDNF1 et BDNF2, 266mg de Fmoc-F-Ctl pour la synthèse de BDNF3/4 et 393mg de Fmoc-R(Pbf)-Wang pour la synthèse de BDNF5.



Peptide	AA	AA(prot)	R après purification
BDNF1	MIDRNLRDGEKAAREV	MID(OtBu)R(Pbf)N(Trt)LR(Pbf)E(OtBu)D(OtBu)GE(OtBu)K(Boc)AAR(Pbf)E(OtBu)V	22%
BDNF2	YSEEECKQYKAVV	Y(tBu)S(tBu)E(OtBu)E(OtBu)E(OtBu)C(Trt)K(Boc)Q(Trt)Y(tBu)K(Boc)AVV	27%
BDNF3	FTFKDLHF	FT(tBu)FK(Boc)D(OtBu)LH(Trt)F	74%
BDNF5	TREGNVRVSR	T(tBu)R(Pbf)E(OtBu)GN(Trt)VR(Pbf)VS(tBu)R(Pbf)	50%

Figure 48 : Schéma réactionnel de la synthèse des peptides linéaires BDNF1, BDNF2, BDNF3 et BDNF5.

### 3 Résultats

A la sortie du robot synthétiseur les résines ont été directement traitées pour décrocher les peptides et retirer les protections des chaînes latérales. Pour les peptides BDNF3/4, la résine a été séparée en deux fractions égales : une pour le peptide linéaire BDNF3 et l'autre pour le peptide cyclique BDNF4. Pour le peptide BDNF1, 440mg de produit a été récupéré dont 167mg a été ensuite purifié ; 37mg de produit pur a été ainsi obtenu pour un rendement final de 22%. Pour le peptide BDNF2, 396,9mg de produit a été récupéré dont 156mg a été ensuite purifié ; 39mg de produit pur a été ainsi obtenu pour un rendement final de 27%. Pour le peptide BDNF3, 123,5mg de produit a été récupéré dont l'intégralité a été ensuite purifiée ; 98mg de produit pur a été ainsi obtenu pour un rendement final de 74%. Pour le peptide BDNF5, 250mg de produit a été récupéré dont 150mg a été ensuite purifié ; 88mg de produit pur a été ainsi obtenu pour un rendement final de 50%.

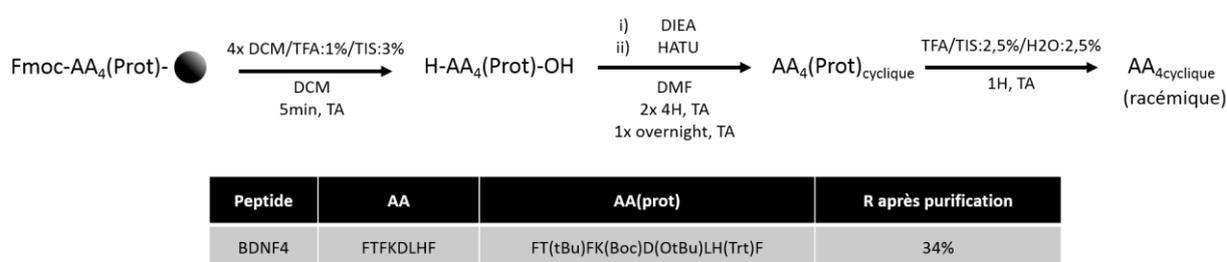


Figure 49 : Schéma réactionnel de la synthèse du peptide cyclique BDNF4.

Pour le peptide BDNF4, après le décrochage de la résine, 1585,6mg de produit a été récupéré, cette masse élevée s'expliquant par la présence d'une impureté liquide. 755mg de produit a été alors utilisé pour la cyclisation avec des quantités de réactifs calculées en supposant une quantité de 0,0625mmol de peptide (rendement maximum) soit 25mg de DIEA et 24mg de HATU. Au terme des 4H de réaction, le produit de départ est toujours présent. Le même nombre d'équivalents de réactifs a été alors ajouté, et la réaction a été laissée pour 4H supplémentaires. La réaction n'étant toujours pas terminée, le même nombre d'équivalents de réactifs a été une dernière fois ajouté, et la réaction a été continuée toute la nuit. La conversion du peptide linéaire a été ainsi complète et 908mg de produit a été obtenu sous la forme d'une huile. Toutefois, la racémisation d'une phénylalanine a été observée en HPLC avec un ratio 1:1 entre les stéréo-isomères, ainsi que des traces du dimère cyclique, dû aux conditions brutes de la réaction. Après la déprotection des chaînes latérales, la disparition de l'impureté liquide a été observée, et 89mg de produit a été obtenu et intégralement purifié. Puisque les phénylalanines ne sont pas impliquées à l'interface  $G_i:G_q$  dans nos simulations, la racémisation n'est pas un problème en soi. Les stéréo-isomères ont été toutefois séparés pour donner 12mg du peptide

BDNF4a et 10mg du peptide BDNF4b pour un rendement final de 34%. Il faudra par la suite déterminer lequel est le peptide isomérisé.

### 3.3.2.c Test biologiques d'inhibition

Pour tester le pouvoir inhibiteur des peptides sur la formation du dimère, j'ai participé aux tests biologiques qui consistaient à mesurer l'activité de  $G_i$ . En effet, lorsque  $G_i$  est dans le complexe DRD2/ $G_i$ :GHSR-1a/ $G_q$ , son activité, c'est-à-dire sa capacité à lier le GTP, augmente en présence de  $G_q$ . Cette modification de son activité peut être mesurée directement par un changement de sa fluorescence intrinsèque. Nos peptides ont donc été testés sur des nanodisques, en présence des récepteurs DRD2 et GHSR-1a dimérisés, de la Dopamine (Dopa), et de la protéine  $G_i$ . Une première moitié des échantillons a été testée sans  $G_q$ . Sans  $G_q$ , on s'attend à ne voir aucun effet des peptides. La seconde moitié a été testée en présence de  $G_q$ . Si les peptides ont bien un effet inhibiteur des interactions entre  $G_q$  et  $G_i$ , alors on s'attend ici à voir une diminution de l'augmentation de fluorescence induite par  $G_q$ . Chacun des peptides a été testé seul. Un mélange des cinq peptides synthétisés a également été testé (BDNF1, BDNF2, BDNF3, BDNF4a et BDNF5). En effet, les interactions protéine:protéine montrent souvent une grande surface d'interaction, ce qui rend difficile leur inhibition par un peptide de petite taille. Aussi, en testant un mélange, nous espérions voir un effet coopératif entre les différents peptides. La liste et la composition des échantillons testés sont données dans le **Tableau 11**. Le GTP a été rajouté en dernier puisqu'il déclenche les changements conformationnels de  $G_i$ . Au bout de 10min la réaction a été stoppée dans la glace et la fluorescence mesurée.

### 3 Résultats

Echantillon	Peptide	R 100nM (μL)	Dopa 0.5nM (μL)	G <sub>i</sub> 1μM (μL)	G <sub>q</sub> 1μM (μL)	Peptide 1mM (μL)	Tampon (μL)	GTPγS 5μM (μL)
1	/	10	1	5	0	0	33	1
2	BDNF1	10	1	5	0	5	28	1
3	BDNF2	10	1	5	0	5	28	1
4	BDNF3	10	1	5	0	5	28	1
5	BDNF4a	10	1	5	0	5	28	1
6	BDNF4b	10	1	5	0	5	28	1
7	BDNF5	10	1	5	0	5	28	1
8	Mélange	10	1	5	0	5	28	1
9	/	10	1	5	5	0	28	1
10	BDNF1	10	1	5	5	5	23	1
11	BDNF2	10	1	5	5	5	23	1
12	BDNF3	10	1	5	5	5	23	1
13	BDNF4a	10	1	5	5	5	23	1
14	BDNF4b	10	1	5	5	5	23	1
15	BDNF5	10	1	5	5	5	23	1
16	Mélange	10	1	5	5	5	23	1

Tableau 11 : Composition des échantillons préparés pour les tests d'activité de G<sub>i</sub> en présence ou l'absence de G<sub>q</sub> et des différents peptides.

Tout d'abord, les contrôles effectués nous permettent de montrer que les peptides n'ont pas d'effet sur l'activité propre de G<sub>i</sub> (en l'absence de G<sub>q</sub>). Nos analyses de fluorescence (**Figure 50**) montrent que les peptides BDNF1, BDNF3 et les deux peptides BDNF4 baissent l'augmentation de l'activité de G<sub>i</sub> induite par G<sub>q</sub> de 15%, 37%, 26% et 11% respectivement alors que les peptides BDNF2 et BDNF5 n'ont aucun effet. Le résultat le plus encourageant vient des tests effectués avec le mélange des cinq peptides qui semble inhiber complètement la formation du dimère. En effet, en présence du mélange, nous n'observons aucune augmentation de la fluorescence de G<sub>i</sub> induite par G<sub>q</sub>.

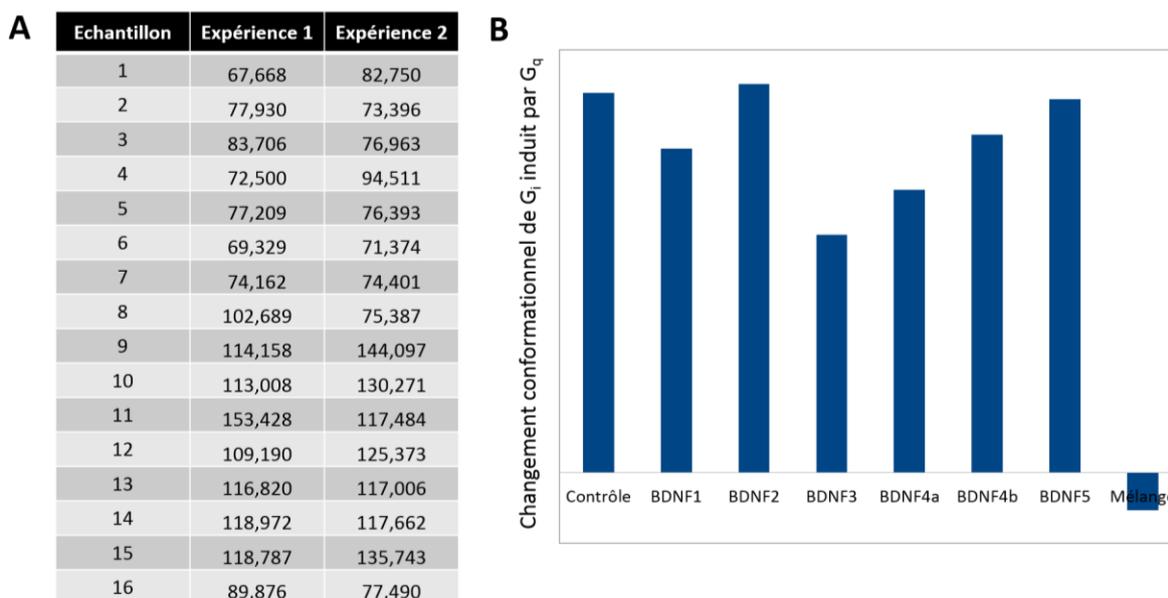


Figure 50 : A) Fluorescence mesurée pour tous les échantillons à  $t = 10\text{min}$  après l'ajout du  $\text{GTP}\gamma\text{S}$  ; B) mesure de l'activité de  $G_i$  induite par  $G_q$ , calculée par la différence de fluorescence entre les échantillons avec et sans  $G_q$  (moyenne avec  $n=2$ ).

Ces résultats très encourageants qu'il faut encore reproduire pour les confirmer, apportent d'ores et déjà une preuve expérimentale supplémentaire validant les interfaces protéine:protéine proposées dans notre modèle de dimère. Ce travail ouvre également d'intéressantes perspectives. Dans un premier temps, il nous faudra déterminer précisément quels peptides du mélange sont responsables de l'inhibition complète du dimère. Si l'on peut supposer l'implication des peptides BDNF1, BDNF3 et BDNF4, il n'est pas à exclure que les peptides BDNF2 et BDNF5, inefficaces seuls, participent malgré tout par un effet synergique. Tester les différentes combinaisons possibles de peptides nous donnera rapidement la réponse à cette question. Dans un second temps, nous envisageons d'optimiser les séquences de ces peptides afin d'obtenir de meilleurs résultats d'inhibition. Des travaux commencent tout juste dans l'équipe visant à caractériser toutes les interfaces possibles protéine-G:protéine-G *via* la construction exhaustive de tous les modèles de dimères possibles impliquant  $G_i$ ,  $G_q$ , mais aussi  $G_s$ ,  $G_o$ , etc... Une autre piste intéressante consiste à cribler des bibliothèques de petites molécules (dans un premier temps commerciales), afin d'identifier de possibles molécules candidates présentant les prérequis pharmacophoriques pour inhiber nos dimères. L'inhibition de ce (ces) dimère(s) de protéines-G pourrait à long terme être utilisée pour le développement de tests biologiques. La méthode pourrait également être appliquée à d'autres dimères de protéines-G liés à différents dimères de récepteurs afin de comprendre plus précisément le rôle joué par ces dimères dans leur fonction (régulation).

### 3.3.3 Prédiction du mode de dimérisation / oligomérisation des RCPGs

#### 3.3.3.a Contexte

Notre étude de la dimérisation de DRD2:GHSR-1a a montré la présence de deux récepteurs supplémentaires formant des homo-dimères avec leur homologue du dimère central (**Figure 51**). Pour essayer de placer ces récepteurs dans le complexe, nous disposons d'une seule contrainte supplémentaire, mesurée en LRET et indiquant une distance de 60Å entre un résidu de  $G_{\alpha i}$  et un résidu de l'ICL2 du deuxième protomère de GHSR-1a. Nous avons essayé différents types de simulation sous contrainte pour essayer de résoudre la structure de ce macro-complexe. Cette information de distance ne nous a malheureusement pas permis de réduire suffisamment les degrés de liberté pour proposer un modèle final (**Figure S16** de l'article en 3.3.1). Aussi, nous avons utilisé notre approche CG-REMD, sans appliquer de contrainte, pour essayer de valider, *a posteriori*, la position des 3<sup>ème</sup> et 4<sup>ème</sup> récepteurs au sein de ce complexe. Pour cela nous avons construit de nombreux systèmes différents, et réalisé des simulations CG-REMD utilisant les mêmes paramètres que ceux décrits pour la validation de la méthode (voir partie 3.1.1), avec des temps de simulation d'environ 5 $\mu$ s pour chacune des 50 répliques échelonnées entre 300 et 447K.

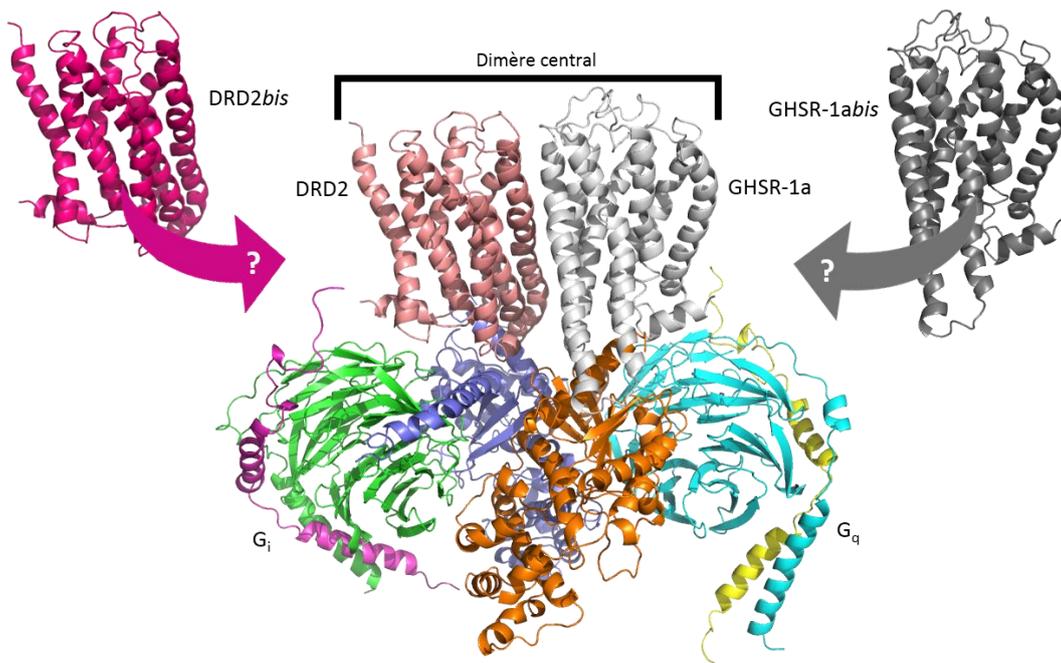


Figure 51 : Nomenclature du tétramère.

La première simulation a été réalisée en partant de notre modèle de dimère complet, avec les récepteurs « *bis* » placés dans la membrane à une distance supérieure au cut-off. Lors de ces simulations, nous avons observé un phénomène très surprenant de courbure anormale de la membrane permettant aux récepteurs « *bis* » de venir faire de nombreux contacts avec les deux protéines-G. Ce phénomène peut s'expliquer peut-être par la taille réduite de la boîte de simulation utilisée et des effets de « bords » très marqués. Quoi qu'il en soit, les résultats expérimentaux obtenus au sein de notre équipe montrent que le tétramère de RCPGs peut se former en l'absence des protéines-G. Aussi, nous avons réalisé d'autres simulations, en l'absence de protéines-G visant à déterminer s'il est possible de prédire par nos calculs le mode de dimérisation des deux récepteurs centraux (hétéro-dimère), et le positionnement des deux autres récepteurs dans le complexe, avec l'hypothèse que celui-ci doit être symétrique.

#### **3.3.3.b Résultats**

### 3 Résultats

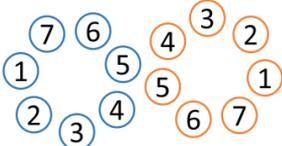
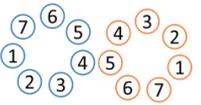
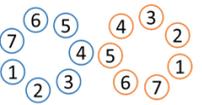
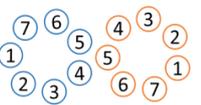
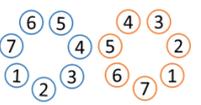
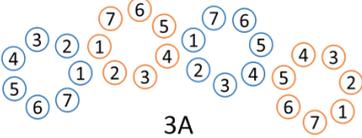
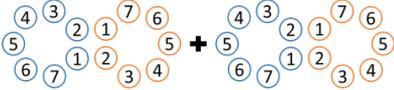
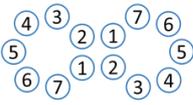
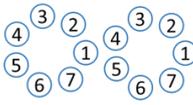
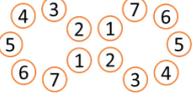
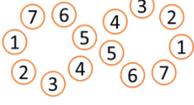
Système étudié	Conformations obtenues après clustering de la trajectoire à 300K			
 <p>Modèle de référence</p>				
 <p>1</p>	 <p>1A</p>	 <p>1B</p>	 <p>1C</p>	 <p>1D</p>
 <p>2</p>	 <p>2A                      2B</p>			
 <p>3</p>	 <p>3A</p>			
 <p>4</p>	 <p>4A</p>			
 <p>5</p>	 <p>5A</p>	 <p>5B</p>		
 <p>6</p>	 <p>6A</p>	 <p>6B</p>		

Tableau 12 : Représentation simplifiée des principales interfaces de dimérisation (par ordre décroissant de population) obtenues pour chaque simulation CG-REMD réalisée; le récepteur DRD2 est représenté en bleu, le récepteur GHSR-1a en orange.

Les simulations 1 et 2 comportent un exemplaire de chaque récepteur déjà dimérisés au début de la trajectoire, ou libres respectivement. Dans la simulation 1, le dimère n'est pas stable et très vite nous avons observé une rotation autour de l'interface TM4 de DRD2 et TM5 de GHSR-1a (**Figure 52**). Les récepteurs ne se séparent pas pour autant, malgré le protocole REMD. Les résultats très différents obtenus dans la simulation 2 qui contient pourtant les mêmes protomères, suggèrent un problème de convergence des données, et la nécessité

d'effectuer des simulations bien plus longues pour ce type d'application. Dans la simulation 2 en effet, l'interface notée « 2A » la plus retrouvée implique TM1/TM2/H8-TM1/TM2/H8. Il s'agit ici d'une des deux principales interfaces de RCPGs identifiées dans la littérature [Meng, Mezei, et Cui 2014; Marsango et al. 2015; Johnston et al. 2012; Ferré et al. 2014]. L'interface TM4/TM5-TM4/TM5, qui correspond pourtant à l'interface validée expérimentalement pour le dimère central n'est pas retrouvée. Un second dimère noté « 2B » est également observé en bien plus faible proportion avec une interface TM1-TM4/TM5, peu probable d'après la littérature.

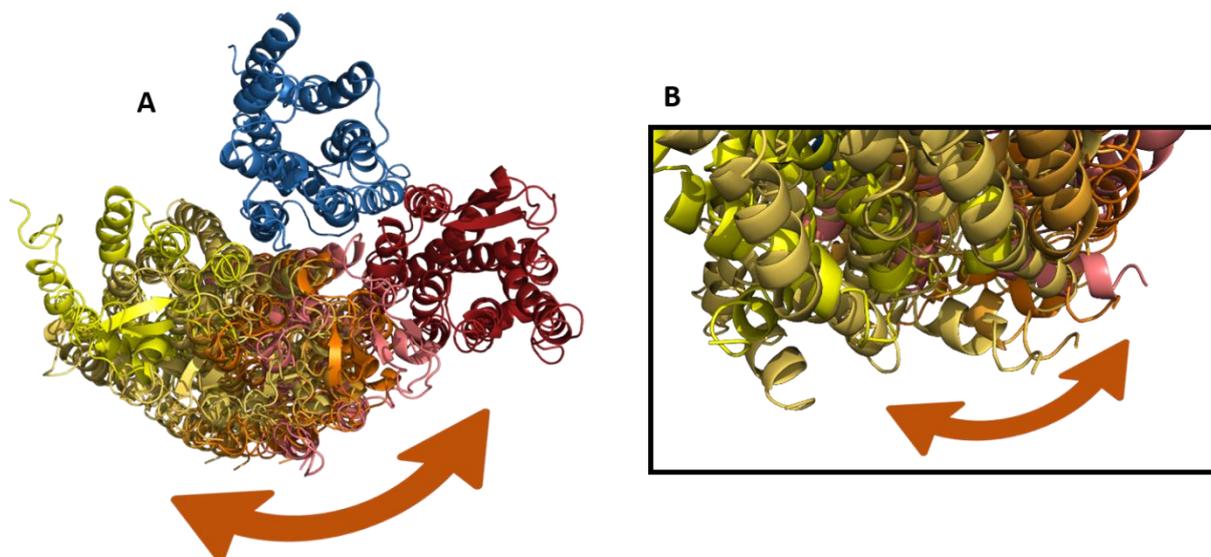


Figure 52 : Dimères obtenus lors de la simulation utilisant le dimère DRD2:GHSR-1a central comme point de départ, mettant en évidence la rotation des récepteurs autour de l'interface TM4-TM5, les conformations sont alignées sur le récepteur DRD2 représenté en bleu, les récepteurs GHSR-1a sont représenté dans un gradient de rouge à jaune ; A) Vue de la face extra-cellulaire des dimères ; B) Gros plan sur les hélices 8.

Dans les simulations 3 et 4 nous sommes partis de deux exemplaires de chacun des récepteurs : deux déjà dimérisés et deux libres (simulation 3), ou bien tous libres (simulation 4). Dans la simulation 3, le même phénomène d'instabilité du dimère de départ est observé, sans séparation complète, comme déjà observé dans la simulation 1. On observe alors un tétramère alternant DRD2 et GHSR-1a, en désaccord avec les résultats expérimentaux. Dans cette simulation, les hétéromères impliquent les interfaces TM4-TM5 et TM1/TM2/H8-TM1/TM2/H8 en accord avec la simulation 2. L'hétéromère central montre une interface TM1-TM4/TM5, là aussi déjà observée dans les autres simulations. En partant de récepteurs tous libres dans la simulation 4, on observe cette fois la formation de deux hétéro-dimères avec l'interface TM1/TM2/H8-TM1/TM2/H8. Les simulations semblent donc converger vers deux

éléments : l'hétéro-dimérisation et l'interface TM1/TM2/H8-TM1/TM2/H8 qui sont largement favorisées.

Les simulations 5 et 6 visaient à prédire quant à elles le mode de dimérisation des homodimères en partant de deux récepteurs libres de DRD2 et de GHSR-1a respectivement. Dans les deux simulations, l'interface la plus observée est à nouveau l'interface TM1/TM2/H8-TM1/TM2/H8. Dans la simulation 5, DRD2 se dimérise également en utilisant une interface TM1-TM4/TM5 déjà observée plus tôt. Dans la simulation 6, GHSR-1a montre une dimérisation possible impliquant l'interface TM4/TM5-TM4/TM5 presque identique au mode de dimérisation du dimère central. Puisque cette face de GHSR-1a est déjà utilisée dans l'hétéromère central, ce mode d'homodimérisation dans le tétramère peut être directement exclu. Ces résultats sont assez intéressants. En effet, le modèle de dimérisation noté « 5B » de DRD2, une fois aligné sur le dimère complet, présente une distance  $G_{\alpha q}$ -DRD2 (entre les résidus équivalents de la distance  $G_{\alpha i}$ -GHSR-1a) de 60Å exactement (**Figure 53B**). Toutefois, pour obtenir cette distance, il est nécessaire d'avoir une interface non décrite dans la littérature. Ceci n'est pas à exclure car des contraintes supplémentaires doivent exister dans le tétramère, favorisant possiblement la création d'autres interfaces, moins favorable dans un dimère.

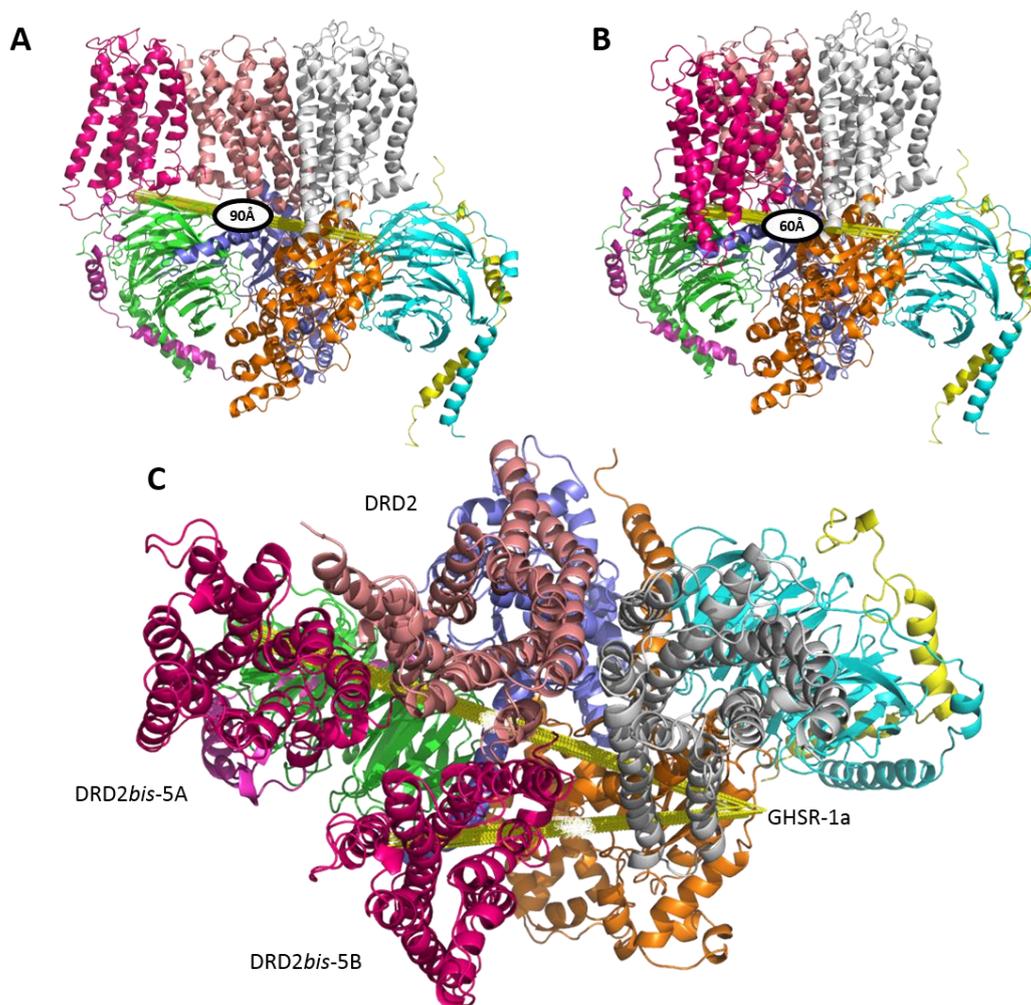


Figure 53 : Alignement des homo-dimères DRD2 obtenus sur le dimère DRD2:GHSR-1a complet et mesure de la distance DRD2bis:G $\alpha_q$ , les récepteurs DRD2 sont représentés en rose, les récepteurs GHSR-1a sont représentés en gris, la protéine G $i$   $\alpha$ : $\beta$ : $\gamma$  est représentée en mauve:vert:magenta, la protéine G $q$   $\alpha$ : $\beta$ : $\gamma$  est représentée en orange:cyan:jaune ; A) Modèle « 5A » aligné sur le dimère DRD2:GHSR-1a ; B) Modèle « 5B » aligné sur le dimère DRD2:GHSR-1a ; C) Modèles « 5A » et « 5B » alignés sur le dimère DRD2:GHSR-1a, vue de la face extracellulaire

Si on se contente de regarder les interfaces connues en cristallographie, alors la seule compatible avec la position de nos deux récepteurs « bis » dans le tétramère, serait une interface TM1/TM2/H8-TM1/TM2/H8, telle que montrée par le modèle noté « 5A » (**Figure 53A**). La mesure de la distance G $\alpha_i$ -GHSR-1a serait alors de 90Å, ce qui n'est toutefois pas complètement en désaccord avec la valeur mesurée par LRET. En effet, 60Å est à la limite de la précision de la méthode utilisée, et une distance de 90Å pourrait donner le même signal. Les récepteurs « bis » se placeraient donc comme décrit dans le modèle construit grâce aux modèles notés « 5A » et « 6A » (**Figure 54**).

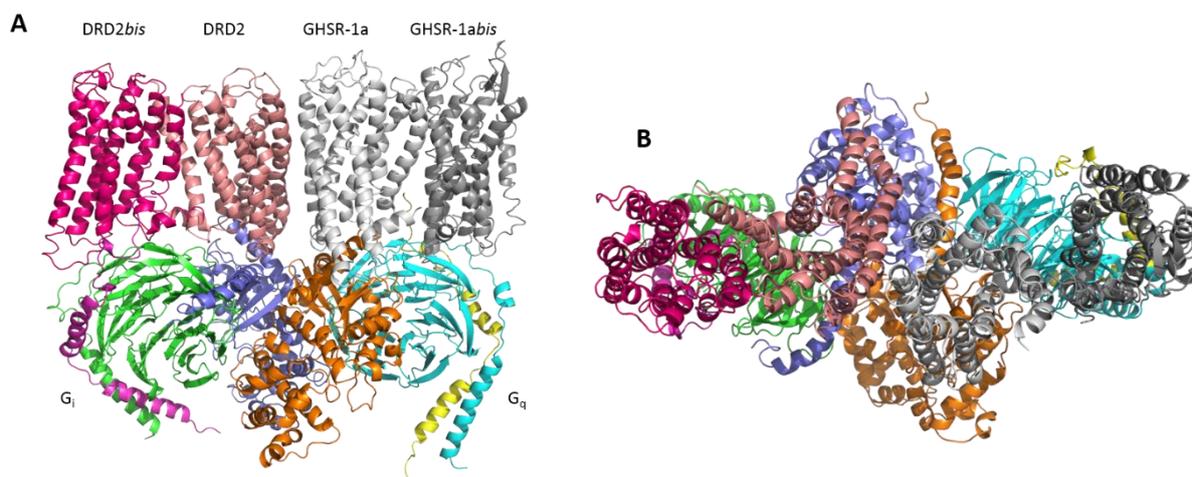


Figure 54 : Modèle du tétramère proposé par le résultats des simulations CG-REMD, les récepteurs DRD2 sont représentés en rose, les récepteurs GHSR-1a sont représentés en gris, la protéine  $G_i \alpha:\beta:\gamma$  est représentée en mauve:vert:magenta, la protéine  $G_q \alpha:\beta:\gamma$  est représentée en orange:cyan:jaune ; A) Vue du tétramère dans l'axe de la membrane ; B) Vue de la face extra-cellulaire du tétramère.

Il est à noter tout de même une étude sur la dimérisation de DRD2 où, par docking et dynamique moléculaire, une interface TM1-TM4/TM5 a été proposée [Agnieszka A. Kaczor, Jörg, et Capuano 2016]. Toutefois, dans cette étude basée sur du docking protéine-protéine utilisant un grand nombre de critère de score et des modèles construits par homologie, les auteurs obtiennent un certain nombre d'interfaces possibles DRD2:DRD2 dont principalement TM2/TM3-TM2/TM3, exclue car non retrouvée dans la littérature et TM1/TM7-TM4/TM5, justifiée par des études expérimentales sur le récepteur DRD3 [Marsango et al. 2015; Guitart et al. 2014] et qui montrent que ce récepteur peut se dimériser *via* une des interfaces TM7/TM1/TM2-TM7/TM1/TM2 et TM4/TM5-TM4/TM5, ce qui impliquait que DRD3, et par analogie DRD2, pouvait réaliser une interface hybride. Cette extrapolation est très discutable, et ne valide en aucun point l'interface TM1/TM7-TM4/TM5.

### 3.3.3.c Perspectives

Nos calculs montrent certaines limites ici ; il n'a en effet pas été possible de retrouver le mode de dimérisation des récepteurs centraux du complexe en partant des deux récepteurs isolés, probablement à cause de problèmes de convergence des données. Néanmoins, un modèle intéressant du placement des récepteurs « *bis* » a pu être proposé.

Nos simulations permettent ainsi de retrouver une interface décrite dans la littérature et finalement de proposer un modèle compatible avec les données expérimentales. Les calculs

réalisés en présence des protéines-G, présentent vraisemblablement aussi des problèmes de convergence, ajoutés à celui de la distorsion de la membrane. D'autres calculs seront nécessaires pour préciser le mode et le mécanisme d'assemblage de ces gros complexes à l'échelle moléculaire. Il nous faudra notamment lancer des simulations bien plus longues. Nous réfléchissons également à la meilleure manière de modifier le champ de forces, en recalculant notamment un réseau élastique « maison » pour ces objets, et basés sur des simulations tout-atomes des différents partenaires isolés, comme montré dans des travaux de 2013 [Globisch et al. 2013]. L'objectif à moyen terme sera d'utiliser notre protocole CG-REMD pour prédire les interactions des RCPGs avec leurs différents partenaires, sans aucun *a priori* sur leurs dynamiques intrinsèques. L'utilisation de notre protocole à des fins purement prédictives est montrée, à travers quelques exemples dans le prochain et dernier chapitre (voir partie 3.4).

### 3.4 Autres applications

Après l'étape de validation du protocole CG-REMD, nous l'avons appliqué à un certain nombre d'autres systèmes. Parmi ces autres projets, réalisés en grande partie en collaboration avec différents laboratoires, quatre sont suffisamment aboutis pour être discutés dans ce manuscrit. Le premier porte sur la prédiction de la liaison d'une toxine de venin de serpent au récepteur de la Vasopressine de type-2, étude dont j'ai réalisé la majorité des analyses. Les second et troisième projets portent sur la liaison des peptides Ghréline et LEAP2 au récepteur de la Ghréline GHSR-1a. Ici la très grande majorité des analyses a été réalisée par Maxime LOUET, ATER au sein de notre équipe. Enfin, un dernier projet portant sur la prédiction de la liaison et la sélectivité d'interaction de différents RCPGs avec des peptides issus de l'hélice C-terminale des protéines-G a récemment été démarré au sein de l'équipe et réalisé en grande partie par Nicolas FLOQUET, mon directeur de thèse.

#### 3.4.1 Prédiction de la liaison d'une toxine aux récepteurs de la Vasopressine

##### 3.4.1.a Contexte

Ce travail a été réalisé en collaboration avec le groupe de Nicolas GILLES du Laboratoire de Toxinologie Moléculaire et Biotechnologies (LTMB, CEA Saclay) et Bernard MOUILLAC de l'Institut de Génomique Fonctionnelle (IGF, Montpellier). L'objectif de ce projet a été de définir les interactions faites par une toxine possédant un repliement de type « Kunitz » avec son récepteur, le récepteur de la Vasopressine de type-2 (V2R). Cette toxine, la Mambaquarétine-1 (MQ-1), est l'une des neuf toxines isolées du venin du Mamba Vert, et agit comme un antagoniste très sélectif du V2R, présentant une affinité de l'ordre du nanomolaire pour son récepteur [Droctové 2018]. Ce travail réalisé à l'interface modélisation-biochimie-biologie moléculaire est la première application de notre protocole à l'étude d'interactions entre une protéine extra-cellulaire et son récepteur.

##### 3.4.1.b Résultats

De nombreuses simulations ont été réalisées sur ce système, et d'autres sont encore en cours. Lors de nos premières simulations CG-REMD de ce système V2R:MQ-1, nous avons observé un problème d'agrégation des deux toxines simulées, dû à la taille de la toxine d'une

part, et à ses nombreux résidus chargés à sa surface. Ainsi, par la suite, nous avons décidé de ne garder qu'un exemplaire de la toxine, mais toujours en présence de deux récepteurs. Les temps de simulation obtenus varient de  $2\mu\text{s}$  à  $10\mu\text{s}$  pour chacune des 50 températures utilisées, pour un temps de simulation total de l'ordre de  $600\mu\text{s}$ . Après le clustering des conformations issues de toutes les trajectoires, nous avons obtenu deux orientations majoritaires de la MQ-1 dans la poche de son récepteur (**Figure 55**). Ces deux conformations, même si symétriquement opposées, présentent un certain nombre de points communs. Tout d'abord, la liaison de la toxine nécessite une ouverture importante du récepteur entraînant d'importantes surfaces de contact ( $4600\text{\AA}^2$  pour le modèle 1 et  $4700\text{\AA}^2$  pour le modèle 2). Dans ces deux orientations, la même boucle de MQ-1 [15-18] (Asn-Gly-Phe-Phe) plonge au fond de la poche orthostérique du V2R, effectuant un contact avec le résidu Q291 ; du fait de cette interaction, les extrémités N et C-terminales de la toxine se retrouvent totalement accessibles au solvant, sans aucun contact avec le récepteur (attention toutefois à ne pas oublier que dans notre modèle le récepteur est tronqué de 35 résidus à son extrémité N-terminale). Enfin, deux résidus du récepteur établissent des contacts ioniques forts avec deux résidus de la MQ1 : E184 (V2R) avec K10 (MQ-1) et E299 (V2R) avec R44 (MQ-1) pour le modèle 1 et, à l'inverse E184 (V2R) avec R44 (MQ-1) et E299 (V2R) avec K10 (MQ-1) pour le modèle 2. Des expériences de double-mutants récepteur:toxine sont actuellement en cours pour tenter de valider l'un de ces contacts.

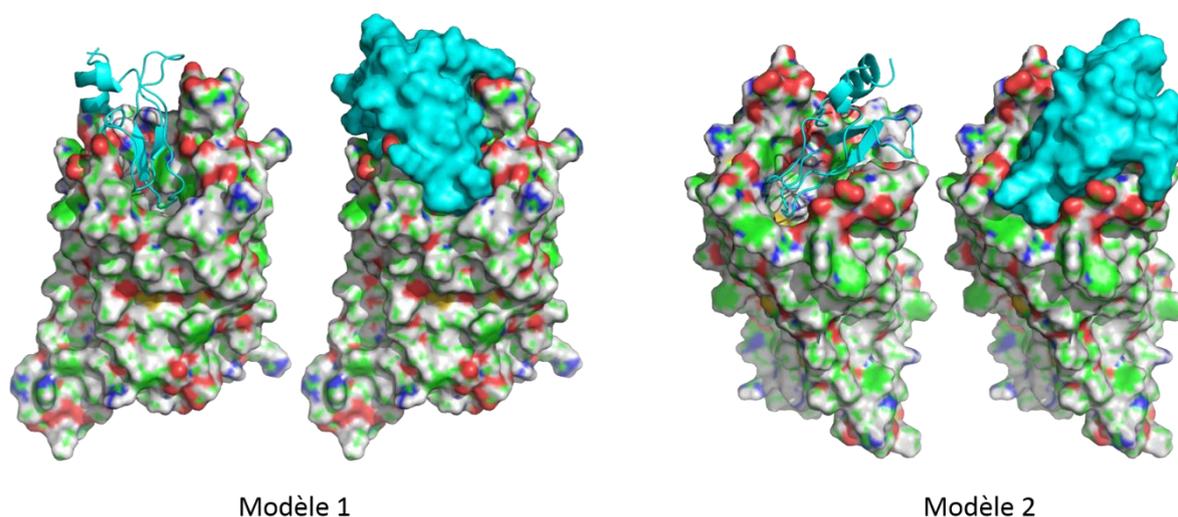


Figure 55 : Modèles de la MQ-1 liée au récepteur V2R obtenus après nos simulations CG-REMD, le récepteur est représenté en surface (couleurs selon le type d'atomes) et MQ-1 est représentée en « cartoon » ou en surface (en bleu).

La conservation de ces interactions entre les modèles 1 et 2 semblent montrer que nos calculs ont convergés vers cette solution. En revanche, la symétrie des deux conformations obtenues a rendu particulièrement difficile la validation par l'expérience de nos résultats et le choix de l'un ou l'autre des modèles. En effet, nos modèles ont permis de proposer des mutants (**Tableau 13**) validant l'orientation générale de la toxine dans son récepteur mais ne permettant pas de discriminer les deux orientations. Il est intéressant de noter ici que la plupart des mutations en alanine n'entraînent pas d'effet sur l'affinité de MQ-1, et que seules des inversions de charge provoquent un effet significatif, soulignant un possible effet compensatoire entre les différents contacts. Ces mutations ont permis toutefois de valider expérimentalement l'orientation générale de la toxine dans son récepteur et l'importance des résidus K10 et R44 de MQ-1 ainsi que E184 et E299 de V2R. Là aussi en accord avec nos modèles, une délétion des extrémités N ou C-terminales de la toxine n'ont aucun effet sur son affinité. Au fond de la poche du récepteur, la différence principale entre nos deux modèles réside dans l'interaction avec le résidu Q291. Dans le modèle 1, ce résidu fait un contact avec les phénylalanines F17 et F18 de la toxine dont la mutation fait perdre énormément en affinité ; dans le modèle 2, ces deux phénylalanines sont plutôt orientées vers une poche hydrophobe, laissant le résidu N15 faire le contact avec le résidu Q291 du récepteur. La mutation de N15 entraîne également une perte d'affinité importante de la toxine. Dans le modèle 1, N15 établit aussi des contacts polaires avec le récepteur, impliquant les résidus K100 et E40 du récepteur, expliquant là aussi possiblement son importance. En résumé, tout le travail réalisé autour des mutants impliquant la boucle de la toxine n'a pas permis de valider l'un ou l'autre de nos modèles.

	Mutation	Effet	Contact modèle 1	Contact modèle 2
MQ-1	Delta N-ter	Aucun	Aucun	Aucun
	K10A	Aucun	E184	E299
	K10E	Perd 10 fois en affinité		
	N15K-G16A	Perd 10 fois en affinité	E40 K100	Q291
	F17A	Perd beaucoup en affinité	Q291	L111
	F18A	Perd beaucoup en affinité		
	R44A	Aucun	E299	E184
	R44E	Perd 10 fois en affinité		
	Delta C-ter	Aucun	Aucun	Aucun
V2R	E184K	Perd 10 fois en affinité	K10	R44
	Q291	Perd 20 fois en affinité	F17 F18	N15
	E299K	Perd 10 fois en affinité	R44	K10

Tableau 13 : Principales mutations réalisées sur la toxine et le récepteur pour valider notre modèle et leurs effets sur l'affinité mesurée de la toxine.

Enfin, l'absence de l'extrémité N-terminale du récepteur est aussi importante à discuter. En effet, nos simulations permettent de décrire la dynamique de liaison de la toxine à son récepteur. L'analyse des trajectoires continues nous a permis de montrer que le positionnement final de la MQ-1 dépend fortement de son approche et de son positionnement à la surface du récepteur (**Figure 56A**). Pour obtenir l'orientation décrite par le modèle 2, la toxine doit nécessairement passer par-dessus le TM1. Aussi, on pourrait penser ici que l'extension du segment N-terminal du récepteur pourrait alors entraîner une gêne stérique et donc être en défaveur du modèle 2. Toutefois, on ne peut exclure que ces résidus manquants pourraient également avoir l'effet inverse, c'est-à-dire interagir directement avec la toxine et la diriger/la maintenir dans la poche de liaison du récepteur. En effet, l'extrémité N-terminale de V2R comporte un certain nombre de résidus chargés (MLMASTTS~~AVPGHPSLPSLPSNSQERPLDTRDPL~~) qui pourraient jouer ce rôle. De plus, la toxine dans le modèle 2 présente une cavité à sa surface qui est dans le prolongement direct du TM1, ce qui pourrait suggérer une complémentarité supplémentaire entre les deux objets (**Figure 56B**).

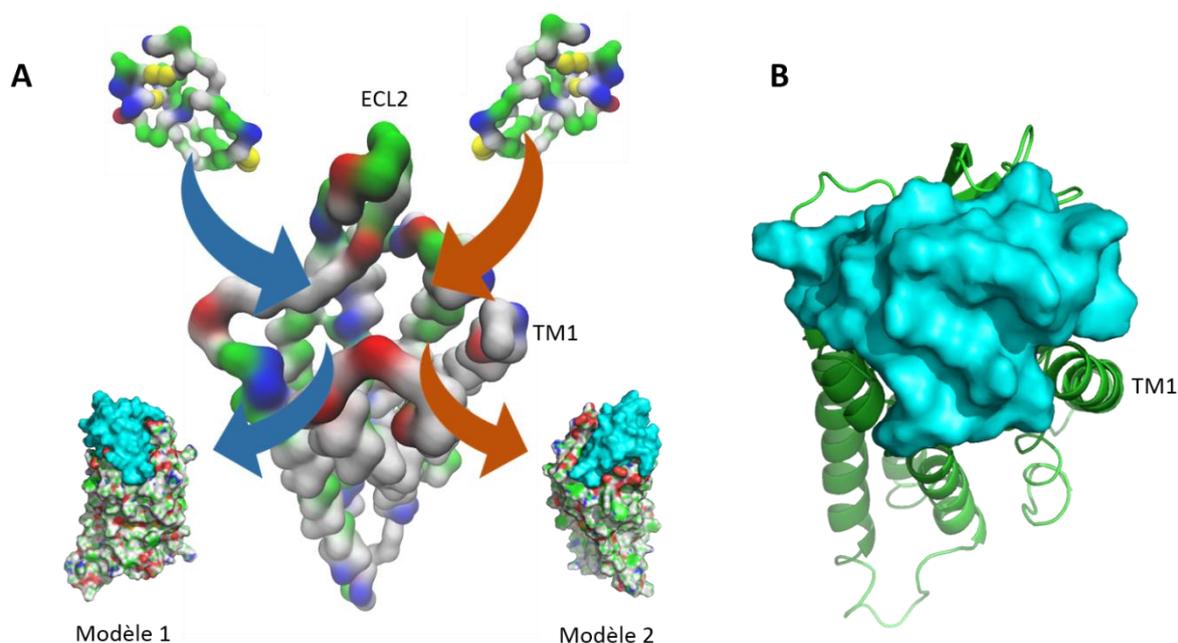


Figure 56 : A) Schématisation du phénomène observé lors de la liaison de la toxine à son récepteur: la face d'approche de MQ-1 à la boucle ECL2 détermine sa conformation liée finale ; B) Modèle 2 de MQ-1 liée au V2R qui présente une cavité à la surface de la toxine dans le prolongement du TM1 tronqué dans sa partie N-terminale.

Toujours, dans l'optique de valider l'un ou l'autre de nos modèles, nous sommes revenus vers l'alignement de séquences des nombreuses toxines présentant un repliement de type « kunitz » et testées par nos collaborateurs (**Tableau 14**). On remarque que le résidu K10 n'est pas essentiel pour l'affinité de la toxine et peut être remplacé par une alanine, sans perte apparente d'affinité, comme montré par mutagenèse dirigée. Le résidu R44 quant à lui est présent dans la plupart des toxines, qu'elles se lient ou non au V2R et n'est donc pas essentiel pour la sélectivité. Enfin, la boucle N15-G16-F17-F18 de la toxine qui plonge au fond de la poche dans nos modèles, semble très importante pour l'affinité des toxines. En effet, toutes les toxines non-affines au V2R présentent des résidus chargés dans cette zone. Il s'agit donc d'un élément supplémentaire en faveur de nos modèles et qui corrèle avec la mutation N15K-G16A qui diminue fortement l'affinité de MQ-1. Il est également intéressant de noter que le résidu N15 peut être remplacé par une serine ou une méthionine dans les différentes mambaquarétines sans affecter l'affinité. Cette information va donc dans le sens du modèle 1 puisque dans le modèle 2, N15 fait un contact avec Q291, pourtant essentiel pour la liaison.

Toxine	Séquence	K <sub>i</sub> (nM)
<b>MQ1</b>	RPSFCNLPVKPGPCNGFFSAFYYSQKTNKCHSFTYGGCKGNANRFSTIEKCRRTC VG	5,02
<b>MQ2</b>	RPSFCNLPVKPGPCNGFFSAFYYSQKTNKCHSFTYGGCRGNANRFSTIEE CRRTC VG	8,16
<b>MQ3</b>	RPSFCNLPVKPGPCNGFFSAFYYSQK <b>KN</b> KNKCHSFTYGGCKGNANRFSTIEE CRRTC VG	13,28
<b>MQ4</b>	RPSFCNLPVKPGPCSGFFSAFYYSQKTNKCHSFTYGGCRGN <b>GN</b> RFRTIEE CRRTC VG	7,34
<b>MQ5</b>	RPSFCNLPVKPGPCSGFFSAFYYSQKTNKCHSFTY <b>SG</b> CRGN <b>GN</b> RFRTIEE CRRTC VG	3,50
<b>MQ6</b>	RPSFCNLPVKPGPCSGFFSAFYYSQKTNKCHSFTYGGCR <b>PG</b> NRFR <b>T</b> IEE CRRTC VG	21,49
<b>MQ7</b>	RP <b>Y</b> ACELIVAAGPC <b>M</b> FFISAFYYS <b>K</b> GANK <b>C</b> YPFTY <b>SG</b> CRGNANRF <b>K</b> TIEE CRRTC <b>V</b> V	3,14
<b>MQ8</b>	RP <b>Y</b> ACELIVAAGPC <b>M</b> FFISAFYYS <b>K</b> GSNK <b>C</b> YPFTY <b>SG</b> CRGNANRF <b>K</b> TIEE CRRTC <b>V</b> V	3,14
<b>MQ9</b>	RP <b>Y</b> ACEL <b>T</b> VAAGPC <b>L</b> RF-SAFYYS <b>K</b> GAN <b>Q</b> CYP <b>F</b> NY <b>S</b> GCGGNANRF <b>S</b> T <b>N</b> Q <b>K</b> CRRTC <b>V</b> V	45,05
<b>P19859</b>	RPRFCE <b>L</b> APSAGSC <b>F</b> GFVSSY <b>Y</b> Y <b>N</b> RYSNTCHSFTY <b>SG</b> CGK <b>N</b> ANRF <b>R</b> T <b>I</b> DECNRTC <b>V</b> V	7,87
<b>Q5ZPJ7</b>	RPRFCE <b>L</b> APSAGSC <b>F</b> A <b>F</b> V <b>P</b> SY <b>Y</b> Y <b>N</b> Q <b>Y</b> SNTCHSFTY <b>SG</b> CG <b>G</b> NANRF <b>R</b> T <b>I</b> DECNRTC <b>V</b> V	111,67
<b>DTX-E</b>	<b>L</b> QHRT <b>F</b> CK <b>L</b> PAEPG <b>P</b> CK <b>A</b> S <b>I</b> PAFY <b>Y</b> N <b>W</b> A <b>A</b> KK <b>Q</b> LFHYGGCKGNANRFSTIEKCR <b>H</b> ACVG	>>1000
<b>DTX-R55</b>	<b>L</b> QHRT <b>F</b> CK <b>L</b> PAEPG <b>P</b> CK <b>A</b> S <b>I</b> PAFY <b>Y</b> N <b>W</b> A <b>A</b> KK <b>Q</b> LFHYGGCKGNANRFSTIEKCR <b>R</b> ACVG	>>1000
<b>δ-DTX</b>	<b>A</b> AKYCK <b>L</b> PVRYG <b>P</b> CK <b>K</b> K <b>I</b> PSFY <b>Y</b> K <b>W</b> K <b>A</b> K <b>Q</b> CLPF <b>D</b> Y <b>S</b> GCGGNANRF <b>K</b> TIEE CRRTC <b>V</b> V	>>1000
<b>Dtx-K</b>	<b>A</b> AKYCK <b>L</b> PLRIG <b>P</b> CK <b>R</b> R <b>K</b> I <b>P</b> SFY <b>Y</b> K <b>W</b> K <b>A</b> K <b>Q</b> CLPF <b>D</b> Y <b>S</b> GCGGNANRF <b>K</b> TIEE CRRTC <b>V</b> V	>>1000
<b>α-Dtx</b>	<b>Q</b> PRR <b>K</b> L <b>C</b> ILHR <b>N</b> PGRCY <b>D</b> K <b>I</b> PAFY <b>Y</b> N <b>Q</b> K <b>K</b> Q <b>C</b> ER <b>F</b> D <b>W</b> S <b>G</b> CGG <b>N</b> SNRF <b>K</b> TIEE CRRTC <b>I</b> G	>>1000
<b>DTX-I</b>	<b>Q</b> PLR <b>K</b> L <b>C</b> ILHR <b>N</b> PGRCY <b>Q</b> K <b>I</b> PAFY <b>Y</b> N <b>Q</b> K <b>K</b> Q <b>C</b> E <b>G</b> F <b>T</b> W <b>S</b> GCGG <b>N</b> SNRF <b>K</b> TIEE CRRTC <b>I</b> R <b>K</b>	>>1000
<b>Calci</b>	<b>W</b> QPP <b>W</b> Y <b>C</b> KE <b>P</b> VRIG <b>S</b> CK <b>K</b> Q <b>F</b> SSFY <b>F</b> K <b>W</b> T <b>A</b> K <b>C</b> L <b>P</b> LF <b>S</b> GCGG <b>N</b> ANRF <b>Q</b> T <b>I</b> GE <b>C</b> R <b>K</b> K <b>L</b> G <b>K</b>	>>1000
<b>E7FL11</b>	RPRFCE <b>L</b> PA <b>D</b> PGPC <b>N</b> GL <b>F</b> QAFY <b>Y</b> N <b>P</b> V <b>Q</b> R <b>K</b> CL <b>K</b> FRYGGCKGN <b>P</b> NT <b>F</b> K <b>T</b> IEE <b>C</b> K <b>R</b> T <b>C</b> AA	>>1000
<b>B4ESA3</b>	RPPFC <b>N</b> L <b>L</b> PE <b>P</b> GR <b>C</b> NA <b>I</b> VRAFY <b>Y</b> NS <b>R</b> PR <b>K</b> CL <b>E</b> FPYGGCGGNAN <b>N</b> F <b>K</b> TIEE <b>C</b> Q <b>R</b> T <b>C</b> AG	>>1000
<b>D5J9Q8</b>	RP <b>K</b> Y <b>C</b> N <b>L</b> P <b>P</b> EPG <b>P</b> CH <b>G</b> R <b>K</b> FAFY <b>Y</b> H <b>P</b> AS <b>N</b> K <b>C</b> KE <b>F</b> VYGGCGG <b>N</b> D <b>N</b> N <b>F</b> K <b>T</b> K <b>D</b> K <b>C</b> Q <b>R</b> AC <b>S</b> G	>>1000
<b>F8J2F6</b>	RP <b>H</b> F <b>C</b> H <b>L</b> PA <b>D</b> PGRC <b>N</b> AL <b>S</b> EAFY <b>Y</b> N <b>P</b> V <b>Q</b> R <b>K</b> CL <b>K</b> FRYGGCK <b>A</b> NAN <b>T</b> F <b>K</b> T <b>I</b> DE <b>C</b> K <b>R</b> T <b>C</b> AA	>>1000
<b>C1IC51</b>	RPR <b>L</b> CE <b>L</b> PA <b>E</b> S <b>G</b> LC <b>N</b> AY <b>I</b> PSFY <b>Y</b> N <b>P</b> H <b>S</b> H <b>K</b> Q <b>K</b> F <b>M</b> YGGCGG <b>N</b> AN <b>N</b> F <b>K</b> T <b>I</b> V <b>E</b> C <b>H</b> R <b>T</b> C <b>V</b> G	>>1000
<b>BPTI</b>	RP <b>D</b> FC <b>L</b> EP <b>P</b> Y <b>T</b> GP <b>C</b> K <b>A</b> R <b>I</b> I <b>R</b> Y <b>F</b> Y <b>N</b> A <b>K</b> AG <b>L</b> C <b>Q</b> T <b>F</b> VYGG <b>C</b> R <b>A</b> K <b>R</b> N <b>N</b> F <b>K</b> S <b>A</b> E <b>D</b> C <b>M</b> R <b>T</b> C <b>G</b> GA	>>1000

Tableau 14 : Alignement et K<sub>i</sub> envers V2R de différentes toxines possédants un repliement de type « kunitz ».

Pour compléter cette étude, nous avons réalisé des simulations sur d'autres systèmes. D'une part nous avons testé la liaison de MQ-1 sur le récepteur V1aR, pour lequel elle ne présente aucune affinité, afin de vérifier si notre protocole pouvait retranscrire cette sélectivité et essayer de l'expliquer. D'autre part, nous avons simulé deux autres toxines possédant également un repliement de type « kunitz » mais n'ayant pas d'affinité pour le V2R : E7FL11 et DTX-E (voir alignement de séquences ci-dessus). Le protocole utilisé est identique en tout point à celui utilisé pour le couple MQ-1:V2R. Nous nous sommes particulièrement focalisé sur la présence de chacune de ces toxines dans la poche de liaison en fonction de la température comme nous avons fait pour la série de peptides sélectifs des récepteurs NTR1 et NTR2 (Figure 57). La première conclusion est que MQ-1 ne se lie pas au V1aR, en accord avec les résultats expérimentaux. Toutefois, E7FL11 et surtout DTX-E se lient de manière significative au V2R. Dans le cas d'E7FL11, on observe une courbe plus basse que MQ-1 et le clustering ne donne pas de conformation dans la poche de liaison parmi les dix clusters les plus peuplés. Comme observé pour les peptides non-affins des récepteurs de la Neurotensine (voir partie 3.2.2), les simulations CG-REMD sont un peu trop permissives, ce résultat n'est donc pas incohérent avec les résultats précédemment obtenus. Pour DTX-E la courbe obtenue est plus

haute que celle correspondant à MQ-1 mais le clustering ne donne qu'un cluster dans la poche parmi les dix premiers en population, montrant une conformation très proche du **modèle 2** observé pour MQ-1:V2R (RMSD de 2,49Å sur le squelette des toxines). Ici, le problème pourrait provenir de la méthode utilisée pour mesurer la liaison de la toxine et qui se base sur un critère de distance à certains résidus de la poche. En effet, pour ce type de système où le récepteur subit une large ouverture pour laisser entrer la toxine, on observe un écartement des résidus de la poche pouvant ainsi fausser nos mesures. Aussi, nous cherchons actuellement comment utiliser une autre variable comme par exemple la mesure d'une surface de contact toxine:récepteur qui pourrait s'avérer plus précise pour ces systèmes.

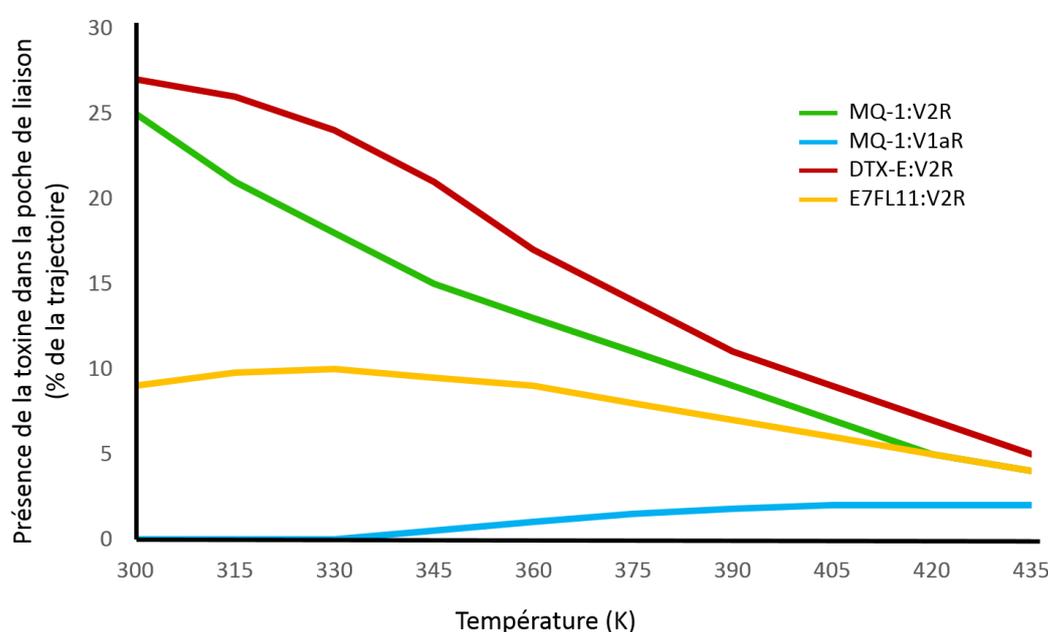


Figure 57 : Evolution du pourcentage de la liaison des différentes toxines en fonction de la température lors de nos simulations CG-REMD.

#### 3.4.1.c Comparaison avec la Vasopressine

Dans le cadre d'un autre projet cette fois en collaboration avec Hélène DEMENE du Centre de Biochimie Structurale (CBS, Montpellier), nous avons simulé la liaison de la Vasopressine au même récepteur V2R. La conformation de départ du peptide a été construite sous contraintes RMN. Cette conformation correspond à la forme libre du peptide (dans l'eau). L'utilisation du même protocole CG-REMD a conduit à l'obtention de deux conformations très nettement majoritaires (**Figure 58**). Ces deux conformations sont relativement proches l'une de l'autre (RMSD de 5 Å sur l'ensemble du peptide) et sont capables de s'échanger lors de

dynamiques moléculaires CG à l'échelle de la  $\mu$ s. Dans ces deux orientations, le pont disulfure intrinsèque de la vasopressine est accessible au solvant et le résidu F3 de la vasopressine plonge au fond de la poche orthostérique du récepteur, pour établir un contact avec le résidu Q291 déjà décrit dans notre modèle d'interaction MQ-1:V2R. Des analyses sont en cours, notamment visant à confronter nos modèles à des résultats de RMN STD (« Saturation-Transfer Difference »), validant cette orientation de la Vasopressine dans son récepteur (travaux non présentés dans ce manuscrit).

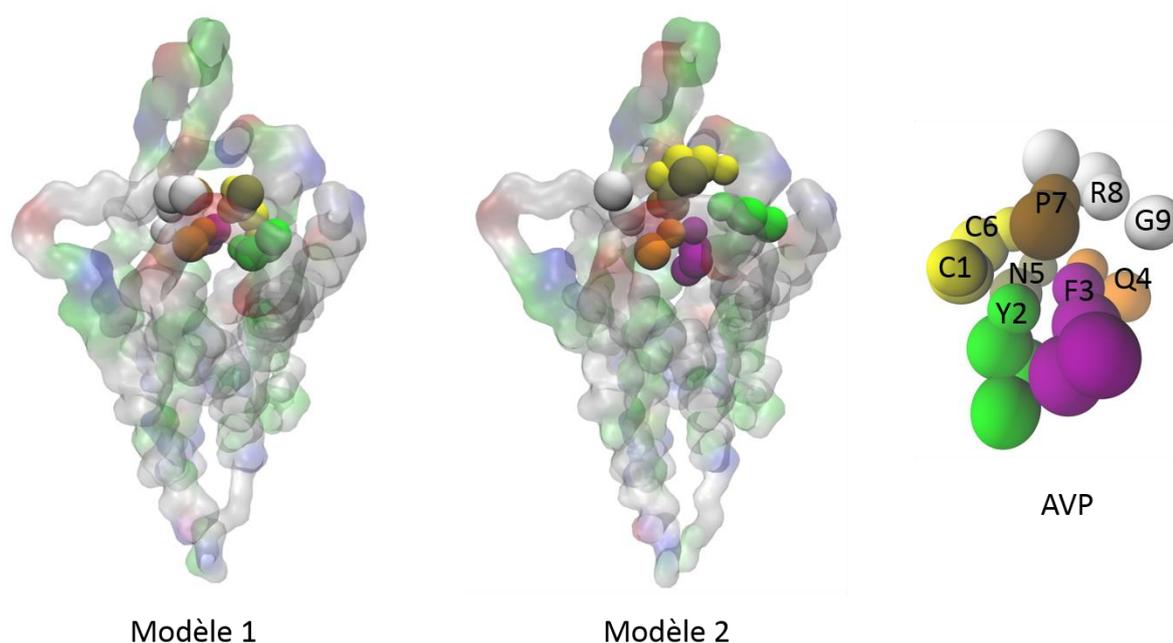


Figure 58 : Modèles de la Vasopressine liée au V2R obtenus après nos simulations CG-REMD.

#### 3.4.1.d Perspectives

Concernant les toxines, nos modèles n'ont pas permis de définir une orientation précise. En effet, l'approche combinée modélisation-biochimie-biologie moléculaire n'a abouti à aucune mutation permettant de clairement discriminer les deux conformations obtenues, à cause notamment de la grande surface d'interaction impliquée et à la symétrie des deux modèles. Toutefois, la comparaison des séquences de différentes toxines semble tendre vers le modèle 1. Mais la grande variété des séquences et les effets compensatoires possibles rendent difficile toute conclusion sur l'effet de résidu spécifique sur l'affinité et la sélectivité des différentes toxines. L'étude faite sur la Vasopressine semble également diriger notre choix vers le modèle 1 puisque que nous retrouvons des contacts similaires entre MQ-1:V2R et AVP:V2R et dont les mutations affectent énormément l'affinité des ligands au récepteur.

Au moment de l'écriture de ce manuscrit, notre protocole a permis de limiter les modèles à deux conformations possibles possédant des caractéristiques communes importantes et de valider une orientation générale de la MQ-1 dans son récepteur. Notre protocole appliqué au couple MQ-1:V1aR a également permis d'observer la sélectivité de liaison de MQ-1 au V2R. Toutefois, il faudra étudier le cas de DTX-E plus en détail. Il pourrait ici s'agir d'une limitation du modèle à prédire correctement la sélectivité de protéines capables d'établir un grand nombre de contacts à la surface de ces récepteurs. Comme déjà décrit plus haut pour la prédiction de la formation de dimères/oligomères de RCPGs, nos analyses souffrent peut-être là aussi d'un manque d'échantillonnage. Pour le vérifier, d'autres calculs portant sur le couple MQ-1:V2R ont été relancés sur le calculateur IRENE au TGCC, visant à obtenir au total plus de 5ms de dynamique CG de ce système.

## 3.4.2 Prédiction de la liaison de peptides au récepteur de la Ghréline

### 3.4.2.a La Ghréline

Ce travail a été réalisé dans le cadre d'une ANR dont le but est de comprendre l'allostérie, la dynamique conformationnelle et la signalisation du récepteur de la Ghréline. Une grande partie des analyses a été faite par Maxime LOUET au sein de notre équipe. Ce projet implique pour la partie RMN les groupes d'Alain MILON de l'Institut de Pharmacologie et de Biologie Structurale (IPBS, Toulouse) et de Laurent CATOIRE du Laboratoire de Biologie Physico-Chimique des Protéines Membranaires (LBPCPM) (IBPC, Paris). Dans le travail présenté ci-après, nous avons comparé nos résultats de CG-REMD visant à déterminer le mode et la dynamique de liaison du peptide de la Ghréline à son récepteur, à ceux obtenus en RMN dans le groupe d'Alain MILON. L'approche RMN utilisée est la même que celle décrite dans leurs travaux récemment publiés sur la Dynorphine, un peptide se liant au récepteur  $\kappa$ -Opioïde [O'Connor et al. 2015]. Ainsi, des contraintes (nOes notamment) ont pu être déterminées permettant de résoudre la structure du peptide Ghréline lié à son récepteur. Lors de nos calculs, nous sommes partis du conformère le plus étendu du peptide, en nous limitant aux 12 premiers résidus ; en effet, la structure obtenue (**Figure 59**) montre une grande flexibilité de sa partie C-terminale.

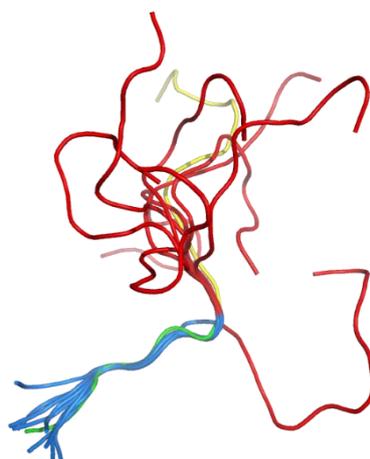
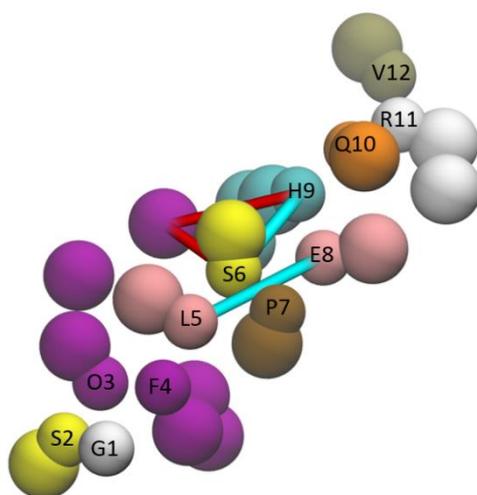


Figure 59 : Représentation de la structure RMN du peptide de la Ghréline lorsqu'il est lié à son récepteur, les résidus [1-7] sont représentés en bleu et les résidus [8-18] en rouge, la conformation utilisée comme point de départ de nos simulations est représentée en vert et jaune.

### 3 Résultats

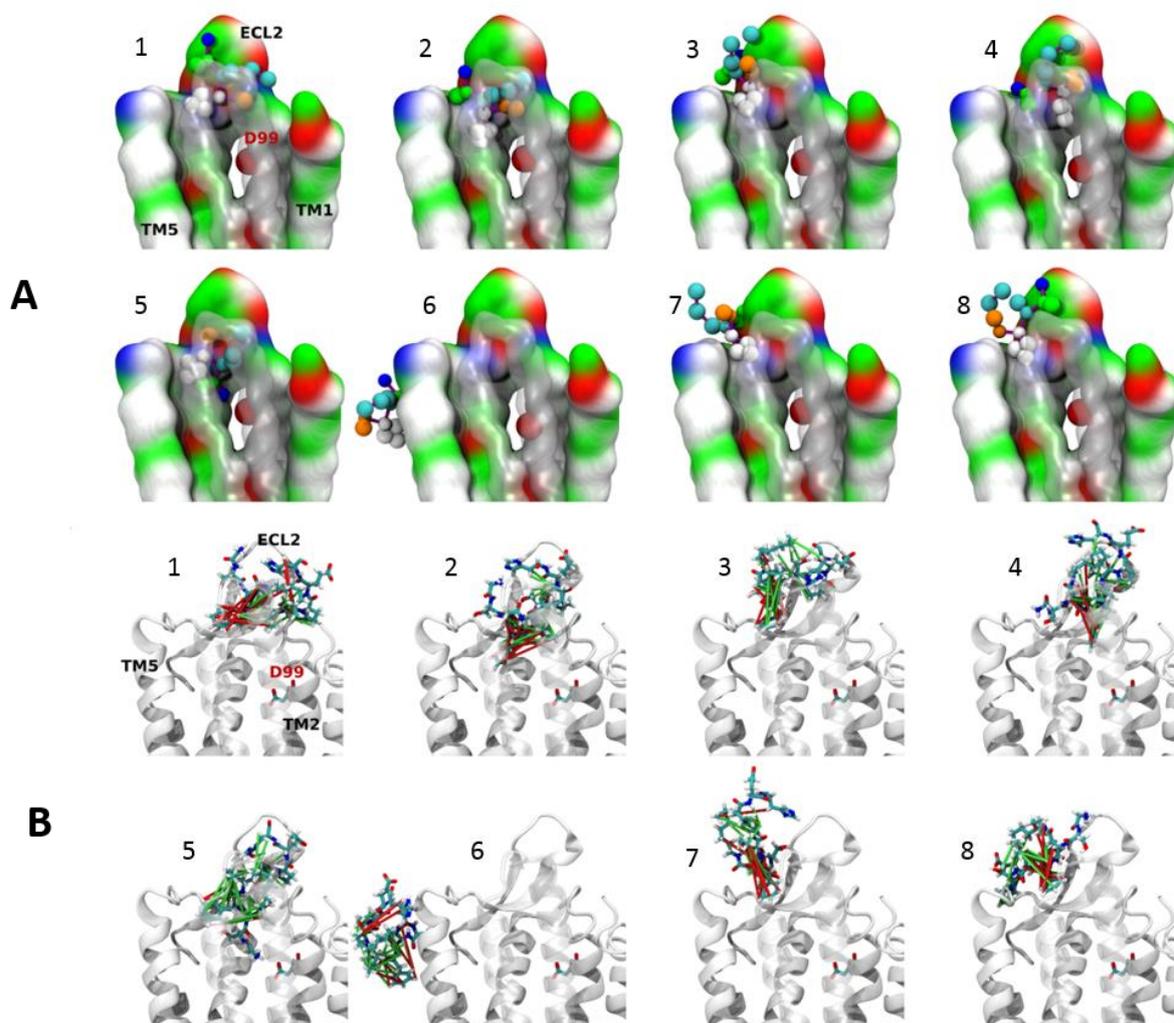
Après conversion en gros-grains, la conformation utilisée comme point de départ de nos simulations présente deux élastiques (entre L5:E8 et S6:H9) (**Figure 60**). Sans aucune autre contrainte supplémentaire appliquée, le peptide Ghréline s'ancre très rapidement à la membrane *via* l'insertion de son groupement Octanoyl dans la bicouche de POPC, rendant ainsi impossible (ou très difficile) sa liaison à son récepteur. De la même façon, des simulations réalisées sur le peptide en introduisant toutes les contraintes nOes issues de la RMN n'ont donné aucun résultat, vraisemblablement parce que ces contraintes empêchent le peptide d'adopter à la surface du récepteur les conformations intermédiaires nécessaires à sa liaison. Ainsi, pour les simulations dont les résultats sont présentés ci-après, nous avons utilisé un modèle de peptide incluant deux élastiques supplémentaires et maintenant la chaîne Octanoyl du peptide au contact de sa chaîne principale (entre la dernière sphère de l'Octanoyl et S6/H9) (**Figure 60**).



*Figure 60 : Représentation des élastiques MARTINI ELNeDyn appliqués au peptide de la Ghréline pendant nos simulations, en bleu les élastiques appliqués lors de la conversion du modèle tout-atomes au gros-grain, en rouge les élastiques ajoutés pour maintenir la chaîne octanoylée le long du squelette du peptide.*

Pour cette étude, trois simulations CG-REMD (50 températures ; 10 $\mu$ s) ont été réalisées pour un temps total de simulation de 1,5ms. Après clustering de toutes les conformations obtenues, nous avons centré notre analyse sur huit conformations, systématiquement retrouvées parmi les 10 clusters les plus peuplés dans les trois simulations indépendantes (**Figure 61A**). Ces conformations décrivent principalement des positions intermédiaires du peptide (clusters 1, 2, 3, 4, 7 et 8), une position du peptide dans la membrane et au contact des TM4 et TM5 du récepteur (cluster 6), mais aussi une conformation montrant le peptide dans la poche de liaison (cluster 5). Bien que dans nos simulations le peptide soit libre d'adopter un grand panel de conformations différentes, on a pu noter que cette dernière conformation présente un RMSD

moyen à la structure RMN d'environ 1,2Å. Une analyse plus détaillée a permis de montrer que parmi les 8 clusters les plus représentatifs, cette orientation au fond de la poche est celle qui respecte le mieux les contraintes nOes issues de la RMN (**Figure 61B**).



*Figure 61 : Représentation des conformations (A : gros grains ; B : tout-atomes) les plus représentatives obtenues lors des trois simulations CG-REMD effectuées sur le peptide octanoylé de la Ghréline ; seuls les cinq premiers résidus de la Ghréline sont représentés ici. (A) les TM6 et TM7 ont été mis en transparence pour un souci de clarté de la figure ; (B) Représentation du respect/violation (vert/rouge) des contraintes nOes issues de la RMN.*

L'analyse des contacts effectués par le peptide Ghréline au fond de la poche montre une très bonne corrélation avec les données de mutagenèse dirigée décrites dans la littérature (**Figure 62**). Le contact principal qui ressort de notre modèle est celui impliquant l'extrémité N-terminale du peptide et le résidu D99 du récepteur GHSR-1a. La mutation D99K confirme un rôle important de ce résidu dans la liaison de la Ghréline [Holst et al. 2007, 2009]. De la même manière, l'extrémité N-terminale et le résidu S2 du peptide interagissent avec Q120 dont

### 3 Résultats

---

la mutation a également un effet important sur l'efficacité de la Ghréline [Holst et al. 2009]. Dans notre modèle, le peptide interagit également avec le résidu R283 qui forme un pont salin avec E124 et dont la mutation entraîne une perte importante de l'activité du récepteur [Holst et al. 2007, 2009]. Les résidus F4 et L5 du peptide se placent dans des poches hydrophobes du récepteur, les résidus S6 et R8 étant orientés quant à eux au contact de résidus chargés du récepteur localisés à l'entrée du site de liaison et dans les boucles ECL2 et ECL3.

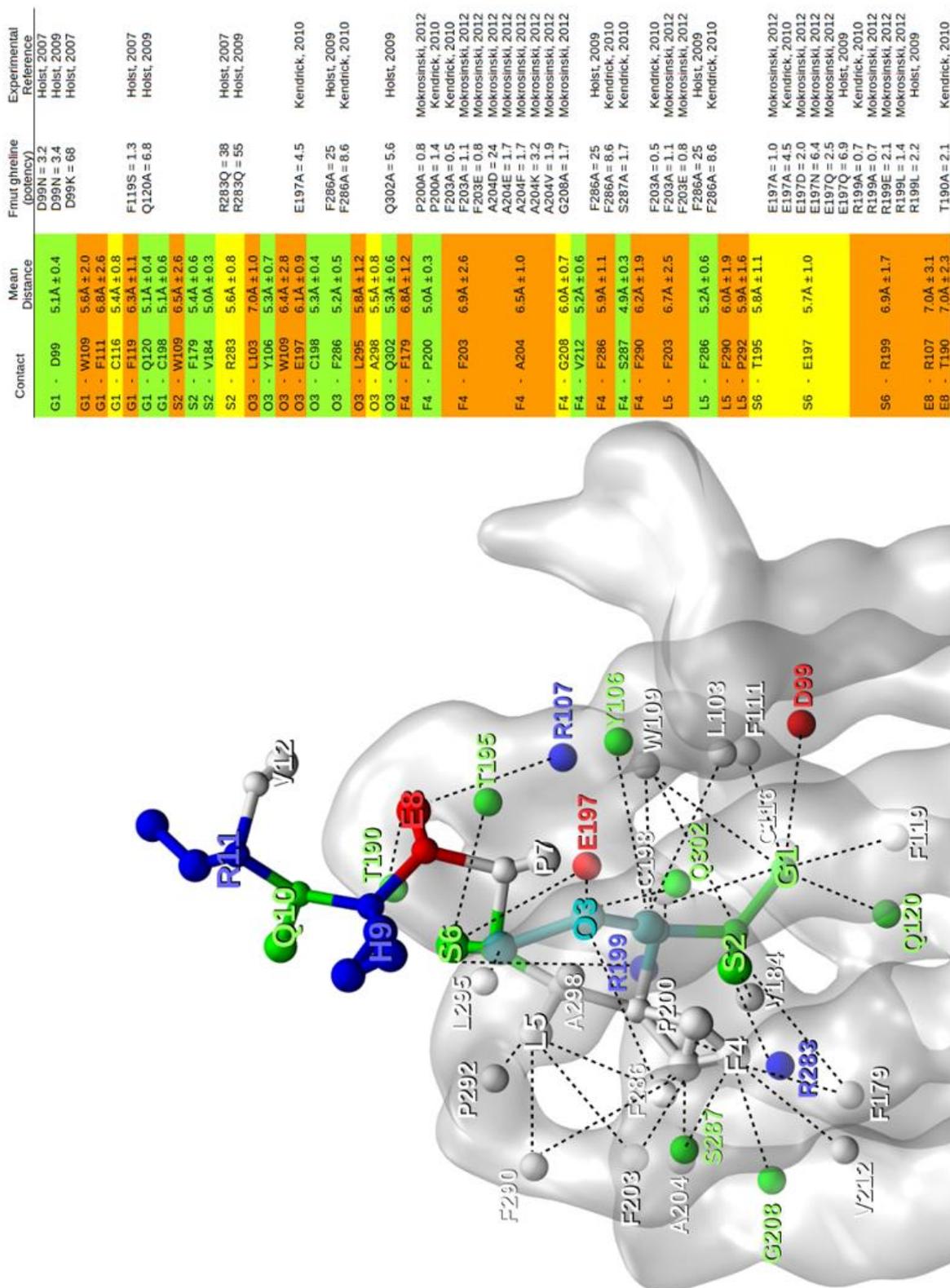


Figure 62 : Représentation des contacts entre la Ghreline et son récepteur GHSR-1a; effets de la mutation des résidus ainsi identifiés comme décrits dans la littérature.

### 3 Résultats

Comme énoncé plus tôt, nos simulations ont été réalisées en « triplicate » et nous ont permis de montrer une bonne convergence des données (**Figure 63**), concernant à la fois les clusters les plus peuplés après clustering, mais aussi le comportement même de ces clusters en fonction de la température. Des analyses complémentaires ont permis de définir qu'un temps d'environ 5 $\mu$ s de simulation était strictement requis pour obtenir une bonne convergence des données pour ce système. D'autres simulations de ce même système sont encore en cours au laboratoire, incluant des simulations CG-REMD en partant de la conformation au fond de la poche pour pouvoir discuter de l'importance du modèle de départ dans cette convergence. Nous réfléchissons également à d'autres simulations dans lesquelles le réseau élastique du récepteur et/ou du peptide seront définis à partir de cette conformation.

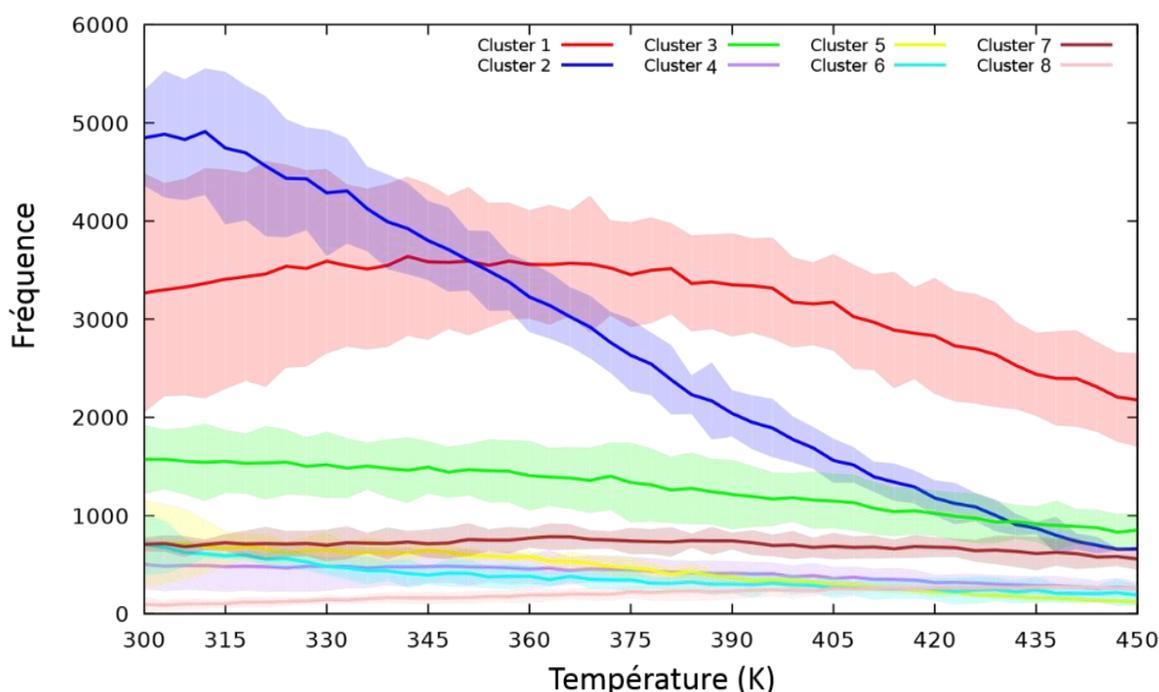


Figure 63 : Courbes de fréquence d'apparition des différents clusters en fonction de la température, chaque courbe représente la valeur moyenne observée pour les trois simulations indépendantes. La zone de couleur transparente indique l'écart entre les valeurs minimale et maximale observées pour chaque point.

Enfin, et pour compléter cette étude qui sera bientôt soumise pour publication, des simulations ont également été réalisées sur le peptide Ghréline en l'absence de son Octanoyl. La littérature montre que, sans son Octanoyl, le peptide de la Ghréline ne se lie que très peu à son récepteur (perte d'un facteur 1000 en affinité) [Bednarek et al. 2000]. En accord avec l'expérience, nos simulations montrent la disparition des clusters montrant le peptide au fond de la poche ou inséré dans la membrane.

### 3.4.2.b Le LEAP2

Le peptide LEAP2 a récemment été identifié comme ligand endogène du récepteur de la Ghréline [Ge et al. 2018]. Des travaux en cours, réalisés dans notre équipe, montrent que ce peptide aurait plutôt un effet agoniste inverse sur le GHSR-1a. Nous nous sommes également intéressés à ce peptide afin de pouvoir comparer les modes de liaison de la Ghréline (agoniste) et du LEAP2 (agoniste inverse). Des études de structure-activité réalisées ces derniers mois au laboratoire par Céline M'KADMI sur LEAP2 ont montrées que le segment [15-40] du peptide n'avait aucune affinité pour le GHSR-1a, et que les 8 premiers résidus du peptide étaient essentiels pour son affinité (**Figure 64**). En accord avec ces résultats, nous avons lancé des calculs de dynamique sur les peptides LEAP2 [1-40], [1-12] et [15-40].

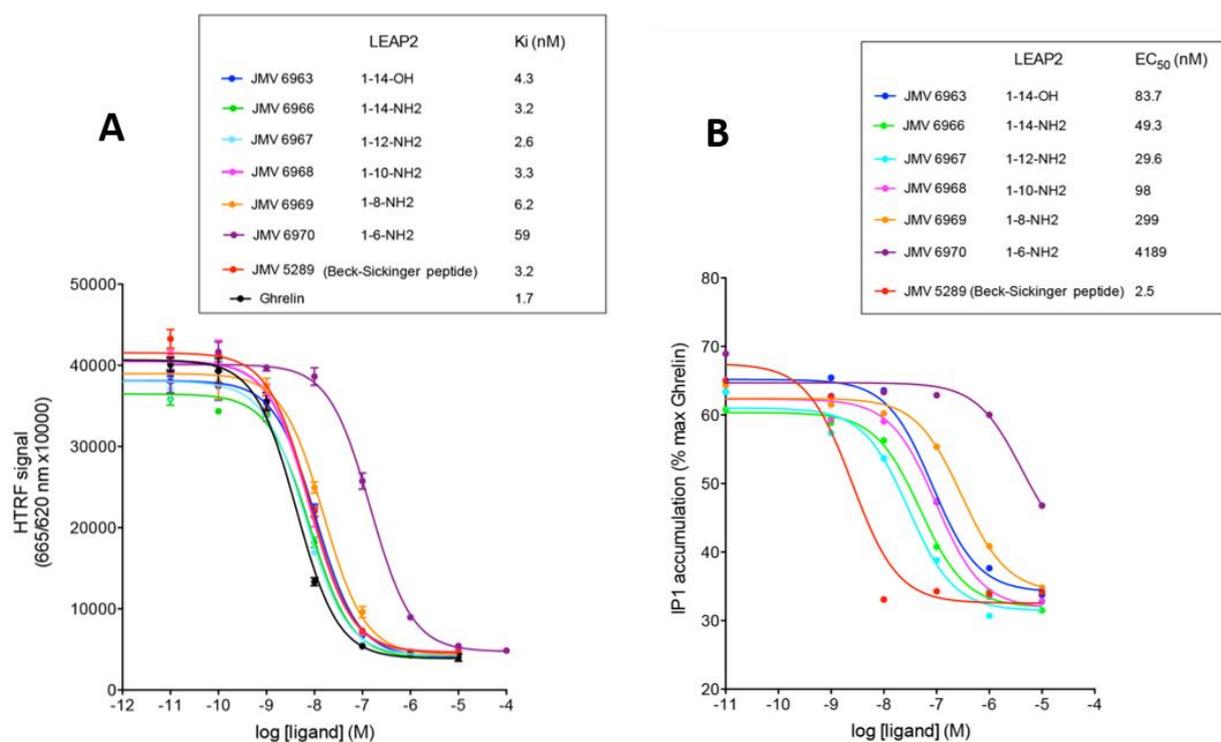


Figure 64 : Etudes structure-activité du LEAP2 réalisés au sein de notre équipe ; courbes de liaison (A) et d'activité (B) des peptides LEAP2 sur des cellules HEK293 exprimant GHSR-1a.

Le même protocole CG-REMD a été ainsi appliqué aux peptides [1-12], [1-40] et [15-40] pour un temps de simulation total de 500 $\mu$ s pour chacun des systèmes. Comme observé expérimentalement, on observe une liaison dans la poche pour les peptides [1-14] et [1-40], mais pas pour le peptide [15-40]. Plus intéressant, les peptides [1-12] et [1-40] adoptent une conformation liée au site orthostérique quasiment identique (**Figure 65A**) et très proche de la conformation déjà prédite plus haut pour la Ghréline. Cette conformation liée fait intervenir

principalement les 6 résidus N-terminaux de LEAP2, le résidu N-terminal faisant une interaction avec le résidu D99 du récepteur. La comparaison des conformations liées des deux peptides LEAP2 et Ghréline, montre une interaction forte au niveau du récepteur et impliquant les résidus E8 de la Ghréline et R6 du LEAP2 (**Figure 65B**). Ainsi, des calculs sont en cours actuellement afin d'essayer de comprendre plus précisément le mécanisme de liaison de ces peptides à leur récepteur.

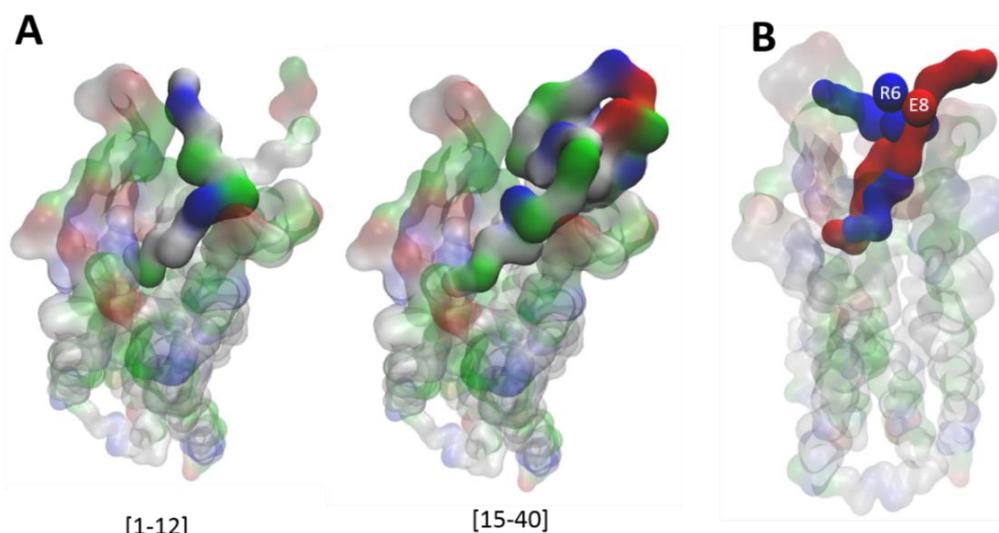


Figure 65 : A) Modèles liés des peptides LEAP2 [1-12] et [1-40] obtenus après simulations CG-REMD ; B) Comparaison du LEAP2 [1-12] lié (en bleu) avec la conformation liée prédite pour la Ghréline (en rouge); les résidus chargés des deux peptides sont mis en évidence (LEAP2:R6 et Ghréline:E8).

#### 3.4.2.c Perspectives

A nouveau, on a pu démontrer la force du protocole CG-REMD développé lors de cette thèse. On a ainsi pu observer la liaison des peptides Ghréline, dans une conformation en accord avec la RMN, et LEAP2. De plus, les peptides Ghréline non-octanoylés et LEAP2 [15-40] ne se lient pas au GHSR-1a, en accord avec les résultats expérimentaux obtenus au laboratoire.

En se basant sur la comparaison des conformations liées de la Ghréline et du LEAP2, des hypothèses ont pu être faites et vérifiées par le calcul quant au mécanisme de liaison de ces peptides à leur récepteur. Selon nos résultats, une simple inversion de charge devrait notamment permettre de concevoir des peptides dérivés de la Ghréline capables de se lier au GHSR-1a en l'absence d'Octanoyl. Ces hypothèses sont en cours de vérification expérimentale avec la synthèse par Jean-Alain FEHRENTZ de l'Institut des Biomolécules Max Mousseron (IBMM, Montpellier) de différents peptides dérivés de la Ghréline non-octanoylée et de LEAP2 dont la charge a été inversée (Ghréline:E8R et LEAP2:R6E), mais également déplacée (Ghréline:S6R/E8V et LEAP2:R6S/V8E).

### 3.4.3 Prédiction de la liaison et de la sélectivité de peptides issus de l'hélice C-terminale de Protéines-G

#### 3.4.3.a Contexte

Dans une étude de 2016 [Semack et al. 2016], il a été démontré que les 26 derniers résidus de la sous-unité  $\alpha$  des protéines-G (structurés en hélice) étaient suffisants pour se lier et conserver leur sélectivité vis-à-vis des RCPGs. De nombreuses études ont visé ces dernières années à réduire la taille des protéines-G, afin de mieux comprendre leur sélectivité de couplage [Carpenter et Tate 2016; Semack et al. 2016]. Nous avons testé ces derniers mois notre protocole pour voir s'il était capable, comme pour les peptides dérivés de la Neurotensine, de prédire la sélectivité de liaison de peptides, mais cette fois du côté intra-cellulaire.

De nombreuses structures ont été résolues ces derniers mois montrant l'interaction de différents récepteurs avec différentes protéines-G. Pour tester notre protocole, nous avons utilisé les structures Cryo-EM du récepteur de l'Adénosine A1R en complexe avec  $G_i$  [Drapier-Joyce et al. 2018], et cristallographique du récepteur de l'Adénosine A2AR complexé à la mini- $G_s$  [Carpenter et al. 2016].

#### 3.4.3.b Résultats

Dans un premier temps, des dynamiques CG libres de 10 $\mu$ s ont été réalisés en utilisant les structures résolues de chacun des complexes comme point de départ. Sans trop de surprise les récepteurs présentent une très faible flexibilité contrairement aux peptides C-terminaux (**Figure 66A** et **66B**). Toutefois, la flexibilité des peptides est principalement expliquée par leur segment n'étant pas en contact avec le récepteur. En effet, les valeurs de RMSD calculées sur les dix derniers résidus (ceux réellement impliqués dans les contacts) montrent une bien meilleure stabilité de ces contacts dans la poche de chacun des récepteurs (**Figure 66C**). Il est toutefois important de garder en tête que notre modèle n'inclut que la partie C-terminale de la sous-unité  $\alpha$ . Manquent ainsi plus de 300 résidus de la sous-unité  $\alpha$  ainsi que les sous-unités  $\beta$  et  $\gamma$ . Cette flexibilité pourrait donc être due à la troncature des modèles.

### 3 Résultats

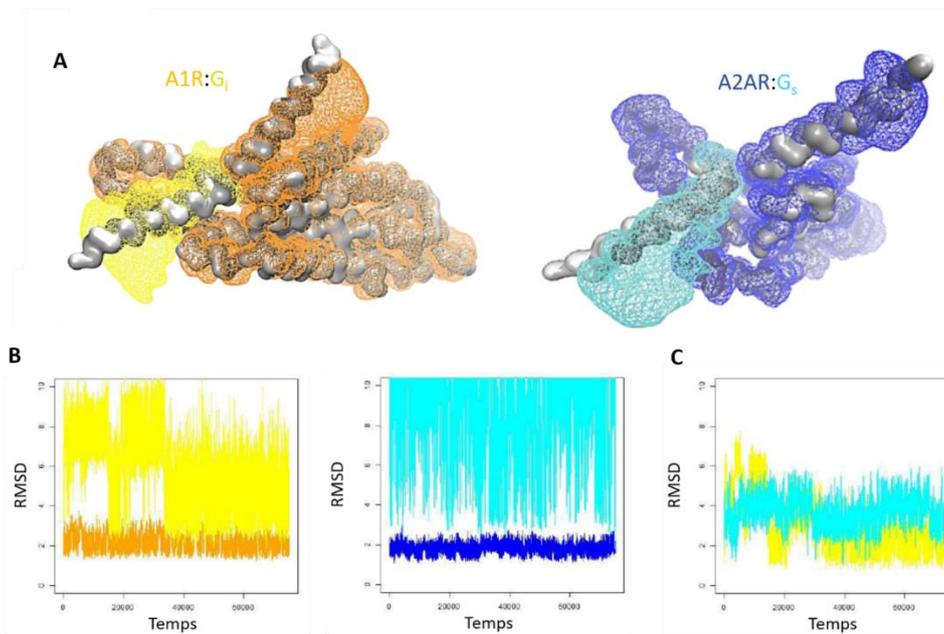


Figure 66 : Flexibilité des récepteurs et des peptides le long des trajectoires de simulations CG en partant des structures résolues ; Pour chaque couple, sont présentées A) une carte de densité (« volmap ») et B) les valeurs de RMSD calculées par rapport à la structure de départ (peptide et récepteur) ; C) Les RMSD calculés sur les dix derniers résidus des peptides montrent une bien plus faible flexibilité.

Le résultat important de ces dynamiques de référence est la conservation de l'interface récepteur:protéine-G avec des contacts forts impliquant les cinq derniers résidus des peptides C-terminaux, et retrouvés à plus de 99% du temps des trajectoires (**Figure 67**). Ces contacts impliquant des résidus des récepteurs identiques, et surtout situés structurellement au même endroit pour les récepteurs A1R et A2AR, nous ont permis de définir des critères déterminant si, oui ou non, le peptide est situé dans la poche intra-cellulaire du récepteur.

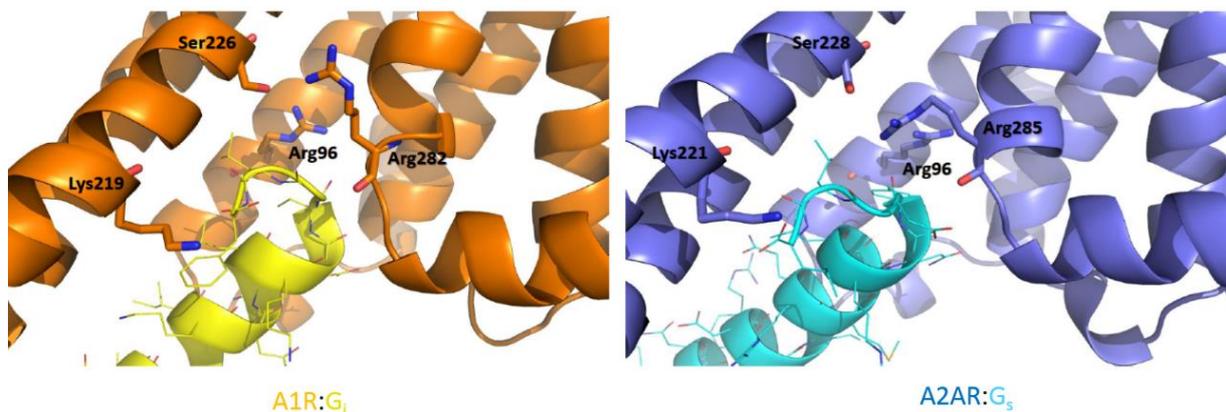


Figure 67 : Contacts peptide:récepteur non spécifiques présents pendant plus de 99% de la simulation CG utilisant la structure résolue des récepteurs A1R et A2AR liés comme point de départ.

Des simulations de liaison ont ensuite été réalisées pour ces deux récepteurs, en testant pour chaque récepteur, chacun des trois peptides  $G_i$ ,  $G_s$  et  $G_q$  construits. Pour les six complexes simulés (3 peptides *versus* 2 récepteurs) des simulations CG-REMD de 10 $\mu$ s pour chacune des 50 températures utilisées et séparées de 3K (300 à 447K) ont été réalisées soit un temps de simulation total de 500 $\mu$ s par couple. En utilisant les critères de contacts définis par les trajectoires de référence, nous avons mesuré la présence des différents peptides dans la poche de chacun des récepteurs. On observe qu'à 300K, A1R se lie préférentiellement à  $G_q$  et  $G_i$  par rapport à  $G_s$ , et A2AR se lie préférentiellement à  $G_s$  et  $G_i$  par rapport à  $G_q$  (**Figure 68A**). On observe ainsi deux comportements très différents, montrant que le modèle permet aux récepteurs de bien différencier les peptides et ce malgré la ressemblance des résidus de la poche intracellulaire. Dans une étude de 2017 [Flock et al. 2017] utilisant une comparaison phylogénétique des différents RCPG et protéines-G, la séquence des protéines-G a été assimilée à un « code barre » pouvant être reconnu par le récepteur, permettant de prédire les préférences de couplage aux protéines-G de nombreux RCPGs. Il a été ainsi proposé que A1R se lie aussi bien à  $G_i$ ,  $G_q$  que  $G_s$ , ce qui n'est pas en accord avec nos résultats. Pour A2AR, il a été proposé une liaison préférentielle à  $G_i$  et  $G_s$  mais pas à  $G_q$ , ce qui cette fois est en accord avec nos résultats. Malgré tout, le calcul du RMSD du peptide vis-à-vis des références (**Figure 68B**) nous montre que les conformations des structures résolues sont bien obtenues dans nos simulations.

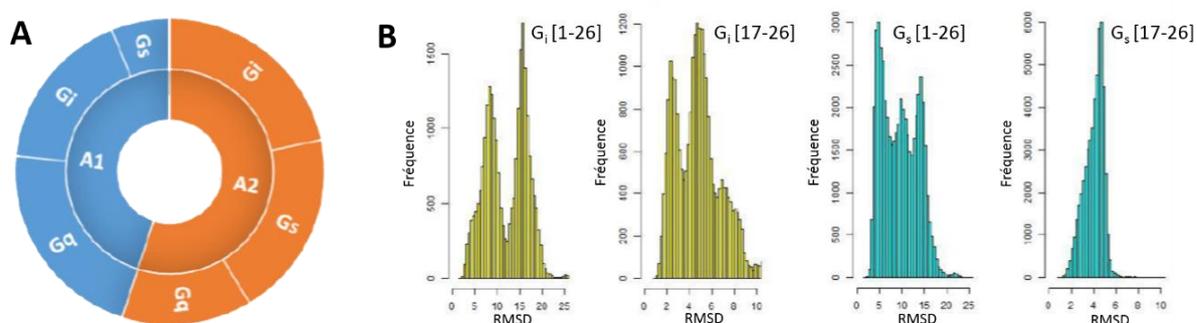


Figure 68 : Résultats des simulations CG-REMD réalisées sur les peptides C-terminaux  $G_i$ ,  $G_s$  et  $G_q$  ; A) Proportion de la liaison des peptides C-terminaux à A1R et A2AR ; B) Fréquences des RMSD des peptides C-terminaux pour les simulations A1R: $G_i$  et A2AR: $G_s$  calculés par rapport aux structures résolues, le RMSD est à chaque fois calculé sur tout le peptide et sur les dix derniers résidus.

### 3.4.3.c Perspectives

Cette étude en est encore assez récente mais montre déjà des résultats très encourageants. Si les peptides C-terminaux présentent une flexibilité assez importante, probablement due à l'absence de la Protéine-G complète, ils présentent à l'interface avec le récepteur des contacts forts impliquant les cinq derniers résidus des peptides. On a également pu reproduire partiellement la sélectivité des différents peptides vis-à-vis des différents récepteurs. Des simulations et analyses supplémentaires pourraient nous permettre d'affiner nos résultats, avec par exemple l'application d'un réseau élastique plus souple permettant de reproduire au mieux la dynamique du récepteur sur sa face intra-cellulaire.

Pour la suite de cette étude, des analyses plus précises devront être mises en place afin de comprendre cette sélectivité, avec notamment l'implication des derniers résidus des peptides. Toutefois, l'étude d'autres récepteurs sera nécessaire. En effet, si les récepteurs de l'Adénosine nous ont permis de valider la liaison des peptides, il ne s'agit pas de récepteurs idéaux pour la prédiction de la sélectivité puisque d'après l'étude précédemment citée [Flock et al. 2017], ces récepteurs se lient à la plupart des protéines-G. Des simulations sur des récepteurs se liant spécifiquement à une seule protéine-G seront ainsi bientôt lancées.

Il sera aussi intéressant par la suite de réaliser les mêmes calculs mais en utilisant la protéine-G complète voir avec des dimères de RCPGs comme celui présenté dans ce manuscrit (voir partie 3.3.1). A terme, on pourrait imaginer réaliser un « criblage » des différentes protéines-G pour un nombre plus conséquent de RCPGs, natifs ou bien mutés, ce qui aurait un vaste champ d'applications très intéressant notamment pour la compréhension des mécanismes impliqués dans le couplage de la protéine-G.

## 4 Conclusions et Perspectives

L'objectif de cette thèse était de développer un protocole capable de prédire la dynamique d'interaction des Récepteurs Couplés aux Protéines-G avec leurs différents partenaires, en tirant à la fois partie de la simplification des systèmes simulés grâce à l'utilisation de modèles gros-grains et de l'amélioration de l'échantillonnage apportée par la dynamique moléculaire « replica-exchange ». En effet, la prédiction de la liaison d'une petite molécule organique à son récepteur reste difficile à réaliser, et nécessitait jusqu'alors soit d'utiliser une méthode de simulation « biaisée », soit d'avoir accès à des machines de calcul extrêmement coûteuses.

Quand je suis arrivé au laboratoire, les tests préliminaires réalisés par mon directeur de thèse et Landry CHARLIER (ancien post-doc dans l'équipe) montraient qu'il était possible a priori de prédire la liaison du peptide de la Neurotensine à son récepteur, sans aucune contrainte, en utilisant un modèle gros-grains (MARTINI) et un cluster de laboratoire de taille raisonnable. Ce résultat ne fût néanmoins observé qu'une seule fois, parmi de très nombreuses trajectoires. L'idée d'utiliser ce champ de forces dans des dynamiques de REMD donna lieu à mon sujet de stage, puis de thèse.

La première étape de mon travail a vu la validation de la méthode sur des complexes peptide:récepteur connus en cristallographie. Les deux seuls exemples connus à l'époque étaient le peptide Neurotensine qui avait servi aux études préliminaires, et le peptide CVX15, antagoniste du récepteur de la Chémokine C-X-C de type-4. Nous avons pu ainsi démontrer que nos simulations CG-REMD permettaient, sans aucune contrainte, d'observer presque systématiquement la liaison des peptides dans le site orthostérique de leur récepteur, mais aussi d'obtenir des conformations liées très proches de la cristallographie. A l'époque, la simulation d'un système nécessitait environ deux semaines de calcul sur notre cluster pour un temps de simulation total de plus de 100 $\mu$ s.

Ces résultats très intéressants nous permirent ensuite de réaliser un très grand nombre de collaborations, au sein du laboratoire tout comme à l'extérieur. Dans un premier temps, grâce aux observations réalisées lors de la validation du protocole sur la Neurotensine, il fut possible de concevoir des peptides dérivés de la Neurotensine, et sélectifs du récepteur Neurotensine-2 humain. En effet, lors de simulations de ces récepteurs, construits par homologie, utilisant le peptide lié comme point de départ, nous avons suggéré le rôle clé d'un contact entre le peptide

et le récepteur dans cette sélectivité. Grâce à cette étude menée en parallèle, et la confirmation de notre hypothèse par la synthèse et tests de peptides dérivés, nous avons pu utiliser ces nouvelles données expérimentales comme point de départ pour la suite de notre travail. Ainsi, nous avons pu démontrer la capacité de notre protocole non seulement à prédire la position liée de ces peptides dans leurs récepteurs, mais aussi leur sélectivité sur le plan qualitatif. La vue dynamique apportée par nos simulations nous a aussi permis de proposer un mécanisme détaillé d'entrée des peptides dans la poche de liaison de leurs récepteurs.

Ces résultats très encourageants nous ont amené à tester notre protocole sur différents autres systèmes. Très rapidement, nous avons travaillé sur la prédiction de la liaison d'une toxine de venin de serpent, la Mambaquarétine-1 au récepteur de la Vasopressine de type-2. Nous avons ainsi pu proposer un modèle général de la liaison de cette toxine à son récepteur, largement validé par les données expérimentales de mutagénèse dirigée, et qui est en accord avec d'autres calculs réalisés sur le peptide de la Vasopressine. Plus récemment, nous avons étudié la liaison des peptides Ghréline et LEAP2 au récepteur de la Ghréline. L'intérêt de ce travail était notamment de confronter nos résultats à ceux obtenus en RMN. Ces résultats très prometteurs et la comparaison du mode de liaison de la Ghréline et du LEAP2 nous a permis de mettre en évidence des résidus importants de ces peptides, et de proposer la synthèse de nouveaux peptides modifiés.

Après avoir validé notre protocole sur des interactions faisant intervenir différents partenaires extra-cellulaires, nous nous sommes logiquement intéressés à des partenaires intra-membranaires, puis à des partenaires intra-cellulaires. En effet, la dimérisation des RCPGs est un sujet très étudié au sein de l'équipe. Il a ainsi été possible de proposer un modèle de dimérisation inédit, entre le récepteur de la Ghréline et le récepteur de la Dopamine D2, validé par des mesures expérimentales, et qui implique deux protéines-G parfaitement imbriquées. Ces protéines-G forment d'ailleurs une interface très complémentaire qui a pu être validée par la synthèse de peptides empêchant la formation de ce dimère. Le protocole CG-REMD a ensuite permis d'étudier la dynamique de dimérisation de ces récepteurs afin notamment de prédire la position des deux récepteurs supplémentaires observés expérimentalement.

Enfin, nous avons il y a quelques mois démarré l'étude des interactions des RCPGs avec leurs partenaires intra-cellulaires favoris, les protéine-G. Pour commencer, nous nous sommes intéressés à des peptides issus de l'extrémité C-terminale de la sous-unité  $\alpha$  de la protéine-G, et qui conservent suffisamment d'information pour présenter une sélectivité pour différents

récepteurs. Nos premiers résultats sont assez encourageants encore une fois, et ont permis de montrer que notre protocole était capable de prédire la conformation liée de ces peptides. Pour la prédiction de la sélectivité, des simulations supplémentaires seront bientôt lancées pour une étude plus complète.

Au final, le protocole CG-REMD développé lors de ma thèse a permis l'étude d'un grand nombre de systèmes. Liaison d'un ligand extra-cellulaire, qu'il s'agisse d'un petit peptide ou d'une protéine, liaison d'un partenaire intra-cellulaire mimant la protéine-G, ou encore la dimérisation des RCPGs. En plus de l'efficacité de la méthode, les temps de calculs qui y sont associés sont très courts puisque nous réalisons actuellement 500 $\mu$ s de simulations en 3-5 jours sur divers centres de calculs. La variété des objets étudiables nous a permis de réaliser de nombreuses applications qui ont chacune apportées leur pierre à l'édifice en renforçant notre manière de procéder, que ce soit au niveau de la construction des systèmes, des analyses effectuées ou encore de la compréhension générale des résultats obtenus.

Mais c'est surtout les perspectives offertes par ces applications qu'il faut retenir. Nous pensons que ce protocole pourra être étendu à d'autres systèmes de plus en plus importants en commençant par des simulations impliquant le récepteur et la protéine-G. Pour ce faire, nous sommes en train de développer un réseau élastique modifié afin de simuler le plus fidèlement possible la dynamique intrinsèque de chacun de partenaires impliqués.

La simplicité et la rapidité de la méthode pourraient ainsi permettre de réaliser un criblage des interactions des RCPGs avec leurs différents partenaires à une plus grande échelle.

## 5 Références

- Abraham, Mark James, Teemu Murtola, Roland Schulz, Szilárd Páll, Jeremy C. Smith, Berk Hess, et Erik Lindahl. 2015. « GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers ». *SoftwareX* 1-2 (septembre): 19-25.
- Adler, E., M. A. Hoon, K. L. Mueller, J. Chandrashekar, N. J. Ryba, et C. S. Zuker. 2000. « A Novel Family of Mammalian Taste Receptors ». *Cell* 100 (6): 693-702.
- Akkermans, Reinier L. C., So/ren Toxvaerd, et W. J. Briels. 1998. « Molecular dynamics of polymer growth ». *The Journal of Chemical Physics* 109 (7): 2929-40.
- Alkhatib, Ghalib, et Edward A. Berger. 2007. « HIV Coreceptors: From Discovery and Designation to New Paradigms and Promise ». *European Journal of Medical Research* 12 (9): 375-84.
- Allavena, P., A. Sica, A. Vecchi, M. Locati, S. Sozzani, et A. Mantovani. 2000. « The Chemokine Receptor Switch Paradigm and Dendritic Cell Migration: Its Significance in Tumor Tissues ». *Immunological Reviews* 177 (octobre): 141-49.
- Allen, John A., et Bryan L. Roth. 2011. « Strategies to Discover Unexpected Targets for Drugs Active at G Protein-Coupled Receptors ». *Annual Review of Pharmacology and Toxicology* 51: 117-44.
- André, Nicolas, Nadia Cherouati, Cécile Prual, Tania Steffan, Gabrielle Zeder-Lutz, Thierry Magnin, Franc Pattus, Hartmut Michel, Renaud Wagner, et Christoph Reinhart. 2006. « Enhancing Functional Production of G Protein-Coupled Receptors in *Pichia Pastoris* to Levels Required for Structural Studies via a Single Expression Screen ». *Protein Science* 15 (5): 1115-26.
- Andrews, Donald H. 1930. « The Relation Between the Raman Spectra and the Structure of Organic Molecules ». *Physical Review* 36 (3): 544-54.
- Arora, K. K., A. Sakai, et K. J. Catt. 1995. « Effects of Second Intracellular Loop Mutations on Signal Transduction and Internalization of the Gonadotropin-Releasing Hormone Receptor ». *The Journal of Biological Chemistry* 270 (39): 22820-26.
- Attwood, T. K., et J. B. Findlay. 1993. « Design of a Discriminating Fingerprint for G-Protein-Coupled Receptors ». *Protein Engineering* 6 (2): 167-76.
- Azpiazu, Inaki, et N. Gautam. 2004. « A Fluorescence Resonance Energy Transfer-Based Sensor Indicates That Receptor Access to a G Protein Is Unrestricted in a Living Mammalian Cell ». *Journal of Biological Chemistry* 279 (26): 27709-18.
- Barbeau, A. 1970. « Dopamine and disease. » *Canadian Medical Association Journal* 103 (8): 824-32.
- Beaulieu, Jean-Martin, et Raul R. Gainetdinov. 2011. « The Physiology, Signaling, and Pharmacology of Dopamine Receptors ». *Pharmacological Reviews* 63 (1): 182-217.
- Bednarek, M. A., S. D. Feighner, S. S. Pong, K. K. McKee, D. L. Hreniuk, M. V. Silva, V. A. Warren, A. D. Howard, L. H. Van Der Ploeg, et J. V. Heck. 2000. « Structure-Function Studies on the New Growth Hormone-Releasing Peptide, Ghrelin: Minimal Sequence of Ghrelin Necessary for Activation of Growth Hormone Secretagogue Receptor 1a ». *Journal of Medicinal Chemistry* 43 (23): 4370-76.

- Belmeguenai, A., H. Vaudry, J. Leprince, B. Vivet, F. Cavelier, J. Martinez, et E. Louiset. 2000. « Neurotensin Modulates the Electrical Activity of Frog Pituitary Melanotropes via Activation of a G-Protein-Coupled Receptor Pharmacologically Related to Both the NTS1 and Nts2 Receptors of Mammals ». *Neuroendocrinology* 72 (6): 379-91.
- Berendsen, H. J. C., D. van der Spoel, et R. van Drunen. 1995. « GROMACS: A message-passing parallel molecular dynamics implementation ». *Computer Physics Communications* 91 (1): 43-56.
- Beuming, Thijs, et Woody Sherman. 2012. « Current Assessment of Docking into GPCR Crystal Structures and Homology Models: Successes, Challenges, and Guidelines ». *Journal of Chemical Information and Modeling* 52 (12): 3263-77.
- Bockaert, J., et J. P. Pin. 1999. « Molecular Tinkering of G Protein-Coupled Receptors: An Evolutionary Success ». *The EMBO Journal* 18 (7): 1723-29.
- Bolivar, Juan H., Juan C. Muñoz-García, Tomas Castro-Dopico, Patricia M. Dijkman, Phillip J. Stansfeld, et Anthony Watts. 2016. « Interaction of Lipids with the Neurotensin Receptor 1 ». *Biochimica Et Biophysica Acta* 1858 (6): 1278-87.
- Boules, Mona, Hannah Johnston, Jessica Tozy, Kristin Smith, Zhimin Li, et Elliott Richelson. 2011. « Analgesic Synergy of Neurotensin Receptor Subtype 2 Agonist NT79 and Morphine ». *Behavioural Pharmacology* 22 (5-6): 573-81.
- Boules, Mona, Yanqi Liang, Siobhan Briody, Tomofumi Miura, Irfan Fauq, Alfredo Oliveros, Mina Wilson, et al. 2010. « NT79: a Novel Neurotensin Analog with Selective Behavioral Effects ». *Brain research* 1308 (janvier): 35-46.
- Bouvier, M., W. P. Hausdorff, A. De Blasi, B. F. O'Dowd, B. K. Kobilka, M. G. Caron, et R. J. Lefkowitz. 1988. « Removal of Phosphorylation Sites from the Beta 2-Adrenergic Receptor Delays Onset of Agonist-Promoted Desensitization ». *Nature* 333 (6171): 370-73.
- Bredeloux, Pierre, Florine Cavelier, Isabelle Dubuc, Bertrand Vivet, Jean Costentin, et Jean Martinez. 2008. « Synthesis and Biological Effects of c(Lys-Lys-Pro-Tyr-Ile-Leu-Lys-Lys-Pro-Tyr-Ile-Leu) (JMV2012), a New Analogue of Neurotensin That Crosses the Blood-Brain Barrier ». *Journal of Medicinal Chemistry* 51 (6): 1610-16.
- Brooks, B. R., C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, et al. 2009. « CHARMM: The Biomolecular Simulation Program ». *Journal of Computational Chemistry* 30 (10): 1545-1614.
- Brooks, Bernard R., Robert E. Bruccoleri, Barry D. Olafson, David J. States, S. Swaminathan, et Martin Karplus. 1983. « CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations ». *Journal of Computational Chemistry* 4 (2): 187-217.
- Bymaster, Frank P., Christian Felder, Saeed Ahmed, et David McKinzie. 2002. « Muscarinic Receptors as a Target for Drugs Treating Schizophrenia ». *Current Drug Targets. CNS and Neurological Disorders* 1 (2): 163-81.
- Carpenter, Byron, Rony Nehmé, Tony Warne, Andrew G. W. Leslie, et Christopher G. Tate. 2016. « Structure of the Adenosine A(2A) Receptor Bound to an Engineered G Protein ». *Nature* 536 (7614): 104-7.

- Carpenter, Byron, et Christopher G. Tate. 2016. « Engineering a Minimal G Protein to Facilitate Crystallisation of G Protein-Coupled Receptors in Their Active Conformation ». *Protein Engineering, Design & Selection: PEDS* 29 (12): 583-94.
- Carraway, R., et S. E. Leeman. 1973. « The Isolation of a New Hypotensive Peptide, Neurotensin, from Bovine Hypothalami ». *The Journal of Biological Chemistry* 248 (19): 6854-61.
- Casciari, Daniele, Daniele Dell'Orco, et Francesca Fanelli. 2008. « Homodimerization of Neurotensin 1 Receptor Involves Helices 1, 2, and 4: Insights from Quaternary Structure Predictions and Dimerization Free Energy Estimations ». *Journal of Chemical Information and Modeling* 48 (8): 1669-78.
- Casiraghi, Marina, Marjorie Damian, Ewen Lescop, Elodie Point, Karine Moncoq, Nelly Morellet, Daniel Levy, et al. 2016. « Functional Modulation of a G Protein-Coupled Receptor Conformational Landscape in a Lipid Bilayer ». *Journal of the American Chemical Society* 138 (35): 11170-75.
- Cauchy, Augustin-Louis. 1847. « Méthode générale pour la résolution des systèmes d'équations simultanées ». *Œuvres complètes* série 1 (tome 10): 399-402.
- Ceruso, Marc A., Xavier Periole, et Harel Weinstein. 2004. « Molecular Dynamics Simulations of Transducin: Interdomain and Front to Back Communication in Activation and Nucleotide Exchange ». *Journal of Molecular Biology* 338 (3): 469-81.
- Chandrashekar, J., K. L. Mueller, M. A. Hoon, E. Adler, L. Feng, W. Guo, C. S. Zuker, et N. J. Ryba. 2000. « T2Rs Function as Bitter Taste Receptors ». *Cell* 100 (6): 703-11.
- Charlier, Landry, Maxime Louet, Laurent Chaloin, Patrick Fuchs, Jean Martinez, Delphine Muriaux, Cyril Favard, et Nicolas Floquet. 2014. « Coarse-Grained Simulations of the HIV-1 Matrix Protein Anchoring: Revisiting Its Assembly on Membrane Domains ». *Biophysical Journal* 106 (3): 577-85.
- CHARMM-GUI. s. d. charmm-gui.org
- Chatterjee, Samit, Babak Behnam Azad, et Sridhar Nimmagadda. 2014. « The Intricate Role of CXCR4 in Cancer ». *Advances in cancer research* 124: 31-82.
- Cherezov, Vadim, Daniel M. Rosenbaum, Michael A. Hanson, Søren G. F. Rasmussen, Foon Sun Thian, Tong Sun Kobilka, Hee-Jung Choi, et al. 2007. « High-Resolution Crystal Structure of an Engineered Human Beta2-Adrenergic G Protein-Coupled Receptor ». *Science (New York, N.Y.)* 318 (5854): 1258-65.
- Cherfils, J., J. Ménétreay, G. Le Bras, I. Janoueix-Lerosey, J. de Gunzburg, J. R. Garel, et I. Auzat. 1997. « Crystal Structures of the Small G Protein Rap2A in Complex with Its Substrate GTP, with GDP and with GTPgammaS ». *The EMBO Journal* 16 (18): 5582-91.
- Cherfils, Jacqueline, et Marc Chabre. 2003. « Activation of G-protein G $\alpha$  subunits by receptors through G $\alpha$ -G $\beta$  and G $\alpha$ -G $\gamma$  interactions ». *Trends in Biochemical Sciences* 28 (1): 13-17.

- Cieplak, Piotr, James W. Caldwell, et Peter A. Kollman. 2001. « Molecular Mechanical Models for Organic and Biological Systems Going Beyond the Atom Centered Two Body Additive Approximation: Aqueous Solution Free Energies of Methanol and N-Methyl Acetamide, Nucleic Acid Base, and Amide Hydrogen Bonding and Chloroform/Water Partition Coefficients of the Nucleic Acid Bases ». *Journal of Computational Chemistry* 22 (juillet): 1048-57.
- Ciolek, Justyna, Helen Reinfrank, Loïc Quinton, Say Viengchareun, Enrico A. Stura, Laura Vera, Sabrina Sigismeu, et al. 2017. « Green Mamba Peptide Targets Type-2 Vasopressin Receptor against Polycystic Kidney Disease ». *Proceedings of the National Academy of Sciences of the United States of America* 114 (27): 7154-59.
- Civelli, Olivier, Yumiko Saito, Zhiwei Wang, Hans-Peter Nothacker, et Rainer K. Reinscheid. 2006. « Orphan GPCRs and Their Ligands ». *Pharmacology & Therapeutics* 110 (3): 525-32.
- Coleman, D. E., A. M. Berghuis, E. Lee, M. E. Linder, A. G. Gilman, et S. R. Sprang. 1994. « Structures of Active Conformations of Gi Alpha 1 and the Mechanism of GTP Hydrolysis ». *Science (New York, N.Y.)* 265 (5177): 1405-12.
- Comps-Agrar, Laëtitia, Julie Kniazeff, Lenea Nørskov-Lauritsen, Damien Maurel, Martin Gassmann, Nathalie Gregor, Laurent Prézeau, et al. 2011. « The oligomeric state sets GABAB receptor signalling efficacy ». *The EMBO Journal* 30 (12): 2336-49.
- Congreve, Miles, Stephen P. Andrews, Andrew S. Dore, Kaspar Hollenstein, Edward Hurrell, Christopher J. Langmead, Jonathan S. Mason, et al. 2012. « Discovery of 1,2,4-Triazine Derivatives as Adenosine A(2A) Antagonists Using Structure Based Drug Design. » *Journal of Medicinal Chemistry* 55 (5): 1898-1903.
- Cordomí, Arnau, Gemma Navarro, María S. Aymerich, et Rafael Franco. 2015. « Structures for G-Protein-Coupled Receptor Tetramers in Complex with G Proteins ». *Trends in Biochemical Sciences* 40 (10): 548-51.
- Costa, Mauricio G. S., Paulo R. Batista, Paulo M. Bisch, et David Perahia. 2015. « Exploring Free Energy Landscapes of Large Conformational Changes: Molecular Dynamics with Excited Normal Modes ». *Journal of Chemical Theory and Computation* 11 (6): 2755-67.
- Cowley, Michael A., Roy G. Smith, Sabrina Diano, Matthias Tschöp, Nina Pronchuk, Kevin L. Grove, Christian J. Strasburger, et al. 2003. « The Distribution and Mechanism of Action of Ghrelin in the CNS Demonstrates a Novel Hypothalamic Circuit Regulating Energy Homeostasis ». *Neuron* 37 (4): 649-61.
- Csépányi-Kömi, Roland, Magdolna Lévy, et Erzsébet Ligeti. 2012. « Small G Proteins and Their Regulators in Cellular Signalling ». *Molecular and Cellular Endocrinology* 353 (1-2): 10-20.
- Damian, Marjorie, Jacky Marie, Jean-Philippe Leyris, Jean-Alain Fehrentz, Pascal Verdié, Jean Martinez, Jean-Louis Banères, et Sophie Mary. 2012. « High Constitutive Activity Is an Intrinsic Feature of Ghrelin Receptor Protein ». *The Journal of Biological Chemistry* 287 (6): 3630-41.
- Damian, Marjorie, Sophie Mary, Mathieu Maingot, Céline M'Kadmi, Didier Gagne, Jean-Philippe Leyris, Séverine Denoyelle, et al. 2015. « Ghrelin Receptor Conformational Dynamics Regulate the Transition from a Preassembled to an Active Receptor:Gq Complex ». *Proceedings of the National Academy of Sciences* 112 (5): 1601-6.

- Damian, Marjorie, Véronique Pons, Pedro Renault, Céline M'Kadmi, Bartholomé Delort, Lucie Hartmann, Ali I. Kaya, et al. 2018. « GHSR-D2R Heteromerization Modulates Dopamine Signaling through an Effect on G Protein Conformation ». *Proceedings of the National Academy of Sciences of the United States of America* 115 (17): 4501-6.
- Daura, X., K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren, et A. E. Mark. 1999. « Peptide folding: When simulation meets experiment ». *Angewandte Chemie-International Edition in English* 38 (1-2): 236-40.
- Dawaliby, Rosie, Cataldo Trubbia, Cédric Delporte, Matthieu Masureel, Pierre Van Antwerpen, Brian K. Kobilka, et Cédric Govaerts. 2016. « Allosteric Regulation of G Protein-Coupled Receptor Activity by Phospholipids ». *Nature Chemical Biology* 12 (1): 35-39.
- De Lean, A., J. M. Stadel, et R. J. Lefkowitz. 1980. « A Ternary Complex Model Explains the Agonist-Specific Binding Properties of the Adenylate Cyclase-Coupled Beta-Adrenergic Receptor ». *The Journal of Biological Chemistry* 255 (15): 7108-17.
- Delort, Bartholomé, Pedro Renault, Landry Charlier, Florent Raussin, Jean Martinez, et Nicolas Floquet. 2017. « Coarse-Grained Prediction of Peptide Binding to G-Protein Coupled Receptors ». *Journal of Chemical Information and Modeling* 57 (3): 562-71.
- DeMarco, Steven J., Heiko Henze, Alexander Lederer, Kerstin Moehle, Reshmi Mukherjee, Barbara Romagnoli, John A. Robinson, et al. 2006. « Discovery of Novel, Highly Potent and Selective Beta-Hairpin Mimetic CXCR4 Inhibitors with Excellent Anti-HIV Activity and Pharmacokinetic Profiles ». *Bioorganic & Medicinal Chemistry* 14 (24): 8396-8404.
- Devesa, Jesús, Cristina Almengló, et Pablo Devesa. 2016. « Multiple Effects of Growth Hormone in the Body: Is it Really the Hormone for Growth? ». *Clinical Medicine Insights. Endocrinology and Diabetes* 9 (octobre): 47-71.
- DeVree, Brian T., Jacob P. Mahoney, Gisselle A. Vélez-Ruiz, Soren G. F. Rasmussen, Adam J. Kuszak, Elin Edwald, Juan-Jose Fung, et al. 2016. « Allosteric Coupling from G Protein to the Agonist-Binding Pocket in GPCRs ». *Nature* 535 (7610): 182-86.
- Digby, Gregory J, Pooja R Sethi, et Nevin A Lambert. 2008. « Differential dissociation of G protein heterotrimers ». *The Journal of Physiology* 586 (Pt 14): 3325-35.
- Dijksterhuis, J. P., J. Petersen, et G. Schulte. 2014. « WNT/Frizzled Signalling: Receptor-Ligand Selectivity with Focus on FZD-G Protein Signalling and Its Physiological Relevance: IUPHAR Review 3 ». *British Journal of Pharmacology* 171 (5): 1195-1209.
- Dixon, R. A., B. K. Kobilka, D. J. Strader, J. L. Benovic, H. G. Dohlman, T. Frielle, M. A. Bolanowski, et al. 1986. « Cloning of the Gene and cDNA for Mammalian Beta-Adrenergic Receptor and Homology with Rhodopsin ». *Nature* 321 (6065): 75-79.
- Dobner, Paul R. 2006. « Neurotensin and Pain Modulation ». *Peptides* 27 (10): 2405-14.
- Dore, Andrew S., Nathan Robertson, James C. Errey, Irene Ng, Kaspar Hollenstein, Ben Tehan, Edward Hurrell, et al. 2011. « Structure of the Adenosine A(2A) Receptor in Complex with ZM241385 and the Xanthines XAC and Caffeine. ». *Structure (London, England : 1993)* 19 (9): 1283-93.

- Dörr, Jonas M., Stefan Scheidelaar, Martijn C. Koorengel, Juan J. Dominguez, Marre Schäfer, Cornelis A. van Walree, et J. Antoinette Killian. 2016. « The styrene–maleic acid copolymer: a versatile tool in membrane research ». *European Biophysics Journal* 45: 3-21.
- Draper-Joyce, C. J., M. Khoshouei, D. M. Thal, Y. L. Liang, A. T. N. Nguyen, S. G. B. Furness, H. Venugopal, et al. 2018. « Structure of the Adenosine-Bound Human Adenosine A1receptor-Gicomplex. » *Nature* 558: 559-63.
- Dröctové, Laura. 2018. « Premières toxines Kunitz antagonistes du récepteur de type 2 à la vasopressine : étude pharmacodynamique et relations structure-activité ».
- Dror, Ron O., Daniel H. Arlow, Paul Maragakis, Thomas J. Mildorf, Albert C. Pan, Huafeng Xu, David W. Borhani, et David E. Shaw. 2011. « Activation Mechanism of the B2-Adrenergic Receptor ». *Proceedings of the National Academy of Sciences of the United States of America* 108 (46): 18684-89.
- Dror, Ron O., Robert M. Dirks, J. P. Grossman, Huafeng Xu, et David E. Shaw. 2012. « Biomolecular Simulation: A Computational Microscope for Molecular Biology ». *Annual Review of Biophysics* 41: 429-52.
- Dror, Ron O., Hillary F. Green, Celine Valant, David W. Borhani, James R. Valcourt, Albert C. Pan, Daniel H. Arlow, et al. 2013. « Structural Basis for Modulation of a G-Protein-Coupled Receptor by Allosteric Drugs ». *Nature* 503 (7475): 295-99.
- Dror, Ron O., Albert C. Pan, Daniel H. Arlow, David W. Borhani, Paul Maragakis, Yibing Shan, Huafeng Xu, et David E. Shaw. 2011. « Pathway and Mechanism of Drug Binding to G-Protein-Coupled Receptors ». *Proceedings of the National Academy of Sciences of the United States of America* 108 (32): 13118-23.
- Dudek, Michael J., et Jay W. Ponder. 1995. « Accurate Modeling of the Intramolecular Electrostatic Energy of Proteins ». *Journal of Computational Chemistry* 16 (7): 791-816.
- Duin, Adri C. T. van, Siddharth Dasgupta, Francois Lorant, et William A. Goddard. 2001. « ReaxFF: A Reactive Force Field for Hydrocarbons ». *The Journal of Physical Chemistry A* 105 (41): 9396-9409.
- Egloff, Pascal, Matthias Hillenbrand, Christoph Klenk, Alexander Batyuk, Philipp Heine, Stefanie Balada, Karola M. Schlinkmann, Daniel J. Scott, Marco Schütz, et Andreas Plückthun. 2014. « Structure of Signaling-Competent Neurotensin Receptor 1 Obtained by Directed Evolution in Escherichia Coli ». *Proceedings of the National Academy of Sciences of the United States of America* 111 (6): E655-662.
- Einsiedel, Jürgen, Cornelia Held, Maud Hervet, Manuel Plomer, Nuska Tschammer, Harald Hübner, et Peter Gmeiner. 2011. « Discovery of Highly Potent and Neurotensin Receptor 2 Selective Neurotensin Mimetics ». *Journal of Medicinal Chemistry* 54 (8): 2915-23.
- Engel, Stanislav, Amanda P. Skoumbourdis, John Childress, Susanne Neumann, Jeffrey R. Deschamps, Craig J. Thomas, Anny-Odile Colson, Stefano Costanzi, et Marvin C. Gershengorn. 2008. « A Virtual Screen for Diverse Ligands: Discovery of Selective G Protein-Coupled Receptor Antagonists ». *Journal of the American Chemical Society* 130 (15): 5115-23.

- Eps, Ned Van, Anita M. Preininger, Nathan Alexander, Ali I. Kaya, Scott Meier, Jens Meiler, Heidi E. Hamm, et Wayne L. Hubbell. 2011. « Interaction of a G Protein with an Activated Receptor Opens the Interdomain Interface in the Alpha Subunit ». *Proceedings of the National Academy of Sciences* 108 (23): 9420-24.
- Fanelli, Francesca, et Angelo Felling. 2011. « Dimerization and Ligand Binding Affect the Structure Network of A(2A) Adenosine Receptor ». *Biochimica Et Biophysica Acta* 1808 (5): 1256-66.
- Fanelli, Francesca, Michele Seeber, Angelo Felling, Daniele Casciari, et Francesco Raimondi. 2013. « Quaternary Structure Predictions and Structural Communication Features of GPCR Dimers ». *Progress in Molecular Biology and Translational Science* 117: 105-42.
- Fanelli, Roberto, Élie Besserer-Offroy, Adeline René, Jérôme Côté, Pascal Tétreault, Jasmin Collettere-Tremblay, Jean-Michel Longpré, et al. 2015. « Synthesis and Characterization in Vitro and in Vivo of (1)-(Trimethylsilyl)Alanine Containing Neurotensin Analogues ». *Journal of Medicinal Chemistry* 58 (19): 7785-95.
- Fanelli, Roberto, Nicolas Floquet, Élie Besserer-Offroy, Bartholomé Delort, Mélanie Vivancos, Jean-Michel Longpré, Pedro Renault, Jean Martinez, Philippe Sarret, et Florine Cavelier. 2017. « Use of Molecular Modeling to Design Selective NTS2 Neurotensin Analogues ». *Journal of Medicinal Chemistry* 60 (8): 3303-13.
- Fätkenheuer, Gerd, Anton L. Pozniak, Margaret A. Johnson, Andreas Plettenberg, Schlomo Staszewski, Andy I. M. Hoepelman, Michael S. Saag, et al. 2005. « Efficacy of Short-Term Monotherapy with Maraviroc, a New CCR5 Antagonist, in Patients Infected with HIV-1 ». *Nature Medicine* 11 (11): 1170-72.
- Felts, Anthony K., Yuichi Harano, Emilio Gallicchio, et Ronald M. Levy. 2004. « Free Energy Surfaces of Beta-Hairpin and Alpha-Helical Peptides Generated by Replica Exchange Molecular Dynamics with the AGBNP Implicit Solvent Model ». *Proteins* 56 (2): 310-21.
- Ferré, Sergi, Vicent Casadó, Lakshmi A. Devi, Marta Filizola, Ralf Jockers, Martin J. Lohse, Graeme Milligan, Jean-Philippe Pin, et Xavier Guitart. 2014. « G Protein-Coupled Receptor Oligomerization Revisited: Functional and Pharmacological Perspectives ». *Pharmacological Reviews* 66 (2): 413-34.
- Flock, Tilman, Alexander S. Hauser, Nadia Lund, David E. Gloriam, Santhanam Balaji, et M. Madan Babu. 2017. « Selectivity Determinants of GPCR-G-Protein Binding ». *Nature* 545 (7654): 317-22.
- Flock, Tilman, Charles N. J. Ravarani, Dawei Sun, A. J. Venkatakrishnan, Melis Kayikci, Christopher G. Tate, Dmitry B. Veprintsev, et M. Madan Babu. 2015. « Universal Allosteric Mechanism for G $\alpha$  Activation by GPCRs ». *Nature* 524 (7564): 173-79.
- Fredholm, Bertil B., Adriaan P. IJzerman, Kenneth A. Jacobson, Joel Linden, et Christa E. Müller. 2011. « International Union of Basic and Clinical Pharmacology. LXXXI. Nomenclature and Classification of Adenosine Receptors—An Update ». *Pharmacological Reviews* 63 (1): 1-34.
- Fredriksson, Robert, Malin C. Lagerström, Lars-Gustav Lundin, et Helgi B. Schiöth. 2003. « The G-Protein-Coupled Receptors in the Human Genome Form Five Main Families. Phylogenetic Analysis, Paralogue Groups, and Fingerprints ». *Molecular Pharmacology* 63 (6): 1256-72.

- Galandrin, Ségolène, et Michel Bouvier. 2006. « Distinct Signaling Profiles of Beta1 and Beta2 Adrenergic Receptor Ligands toward Adenylyl Cyclase and Mitogen-Activated Protein Kinase Reveals the Pluridimensionality of Efficacy ». *Molecular Pharmacology* 70 (5): 1575-84.
- Galés, Céline, R. Victor Rebois, Mireille Hogue, Phan Trieu, Andreas Breit, Terence E. Hébert, et Michel Bouvier. 2005. « Real-Time Monitoring of Receptor and G-Protein Interactions in Living Cells ». *Nature Methods* 2 (3): 177-84.
- García-Nafraía, Javier, Yang Lee, Xiaochen Bai, Byron Carpenter, et Christopher G. Tate. 2018. « Cryo-EM Structure of the Adenosine A2A Receptor Coupled to an Engineered Heterotrimeric G Protein ». *ELife* 7 (mai).
- Ge, Xuecai, Hong Yang, Maria A. Bednarek, Hadas Galon-Tilleman, Peirong Chen, Michael Chen, Joshua S. Lichtman, et al. 2018. « LEAP2 Is an Endogenous Antagonist of the Ghrelin Receptor ». *Cell Metabolism* 27 (2): 461-469.e6.
- Ghosh, Anirban, Uddhaveson Sonavane, et Rajendra Joshi. 2014. « Multiscale Modelling to Understand the Self-Assembly Mechanism of Human B2-Adrenergic Receptor in Lipid Bilayer ». *Computational Biology and Chemistry* 48 (février): 29-39.
- Gilman, A. G. 1987. « G Proteins: Transducers of Receptor-Generated Signals ». *Annual Review of Biochemistry* 56: 615-49.
- Gimpl, G., et F. Fahrenholz. 2001. « The Oxytocin Receptor System: Structure, Function, and Regulation ». *Physiological Reviews* 81 (2): 629-83.
- Globisch, Christoph, Venkatramanan Krishnamani, Markus Deserno, et Christine Peter. 2013. « Optimization of an Elastic Network Augmented Coarse Grained Model to Study CCMV Capsid Deformation ». *PloS One* 8 (4): e60582.
- González, Angel, Tomas Perez-Acle, Leonardo Pardo, et Xavier Deupi. 2011. « Molecular Basis of Ligand Dissociation in  $\beta$ -Adrenergic Receptors ». *PloS One* 6 (9): e23815.
- GPCRdb. s. d. [gpcrdb.org](http://gpcrdb.org)
- Grace, Christy R. R., Marilyn H. Perrin, Michael R. DiGrucchio, Charleen L. Miller, Jean E. Rivier, Wylie W. Vale, et Roland Riek. 2004. « NMR Structure and Peptide Hormone Binding Site of the First Extracellular Domain of a Type B1 G Protein-Coupled Receptor. ». *Proceedings of the National Academy of Sciences of the United States of America* 101 (35): 12836-41.
- Granier, C., J. van Rietschoten, P. Kitabgi, C. Poustis, et P. Freychet. 1982. « Synthesis and Characterization of Neurotensin Analogues for Structure/Activity Relationship Studies. Acetyl-Neurotensin-(8--13) Is the Shortest Analogue with Full Binding and Pharmacological Activities ». *European Journal of Biochemistry* 124 (1): 117-24.
- Großbauer, Jörg, Simone Kosol, Evelyne Schrank, et Klaus Zangger. 2010. « The peptide hormone ghrelin binds to membrane-mimetics via its octanoyl chain and an adjacent phenylalanine ». *Bioorganic & Medicinal Chemistry* 18 (15): 5483-88.
- Guitart, Xavier, Gemma Navarro, Estefania Moreno, Hideaki Yano, Ning-Sheng Cai, Marta Sánchez-Soto, Sandeep Kumar-Barodia, et al. 2014. « Functional Selectivity of Allosteric Interactions within G Protein-Coupled Receptor Oligomers: The Dopamine D1-D3 Receptor Heterotetramer ». *Molecular Pharmacology* 86 (4): 417-29.

- Gunsteren, Wilfred van, SR Billeter, AA Eising, Philippe Hünenberger, P Krüger, Alan Mark, WRP Scott, et Ilario Tironi. 1996. *Biomolecular Simulation: The {GROMOS96} manual and user guide*.
- Guo, Hui, Su An, Richard Ward, Yang Yang, Ying Liu, Xiao-Xi Guo, Qian Hao, et Tian-Rui Xu. 2017. « Methods used to study the oligomeric structure of G-protein-coupled receptors ». *Bioscience Reports* 37 (2).
- Gurevich, Eugenia V., et Vsevolod V. Gurevich. 2006. « Arrestins: Ubiquitous Regulators of Cellular Signaling Pathways ». *Genome Biology* 7 (9): 236.
- Gurevich, Vsevolod V., et Eugenia V. Gurevich. 2017. « Molecular Mechanisms of GPCR Signaling: A Structural Perspective ». *International Journal of Molecular Sciences* 18 (12).
- Haack, Karla K.V., et Nael A. McCarty. 2011. « Functional Consequences of GPCR Heterodimerization: GPCRs as Allosteric Modulators ». *Pharmaceuticals* 4 (3): 509-23.
- Handbook of Cell Signaling. 2003. Elsevier.
- Hapău, Denisa, Emmanuelle Rémond, Roberto Fanelli, Mélanie Vivancos, Adeline René, Jérôme Côté, Élie Besserer-Offroy, et al. 2016. « Stereoselective Synthesis of  $\beta$ -(5-Arylthiazolyl)  $\alpha$ -Amino Acids and Use in Neurotensin Analogues ». *European Journal of Organic Chemistry* 2016 (5): 1017-24.
- Hargrave, P. A., J. H. McDowell, D. R. Curtis, J. K. Wang, E. Juszczak, S. L. Fong, J. K. Rao, et P. Argos. 1983. « The Structure of Bovine Rhodopsin ». *Biophysics of Structure and Mechanism* 9 (4): 235-44.
- Harrison, C, et J. R Traynor. 2003. « The [35S]GTP $\gamma$ S binding assay: approaches and applications in pharmacology ». *Life Sciences* 74 (4): 489-508.
- Hauser, Alexander S., Misty M. Attwood, Mathias Rask-Andersen, Helgi B. Schiöth, et David E. Gloriam. 2017. « Trends in GPCR Drug Discovery: New Agents, Targets and Indications ». *Nature Reviews Drug Discovery* 16 (12): 829-42.
- Henriques, Sónia Troeira, Chia Chia Tan, David J. Craik, et Richard J. Clark. 2010. « Structural and Functional Analysis of Human Liver-Expressed Antimicrobial Peptide 2 ». *Chembiochem: A European Journal of Chemical Biology* 11 (15): 2148-57.
- Hino, Tomoya, Takatoshi Arakawa, Hiroko Iwanari, Takami Yurugi-Kobayashi, Chiyo Ikeda-Suno, Yoshiko Nakada-Nakura, Osamu Kusano-Arai, et al. 2012. « G-Protein-Coupled Receptor Inactivation by an Allosteric Inverse-Agonist Antibody. » *Nature* 482 (7384): 237-40.
- Holst, Birgitte, Thomas M. Frimurer, Jacek Mokrosinski, Tine Halkjaer, Karina B. Cullberg, Christina R. Underwood, et Thue W. Schwartz. 2009. « Overlapping Binding Site for the Endogenous Agonist, Small-Molecule Agonists, and Ago-Allosteric Modulators on the Ghrelin Receptor ». *Molecular Pharmacology* 75 (1): 44-59.
- Holst, Birgitte, Jacek Mokrosinski, Manja Lang, Erik Brandt, Rie Nygaard, Thomas M. Frimurer, Annette G. Beck-Sickinger, et Thue W. Schwartz. 2007. « Identification of an Efficacy Switch Region in the Ghrelin Receptor Responsible for Interchange between Agonism and Inverse Agonism ». *The Journal of Biological Chemistry* 282 (21): 15799-811.

- Hornak, Viktor, Shivani Ahuja, Markus Eilers, Joseph A. Goncalves, Mordechai Sheves, Philip J. Reeves, et Steven O. Smith. 2010. « Light activation of rhodopsin: insights from molecular dynamics simulations guided by solid-state NMR distance restraints ». *Journal of molecular biology* 396 (3): 510.
- Huang, Chia Ying, Vincent Olieric, Pikyee Ma, Nicole Howe, Lutz Vogeley, Xiangyu Liu, Rangana Warshamanage, et al. 2016. « In Meso in Situ Serial X-Ray Crystallography of Soluble and Membrane Proteins at Cryogenic Temperatures ». *Acta Crystallographica. Section D, Structural Biology* 72 (Pt 1): 93-112.
- Huang, Jianyun, Shuai Chen, J. Jillian Zhang, et Xin-Yun Huang. 2013. « Crystal Structure of Oligomeric Beta1-Adrenergic G Protein-Coupled Receptors in Ligand-Free Basal State. » *Nature Structural & Molecular Biology* 20 (4): 419-25.
- Huang, Sheng-You, et Xiaoqin Zou. 2010. « Advances and Challenges in Protein-Ligand Docking ». *International Journal of Molecular Sciences* 11 (8): 3016-34.
- Huang, Xi-Ping, Joel Karpiak, Wesley K. Kroeze, Hu Zhu, Xin Chen, Sheryl S. Moy, Kara A. Saddoris, et al. 2015. « Allosteric Ligands for the Pharmacologically Dark Receptors GPR68 and GPR65 ». *Nature* 527 (7579): 477-83.
- Humphrey, William, Andrew Dalke, et Klaus Schulten. 1996. « VMD: Visual molecular dynamics ». *Journal of Molecular Graphics* 14 (1): 33-38.
- Iiri, T., Z. Farfel, et H. R. Bourne. 1998. « G-Protein Diseases Furnish a Model for the Turn-on Switch ». *Nature* 394 (6688): 35-38.
- Inagaki, Sayaka, Rodolfo Ghirlando, Jim F. White, Jelena Gvozdenovic-Jeremic, John K. Northup, et Reinhard Grisshammer. 2012. « Modulation of the Interaction between Neurotensin Receptor NTS1 and Gq Protein by Lipid ». *Journal of Molecular Biology* 417 (1-2): 95-111.
- Javanainen, Matti, Hector Martinez-Seara, et Ilpo Vattulainen. 2017. « Excessive Aggregation of Membrane Proteins in the Martini Model ». *PLOS ONE* 12 (11): e0187936.
- Jiang, Wei, David J. Hardy, James C. Phillips, Alexander D. Mackerell, Klaus Schulten, et Benoît Roux. 2011. « High-Performance Scalable Molecular Dynamics Simulations of a Polarizable Force Field Based on Classical Drude Oscillators in NAMD ». *The Journal of Physical Chemistry Letters* 2 (2): 87-92.
- Jo, Sunhwan, Taehoon Kim, Vidyashankara G. Iyer, et Wonpil Im. 2008. « CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM ». *Journal of Computational Chemistry* 29 (11): 1859-65.
- Johnston, Jennifer M., Hao Wang, Davide Provasi, et Marta Filizola. 2012. « Assessing the Relative Stability of Dimer Interfaces in G Protein-Coupled Receptors ». *PLOS Computational Biology* 8 (8): e1002649.
- Jorgensen, W. L., et J. Tirado-Rives. 1988. « The OPLS [Optimized Potentials for Liquid Simulations] Potential Functions for Proteins, Energy Minimizations for Crystals of Cyclic Peptides and Crambin ». *Journal of the American Chemical Society* 110 (6): 1657-66.
- Julius, David, et Jeremy Nathans. 2012. « Signaling by Sensory Receptors ». *Cold Spring Harbor Perspectives in Biology* 4 (1).

- Juul, Kristian Vinter, Daniel Georges Bichet, Søren Nors Nielsen, et Jens Peter Nørgaard. 2014. « The physiological and pathophysiological functions of renal and extrarenal vasopressin V2 receptors. » undefined.
- Kaczor, A. A., et J. Selent. 2011. « Oligomerization of G Protein-Coupled Receptors: Biochemical and Biophysical Methods ». *Current Medicinal Chemistry* 18 (30): 4606-34.
- Kaczor, Agnieszka A., Manuela Jörg, et Ben Capuano. 2016. « The dopamine D2 receptor dimer and its interaction with homobivalent antagonists: homology modeling, docking and molecular dynamics ». *Journal of Molecular Modeling* 22 (9).
- Kang, Yanyong, X. Edward Zhou, Xiang Gao, Yuanzheng He, Wei Liu, Andrii Ishchenko, Anton Barty, et al. 2015. « Crystal Structure of Rhodopsin Bound to Arrestin by Femtosecond X-Ray Laser. » *Nature* 523 (7562): 561-67.
- Kappel, Kalli, Yinglong Miao, et J. Andrew McCammon. 2015. « Accelerated Molecular Dynamics Simulations of Ligand Binding to a Muscarinic G-protein Coupled Receptor ». *Quarterly reviews of biophysics* 48 (4): 479-87.
- Katritch, Vsevolod, Vadim Cherezov, et Raymond C. Stevens. 2012. « Diversity and Modularity of G Protein-Coupled Receptor Structures ». *Trends in Pharmacological Sciences* 33 (1): 17-27.
- Katritch, Vsevolod, Gustavo Fenalti, Enrique E. Abola, Bryan L. Roth, Vadim Cherezov, et Raymond C. Stevens. 2014. « Allosteric sodium in class A GPCR signaling ». *Trends in biochemical sciences* 39 (5): 233-44.
- Keane, Michael P., John A. Belperio, Ying Y. Xue, Marie D. Burdick, et Robert M. Strieter. 2004. « Depletion of CXCR2 Inhibits Tumor Growth and Angiogenesis in a Murine Model of Lung Cancer ». *Journal of Immunology (Baltimore, Md.: 1950)* 172 (5): 2853-60.
- Kendrick, Rachel. 2011. « The Structure and Function of the Human Ghrelin Receptor ». D\_ph, University of Birmingham.
- Kent, R. S., A. De Lean, et R. J. Lefkowitz. 1980. « A Quantitative Analysis of Beta-Adrenergic Receptor Interactions: Resolution of High and Low Affinity States of the Receptor by Computer Modeling of Ligand Binding Data ». *Molecular Pharmacology* 17 (1): 14-23.
- Kern, Andras, Rosie Albarran-Zeckler, Heidi E. Walsh, et Roy G. Smith. 2012. « Apo-Ghrelin Receptor Forms Heteromers with DRD2 in Hypothalamic Neurons and Is Essential for Anorexigenic Effects of DRD2 Agonism ». *Neuron* 73 (2): 317-32.
- Khorana, H G, G E Gerber, W C Herlihy, C P Gray, R J Andereg, K Nihei, et K Biemann. 1979. « Amino acid sequence of bacteriorhodopsin. » *Proceedings of the National Academy of Sciences of the United States of America* 76 (10): 5046-50.
- Kimple, Adam J., Meera Soundararajan, Stephanie Q. Hutsell, Annette K. Roos, Daniel J. Urban, Vincent Setola, Brenda R. S. Temple, et al. 2009. « Structural Determinants of G-Protein Alpha Subunit Selectivity by Regulator Of ». *The Journal of Biological Chemistry* 284 (29): 19402-11.
- Kleuss, C, A S Raw, E Lee, S R Sprang, et A G Gilman. 1994. « Mechanism of GTP hydrolysis by G-protein alpha subunits. » *Proceedings of the National Academy of Sciences of the United States of America* 91 (21): 9828-31.

- Koehl, Antoine, Hongli Hu, Shoji Maeda, Yan Zhang, Qianhui Qu, Joseph M. Paggi, Naomi R. Latorraca, et al. 2018. « Structure of the M-Opioid Receptor-Gi Protein Complex ». *Nature* 558 (7711): 547-52.
- Kofuku, Yutaka, Takumi Ueda, Junya Okude, Yutaro Shiraishi, Keita Kondo, Masahiro Maeda, Hideki Tsujishita, et Ichio Shimada. 2012. « Efficacy of the B<sub>2</sub>-Adrenergic Receptor Is Determined by Conformational Equilibrium in the Transmembrane Region ». *Nature Communications* 3: 1045.
- Kojima, M., H. Hosoda, Y. Date, M. Nakazato, H. Matsuo, et K. Kangawa. 1999. « Ghrelin Is a Growth-Hormone-Releasing Acylated Peptide from Stomach ». *Nature* 402 (6762): 656-60.
- Kolakowski, L. F. 1994. « GCRDb: A G-Protein-Coupled Receptor Database ». *Receptors & Channels* 2 (1): 1-7.
- Kolb, Peter, Daniel M. Rosenbaum, John J. Irwin, Juan José Fung, Brian K. Kobilka, et Brian K. Shoichet. 2009. « Structure-Based Discovery of Beta2-Adrenergic Receptor Ligands ». *Proceedings of the National Academy of Sciences of the United States of America* 106 (16): 6843-48.
- Komolov, Konstantin E., Anshul Bhardwaj, et Jeffrey L. Benovic. 2015. « Atomic Structure of GRK5 Reveals Distinct Structural Features Novel for G Protein-Coupled Receptor Kinases. » *The Journal of Biological Chemistry* 290 (34): 20629-47.
- Korbonits, Márta, Anthony P. Goldstone, Maria Gueorguiev, et Ashley B. Grossman. 2004. « Ghrelin--a Hormone with Multiple Functions ». *Frontiers in Neuroendocrinology* 25 (1): 27-68.
- Koshimizu, Taka-aki, Kazuaki Nakamura, Nobuaki Egashira, Masami Hiroyama, Hiroshi Nonoguchi, et Akito Tanoue. 2012. « Vasopressin V1a and V1b Receptors: From Molecules to Physiological Systems ». *Physiological Reviews* 92 (4): 1813-64.
- Kozasa, Tohru. 2004. « Purification of G Protein Subunits from Sf9 Insect Cells Using Hexahistidine-Tagged Alpha and Beta Gamma Subunits ». *Methods in Molecular Biology (Clifton, N.J.)* 237: 21-38.
- Krause, Alexander, Rannar Sillard, Burkhard Kleemeier, Enno Klüver, Erik Maronde, José Ramon Conejo-García, Wolf Georg Forssmann, et al. 2003. « Isolation and Biochemical Characterization of LEAP-2, a Novel Blood Peptide Expressed in the Liver ». *Protein Science: A Publication of the Protein Society* 12 (1): 143-52.
- Krupnick, J. G., et J. L. Benovic. 1998. « The Role of Receptor Kinases and Arrestins in G Protein-Coupled Receptor Regulation ». *Annual Review of Pharmacology and Toxicology* 38: 289-319.
- Kruse, Andrew C., Jianxin Hu, Albert C. Pan, Daniel H. Arlow, Daniel M. Rosenbaum, Erica Rosemond, Hillary F. Green, et al. 2012. « Structure and Dynamics of the M3 Muscarinic Acetylcholine Receptor ». *Nature* 482 (7386): 552-56.
- Kruse, Andrew C., Aaron M. Ring, Aashish Manglik, Jianxin Hu, Kelly Hu, Katrin Eitel, Harald Hübner, et al. 2013. « Activation and Allosteric Modulation of a Muscarinic Acetylcholine Receptor ». *Nature* 504 (7478): 101-6.
- Kumar, Shankar, Djamel Bouzida, Robert H. Swendsen, Peter A. Kollman, et John M. Rosenberg. 1992. « The Weighted Histogram Analysis Method for Free-energy Calculations on Biomolecules. I: The Method ». *J. Comput. Chem.* 13 (8): 1011–1021.

- Kunishima, N., Y. Shimada, Y. Tsuji, T. Sato, M. Yamamoto, T. Kumasaka, S. Nakanishi, H. Jingami, et K. Morikawa. 2000. « Structural Basis of Glutamate Recognition by a Dimeric Metabotropic Glutamate Receptor ». *Nature* 407 (6807): 971-77.
- Lagerström, Malin C., et Helgi B. Schiöth. 2008. « Structural Diversity of G Protein-Coupled Receptors and Significance for Drug Discovery ». *Nature Reviews. Drug Discovery* 7 (4): 339-57.
- Lambert, Nevin A., Christopher A. Johnston, Steven D. Cappell, Sudhakaranmayi Kuravi, Adam J. Kimple, Francis S. Willard, et David P. Siderovski. 2010. « Regulators of G-Protein Signaling Accelerate GPCR Signaling Kinetics and Govern Sensitivity Solely by Accelerating GTPase Activity ». *Proceedings of the National Academy of Sciences of the United States of America* 107 (15): 7066-71.
- Lambright, D. G., J. Sondek, A. Bohm, N. P. Skiba, H. E. Hamm, et P. B. Sigler. 1996. « The 2.0 Å Crystal Structure of a Heterotrimeric G Protein ». *Nature* 379 (6563): 311-19.
- Langley, J. N. 1905. « On the Reaction of Cells and of Nerve-Endings to Certain Poisons, Chiefly as Regards the Reaction of Striated Muscle to Nicotine and to Curari ». *The Journal of Physiology* 33 (4-5): 374-413.
- Latorraca, Naomi R., A. J. Venkatakrishnan, et Ron O. Dror. 2017. « GPCR Dynamics: Structures in Motion ». *Chemical Reviews* 117 (1): 139-55.
- Lebon, Guillaume, Tony Warne, Patricia C. Edwards, Kirstie Bennett, Christopher J. Langmead, Andrew G. W. Leslie, et Christopher G. Tate. 2011. « Agonist-Bound Adenosine A2A Receptor Structures Reveal Common Features of GPCR Activation. ». *Nature* 474 (7352): 521-25.
- Lee, Hwanky, et Ronald G. Larson. 2008. « Coarse-Grained Molecular Dynamics Studies of the Concentration and Size Dependence of Fifth- and Seventh-Generation PAMAM Dendrimers on Pore Formation in DMPC Bilayer ». *The Journal of Physical Chemistry B* 112 (26): 7778-84.
- Lefkowitz, Robert J. 2004. « Historical Review: A Brief History and Personal Retrospective of Seven-Transmembrane Receptors ». *Trends in Pharmacological Sciences* 25 (8): 413-22.
- Leioatts, Nicholas, Tod D. Romo, Shairy Azmy Danial, et Alan Grossfield. 2015. « Retinal Conformation Changes Rhodopsin's Dynamic Ensemble ». *Biophysical Journal* 109 (3): 608-17.
- Lemaître, Vincent, Philip Yeagle, et Anthony Watts. 2005. « Molecular Dynamics Simulations of Retinal in Rhodopsin: From the Dark-Adapted State towards Lumirhodopsin », *Biochemistry* 44 (45): 15082-15082.
- Leurs, Rob, Maria Sol Rodriguez Pena, Remko A. Bakker, Astrid E. Alewijnse, et Henk Timmerman. 2000. « Constitutive activity of G protein coupled receptors and drug action ». In *Pharmacology Library*, édité par Ugo Gulini, Mario Gianella, Wilma Quaglia, et Gabriella Marucci, 31:327-31. Receptor Chemistry towards the Third Millennium. Elsevier.
- Levitt, Michael, et Arieh Warshel. 1975. « Computer Simulation of Protein Folding ». *Nature* 253 (5494): 694-98.
- Li, Juan, Yuchen Tian, et Aiguo Wu. 2015. « Neuropeptide Y receptors: a promising target for cancer imaging and therapy ». *Regenerative Biomaterials* 2 (3): 215-19.

- Liebscher, Ines, Torsten Schöneberg, et Simone Prömel. 2013. « Progress in Demystification of Adhesion G Protein-Coupled Receptors ». *Biological Chemistry* 394 (8): 937-50.
- Lifson, S., et A. Warshel. 1968. « Consistent Force Field for Calculations of Conformations, Vibrational Spectra, and Enthalpies of Cycloalkane and n-Alkane Molecules ». *The Journal of Chemical Physics* 49 (11): 5116-29.
- Limbird, Lee E. 1996. *Cell Surface Receptors: A Short Course on Theory and Methods: A Short Course on Theory and Methods*. 2<sup>e</sup> éd. Springer US.
- Lindahl, Erik, Berk Hess, et David van der Spoel. 2001. « GROMACS 3.0: A Package for Molecular Simulation and Trajectory Analysis ». *Molecular Modeling Annual* 7 (8): 306-17.
- Liu, Wenyuan, Chao Wang, Yanbin Li, Yuyang Lao, Yongjian Han, Guang-Can Guo, Yong-Hua Zhao, et Lixin He. 2015. « Replica Exchange Molecular Dynamics Optimization of Tensor Network States for Quantum Many-Body Systems ». *Journal of Physics: Condensed Matter* 27 (8): 085601.
- Logez, Christel, Marjorie Damian, Céline Legros, Clémence Dupré, Mélody Guéry, Sophie Mary, Renaud Wagner, et al. 2016. « Detergent-free Isolation of Functional G Protein-Coupled Receptors into Nanometric Lipid Particles ». *Biochemistry* 55 (1): 38-48.
- Lohse, M. J., J. L. Benovic, J. Codina, M. G. Caron, et R. J. Lefkowitz. 1990. « Beta-Arrestin: A Protein That Regulates Beta-Adrenergic Receptor Function ». *Science (New York, N.Y.)* 248 (4962): 1547-50.
- Lomasney, J. W., L. M. Leeb-Lundberg, S. Cotecchia, J. W. Regan, J. F. DeBernardis, M. G. Caron, et R. J. Lefkowitz. 1986. « Mammalian Alpha 1-Adrenergic Receptor. Purification and Characterization of the Native Receptor Ligand Binding Subunit ». *The Journal of Biological Chemistry* 261 (17): 7710-16.
- Loo, Jason S. E., Abigail L. Emtage, Kar Weng Ng, Alene S. J. Yong, et Stephen W. Doughty. 2018. « Assessing GPCR homology models constructed from templates of various transmembrane sequence identities: Binding mode prediction and docking enrichment ». *Journal of Molecular Graphics and Modelling* 80 (mars): 38-47.
- Louet, Maxime. 2012. « Étude à l'échelle moléculaire des protéines-G couplées à leurs récepteurs. » Thesis, Montpellier 1.
- Louet, Maxime, Landry Charlier, Jean Martinez, et Nicolas Floquet. 2012. « Dissociation of Membrane-Anchored Heterotrimeric G-Protein Induced by G( $\alpha$ ) Subunit Binding to GTP ». *Journal of Chemical Information and Modeling* 52 (11): 3022-27.
- Louet, Maxime, Esra Karakas, Alexandre Perret, David Perahia, Jean Martinez, et Nicolas Floquet. 2013. « Conformational Restriction of G-Proteins Coupled Receptors (GPCRs) upon Complexation to G-Proteins: A Putative Activation Mode of GPCRs? » *FEBS Letters* 587 (16): 2656-61.
- Louet, Maxime, Jean Martinez, et Nicolas Floquet. 2012. « GDP Release Preferentially Occurs on the Phosphate Side in Heterotrimeric G-Proteins ». *PLoS Computational Biology* 8 (7): e1002595.
- Louet, Maxime, David Perahia, Jean Martinez, et Nicolas Floquet. 2011. « A Concerted Mechanism for Opening the GDP Binding Pocket and Release of the Nucleotide in Hetero-Trimeric G-Proteins ». *Journal of Molecular Biology* 411 (1): 298-312.

- Luchetti, Giovanni, Ria Sircar, Jennifer H. Kong, Sigrid Nachtergaele, Andreas Sagner, Eamon Fx Byrne, Douglas F. Covey, Christian Siebold, et Rajat Rohatgi. 2016. « Cholesterol Activates the G-Protein Coupled Receptor Smoothed to Promote Hedgehog Signaling ». *ELife* 5.
- Luitz, Manuel P., et Martin Zacharias. 2014. « Protein-Ligand Docking Using Hamiltonian Replica Exchange Simulations with Soft Core Potentials ». *Journal of Chemical Information and Modeling* 54 (6): 1669-75.
- Maegley, K A, S J Admiraal, et D Herschlag. 1996. « Ras-catalyzed hydrolysis of GTP: a new perspective from model studies. » *Proceedings of the National Academy of Sciences of the United States of America* 93 (16): 8160-66.
- Malnic, B., J. Hirono, T. Sato, et L. B. Buck. 1999. « Combinatorial Receptor Codes for Odors ». *Cell* 96 (5): 713-23.
- Man, Orna, Yoav Gilad, et Doron Lancet. 2004. « Prediction of the Odorant Binding Site of Olfactory Receptor Proteins by Human-Mouse Comparisons ». *Protein Science: A Publication of the Protein Society* 13 (1): 240-54.
- Manglik, Aashish, et Brian Kobilka. 2014. « The Role of Protein Dynamics in GPCR Function: Insights from the B2AR and Rhodopsin ». *Current Opinion in Cell Biology* 27 (avril): 136-43.
- Manglik, Aashish, Andrew C. Kruse, Tong Sun Kobilka, Foon Sun Thian, Jesper M. Mathiesen, Roger K. Sunahara, Leonardo Pardo, William I. Weis, Brian K. Kobilka, et Sébastien Granier. 2012. « Crystal Structure of the M-Opioid Receptor Bound to a Morphinan Antagonist ». *Nature* 485 (7398): 321-26.
- Manna, Moutusi, Miia Niemelä, Joona Tynkkynen, Matti Javanainen, Waldemar Kulig, Daniel J. Müller, Tomasz Rog, et Ilpo Vattulainen. 2016. « Mechanism of Allosteric Regulation of B2-Adrenergic Receptor by Cholesterol ». *ELife*.
- Marrink, Siewert J., H. Jelger Risselada, Serge Yefimov, D. Peter Tieleman, et Alex H. de Vries. 2007. « The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations ». *The Journal of Physical Chemistry. B* 111 (27): 7812-24.
- Marrink, Siewert J., Alex H. de Vries, et Alan E. Mark. 2004. « Coarse Grained Model for Semiquantitative Lipid Simulations ». *The Journal of Physical Chemistry B* 108 (2): 750-60.
- Marsango, Sara, Gianluigi Caltabiano, Chantevy Pou, María José Varela Liste, et Graeme Milligan. 2015. « Analysis of Human Dopamine D3 Receptor Quaternary Structure ». *The Journal of Biological Chemistry* 290 (24): 15146-62.
- Maupetit, Julien, P. Tuffery, et Philippe Derreumaux. 2007. « A Coarse-Grained Protein Force Field for Folding and Structure Prediction ». *Proteins* 69 (2): 394-408.
- Meltzer, H. Y., S. Matsubara, et J. C. Lee. 1989. « Classification of Typical and Atypical Antipsychotic Drugs on the Basis of Dopamine D-1, D-2 and Serotonin2 PKi Values ». *The Journal of Pharmacology and Experimental Therapeutics* 251 (1): 238-46.
- Mendonça, Alexandre de, et Joachim A. Ribeiro. 2001. « Adenosine and Synaptic Plasticity ». *Drug Development Research* 52 (1-2): 283-90.
- Meng, Xuan-Yu, Mihaly Mezei, et Meng Cui. 2014. « Computational Approaches for Modeling GPCR Dimerization ». *Current pharmaceutical biotechnology* 15 (10): 996-1006.

- Meng, Xuan-Yu, Hong-Xing Zhang, Mihaly Mezei, et Meng Cui. 2011. « Molecular Docking: A powerful approach for structure-based drug discovery ». *Current computer-aided drug design* 7 (2): 146-57.
- Meral, Derya, Davide Provasi, Diego Prada-Gracia, Jan Möller, Kristen Marino, Martin J. Lohse, et Marta Filizola. 2018. « Molecular Details of Dimerization Kinetics Reveal Negligible Populations of Transient M-Opioid Receptor Homodimers at Physiological Concentrations ». *Scientific Reports* 8 (1): 7705.
- Mesnier, Danielle, et Jean-Louis Banères. 2004. « Cooperative Conformational Changes in a G-Protein-Coupled Receptor Dimer, the Leukotriene B(4) Receptor BLT1 ». *The Journal of Biological Chemistry* 279 (48): 49664-70.
- Miao, Yinglong, Alisha D. Caliman, et J. Andrew McCammon. 2015. « Allosteric Effects of Sodium Ion Binding on Activation of the M3 Muscarinic G-Protein-Coupled Receptor ». *Biophysical Journal* 108 (7): 1796-1806.
- Miao, Yinglong, Dahlia Anne Goldfeld, Ee Von Moo, Patrick M. Sexton, Arthur Christopoulos, J. Andrew McCammon, et Celine Valant. 2016. « Accelerated Structure-Based Design of Chemically Diverse Allosteric Modulators of a Muscarinic G Protein-Coupled Receptor ». *Proceedings of the National Academy of Sciences of the United States of America* 113 (38): E5675-5684.
- Miao, Yinglong, Sara E. Nichols, Paul M. Gasper, Vincent T. Metzger, et J. Andrew McCammon. 2013. « Activation and Dynamic Network of the M2 Muscarinic Receptor ». *Proceedings of the National Academy of Sciences of the United States of America* 110 (27): 10982-87.
- Miller, R. Lance, Pablo C. Sandoval, Trairak Pisitkun, Mark A. Knepper, et Jason D. Hoffert. 2013. « Vasopressin Inhibits Apoptosis in Renal Collecting Duct Cells ». *American Journal of Physiology. Renal Physiology* 304 (2): F177-188.
- Milligan, Graeme, et Evi Kostenis. 2006. « Heterotrimeric G-Proteins: A Short History ». *British Journal of Pharmacology* 147 Suppl 1 (janvier): S46-55.
- Missale, C., S. R. Nash, S. W. Robinson, M. Jaber, et M. G. Caron. 1998. « Dopamine Receptors: From Structure to Function ». *Physiological Reviews* 78 (1): 189-225.
- Mokrosiński, Jacek, Thomas M. Frimurer, Bjørn Sivertsen, Thue W. Schwartz, et Birgitte Holst. 2012. « Modulation of Constitutive Activity and Signaling Bias of the Ghrelin Receptor by Conformational Constraint in the Second Extracellular Loop ». *The Journal of Biological Chemistry* 287 (40): 33488-502.
- Monticelli, Luca, Senthil K. Kandasamy, Xavier Periolo, Ronald G. Larson, D. Peter Tieleman, et Siewert-Jan Marrink. 2008. « The MARTINI Coarse-Grained Force Field: Extension to Proteins ». *Journal of Chemical Theory and Computation* 4 (5): 819-34.
- Murakami, Midori, et Tsutomu Kouyama. 2008. « Crystal Structure of Squid Rhodopsin. » *Nature* 453 (7193): 363-67.
- Murone, M., A. Rosenthal, et F. J. de Sauvage. 1999. « Sonic Hedgehog Signaling by the Patched-Smoothed Receptor Complex ». *Current Biology: CB* 9 (2): 76-84.
- Nakata, H., K. Kameyama, K. Haga, et T. Haga. 1994. « Location of Agonist-Dependent-Phosphorylation Sites in the Third Intracellular Loop of Muscarinic Acetylcholine Receptors (M2 Subtype) ». *European Journal of Biochemistry* 220 (1): 29-36.

- Navarro, Gemma, Arnau Cordoní, Monika Zelman-Femiak, Marc Brugarolas, Estefania Moreno, David Aguinaga, Laura Perez-Benito, et al. 2016. « Quaternary Structure of a G-Protein-Coupled Receptor Heterotetramer in Complex with Gi and Gs ». *BMC Biology* 14 (avril): 26.
- Nehmé, Rony, Byron Carpenter, Ankita Singhal, Annette Strege, Patricia C. Edwards, Courtney F. White, Haijuan Du, Reinhard Grisshammer, et Christopher G. Tate. 2017. « Mini-G Proteins: Novel Tools for Studying GPCRs in Their Active Conformation ». *PLoS One* 12 (4): e0175642.
- Nelson, G., M. A. Hoon, J. Chandrashekar, Y. Zhang, N. J. Ryba, et C. S. Zuker. 2001. « Mammalian Sweet Taste Receptors ». *Cell* 106 (3): 381-90.
- Nelson, Greg, Jayaram Chandrashekar, Mark A. Hoon, Luxin Feng, Grace Zhao, Nicholas J. P. Ryba, et Charles S. Zuker. 2002. « An Amino-Acid Taste Receptor ». *Nature* 416 (6877): 199-202.
- Nguyen, Anh T. N., Jo-Anne Baltos, Trayder Thomas, Toan D. Nguyen, Laura López Muñoz, Karen J. Gregory, Paul J. White, Patrick M. Sexton, Arthur Christopoulos, et Lauren T. May. 2016. « Extracellular Loop 2 of the Adenosine A1 Receptor Has a Key Role in Orthosteric Ligand Affinity and Agonist Efficacy ». *Molecular Pharmacology* 90 (6): 703-14.
- Nishimura, Akiyuki, Ken Kitano, Jun Takasaki, Masatoshi Taniguchi, Norikazu Mizuno, Kenji Tago, Toshio Hakoshima, et Hiroshi Itoh. 2010. « Structural Basis for the Specific Inhibition of Heterotrimeric Gq Protein by a Small Molecule ». *Proceedings of the National Academy of Sciences of the United States of America* 107 (31): 13666-71.
- Noda, K., Y. Saad, R. M. Graham, et S. S. Karnik. 1994. « The High Affinity State of the Beta 2-Adrenergic Receptor Requires Unique Interaction between Conserved and Non-Conserved Extracellular Loop Cysteines ». *The Journal of Biological Chemistry* 269 (9): 6743-52.
- Nordström, Karl J. V., Markus Sällman Almén, Monika M. Edstam, Robert Fredriksson, et Helgi B. Schiöth. 2011. « Independent HHsearch, Needleman--Wunsch-Based, and Motif Analyses Reveal the Overall Hierarchy for Most of the G Protein-Coupled Receptor Families ». *Molecular Biology and Evolution* 28 (9): 2471-80.
- O'Connor, Casey, Kate L. White, Nathalie Doncescu, Tatiana Didenko, Bryan L. Roth, Georges Czaplicki, Raymond C. Stevens, Kurt Wüthrich, et Alain Milon. 2015. « NMR Structure and Dynamics of the Agonist Dynorphin Peptide Bound to the Human Kappa Opioid Receptor ». *Proceedings of the National Academy of Sciences of the United States of America* 112 (38): 11852-57.
- O'Hara, P. J., P. O. Sheppard, H. Thøgersen, D. Venezia, B. A. Haldeman, V. McGrane, K. M. Houamed, C. Thomsen, T. L. Gilbert, et E. R. Mulvihill. 1993. « The Ligand-Binding Domain in Metabotropic Glutamate Receptors Is Related to Bacterial Periplasmic Binding Proteins ». *Neuron* 11 (1): 41-52.
- Okude, Junya, Takumi Ueda, Yutaka Kofuku, Motohiko Sato, Naoyuki Nobuyama, Keita Kondo, Yutaro Shiraishi, et al. 2015. « Identification of a Conformational Equilibrium That Determines the Efficacy and Functional Selectivity of the  $\mu$ -Opioid Receptor ». *Angewandte Chemie (International Ed. in English)* 54 (52): 15771-76.
- Oldham, William M., et Heidi E. Hamm. 2008. « Heterotrimeric G Protein Activation by G-Protein-Coupled Receptors ». *Nature Reviews. Molecular Cell Biology* 9 (1): 60-71.

- Olivella, Mireia, Angel Gonzalez, Leonardo Pardo, et Xavier Deupi. 2013. « Relation between Sequence and Structure in Membrane Proteins ». *Bioinformatics (Oxford, England)* 29 (13): 1589-92.
- Oostenbrink, Chris, Alessandra Villa, Alan E. Mark, et Wilfred F. van Gunsteren. 2004. « A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: The GROMOS Force-Field Parameter Sets 53A5 and 53A6 ». *Journal of Computational Chemistry* 25 (13): 1656-76.
- Osadchii, Oleg E. 2015. « Emerging Role of Neurotensin in Regulation of the Cardiovascular System ». *European Journal of Pharmacology* 762 (septembre): 184-92.
- Palczewski, K., T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, I. Le Trong, et al. 2000. « Crystal Structure of Rhodopsin: A G Protein-Coupled Receptor ». *Science (New York, N.Y.)* 289 (5480): 739-45.
- Páll, Szilárd, Mark James Abraham, Carsten Kutzner, Berk Hess, et Erik Lindahl. 2014. « Tackling Exascale Software Challenges in Molecular Dynamics Simulations with GROMACS ». In *Solving Software Challenges for Exascale*, 3-27. Lecture Notes in Computer Science. Springer, Cham.
- Park, Jung Hee, Patrick Scheerer, Klaus Peter Hofmann, Hui-Woog Choe, et Oliver Peter Ernst. 2008. « Crystal Structure of the Ligand-Free G-Protein-Coupled Receptor Opsin. » *Nature* 454 (7201): 183-87.
- Patel, Sandeep, et Charles L. Brooks III. 2006. « Fluctuating charge force fields: recent developments and applications from small molecules to macromolecular biological systems ». *Molecular Simulation* 32 (3-4): 231-49.
- Patel, Sunita, Elizabeth Vierling, et Florence Tama. 2014. « Replica Exchange Molecular Dynamics Simulations Provide Insight into Substrate Recognition by Small Heat Shock Proteins ». *Biophysical Journal* 106 (12): 2644-55.
- Periole, Xavier, Marco Cavalli, Siewert-Jan Marrink, et Marco A. Ceruso. 2009. « Combining an Elastic Network With a Coarse-Grained Molecular Force Field: Structure, Dynamics, and Intermolecular Recognition ». *Journal of Chemical Theory and Computation* 5 (9): 2531-43.
- Periole, Xavier, Thomas Huber, Siewert-Jan Marrink, et Thomas P. Sakmar. 2007a. « G Protein-Coupled Receptors Self-Assemble in Dynamics Simulations of Model Bilayers ». *Journal of the American Chemical Society* 129 (33): 10126-32.
- Pert, C. B., et S. H. Snyder. 1973. « Opiate Receptor: Demonstration in Nervous Tissue ». *Science (New York, N.Y.)* 179 (4077): 1011-14.
- Pitcher, J., M. J. Lohse, J. Codina, M. G. Caron, et R. J. Lefkowitz. 1992. « Desensitization of the Isolated Beta 2-Adrenergic Receptor by Beta-Adrenergic Receptor Kinase, CAMP-Dependent Protein Kinase, and Protein Kinase C Occurs via Distinct Molecular Mechanisms ». *Biochemistry* 31 (12): 3193-97.
- Pitera, Jed W., et William Swope. 2003. « Understanding Folding and Design: Replica-Exchange Simulations of ``Trp-Cage`` Miniproteins ». *Proceedings of the National Academy of Sciences* 100 (13): 7587-92.

- Pluhackova, Kristyna, Stefan Gahbauer, Franziska Kranz, Tsjerk A. Wassenaar, et Rainer A. Böckmann. 2016. « Dynamic Cholesterol-Conditioned Dimerization of the G Protein Coupled Chemokine Receptor Type 4 ». *PLOS Computational Biology* 12 (11): e1005169.
- Prada-Gracia, Diego, Kristen Marino, Davide Provasi, et Marta Filizola. 2017. « Kinetics of G Protein-Coupled Receptor Dimerization from Markov State Model Analysis of Coarse-Grained Simulations ». *Biophysical Journal* 112 (3): 27a.
- Prasanna, Xavier, Durba Sengupta, et Amitabha Chattopadhyay. 2016. « Cholesterol-Dependent Conformational Plasticity in GPCR Dimers ». *Scientific Reports* 6 (août): 31858.
- Preininger, Anita M., Jens Meiler, et Heidi Hamm. 2013. « Conformational flexibility and structural dynamics in GPCR-mediated G protein activation: a perspective ». *Journal of molecular biology* 425 (13): 2288-98.
- Provasi, Davide, Marta Camacho Artacho, Ana Negri, Juan Carlos Mobarec, et Marta Filizola. 2011. « Ligand-Induced Modulation of the Free-Energy Landscape of G Protein-Coupled Receptors Explored by Adaptive Biasing Techniques ». *PLOS Computational Biology* 7 (10): e1002193.
- Provasi, Davide, Andrea Bortolato, et Marta Filizola. 2009. « Exploring Molecular Mechanisms of Ligand Recognition by Opioid Receptors with Metadynamics ». *Biochemistry* 48 (42): 10020-29.
- Provasi, Davide, et Marta Filizola. 2010. « Putative Active States of a Prototypic G-Protein-Coupled Receptor from Biased Molecular Dynamics ». *Biophysical Journal* 98 (10): 2347-55.
- Qi, Yifei, Helgi I. Ingólfsson, Xi Cheng, Jumin Lee, Siewert J. Marrink, et Wonpil Im. 2015. « CHARMM-GUI Martini Maker for Coarse-Grained Simulations with the Martini Force Field ». *Journal of Chemical Theory and Computation* 11 (9): 4486-94.
- Qin, Kou, Chunmin Dong, Guangyu Wu, et Nevin A. Lambert. 2011. « Inactive-state preassembly of Gq-coupled receptors and Gq heterotrimers ». *Nature chemical biology* 7 (10): 740-47.
- R Development Core Team (2008). s. d. R. R Foundation for Statistical Computing, Vienna, Austria.
- Raimondi, Francesco, Michele Seeber, Pier G. De Benedetti, et Francesca Fanelli. 2008. « Mechanisms of Inter- and Intramolecular Communication in GPCRs and G Proteins ». *Journal of the American Chemical Society* 130 (13): 4310-25.
- Ramírez-Anguita, Juan Manuel, Ismael Rodríguez-Espigares, Ramon Guixà-González, Agostino Bruno, Mariona Torrens-Fontanals, Alejandro Varela-Rial, et Jana Selent. 2018. « Membrane Cholesterol Effect on the 5-HT<sub>2A</sub> Receptor: Insights into the Lipid-Induced Modulation of an Antipsychotic Drug Target ». *Biotechnology and Applied Biochemistry* 65 (1): 29-37.
- Rasmussen, Søren G. F., Hee-Jung Choi, Daniel M. Rosenbaum, Tong Sun Kobilka, Foon Sun Thian, Patricia C. Edwards, Manfred Burghammer, et al. 2007. « Crystal Structure of the Human Beta<sub>2</sub> Adrenergic G-Protein-Coupled Receptor ». *Nature* 450 (7168): 383-87.

- Rasmussen, Søren G. F., Brian T. DeVree, Yaozhong Zou, Andrew C. Kruse, Ka Young Chung, Tong Sun Kobilka, Foon Sun Thian, et al. 2011. « Crystal Structure of the B2 Adrenergic Receptor-Gs Protein Complex ». *Nature* 477 (7366): 549-55.
- Rathore, Nitin, Manan Chopra, et Juan J. de Pablo. 2004. « Optimal allocation of replicas in parallel tempering simulations ». *The Journal of Chemical Physics* 122 (2): 024111.
- Regan, J. W., H. Nakata, R. M. DeMarinis, M. G. Caron, et R. J. Lefkowitz. 1986. « Purification and Characterization of the Human Platelet Alpha 2-Adrenergic Receptor ». *The Journal of Biological Chemistry* 261 (8): 3894-3900.
- Rémond, Emmanuelle, Charlotte Martin, Jean Martinez, et Florine Cavelier. 2016. « Silicon-Containing Amino Acids: Synthetic Aspects, Conformational Studies, and Applications to Bioactive Peptides ». *Chemical Reviews* 116 (19): 11654-84.
- Renault, Pedro, Maxime Louet, Bartholomé Delort, Jacky Marie, Gilles Labesse, et Nicolas Floquet. 2018. « Molecular Dynamics Simulations of the Allosteric Modulation of the Adenosine A2A Receptor by a Mini-G Protein », 2018.
- Rhee, Young Min, et Vijay S. Pande. 2003. « Multiplexed-Replica Exchange Molecular Dynamics Method for Protein Folding Simulation ». *Biophysical Journal* 84 (2 Pt 1): 775-86.
- Ring, Aaron M., Aashish Manglik, Andrew C. Kruse, Michael D. Enos, William I. Weis, K. Christopher Garcia, et Brian K. Kobilka. 2013. « Adrenaline-Activated Structure of Beta2-Adrenoceptor Stabilized by an Engineered Nanobody. » *Nature* 502 (7472): 575-79.
- Robison, G. A., R. W. Butcher, et E. W. Sutherland. 1967. « Adenyl Cyclase as an Adrenergic Receptor ». *Annals of the New York Academy of Sciences* 139 (3): 703-23.
- Rodbell, M., L. Birnbaumer, S. L. Pohl, et H. M. Krans. 1971. « The Glucagon-Sensitive Adenyl Cyclase System in Plasma Membranes of Rat Liver. V. An Obligatory Role of Guanylnucleotides in Glucagon Action ». *The Journal of Biological Chemistry* 246 (6): 1877-82.
- Roth, Bryan L. 2007. « Drugs and Valvular Heart Disease ». *The New England Journal of Medicine* 356 (1): 6-9.
- Rudack, Till, Fei Xia, Jürgen Schlitter, Carsten Köttig, et Klaus Gerwert. 2012. « The Role of Magnesium for Geometry and Charge in GTP Hydrolysis, Revealed by Quantum Mechanics/Molecular Mechanics Simulations ». *Biophysical Journal* 103 (2): 293-302.
- Sabio, Michael, Kenneth Jones, et Sid Topiol. 2008. « Use of the X-Ray Structure of the Beta2-Adrenergic Receptor for Drug Discovery. Part 2: Identification of Active Compounds ». *Bioorganic & Medicinal Chemistry Letters* 18 (20): 5391-95.
- Salahpour, Ali, et Marc G. Caron. 2012. « Food for Thought: The Physiological Relevance of Ghrelin and Dopamine D2 Receptor Heterodimerization in the Regulation of Appetite ». *Neuron* 73 (2): 210-11.
- Saleh, Nouredin, Passainte Ibrahim, Giorgio Saladino, Francesco Luigi Gervasio, et Timothy Clark. 2017. « An Efficient Metadynamics-Based Protocol To Model the Binding Affinity and the Transition State Ensemble of G-Protein-Coupled Receptor Ligands ». *Journal of Chemical Information and Modeling* 57 (5): 1210-17.

- Saleh, Noureldin, Giorgio Saladino, Francesco Luigi Gervasio, et Timothy Clark. 2017. « Investigating Allosteric Effects on the Functional Dynamics of B2-Adrenergic Ternary Complexes with Enhanced-Sampling Simulations ». *Chemical Science* 8 (5): 4019-26.
- Sali, A., et T. L. Blundell. 1993. « Comparative Protein Modelling by Satisfaction of Spatial Restraints ». *Journal of Molecular Biology* 234 (3): 779-815.
- Sarret, Philippe, Michael J. Esdaile, Amélie Perron, Jean Martinez, Thomas Stroh, et Alain Beaudet. 2005. « Potent Spinal Analgesia Elicited through Stimulation of NTS2 Neurotensin Receptors ». *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 25 (36): 8188-96.
- Scheerer, Patrick, Jung Hee Park, Peter W. Hildebrand, Yong Ju Kim, Norbert Krauss, Hui-Woog Choe, Klaus Peter Hofmann, et Oliver P. Ernst. 2008. « Crystal Structure of Opsin in Its G-Protein-Interacting Conformation. » *Nature* 455 (7212): 497-502.
- Schihada, Hannes, Sylvie Vandenabeele, Ulrike Zabel, Monika Frank, Martin J. Lohse, et Isabella Maiellaro. 2018. « A Universal Bioluminescence Resonance Energy Transfer Sensor Design Enables High-Sensitivity Screening of GPCR Activation Dynamics ». *Communications Biology* 1 (1): 105.
- Schrödinger, LLC. s. d. *The PyMOL Molecular Graphics System, Version 2.0*.
- Sebastião, Ana M, et J. Alexandre Ribeiro. 2000. « Fine-tuning neuromodulation by adenosine ». *Trends in Pharmacological Sciences* 21 (9): 341-46.
- Seifert, Roland, et Katharina Wenzel-Seifert. 2002. « Constitutive Activity of G-Protein-Coupled Receptors: Cause of Disease and Common Property of Wild-Type Receptors ». *Naunyn-Schmiedeberg's Archives of Pharmacology* 366 (5): 381-416.
- Semack, Ansley, Manbir Sandhu, Rabia U. Malik, Nagarajan Vaidehi, et Sivaraj Sivaramakrishnan. 2016. « Structural Elements in the Gas and Gαq C Termini That Mediate Selective G Protein-Coupled Receptor (GPCR) Signaling ». *The Journal of Biological Chemistry* 291 (34): 17929-40.
- Senftle, Thomas P., Sungwook Hong, Md Mahbulul Islam, Sudhir B. Kylasa, Yuanxia Zheng, Yun Kyung Shin, Chad Junkermeier, et al. 2016. « The ReaxFF Reactive Force-Field: Development, Applications and Future Directions ». *Npj Computational Materials* 2 (mars): 15011.
- Sheth, Sandeep, Rafael Brito, Debashree Mukherjea, Leonard P. Rybak, et Vickram Ramkumar. 2014. « Adenosine Receptors: Expression, Function and Regulation ». *International Journal of Molecular Sciences* 15 (2): 2024-52.
- Shi, Qiang, Sergei Izvekov, et Gregory A. Voth. 2006. « Mixed Atomistic and Coarse-Grained Molecular Dynamics: Simulation of a Membrane-Bound Ion Channel ». *The Journal of Physical Chemistry B* 110 (31): 15045-48.
- Shi, Yue, Zhen Xia, Jiajing Zhang, Robert Best, Chuanjie Wu, Jay W. Ponder, et Pengyu Ren. 2013. « The Polarizable Atomic Multipole-Based AMOEBA Force Field for Proteins ». *Journal of Chemical Theory and Computation* 9 (9): 4046-63.
- Shorr, R. G., R. J. Lefkowitz, et M. G. Caron. 1981. « Purification of the Beta-Adrenergic Receptor. Identification of the Hormone Binding Subunit ». *The Journal of Biological Chemistry* 256 (11): 5820-26.

- Sleno, Rory, et Terence E. Hébert. 2018. « The Dynamics of GPCR Oligomerization and Their Functional Consequences ». *International Review of Cell and Molecular Biology* 338: 141-71.
- Soundararajan, Meera, Francis S. Willard, Adam J. Kimple, Andrew P. Turnbull, Linda J. Ball, Guillaume A. Schoch, Carina Gileadi, et al. 2008. « Structural Diversity in the RGS Domain and Its Interaction with Heterotrimeric G Protein Alpha-Subunits ». *Proceedings of the National Academy of Sciences of the United States of America* 105 (17): 6457-62.
- Sperlágh, Beáta, et E. Sylvester Vizi. 2011. « The Role of Extracellular Adenosine in Chemical Neurotransmission in the Hippocampus and Basal Ganglia: Pharmacological and Clinical Aspects ». *Current Topics in Medicinal Chemistry* 11 (8): 1034-46.
- Spindel, Eliot R. 2012. « Muscarinic Receptor Agonists and Antagonists: Effects on Cancer ». *Handbook of experimental pharmacology*, n° 208: 451-68.
- Sriram, Krishna, et Paul A. Insel. 2018. « G Protein-Coupled Receptors as Targets for Approved Drugs: How Many Targets and How Many Drugs? » *Molecular Pharmacology* 93 (4): 251-58.
- Staes, Edith, Pierre-Antoine Absil, Laurence Lins, Robert Brasseur, Magali Deleu, Nathalie Lecouturier, Virginie Fievez, et al. 2010. « Acylated and unacylated ghrelin binding to membranes and to ghrelin receptor: Towards a better understanding of the underlying mechanisms ». *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1798 (11): 2102-13.
- Sugita, Yuji, et Yuko Okamoto. 1999. « Replica-exchange molecular dynamics method for protein folding ». *Chemical Physics Letters* 314 (1): 141-51.
- Syed Ibrahim, B., et V. Pattabhi. 2005. « Trypsin Inhibition by a Peptide Hormone: Crystal Structure of Trypsin-Vasopressin Complex ». *J.Mol.Biol.* 348: 1191-98.
- Taylor, Christina M., Nicole B. Rockweiler, Cassie Liu, Loryn Rikimaru, Anna-Karin Tunemalm, Oleg G. Kisselev, et Garland R. Marshall. 2010. « Using Ligand-Based Virtual Screening to Allosterically Stabilize the Activated State of a GPCR ». *Chemical biology & drug design* 75 (3): 325-32.
- Toshinai, Koji, Yukari Date, Noboru Murakami, Mitsushi Shimada, Muhtashan S. Mondal, Takuya Shimbara, Jian-Lian Guan, et al. 2003. « Ghrelin-Induced Food Intake Is Mediated via the Orexin Pathway ». *Endocrinology* 144 (4): 1506-12.
- Tschöp, M., D. L. Smiley, et M. L. Heiman. 2000. « Ghrelin Induces Adiposity in Rodents ». *Nature* 407 (6806): 908-13.
- Unson, Cecilia G., Cui-Rong Wu, Youwei Jiang, Barney Yoo, Connie Cheung, Thomas P. Sakmar, et R. B. Merrifield. 2002. « Roles of Specific Extracellular Domains of the Glucagon Receptor in Ligand Binding and Signaling ». *Biochemistry* 41 (39): 11795-803.
- Uusitalo, Jaakko J., Helgi I. Ingólfsson, Parisa Akhshi, D. Peter Tieleman, et Siewert J. Marrink. 2015. « Martini Coarse-Grained Force Field: Extension to DNA ». *Journal of Chemical Theory and Computation* 11 (8): 3932-45.
- Vaidehi, Nagarajan, Supriyo Bhattacharya, et Adrien B. Larsen. 2014. « Structure and Dynamics of G-Protein Coupled Receptors ». *Advances in Experimental Medicine and Biology* 796: 37-54.

- Venkatakrishnan, A. J., Xavier Deupi, Guillaume Lebon, Christopher G. Tate, Gebhard F. Schertler, et M. Madan Babu. 2013. « Molecular Signatures of G-Protein-Coupled Receptors ». *Nature* 494 (7436): 185-94.
- Venter, J. C., M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith, et al. 2001. « The Sequence of the Human Genome ». *Science (New York, N.Y.)* 291 (5507): 1304-51.
- Verlet, Loup. 1967. « Computer “Experiments” on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules ». *Physical Review* 159 (1): 98-103.
- Vu, T. K., D. T. Hung, V. I. Wheaton, et S. R. Coughlin. 1991. « Molecular Cloning of a Functional Thrombin Receptor Reveals a Novel Proteolytic Mechanism of Receptor Activation ». *Cell* 64 (6): 1057-68.
- Wacker, Daniel, Gustavo Fenalti, Monica A. Brown, Vsevolod Katritch, Ruben Abagyan, Vadim Cherezov, et Raymond C. Stevens. 2010. « Conserved Binding Mode of Human Beta2 Adrenergic Receptor Inverse Agonists and Antagonist Revealed by X-Ray Crystallography. » *Journal of the American Chemical Society* 132 (33): 11443-45.
- Wall, M. A., D. E. Coleman, E. Lee, J. A. Iñiguez-Lluhi, B. A. Posner, A. G. Gilman, et S. R. Sprang. 1995. « The Structure of the G Protein Heterotrimer Gi Alpha 1 Beta 1 Gamma 2 ». *Cell* 83 (6): 1047-58.
- Walsh, D. A., J. P. Perkins, et E. G. Krebs. 1968. « An Adenosine 3',5'-Monophosphate-Dependant Protein Kinase from Rabbit Skeletal Muscle ». *The Journal of Biological Chemistry* 243 (13): 3763-65.
- Wang, Sheng, Tao Che, Anat Levit, Brian K. Shoichet, Daniel Wacker, et Bryan L. Roth. 2018. « Structure of the D2 Dopamine Receptor Bound to the Atypical Antipsychotic Drug Risperidone ». *Nature* 555 (7695): 269-73.
- Wang, Ting, et Yong Duan. 2007. « Chromophore Channeling in the G-Protein Coupled Receptor Rhodopsin ». *Journal of the American Chemical Society* 129 (22): 6970-71.
- Wang, Ting, et Yong Duan. 2009. « Ligand Entry and Exit Pathways in the Beta2-Adrenergic Receptor ». *Journal of Molecular Biology* 392 (4): 1102-15.
- Wei, Catherine J., Wei Li, et Jiang-Fan Chen. 2011. « Normal and abnormal functions of adenosine receptors in the central nervous system revealed by genetic knockout studies ». *Biochimica et Biophysica Acta (BBA) - Biomembranes, Adenosine Receptors*, 1808 (5): 1358-79.
- Weiner, Paul K., et Peter A. Kollman. 1981. « AMBER: Assisted Model Building with Energy Refinement. A General Program for Modeling Molecules and Their Interactions ». *Journal of Computational Chemistry* 2 (3): 287-303.
- Wheatley, M, D Wootten, MT Conner, J Simms, R Kendrick, RT Logan, DR Poyner, et J Barwell. 2012. « Lifting the lid on GPCRs: the role of extracellular loops ». *British Journal of Pharmacology* 165 (6): 1688-1703.
- White, Jim F., Nicholas Noinaj, Yoko Shibata, James Love, Brian Kloss, Feng Xu, Jelena Gvozdenovic-Jeremic, et al. 2012. « Structure of the Agonist-Bound Neurotensin Receptor ». *Nature* 490 (7421): 508-13.

- Wu, Beili, Ellen Y. T. Chien, Clifford D. Mol, Gustavo Fenalti, Wei Liu, Vsevolod Katritch, Ruben Abagyan, et al. 2010. « Structures of the CXCR4 Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists ». *Science (New York, N.Y.)* 330 (6007): 1066-71.
- Wu, C. K., B. Hu, J. P. Rose, Z. J. Liu, T. L. Nguyen, C. Zheng, E. Breslow, et B. C. Wang. 2001. « Structures of an Unliganded Neurophysin and Its Vasopressin Complex: Implications for Binding and Allosteric Mechanisms ». *Protein Science: A Publication of the Protein Society* 10 (9): 1869-80.
- Wu, Emilia L., Xi Cheng, Sunhwan Jo, Huan Rui, Kevin C. Song, Eder M. Dávila-Contreras, Yifei Qi, et al. 2014. « CHARMM-GUI Membrane Builder toward Realistic Biological Membrane Simulations ». *Journal of Computational Chemistry* 35 (27): 1997-2004.
- Wu, Huixian, Daniel Wacker, Mauro Mileni, Vsevolod Katritch, Gye Won Han, Eyal Vardy, Wei Liu, et al. 2012. « Structure of the Human  $\kappa$ -Opioid Receptor in Complex with JD1c ». *Nature* 485 (7398): 327-32.
- Yen, Hsin-Yung, Kin Kuan Hoi, Idir Liko, George Hedger, Michael R. Horrell, Wanling Song, Di Wu, et al. 2018. « PtdIns(4,5)P2 Stabilizes Active States of GPCRs and Enhances Selectivity of G-Protein Coupling ». *Nature* 559 (7714): 423-27.
- Yi, Min, Hekai Li, Zhiye Wu, Jianyun Yan, Qicai Liu, Caiwen Ou, et Minsheng Chen. 2018. « A Promising Therapeutic Target for Metabolic Diseases: Neuropeptide Y Receptors in Humans ». *Cellular Physiology and Biochemistry: International Journal of Experimental Cellular Physiology, Biochemistry, and Pharmacology* 45 (1): 88-107.
- Yi, Tau-Mu, Hiroaki Kitano, et Melvin I. Simon. 2003. « A Quantitative Characterization of the Yeast Heterotrimeric G Protein Cycle ». *Proceedings of the National Academy of Sciences of the United States of America* 100 (19): 10764-69.
- Zacharias, Martin. 2003. « Protein-Protein Docking with a Reduced Protein Model Accounting for Side-Chain Flexibility ». *Protein Science: A Publication of the Protein Society* 12 (6): 1271-82.
- Zhang, Wei, Chun Wu, et Yong Duan. 2005. « Convergence of Replica Exchange Molecular Dynamics ». *The Journal of Chemical Physics* 123 (15): 154105.
- Zhou, X. Edward, Yuanzheng He, Parker W. de Waal, Xiang Gao, Yanyong Kang, Ned Van Eps, Yanting Yin, et al. 2017. « Identification of Phosphorylation Codes for Arrestin Recruitment by G Protein-Coupled Receptors ». *Cell* 170 (3): 457-469.e13.
- Zou, Y. R., A. H. Kottmann, M. Kuroda, I. Taniuchi, et D. R. Littman. 1998. « Function of the Chemokine Receptor CXCR4 in Haematopoiesis and in Cerebellar Development ». *Nature* 393 (6685): 595-99.
- Zwingmann, C., D. Leibfritz, et A. S. Hazell. 2003. « Energy Metabolism in Astrocytes and Neurons Treated with Manganese: Relation among Cell-Specific Energy Failure, Glucose Metabolism, and Intercellular Trafficking Using Multinuclear NMR-Spectroscopic Analysis. » *Journal of Cerebral Blood Flow and Metabolism : Official Journal of the International Society of Cerebral Blood Flow and Metabolism* 23 (6): 756-71.

