



**HAL**  
open science

# Modélisation du complexe récepteur muscarinique/ toxique MT7 à partir de données thermodynamiques

Guillaume Letellier

► **To cite this version:**

Guillaume Letellier. Modélisation du complexe récepteur muscarinique/ toxique MT7 à partir de données thermodynamiques. Sciences du Vivant [q-bio]. Université Paris-Diderot - Paris VII, 2008. Français. NNT: . tel-00447060

**HAL Id: tel-00447060**

**<https://theses.hal.science/tel-00447060>**

Submitted on 14 Jan 2010

**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.

École doctorale **Inter///Bio**

**Modélisation du complexe récepteur muscarinique /  
toxine MT7 à partir de données thermodynamiques**

Présentée et soutenue publiquement le 8 octobre 2008

Pour l'obtention du grade de

*Docteur de l'université Denis Diderot*

Spécialité

*Analyse des Génomes et Modélisation Moléculaire*

Par

**GUILLAUME LETELLIER**

**Composition du jury :**

*Directeur de thèse :*

**Dr. Bernard GILQUIN**

*Rapporteurs:*

**Dr. Michael NILGES  
Dr. Serge CROUZY**

*Examineurs :*

**Pr. Catherine ETCHEBEST, présidente  
Pr. Marc LE MAIRE  
Dr. Denis SERVENT**

## Résumé

Les récepteurs muscariniques à l'acétylcholine sont des protéines transmembranaires jouant des rôles dans de nombreux processus physiologiques. La toxine MT7 est un puissant modulateur allostérique de ces récepteurs. De plus, cette toxine est le seul ligand connu spécifique du sous-type 1 des récepteur muscariniques (hM1). Nous avons étudié les bases moléculaires de l'interaction entre la toxine MT7 et le récepteur hM1 par modélisation. Tout d'abord un échantillonnage des structures des deux partenaires a été réalisé par dynamique moléculaire. Les mouvements de grande amplitude de la boucle e2 du récepteur ont été prédits par dynamique activée. Puis la toxine a été arrimée sur le récepteur par des calculs de dynamique moléculaire sous contraintes ambiguës dérivées de données de mutagenèse. Ce modèle a ensuite été optimisé par une simulation de dynamique moléculaire libre, en environnement membranaire explicite. Enfin, des calculs en retour des valeurs d'énergie libre de liaison ont été effectués afin de valider le modèle. Nous prédisons que la toxine se lie sur un dimère de récepteurs hM1. Le cœur de l'interaction est localisé sur un des monomères en contact avec les boucles II et III de la toxine. Cette interaction est complétée par des interactions hydrophobes au niveau de la boucle I sur le second monomère. L'analyse de ce modèle apporte des éléments de compréhension à la fois sur bases structurale de la haute affinité de cette toxine ainsi que sur sa sélectivité pour ce sous-type de récepteur. Il apparaît que cette sélectivité est essentiellement portée par la boucle extracellulaire e2 du récepteur.

# Remerciements

*Le travail de recherche présenté dans ce manuscrit a été réalisé au Laboratoire de Biologie Structurale et Radiobiologie du CEA de Saclay. Le CEA a financé en totalité cette thèse.*

*Je remercie Bernard Gilquin de m'avoir accueilli au sein du laboratoire et de m'avoir confié ce sujet de recherche. J'ai appris énormément à son contact au cours de ces trois années.*

*Je tiens à remercier les membres de mon jury, Serge Crouzy, Catherine Etchebest, Marc Le Maire, Michael Nilges et Denis Servent qui ont accepté de lire et de juger ce travail. Je suis particulièrement reconnaissant envers mes rapporteurs qui ont accepté de lire ce manuscrit en pleine période estivale.*

*Je remercie également tous nos collaborateurs du SIMOPRO qui ont réalisé les expériences de biochimie sur lesquelles repose cette thèse, Carole Fruchart, Catherine Marquer, Gille Mourier et Denis Servent. Un grand merci à Catherine pour le temps passé à la relecture de ce manuscrit.*

*Merci à Enrico Stura et André Menez qui ont mené l'étude cristallographique des toxines muscariniques.*

*Je remercie David Perahia pour ses précieux conseils et pour avoir mis à notre disposition les outils de dynamique moléculaires.*

*Je remercie Sophie Zinn-Justin pour la relecture des articles et pour ses remarques pertinentes.*

*Merci à Jacky Husser, Jean-Marc LeFailler et Arnaud Martel du GIPSI qui ont rendu possible ce travail par leur soutien pour l'utilisation des moyens de calcul.*

*Enfin, je remercie mes parents et ma compagne qui m'ont soutenu et supporté, durant cette période.*

*Pour finir, je tiens à rendre hommage à tous les intermittents de la recherche, passés par le laboratoire durant ces trois années. Par ordre de disparition : Guillaume, Cédric, Sandrine, Nathalie, Vincent, Mathilde, Lionel, Anaïs, Emmanuelle, Hocine, Benjamin, Albane, Anne, Anne-Christine, Samira et Marjorie.*

## Abréviations

AR	récepteur adrénergique
Ach	Acétylcholine
ADN	Acide désoxyribonucléique
AFM	Microscopie de force atomique
AMPc	Adénosine-monophosphate cyclique
TCM	Modèle du complexe ternaire
CHO	Chinese hamster ovary
DAG	Diacylglycérol
eGFP	Enhanced green fluorescent protein
FRET	Fluorescence ou Förster resonance energy transfer
FTDock	Fourier transform dock
GDTP	Guanosine di-phosphate
GTP	Guanosine tri-phosphate
HADDOCK	High ambiguity driven biomolecular docking
hM1	Récepteur muscarinique humain de sous-type 1
hM3	Récepteur muscarinique humain de sous-type 3
IP3	Inositol tri-triphosphate
M	molaire ( $\text{mol.l}^{-1}$ )
mAChR	Récepteur muscarinique de l'acétylcholine
MT	Toxine muscarinique
nM	nanomolaire
-NMS	N-méthylscopolamine
ns	nanoseconde
PDB	Protein data bank
PKA	Protéine kinase A
PKC	Protéine kinase C
PLC	Phospholipase de type C
pM	picomolaire
RCPG	Récepteur couplé aux protéines G
RMN	Résonance magnétique nucléaire
RMSD	Ecart quadratique moyen
ROI	Region of interest
SDS-Page	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TCM	Modèle du complexe ternaire
TM	Domaine transmembranaire

# Table des matières

<b>Chapitre 1: Introduction</b>	<b>9</b>
<b>1.1. Préambule</b>	<b>10</b>
<b>1.2. Les récepteurs couplés aux protéines G (RCPG)</b>	<b>11</b>
1.2.1. Introduction à la signalisation cellulaire	11
1.2.2. Découverte et importance des récepteurs couplés aux protéines G	12
1.2.3. Couplage aux protéines G	13
1.2.4. Classification des RCPG	14
1.2.5. Structure des RCPG	15
1.2.5.1. Structure de la rhodopsine	16
1.2.5.2. Structure du récepteur $\alpha$ -adrénergique	18
1.2.5.3. Conclusion	20
1.2.6. Les différents modes de liaison des ligands des RCPG	21
1.2.7. Oligomérisation des RCPG	22
<b>1.3. Modélisation des RCPG</b>	<b>24</b>
1.3.1. Modélisation comparative	24
1.3.2. Retour d'expérience sur les modèles par homologie de RCPG	25
1.3.2.1. Modèles de la rhodopsine	25
1.3.2.2. Modèles de $\alpha_2$ AR	26
1.3.2.3. Criblage in silico de RCPG	27
1.3.3. Modélisation de novo	28
<b>1.4. Les récepteurs muscariniques</b>	<b>29</b>
1.4.1. Sous-types de récepteurs muscariniques	30
1.4.2. Le sous-type 1 des récepteurs muscariniques	33
1.4.3. Caractérisation des sites de liaison des récepteurs muscariniques	33
1.4.3.1. Site orthostérique	33
1.4.3.2. Site allostérique	35
<b>1.5. Les toxines muscariniques</b>	<b>37</b>
1.5.1. Synthèse	37
1.5.2. Structure	38
1.5.3. L'interaction toxine MT7 – récepteur hM1	39
<b>1.6. Prédiction de la structure des complexes protéine-protéine</b>	<b>41</b>
1.6.1. Nature des interfaces des complexes macromoléculaires	42
1.6.1.1. Nature des résidus	42
1.6.1.2. Dimension des interfaces de protéines	42

1.6.1.3. Réarrangements structuraux	43
1.6.2. Prédiction de la structure des complexes de protéine	44
1.6.3. Comparaison des méthodes de docking : le concours CAPRI	45
1.6.4. Méthode des cycles de doubles mutants	49
1.6.4.1. Principe	49
1.6.5. Validation de l'utilisation des données de doubles mutant pour le docking	51
1.6.5.1. Evaluation du modèle Ach- 7/ -cobratoxine	53
1.6.5.2. Evaluation du modèle du canal potassique Kv1.1/BgK	57
1.6.5.3. Evaluation du modèle uPAR/uPA.GFD	59
1.6.6. Conclusion	60

## ***Chapitre 2: Étude structurale des partenaires*** \_\_\_\_\_ **62**

<b>2.1. Le récepteur hM1</b>	<b>63</b>
2.1.1. Séquences peptidique	63
2.1.2. Modélisation du cœur d'hélices	64
2.1.3. Prédiction des boucles	64
<b>2.2. La toxine MT7</b>	<b>67</b>

## ***Chapitre 3: Prédiction de la structure du complexe***

### ***hM1/MT7*** \_\_\_\_\_ **72**

<b>3.1. Mesures thermodynamiques hM1/MT7</b>	<b>73</b>
3.1.1. Modèle du complexe ternaire	73
3.1.2. Alanine scanning des partenaires	75
3.1.2.1. Modifications de la toxine	75
3.1.2.2. Mutagénèse du récepteur	79
3.1.3. Cycles de doubles mutants sur le système hM1/MT7	82
3.1.3.1. Choix des couples de résidus	82
3.1.3.2. Résultats des expériences de cycles de doubles mutants	83
<b>3.2. Exploration de la flexibilité conformationnelle</b>	<b>84</b>
3.2.1. Dynamique moléculaire de la toxine	84
Dynamique moléculaire du récepteur	86
3.2.1.1. La boucle extracellulaire E2	86
3.2.1.2. La méthode PEDC	87
3.2.1.3. Dynamique moléculaire du récepteur	89
<b>3.3. Procédure d'arrimage moléculaire des partenaires</b>	<b>92</b>
3.3.1.1. Dynamique moléculaire	92
3.3.1.2. Échantillonnage des positions de départ	92

3.3.1.3.	Introduction des données expérimentales	93
3.3.1.4.	Utilisation de la dynamique moléculaire en 4D	94
3.3.1.5.	Conclusion	95
<b>3.4.</b>	<b>Topologie des contraintes</b>	<b>96</b>
<b>3.5.</b>	<b>Premiers modèles</b>	<b>101</b>
3.5.1.1.	Vers un dimère de récepteur	103
<b>3.6.</b>	<b>Démonstration expérimentale de l'existence de dimère de hM1</b>	<b>105</b>
3.6.1.	Expériences de gels natif	105
3.6.2.	Transfert de fluorescence par photo-blanchiment de l'accepteur	106
3.6.3.	Expérience d'anisotropie de fluorescence	107
3.6.4.	Conclusion	109
<b>3.7.</b>	<b>Modèles trimériques du complexe dimère de hM1 / MT7</b>	<b>110</b>
3.7.1.	Construction d'un modèle de dimère de hM1	110
<b>Chapitre 4: Validation et optimisation du modèle</b>		<b>114</b>
<b>4.1.</b>	<b>Introduction</b>	<b>115</b>
4.1.1.	Dynamique moléculaire des RCPG	115
4.1.1.1.	Travaux méthodologiques	115
4.1.1.2.	Activation de la rhodopsine	118
4.1.1.3.	Modèles	119
4.1.2.	Effets électrostatiques et protéines	120
4.1.2.1.	Seuils de distance sphériques	121
4.1.2.2.	Particle Mesh Ewald	121
4.1.2.3.	Équation de Poisson Boltzmann	123
4.1.3.	Insertion du complexe dans une membrane	125
<b>4.2.</b>	<b>Dynamique moléculaire du modèle</b>	<b>129</b>
<b>4.3.</b>	<b>Validation du modèle par recalculs d'énergie libre</b>	<b>133</b>
<b>4.4.</b>	<b>Analyse du modèle MT7/dimère de hM1</b>	<b>137</b>
4.4.1.	Interface de dimérisation	137
4.4.2.	Site de liaison de la toxine	138
4.4.2.1.	Interface MT7/hM1A	138
4.4.2.2.	Interface MT7/hM1B	142
<b>Chapitre 5: Discussion</b>		<b>144</b>
<b>5.1.</b>	<b>Apports et limites de la dynamique moléculaire</b>	<b>145</b>
5.1.1.	Influence de la boucle E2	145



5.1.2.	Données non prises en comptes	147
5.1.2.1.	W91 <sub>M1</sub> / Y30 <sub>tox</sub>	147
	R52 <sub>tox</sub> / W400 <sub>M1</sub>	149
<b>5.2.</b>	<b>Dimérisation des récepteurs muscariniques</b>	<b>150</b>
5.2.1.	Nature de la surface d'interaction	150
5.2.2.	Mode d'assemblage du dimère de récepteurs	151
5.2.3.	Oligomères d'ordre supérieurs	154
<b>5.3.</b>	<b>Bases moléculaire des propriétés pharmacologiques de la toxine</b>	<b>156</b>
5.3.1.	Site de liaison	156
5.3.2.	Stoechiométrie NMS / MT7 / hM1	156
5.3.2.1.	Hypothèse 1 : 1 dimère de M1, 1 toxine, 1 NMS	156
5.3.2.2.	Hypothèse 2 : 1 dimère de M1, 2 toxines, 2 NMS	157
5.3.3.	Haute affinité de liaison	158
5.3.4.	Sélectivité pour le sous-type M1 des récepteurs muscariniques	161
<b>Chapitre 6:</b>	<b>Conclusion et Perspectives</b>	<b>164</b>
<b>6.1.</b>	<b>Perspectives : les bases moléculaires de l'allostérie</b>	<b>166</b>
<b>Chapitre 7:</b>	<b>Index des figures</b>	<b>171</b>
<b>Chapitre 8:</b>	<b>Bibliographie</b>	<b>175</b>
<b>Chapitre 9:</b>	<b>Annexes</b>	<b>190</b>
9.1.1.	ligands orthostériques des RCPG	191
9.1.2.	Alignement des séquences du hM1 et de la rhodopsine	192
9.1.3.	Simulation d'arrimage	193
9.1.4.	Simulation de dynamique moléculaire activée	196
9.1.5.	Simulations de dynamique moléculaire dans la membrane	198
9.1.6.	Benchmark des codes de dynamique moléculaire sur la machine TERA10	200

---

# **Chapitre 1: Introduction**

---

### ***1.1. Préambule***

La structure tridimensionnelle des protéines apporte des informations essentielles à la compréhension des mécanismes moléculaires dont dépend la réalisation des processus cellulaires. Dernièrement, on a assisté, parallèlement à l'essor de la génomique qui fournissait des quantités formidables de données de séquence, à celui de grands projets de détermination de structures des protéines à haut débit. Cependant, de nombreuses protéines difficiles à étudier telles que les protéines membranaires, en dépit de leur importance biologique, ne sont représentées que de façon marginale dans les bases de données de structures.

Les informations apportées par les structures à l'échelle atomique ont permis des avancées indéniables dans la compréhension du fonctionnement des protéines. Cependant, une vision réellement globale des processus cellulaires impose de considérer la protéine dans son contexte, c'est-à-dire en complexe avec ses partenaires. Ceci est particulièrement important lorsque l'on s'intéresse aux phénomènes de signalisation et à leurs mécanismes de régulation. Toutefois, ces assemblages macromoléculaires sont par nature instables et donc difficiles à étudier.

On estime que les 30.000 protéines du génome humain (Venter, Adams et al. 2001) possèdent en moyenne 9 partenaires chacune (Chance, Fiser et al. 2004). L'importance de la combinatoire de ces chiffres rend totalement irréaliste l'idée d'étudier de manière systématique les structures des complexes de protéines par des méthodes expérimentales classiques. A l'inverse, les méthodes de modélisation moléculaire en particulier les méthodes d'arrimage *in silico* donnent aujourd'hui des résultats très encourageants. Ces approches de prédiction de la structure de complexes de protéines à partir des coordonnées atomiques des partenaires ont beaucoup progressé ces dernières années et permettent maintenant d'obtenir des modèles proches des structures natives des protéines (Janin, 2003). La disponibilité de données issues de l'expérimentation *in vivo* ou *in vitro* est un facteur déterminant dans la réussite des prédictions de la structure des complexes. On a ainsi vu se développer des méthodes hybrides, s'appuyant sur des données expérimentales d'origines variées. En particulier, des informations très utiles peuvent être obtenues à partir des mesures thermodynamiques associées à des expériences de mutagenèse.

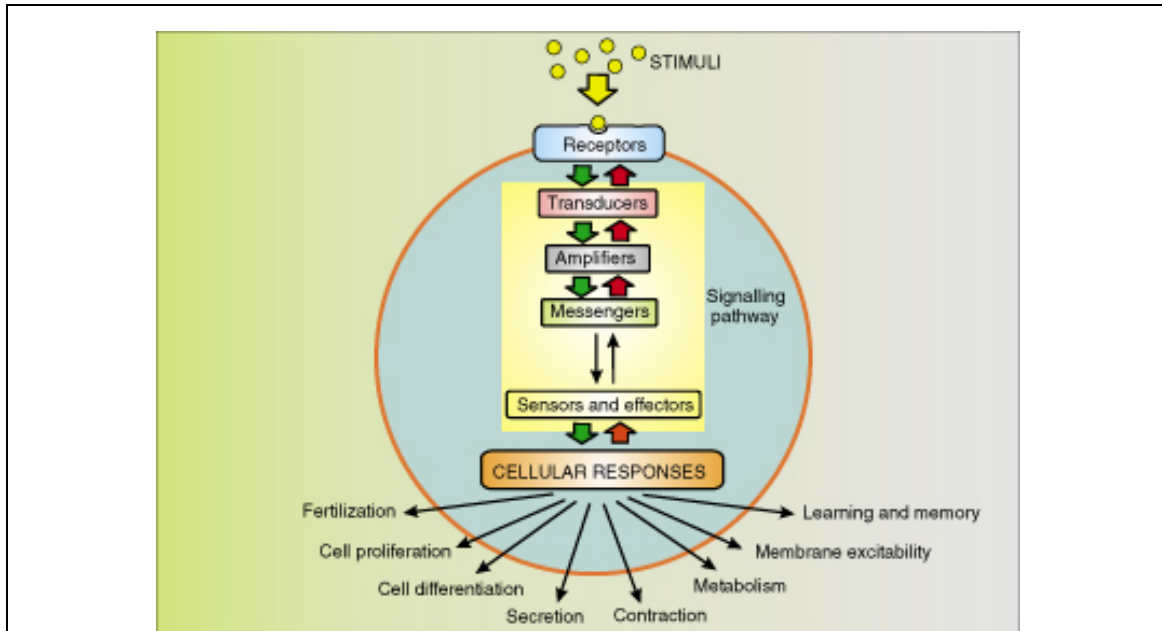
## ***1.2. Les récepteurs couplés aux protéines G (RCPG)***

### **1.2.1. Introduction à la signalisation cellulaire**

Les cellules vivantes ont besoin, pour assurer leur survie, de pouvoir répondre à des stimuli variés de leur environnement. De plus, dans le cas des organismes pluricellulaires, les cellules sont en communication constante les unes avec les autres. La diversité des signaux à laquelle une cellule doit pouvoir répondre est donc très importante. On peut citer par exemple : les ions, les métabolites, les hormones, les facteurs de croissance, les neurotransmetteurs, et même la lumière. Il s'agit donc, dans la grande majorité des cas, de l'émission et de la réception de signaux chimiques.

Cependant, les cellules sont entourées d'une membrane lipidique qui constitue une barrière imperméable maintenant leur intégrité. Certains signaux, tels que des hormones hydrophobes, ont la capacité à diffuser de manière passive au travers des membranes pour atteindre leurs effecteurs dans la cellule mais pour la grande majorité des signaux, des stratégies plus élaborées sont nécessaires afin de transmettre l'information au travers la membrane. La notion de signalisation cellulaire concerne donc l'ensemble des mécanismes impliqués dans la réception d'un stimulus extracellulaire et son relayage (Figure 1) par différents intermédiaires vers l'intérieur de la cellule jusqu'à un effecteur chargé de modifier la réponse cellulaire.

La première étape de la signalisation cellulaire est médiée par un récepteur, situé au niveau de la membrane cytoplasmique. Celui-ci reçoit le stimulus extérieur et initie la cascade de signalisation interne. Il existe une grande variété dans les processus moléculaires mis en jeu par les récepteurs membranaires pour transmettre l'information au travers de la membrane. On distingue trois grands types de récepteurs : les récepteurs à activité enzymatique, souvent à un seul passage transmembranaire, les canaux et les récepteurs couplés aux protéines G.



**Figure 1 : La signalisation cellulaire**

*La liaison d'un agoniste sur la face extracellulaire du récepteur va entraîner une cascade d'activation aboutissant à des réponses cellulaires variées.*

### 1.2.2. Découverte et importance des récepteurs couplés aux protéines G

Il y a plus d'un siècle que l'existence de récepteurs est connue et qu'ils sont étudiés par des médecins et pharmacologistes. Il a pourtant fallu attendre les années 70 pour que l'existence des récepteurs couplés aux protéines G (RCPG) en tant qu'entités physiques indépendantes ne soit définitivement démontrée et admise par la biochimie (Lefkowitz 2007). Leur mise en évidence par les travaux de l'équipe d'Alfred G. Gilman a d'ailleurs valu à celui-ci le prix Nobel de médecine en 1994 (Gilman 1994).

Du point de vue de l'évolution, les RCPG figurent parmi les plus anciens transducteurs de signaux connus. En effet, ils sont présents dans les plantes (Plakidou-Dymock, Dymock et al. 1998), les levures (Dohlman HG 1991), les champignons, (Kim and Devreotes 1994), ainsi que chez les protozoaires et les métazoaires, (Bockaert J 1999). Aujourd'hui, grâce à l'essor de la génomique et au séquençage du génome humain (Venter, Adams et al. 2001), 799 membres de la famille des RCPG ont été identifiés chez l'homme, dont 400 environ sont des récepteurs sensoriels (olfactifs et goût) (Takeda, Kadowaki et al. 2002). Avec près de 2% du génome la superfamille des RCPG est donc une des plus importantes du point de vue de la génomique. Le rôle des RCPG est également important pour la physiologie. En effet ils interviennent dans des

processus aussi variés que la vue, l'odorat, la régulation de l'humeur, du comportement, du système immunitaire et de l'inflammation, ainsi que des systèmes nerveux sympathiques et parasympathiques. Les RCPG répondent donc à une grande variété de signaux : stimuli sensoriels exogènes (molécules olfactives, gustatives, phéromones), petites molécules endogènes (acides aminés, amines, nucléosides et nucléotides, peptides, lipides...) et des protéines (hormones, protéases). On estime que près de 80% des hormones et neurotransmetteurs connus ciblent des RCPG. Enfin, les RCPG sont des protéines essentielles du point de vue médical (Miligan 2007) puisque plus de 50% des médicaments utilisés à l'heure actuelle régulent la fonction d'un RCPG. Parmi eux, 30% les ciblent directement.

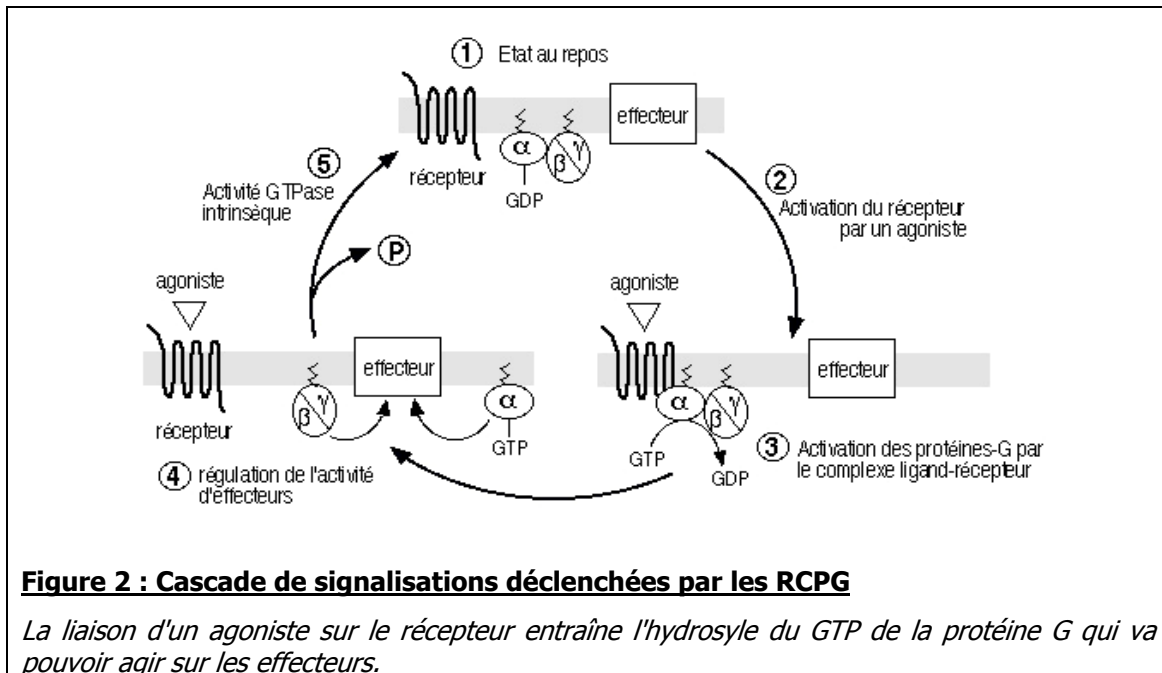
Pourtant, à l'heure actuelle, plus de trois quarts des RCPG sont qualifiés d'orphelins, aucun ligand ni aucune fonction n'ayant pu leur être attribués. Finalement seul un petit nombre de ces récepteurs sont la cible de médicaments, ce qui laisse donc ouvertes des perspectives prometteuses pour la recherche médicale.

### 1.2.3. Couplage aux protéines G

Les RCPG tirent leur nom de leur fonctionnement couplé aux protéines G. L'activation d'un RCPG, par son ligand, entraîne un remaniement de la structure du récepteur, qui devient capable d'activer des protéines-G intracellulaires (Figure 2). Les protéines G sont des protéines hétérotrimériques formées de 3 sous-unités  $\alpha$ ,  $\beta$  et  $\gamma$ . La sous unité  $\alpha$  permet l'ancrage de la protéine G du côté cytosolique de la membrane, grâce à des modifications post-traductionnelles (myristoylation, et l'isoprénylation). Dans sa forme inactive, la protéine G lie le GDP par sa sous unité  $\alpha$ . L'activation de la protéine G par le récepteur s'accompagne donc de la substitution du GDP par un GTP et de la dissociation de la sous unité  $\beta\gamma$ . La sous unité  $\beta\gamma$  active alors à son tour toute une série d'effecteurs intracellulaires (enzymes, canaux ioniques, transporteurs...). Ces effecteurs permettent à leur tour, la modulation des concentrations intracellulaires des seconds messagers tels l'AMPc, l'IP3, le calcium, les diacyl-glycérols (DAG), ...

Le dogme du couplage RCPG / protéine G est aujourd'hui remis en cause par la découverte de voies de signalisation activées par des RCPG, indépendamment des protéines G (Levoye, Dam et al. 2006), (Sun, McGarrigle et al. 2007). C'est par exemple le cas des récepteurs  $\beta_2$  adrénergiques et des récepteurs muscariniques qui présentent la capacité de pouvoir modifier le potentiel transmembranaire du sodium,

indépendamment de tout couplage à une protéine G. D'autre part, en plus de leur capacité à lier des ligands, il apparaît que certains RCPG sont également capables de répondre à des changements de potentiel électrique transmembranaire d'une manière similaire à de nombreux canaux ioniques. Ceci a notamment été observé, pour les récepteurs muscariniques M1 et M2, dont la liaison à l'acétylcholine est modulée par le potentiel de membrane (Yair Ben-Chaim 2006).



#### 1.2.4. Classification des RCPG

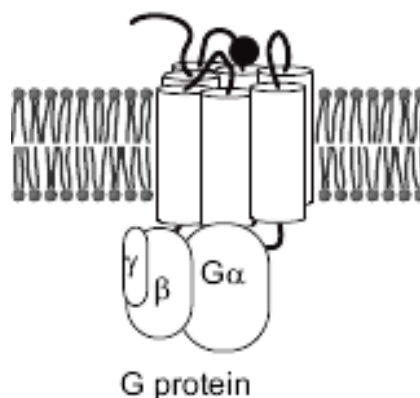
Plusieurs méthodes de classification des RCPG ont été proposées. Historiquement, 6 classes portant les lettres A à F avaient été définies (Bockaert J 1999). Des travaux plus récents ont proposé d'autres classifications, basées par exemple sur des similarités de séquences (Huang, Cai et al. 2004) ou sur l'analyse phylogénétique (Fredriksson R 2003). Dans la suite de cette étude, nous utiliserons cette dernière classification, également connue sous le nom de GRAFS. Elle distingue 5 classes de RCPG.

- **Glutamate** : Cette classe concerne un petit nombre de récepteurs, dont essentiellement les récepteurs au  $\text{Ca}^{++}$  et au glutamate.

- **Rhodopsine** : Il s'agit de la classe la plus peuplée. Elle regroupe entre autre la rhodopsine, les récepteurs  $\alpha$ -adrénergiques, muscariniques et surtout les nombreux récepteurs olfactifs.
- **Adhésion** : Ces récepteurs sont impliqués dans le fonctionnement du système nerveux central, et de l'immunité
- **Frizzled/taste** : Ces récepteurs jouent des rôles respectivement dans le développement embryonnaire et la perception du goût.
- **Secretin** : Il s'agit des récepteurs liant les hormones de haut poids moléculaire, telles que le glucagon, la sécrétine ou la calcitonine.

### 1.2.5. Structure des RCPG

Les RCPG sont des protéines membranaires intégrales dont la masse moléculaire varie de 40 à 85 kDa. Tous les RCPG partagent une organisation commune (Figure 3) à 7 passages transmembranaires structurés en hélice de 30 à 40 résidus environ, interconnectés par des boucles (Baldwin 1993). L'extrémité C-terminale est orientée vers le cytoplasme et la N-terminale vers le milieu extracellulaire. La longueur de l'extrémité N-terminale est très variable, allant de quelques dizaines de résidus pour les récepteurs de la classe B à plusieurs centaines pour ceux de la classe G. Par convention, les trois boucles intracellulaires sont désignées par I1, I2 et I3 et les trois boucles extracellulaires par E1, E2 et E3, numérotées dans le sens  $N_{\text{ter}} \rightarrow C_{\text{ter}}$ .



**Figure 3 : Organisation d'un RCPG**

*Les récepteurs couplés aux protéines G sont constitués d'un domaine transmembranaire en hélice, connectée par des boucles. Les protéines G hétérotrimériques se lient à la face intracellulaire.*

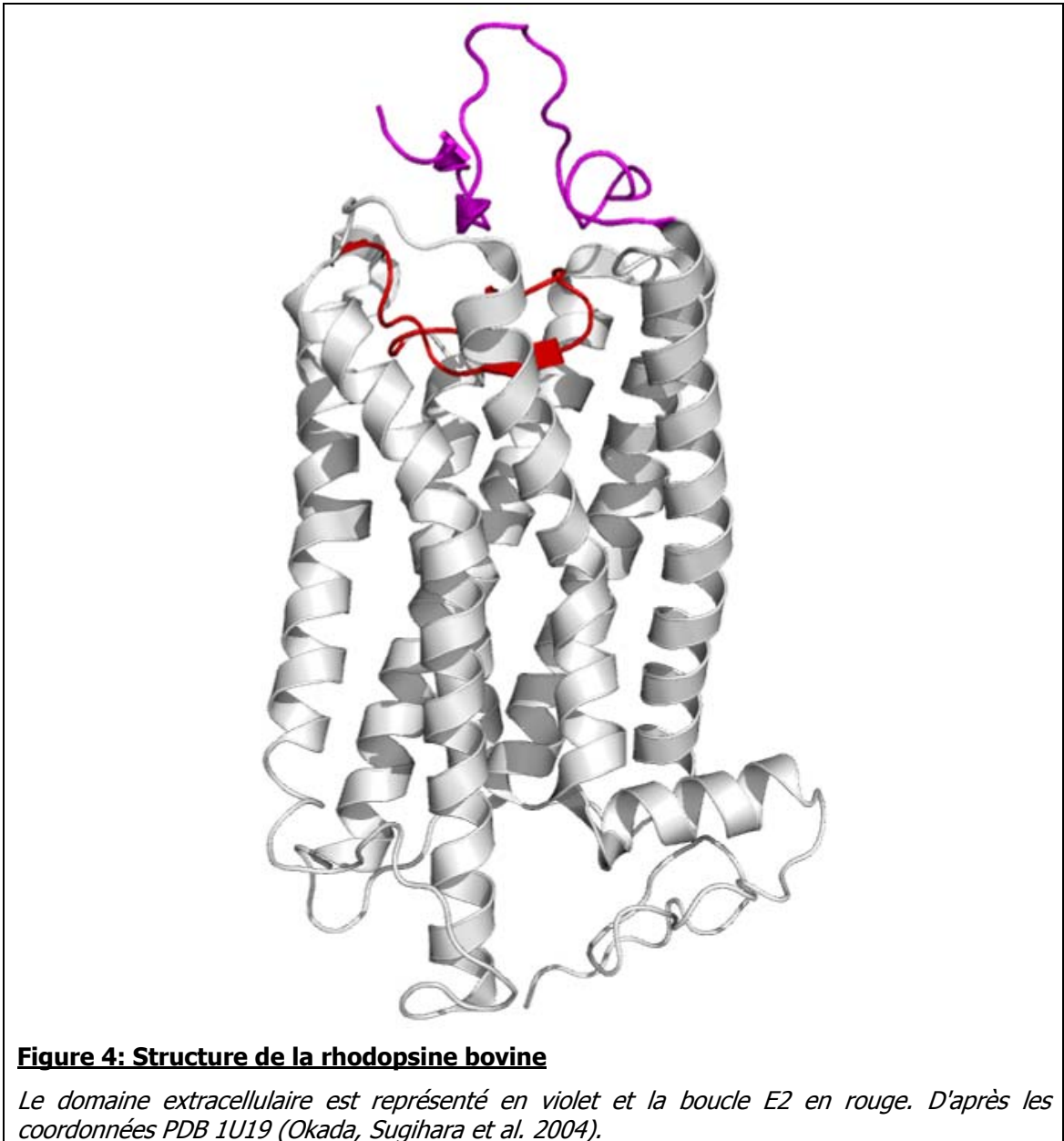


### 1.2.5.1. Structure de la rhodopsine

La rhodopsine est un RCPG impliqué dans le mécanisme de la vision. Présent au niveau des membranes des cellules en bâtonnet de l'œil, il est disponible facilement en grande quantité. Ceci explique que ce RCPG fut le premier dont la structure a été caractérisée et qu'il soit aujourd'hui encore le plus étudié. La rhodopsine est associée à un chromophore, le 11-cis-rétinal, lié de manière covalente à sa lysine 296 par une base de Schiff. Le site de liaison du rétinol se trouve au cœur du domaine transmembranaire, dans une cavité induite par les coudes dus aux prolines des hélices. L'impact d'un photon provoque l'isomérisation du rétinol qui va entraîner des changements conformationnels locaux qui aboutissent finalement à un réarrangement des hélices et à l'activation d'une protéine G, la transducine.

La première structure de RCPG disponible fut celle de la rhodopsine, déterminée par cristallographie des rayons X à une résolution de 2.8 Å en 2000 (Palczewski, Kumasaka et al. 2000). Cette première structure a par la suite été raffinée jusqu'à une résolution de 2.6 Å permettant d'identifier le réseau interne de molécules d'eau (Okada, Fujiyoshi et al. 2002). Puis, de nouvelles conditions de cristallisation ont permis d'atteindre une résolution de 2.2 Å avec un empilement cristallin différent (Okada, Sugihara et al. 2004). A partir de ces structures, par remplacements moléculaires, plusieurs formes de la rhodopsine ont été déterminées: des mutants (Standfuss, Xie et al. 2007), des intermédiaires d'activation (Schertler 2005; Salom, Lodowski et al. 2006) et dernièrement en absence du rétinol (Park, Scheerer et al. 2008).

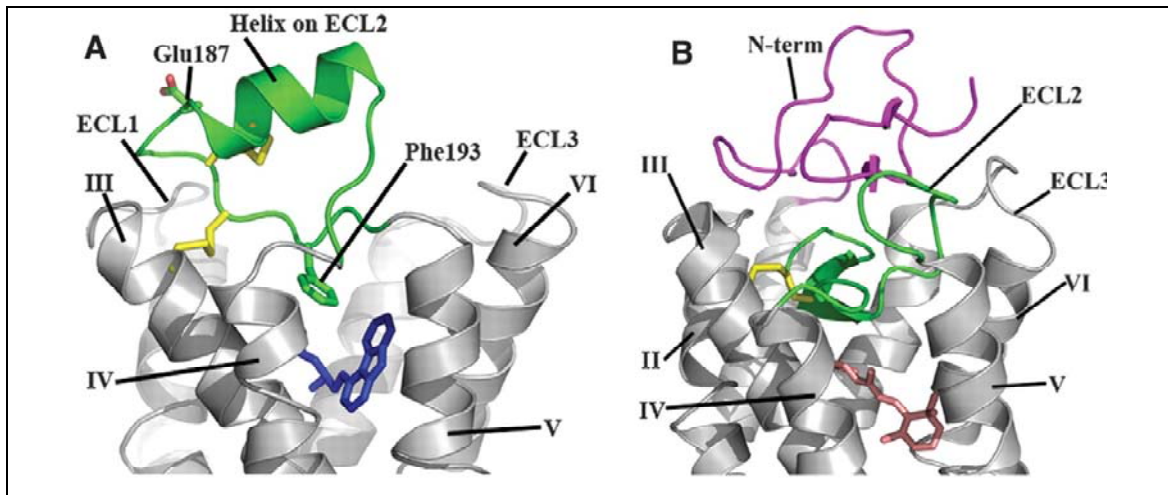
La structure des hélices I et II est régulière tandis que les hélices TM1, TM4, TM6 sont coudées par des prolines et que l'hélice TM7 présente des irrégularités autour du site de liaison du rétinol. Le cœur du domaine transmembranaire est stabilisée par un important réseau de liaisons hydrogène et d'interactions hydrophobes. La structure révèle également l'existence d'une huitième courte hélice cytoplasmique parallèle au plan de la membrane. Tandis que les boucles E1 et E3 sont dans des conformations périphériques de l'apex de la protéine, la boucle E2 adopte une conformation étendue au dessus de la cavité de fixation du rétinol. Cette boucle est maintenue à sa base par un pont disulfure très conservé entre les résidus Cys108 et Cys186, localisés respectivement sur la boucle E2 et au niveau de l'hélice TM4, au cœur du domaine transmembranaire.



### 1.2.5.2. Structure du récepteur $\alpha_2$ -adrénergique

La rhodopsine est un RCPG particulier, puisque ce récepteur hautement spécialisé, présent dans un seul type de tissu, est dédié à une fonction unique de détection de la lumière et ne lie aucun ligand. Fin 2007, une seconde structure de RCPG a été déterminée (Cherezov, Rosenbaum et al. 2007; Rasmussen, Choi et al. 2007), celle du récepteur  $\alpha_2$ -adrénergique. Au contraire de la rhodopsine, le récepteur  $\alpha_2$ -adrénergique est un bon représentant des récepteurs RCPG "classiques". Le  $\alpha_2$ -adrénergique comme la majorité des autres RCPG répond en effet à la liaison d'une grande variété de ligands et est impliqué dans un nombre important de voies de signalisation différentes. Il joue en particulier un rôle important dans la régulation des processus physiologiques cardiovasculaires et pulmonaires. Ces récepteurs lient l'adrénaline et activent la voie de l'adénylate cyclase.

Comme on pouvait s'y attendre, les régions transmembranaires de la rhodopsine et du récepteur  $\alpha_2$ AR sont hautement similaires, avec une déviation des carbones- du squelette de seulement 1.56 Å. La disposition des sept hélices est analogue à celle de la rhodopsine, mais pas strictement identique. En effet le faisceau d'hélices présente une disposition plus ouverte. Une des différences les plus importantes concerne l'hélice I qui ne présente pas de coude à proline et est beaucoup plus régulière que son équivalent dans la rhodopsine. Cette différence a pour conséquence de limiter les possibilités d'interaction entre le N-ter de la protéine et les boucles extracellulaires. La différence la plus évidente entre les deux structures est la conformation de la boucle extracellulaire E2 en feuillet étendu pour la rhodopsine (Figure 5B), alors qu'elle forme une courte hélice contrainte par 2 ponts disulfures pour le récepteur adrénergique (Figure 5A). La conformation de cette boucle chez le récepteur  $\alpha_2$ -adrénergique laisse donc l'accès ouvert à la cavité apicale du récepteur.

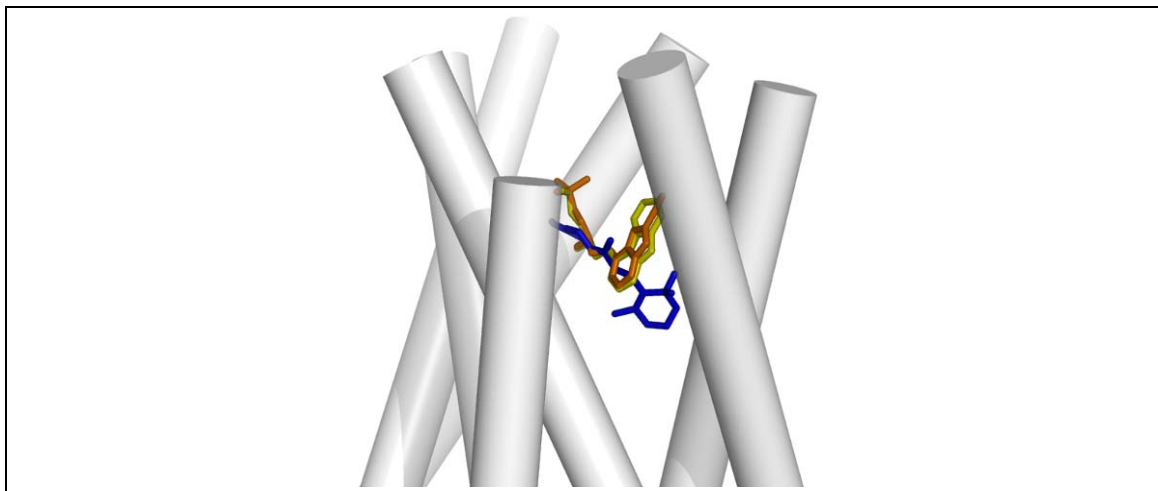


**Figure 5 : Domains extracellulaires de la rhodopsine et du récepteur  $\beta_2$ adrénergique**

*Comparaisons des régions extracellulaires de la rhodopsine (B) et du récepteur  $\beta_2$ AR (A). En vert, les boucles extracellulaires du  $\beta_2$ AR et de la rhodopsine forment respectivement une courte hélice et un feuillet étendu. La rhodopsine comporte un domaine N-terminal (en violet) qui couvre sa région extracellulaire. D'après (Kobilka and Schertler 2008).*

La structure a été déterminée en présence d'un antagoniste qui se lie avec une très forte affinité au récepteur  $\beta_2$ AR, le carbazolol (Cherezov, Rosenbaum et al. 2007). Le site de liaison du  $\beta_2$ AR se situe dans une cavité localisée entre les 7 hélices transmembranaires, dans une région équivalente à celle qui contient le rétinol dans la structure de la rhodopsine (Figure 5A). Cependant, le carbazolol est positionné différemment du rétinol et peut en particulier interagir avec le résidu hautement conservé Trp265 (Figure 6). Ce résidu est connu comme déclencheur de l'activation des RCPG via une modification de l'orientation de sa chaîne latérale (Schwartz, Frimurer et al. 2006).

Les connaissances apportées par les expériences de mutagenèse du récepteur  $\beta_2$ AR combinées à sa structure ont permis de mieux comprendre les mécanismes d'activation du récepteur (Kobilka and Schertler 2008). Il existe par exemple un certain nombre de mutations connues pour conduire à une activité constitutive du récepteur. La structure cristalline du récepteur révèle que ces résidus sont en contact avec une ou plusieurs hélices où ils jouent probablement un rôle de stabilisation dans l'état basal du récepteur (Rosenbaum, Cherezov et al. 2007). Ceci confirme l'hypothèse selon laquelle le processus d'activation du récepteur implique des mouvements de grande amplitude des hélices transmembranaires.



**Figure 6 : Positionnement des ligands orthostériques dans les structures de RCPG**

*Les coordonnées des récepteurs  $\beta_1AR$ ,  $\beta_2AR$  et de la rhodopsine sont superposées. Le cylindre représente les hélices de la rhodopsine. Les ligands orthostérique sont figurés en jaune (carazolo), orange (Cyanopindolol) et bleu (rétinal).*

### 1.2.5.3. Conclusion

Les structures des récepteurs  $\beta_1$  et  $\beta_2$  adrénergiques publiées récemment confirment la grande similarité entre les structures des différents RCPG, au moins pour ceux appartenant à la famille de la rhodopsine. Cette ressemblance explique les succès obtenus avec des modèles de RCPG dérivés de la structure de la rhodopsine en dépit de la relative faiblesse de l'identité de séquence (20 à 30 %). La grande proximité structurale entre ces protéines valide donc l'utilisation de la modélisation comparative pour prédire la structure des récepteurs de la famille de la rhodopsine. La modélisation comparative d'une protéine à partir de la structure d'un unique homologue est par définition limitée par le fait qu'il n'est dans ce cas pas possible de corrélérer les propriétés structurales à des divergences particulières dans les séquences des protéines étudiées. Disposer de nouvelles structures de RCPG permettra donc à l'avenir de pouvoir croiser les informations structurales et de séquences. De tels modèles pourront se révéler très utiles lors des études de criblage à haut débit ou de conception de nouveaux médicaments par exemple. Ces modèles sont également nécessaires à la modélisation de complexes ligand/récepteur et à la compréhension des bases moléculaires de la spécificité de ligands.

### 1.2.6. Les différents modes de liaison des ligands des RCPG

Le premier modèle d'interaction allostérique a été proposé par Monod, Wyman et Changeux en 1965 (Monod, Wyman et al. 1965). Ce modèle qui a toujours cours aujourd'hui (Changeux JP 2005) décrit les effets indirects, coopératifs ou anti-coopératifs, de liaisons de ligands à des sites distincts. Les interactions allostériques supposent que les deux ligands ne soient pas en compétition, et donc ne se lient pas au même site. Leurs interactions entre les deux ligands ont lieu de manière indirecte par l'intermédiaire des changements structuraux qu'ils induisent au niveau du récepteur.

La transmission du signal par un RCPG est déclenchée par la liaison d'un ligand à un site extracellulaire qui entraîne des changements structuraux (Okada, Ernst et al. 2001) dans le faisceau d'hélices du récepteur et qui se propagent à vers la face intracellulaire jusqu'au site de liaison des protéines G. Celles-ci vont alors activer les seconds messagers de la chaîne de signalisation (Miligan 2007). Le fonctionnement des RCPG est donc un mécanisme intrinsèquement allostérique dans la mesure où l'activation du récepteur résulte de l'interaction entre deux sites distincts au travers de changements structuraux.

La notion d'allostérie chez les RCPG est en fait plus complexe car il existe chez les RCPG une troisième catégorie de ligands. Ces ligands se lient à des sites différents de celui de l'activateur naturel ou du site de fixation des effecteurs intracellulaires (Lefkowitz 1978) et vont moduler la réponse des récepteurs. Ces ligands allostériques des RCPG sont généralement de petites molécules, se liant à la face extracellulaire des récepteurs (May and Christopoulos 2003)

### 1.2.7. Oligomérisation des RCPG

De nombreuses classes de récepteurs, notamment à activité enzymatique, sont connues pour leur capacité à s'assembler en oligomères. Cette propriété est même généralement essentielle pour leur fonction (par ex, récepteurs tyrosine kinase).

Pour les RCPG, le modèle monomérique a longtemps prévalu : 1 récepteur / 1 ligand / une protéine G. Pourtant, dès 1982, des résultats suggérant la possibilité d'oligomérisation des RCPG ont été publiés (Conn PM 1982). De nombreuses démonstrations de l'existence de d'oligomères de RCPG sont proposées dans la littérature (Milligan 2004; Terrillon and Bouvier 2004). Il pourrait également exister des équilibres d'homo-oligomérisation, (entre récepteurs de la même famille) ou d'hétéro-oligomérisation (entre récepteurs de familles différentes) ce qui ouvre des perspectives combinatoires extrêmement vastes.

La capacité des RCPG à former des oligomères pourrait être directement liée à leur fonction. L'oligomérisation pourrait également jouer un rôle dans le trafic cellulaire des RCPG, dans l'activation des protéines G ou dans l'inhibition des récepteurs. Plusieurs études suggèrent que la liaison d'un ligand pourrait promouvoir ou inhiber la dimérisation contribuant ainsi aux processus de régulation (Terrillon and Bouvier 2004).

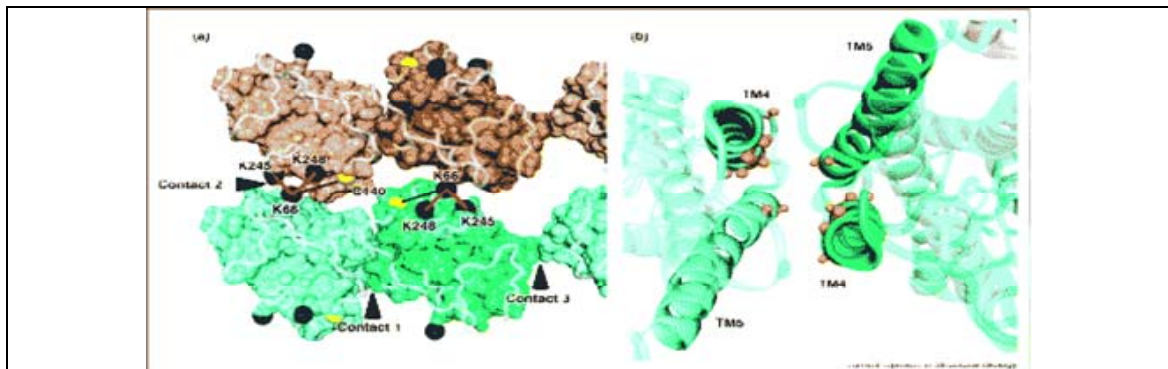
Parallèlement, de nombreuses données indiquent que la liaison d'une seule protéine G suffit pour provoquer l'activation du récepteur. La dimérisation n'est donc pas nécessaire à l'activation du récepteur les protéines G (Chabre and leMaire 2005). Il a également été rapporté récemment le cas de RCPG orphelins n'ayant pas de capacité connue d'activation des protéines G mais pouvant former des dimères avec d'autres RCPG et dont le seul rôle pourrait être de moduler leur réponse (Levoye, Dam et al. 2006).

Généralement l'interprétation des données expérimentales sur les oligomères de RCPG est ambiguë car elle ne permet pas par exemple de différencier les oligomères transitoires et permanents ou de révéler les situations où l'oligomérisation joue un rôle fonctionnel (Gurevich and Gurevich 2008; Gurevich and Gurevich 2008). Il n'existe donc à l'heure actuelle aucune certitude sur le rôle physiologique de cette dimérisation

Les seules données structurales disponibles au sujet de la dimérisation des RCPG concernent des dimères de rhodopsine. En 2003, en microscopie de force

atomique, ont été observé des rangées de rhodopsine dans les membranes de rétine native (D. Fotiadis 2003; Liang, Fotiadis et al. 2003). Les données de microscopie ainsi que la structure cristallographique de la rhodopsine ont ainsi permis de reconstruire un modèle (Fotiadis D 2006) à l'échelle atomique des oligomères de rhodopsine ((Fotiadis D 2006), Figure 7).

Toutefois, la signification biologique de ces résultats reste débattue (Chabre and leMaire 2005), certains auteurs étant convaincus que les rangées observées pourraient n'être qu'un artefact de la microscopie de force atomique (Chabre, Cone et al. 2003). L'examen de ce modèle montre en effet que la surface de l'interface de dimérisation est extrêmement faible. Une simulation par dynamique moléculaire en environnement membranaire explicite de ce modèle (Filizola, Wang et al. 2006) indique que le modèle n'est pas stable et que les deux monomères ont tendance à se rapprocher l'un de l'autre au cours de la simulation. Enfin, des études calorimétriques récentes montrent que la rhodopsine se comporte essentiellement comme un monomère au sein de ces mêmes disques de membranes de rétine native (Edrington, Bennett et al. 2008).



**Figure 7 : Modélisation de rangées de rhodopsines à partir de données AFM**

*Ce modèle est construit par ajustement de la structure cristallographique de la rhodopsine à des données de microscopie de force atomique. Les auteurs observent la formation de rangées de récepteurs dans les disques de membrane native.  
D'après Fotiadis (Fotiadis D 2006)*



### ***1.3. Modélisation des RCPG***

#### **1.3.1. Modélisation comparative**

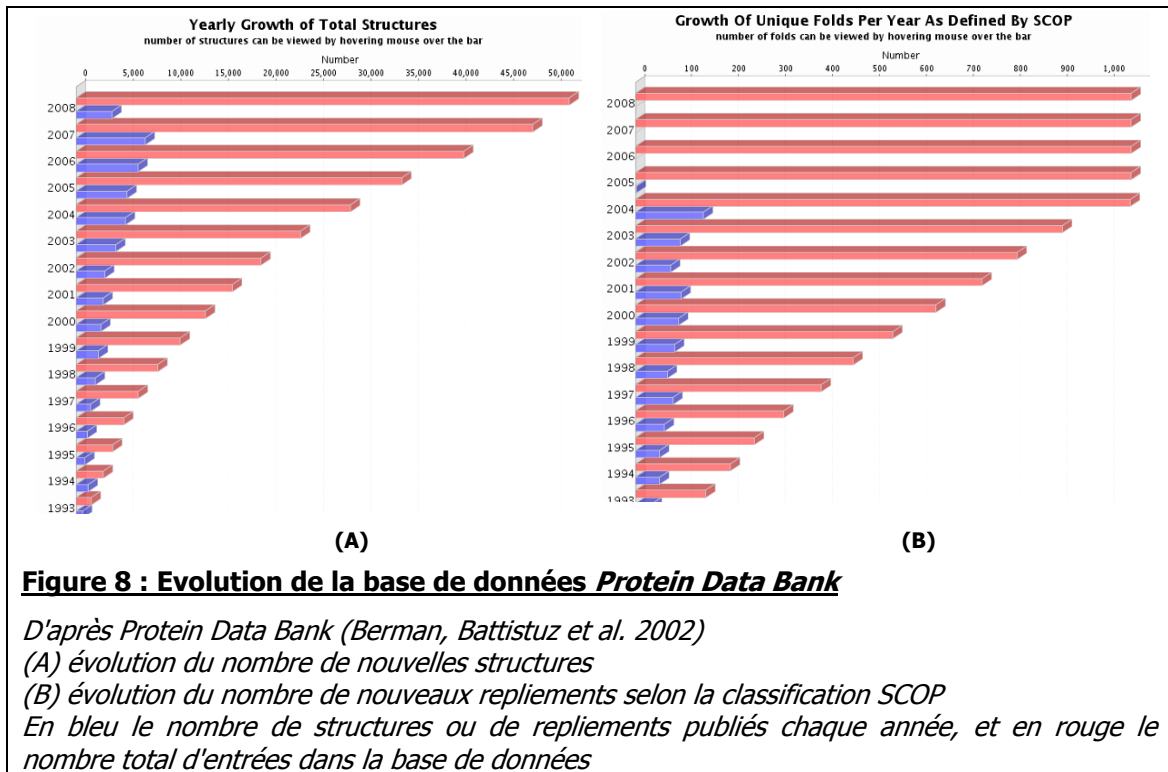
Il est aujourd'hui admis que la structure 3D des protéines d'une même famille est plus conservée que leur séquence primaire (Chothia and Lesk 1986) et donc que des homologues adoptent généralement un repliement similaire. Ceci ouvre la voie de la modélisation comparative. Quand la structure d'une protéine est déterminée expérimentalement, les autres membres de la famille peuvent alors être modélisés sur la base de leur alignement avec la protéine de repliement connue.

Les conditions à remplir pour pouvoir construire un modèle par homologie sont :

- une similarité de séquence suffisante entre la protéine de structure connue et la séquence de celle à prédire. Classiquement, il est admis que la limite basse pour obtenir un modèle correct se situe autour de 30%.
- un alignement fiable des séquences des deux protéines

Compte tenu du fait qu'approximativement 57% de la totalité des séquences connues ont au moins un domaine apparenté de structure connue dans les bases de données (Chance, Fiser et al. 2004), cette approche se révèle donc particulièrement puissante. De plus, alors que la détermination de nouvelles structures ne cesse de s'accélérer, la découverte de repliements originaux a fortement ralenti ces dernières années (Figure 8). La modélisation comparative devient de fait d'autant plus utile et largement applicable.

Il faut toutefois noter que ces chiffres concernent les protéines solubles. Les protéines membranaires sont largement sous-représentées dans les bases de données en regard de leur abondance in vivo. On ne recense en effet parmi les 50.000 structures de protéines disponibles à ce jour que seulement 178 protéines membranaires (Raman, Cherezov et al. 2006).



### 1.3.2. Retour d'expérience sur les modèles par homologie de RCPG

#### 1.3.2.1. Modèles de la rhodopsine

Les premiers essais de modélisation de RCPG ont été réalisés en 1993 sur la base de cartes de diffraction électronique à très basse résolution de la rhodopsine (Baldwin 1993). En 1997, la structure d'une protéine à 7 segments transmembranaires a été publiée : la bactériorhodopsine (Pebay-Peyroula, Rummel et al. 1997). Il s'agit d'une pompe à proton bactérienne activée par la lumière. Bien que ne présentant aucun lien phylogénétique avec les RCPG, cette protéine présente une topologie et une fonction d'une proximité troublante avec la rhodopsine. Elle a donc fut été utilisée comme référence pour modéliser les RCPG par modélisation comparative.

La publication en 2000 de la structure cristallographique de la rhodopsine bovine (Palczewski, Kumasaka et al. 2000) a permis d'avoir un premier retour sur l'efficacité de ces approches de prédiction de la structure des RCPG. Le modèle de Baldwin proposé en 1993 (Baldwin 1993) présente, au niveau des hélices transmembranaires un écart de seulement 3.2Å par rapport à la structure cristallographique. En revanche, les modèles basés sur la bactériorhodopsine sont beaucoup plus éloignés de la structure native de la rhodopsine (Oliveira, Hulsen et al. 2004).

### 1.3.2.2. Modèles de $\beta_2$ AR

Suite à la détermination de la structure de la rhodopsine, de nombreux modèles de RCPG de classe A ont été construits par homologie sur cette base. Le pourcentage d'identité de séquences entre la rhodopsine et les autres RCPG n'est dans la plupart des cas que de l'ordre de 20%. Avec un pourcentage d'identité inférieur à 30%, la difficulté à déterminer un alignement fiable des séquences entre deux homologues limite la qualité des modèles prédits. Les RCPG de classe A font cependant exception à cette règle car ils possèdent, au niveau de chaque hélice, un ou plusieurs résidus très fortement conservés permettant l'alignement non ambiguë de leurs séquences (Baldwin, Schertler et al. 1997).

La publication récente de la structure du  $\beta_2$ AR (Cherezov, Rosenbaum et al. 2007) permet aujourd'hui d'évaluer les divergences entre les modèles publiés et la structure réelle du récepteur. On peut ainsi en tirer des enseignements sur la modélisation comparative des récepteurs couplés aux protéines G (Costanzi 2008). Comme on pouvait s'y attendre, les modèles de  $\beta_2$ AR prédits d'après la structure de la rhodopsine sont en fait plus proches de la structure de la rhodopsine, que de celle du récepteur adrénergique cristallisé. Ceci dit, ces modèles restent d'une très bonne précision, puisque la comparaison des modèles avec la structure cristallographique du  $\beta_2$ AR présentent une divergence de moins de 2.8 Å au niveau des hélices transmembranaires.

La rhodopsine ne possède aucune activité à l'état de repos (en absence de lumière) dans lequel sa structure a été déterminée. Ceci valide donc son utilisation pour construire des modèles de RCPG liés à des molécules antagonistes. Mais la comparaison des deux structures de rhodopsine /  $\beta_2$ AR permet également d'avancer dans la compréhension des relations structure/fonction. Le récepteur  $\beta_2$ AR en effet possède une activité basale, indépendante de la liaison de ligand. La comparaison des deux structures et de leurs différences, notamment dans les régions distinctes des sites de liaison, permettra donc d'obtenir des indices sur les bases structurales de cette différence fonctionnelle (Kobilka and Schertler 2008). Par exemple la superposition des structures sur la base de l'alignement des séquences rhodopsine/  $\beta_2$ AR révèle des

différences structurales autour du motif très conservé E/DRY (coté cytoplasmique de TM3) qui pourrait expliquer les différences dans l'activité basale des deux récepteurs.

#### 1.3.2.3. Criblage in silico de RCPG

Une autre forme de validation des approches de modélisation de RCPG est apportée par les expériences de criblage in silico de bibliothèques de ligands sur des modèles de RCPG. Les expériences de criblage à haut débit consistent à rechercher de nouveaux ligands pour un récepteur donné en testant in vitro de très nombreux composés issus d'une chimiotèque. Afin de réduire le cout de ce type de travaux, il est très intéressant d'effectuer un tri préliminaire in silico de la chimiotèque afin de sélectionner un nombre restreint de composés ayant de bonnes probabilités de liaison au récepteur étudié. La structure cristallographique de la rhodopsine a permis les premiers travaux de criblage de RCPG. En particulier, Bissantz et al (Bissantz, Bernard et al. 2003) ont construit des modèles de récepteurs M1, D3, V1a,  $\mu_2$  et  $\delta$ -opioïd. Ces modèles ont été utilisés pour filtrer des banques de petites molécules chimiques. Cette étude montre que pour certains de ces modèles, il est possible d'enrichir de 30 à 40 fois les jeux de données de molécules chimiques en ligands potentiels.

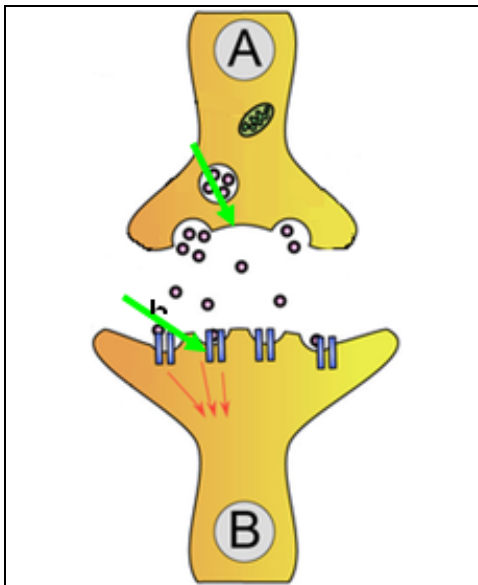
Plus récemment, le même groupe s'est intéressé à l'impact de la boucle extracellulaire E2 sur de telles expériences de criblage in silico (de Graaf, Foata et al. 2008). Ils montrent ainsi que l'utilisation de modèles dans lesquels la boucle E2 est construite avec soin sur la base de données expérimentales permet d'améliorer la capacité de ces modèles à prédire la liaison de ligands.

### 1.3.3. Modélisation de novo

Certains auteurs ont réalisé des modèles de novo de la structure des RCPG en se basant uniquement sur leurs séquences. C'est le cas par exemple de la méthode Membstruck (Vaidehi, Floriano et al. 2002; Li and Goddard 2008) particulièrement focalisée sur le traitement des passages transmembranaires. Après prédiction des régions transmembranaires dans la séquence, des hélices canoniques sont construites avec des chaînes latérales étendues. Ensuite, après une phase de recherche en corps rigides, l'empilement de ces hélices est optimisé par une procédure faisant appel à différentes étapes de minimisation et de dynamique moléculaire, notamment avec des lipides explicites. Les auteurs annoncent des prédictions situées à moins de 3Å de RMSD sur la rhodopsine ou la bactériorhodopsine et de bonnes corrélations avec les données de mutagénèse. Toutefois, il est difficile de juger des performances prédictives réelles de cet outil en raison du manque de données structurales pour valider les prédictions.

Plus récemment, des développements pour la prédiction des protéines transmembranaires ont été implémentés dans Rosetta (Barth, Schonbrun et al. 2007). Cette approche reprend entre autre le traitement par fragments ainsi que le champ de force de Rosetta, adapté afin de modéliser les spécificités de l'environnement membranaire. La membrane est dans ce cas traitée de manière implicite comme un milieu à trois phases : eau, tête polaire des lipides et région hydrophobe. Enfin, les irrégularités dans les structures des hélices telles que les coudes à proline, font l'objet d'un traitement spécifique.

### 1.4. Les récepteurs muscariniques



**Figure 9 : Synapse cholinergique**

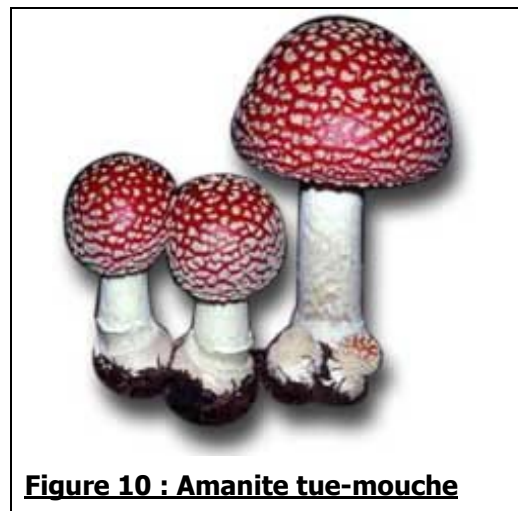
*L'acétylcholine est sécrétée par le neurone afférent A par exocytose. Elle se lie sur les récepteurs localisés au niveau de la dendrite du neurone efférent B*

L'acétylcholine est l'un des neuromédiateurs les plus étudiés, du fait de sa distribution ubiquitaire et de sa capacité à réguler de nombreux processus physiologiques essentiels. Les synapses cholinergiques (Figure 9) sont présentes aussi bien dans le système nerveux central qu'au niveau de la jonction neuromusculaire.

Il existe deux familles de récepteurs liant l'acétylcholine: Les récepteurs nicotiques sont sensibles à la nicotine et appartiennent à la famille des récepteurs ionotropes. Ce sont des canaux membranaires dont l'ouverture est commandée par la liaison de l'acétylcholine. Leur activation va donc influencer sur la perméabilité de la membrane à

certaines ions et modifier ses propriétés électro-physiologiques.

Les récepteurs muscariniques tirent leur nom d'un de leur agoniste, la muscarine, découverte au milieu du 19<sup>ème</sup> siècle, dans le champignon appelé *Amanita muscaria* (Figure 10). C'est d'ailleurs la muscarine, présente dans le chapeau du champignon qui provoque la mort des mouches et lui vaut son nom vernaculaire d'amanite tue-mouches. Les récepteurs muscariniques font partie de la catégorie des récepteurs métabotropes dont l'activation s'accompagne du déclenchement d'une cascade de signalisation impliquant des seconds messagers.

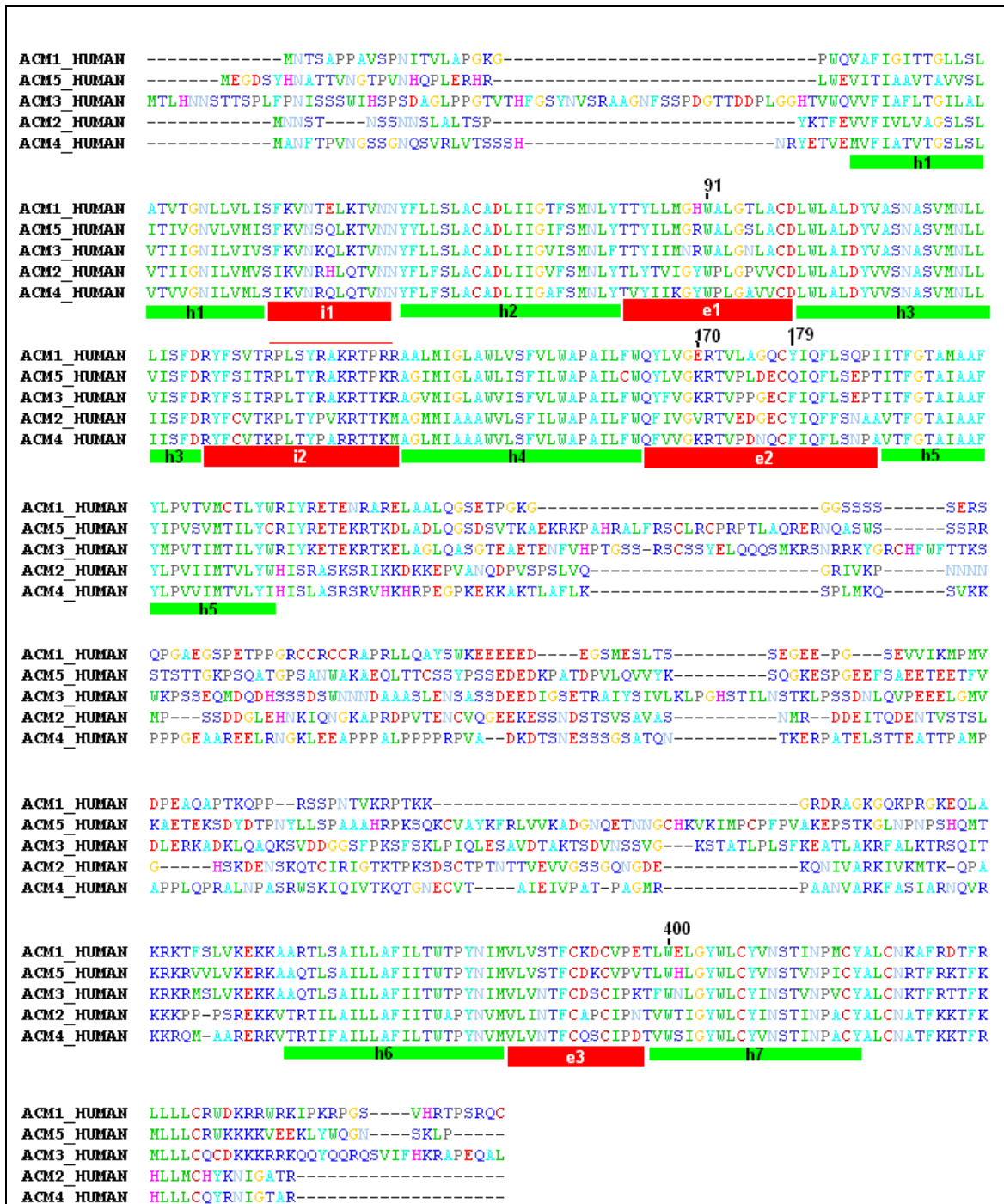


**Figure 10 : Amanite tue-mouche**

### 1.4.1. Sous-types de récepteurs muscariniques

Il est apparu à la suite du séquençage du génome humain qu'il existait chez l'homme plusieurs sous-types de récepteurs muscariniques, avec de très hautes similarités de séquences (Figure 11). A l'heure actuelle, 5 sous-types de récepteurs muscariniques sont connus chez l'homme. Alors que ceux portant des numéros pairs (1, 3, 5) sont préférentiellement couplés à la voie de l'inositol-triphosphate, ceux portant des numéros impairs sont eux plutôt couplés à la voie de l'AMP cyclique.

Localisation des récepteurs muscariniques a été intensivement étudiée (Abrams, Andersson et al. 2006). Ils sont distribués dans de nombreux organes tels que la vessie, l'intestin, les glandes salivaires, le cœur, l'œil, etc. Une attention particulière a été portée à leur distribution dans le cerveau (Volpicelli and Levey 2004) où ils jouent, entre autre, des rôles dans l'excitabilité neuronale, la plasticité synaptique ainsi que les mécanismes de rétroaction sur l'acétylcholine. Les 5 récepteurs sont exprimés dans différentes régions du cerveau. Les sous-types M1, M4 et M5 sont par exemple localisés au niveau du système nerveux central tandis que les sous-types hM2, 3 sont présents au niveau du système nerveux périphérique.



**Figure 11 : Alignement des séquences des 5 sous-types de récepteurs muscariniques**

*Les hélices transmembranaires sont signalées par les bandes vertes*



Une des conséquences de la très forte similarité de séquence dans la famille des récepteurs muscariniques (mAChR) est qu'il n'existe que peu de ligands connus ayant un degré de sélectivité suffisant pour cibler uniquement un sous-type. De plus, les tissus expriment souvent plusieurs sous-types de mAChRs simultanément. Identifier les propriétés pharmacologiques et les fonctions physiologiques précises de chaque sous-type de récepteur est donc un problème difficile. Ceci impose des stratégies d'études indirectes et très lourdes (Wess, Duttaroy et al. 2003). Des travaux récents de knock-out chez la souris (Wess, Eglen et al. 2007) ont cependant permis d'identifier le rôle potentiel de chaque sous-type de récepteur dans diverses pathologies (Figure 12). De nombreux ligands des récepteurs muscariniques utiles dans le traitement de ces maladies sont déjà disponibles à l'heure actuelle. Mais, ces ligands manquant généralement de sélectivité, ils activent ou inhibent plusieurs sous-types de mAChR simultanément. L'usage clinique de ces molécules se trouve donc limité par des effets secondaires.

Clinical condition	Drug
Alzheimer's disease Cognitive impairment	M <sub>1</sub> , M <sub>3</sub> or mixed M <sub>1</sub> /M <sub>3</sub> agonist; M <sub>2</sub> antagonist <sup>b</sup>
Sjögren's syndrome	M <sub>1</sub> , M <sub>3</sub> or mixed M <sub>1</sub> /M <sub>3</sub> agonist
Schizophrenia	M <sub>1</sub> , M <sub>4</sub> or mixed M <sub>1</sub> /M <sub>4</sub> agonist
Parkinson's disease	M <sub>1</sub> , M <sub>4</sub> or mixed M <sub>1</sub> /M <sub>4</sub> antagonist
Type 2 diabetes	M <sub>1</sub> agonist (peripherally acting)
Obesity	M <sub>1</sub> antagonist (centrally acting)
Peptic ulcer disease <sup>c</sup>	M <sub>3</sub> or mixed M <sub>3</sub> /M <sub>5</sub> antagonist
OAB, COPD	M <sub>3</sub> antagonist
Irritable bowel syndrome Gastrointestinal spasms	M <sub>1</sub> or mixed M <sub>1</sub> /M <sub>3</sub> antagonist
Antinociception	M <sub>4</sub> agonist
Wound healing	M <sub>4</sub> agonist, M <sub>1</sub> antagonist
Cerebrovascular insufficiency	M <sub>3</sub> agonist
Drug addiction and withdrawal	M <sub>1</sub> antagonist

**Figure 12 : Implication des récepteurs muscariniques dans les pathologies humaines**

*Les récepteurs muscariniques constituent des cibles potentielles pour le traitement de nombreuses pathologies. D'après Wess (Wess, Eglen et al. 2007)*

### **1.4.2. Le sous-type 1 des récepteurs muscariniques**

Les pathologies telles que les maladies d'Alzheimer ou de Parkinson se caractérisent par des dysfonctionnements des mécanismes neuronaux liés à la réponse à l'acétylcholine. Dans le cas de la maladie d'Alzheimer (Iwata, Higuchi et al. 2005), l'accumulation de  $\beta$ -amyloïde provoque une insensibilité des neurones à l'acétylcholine. Dans le cas de la maladie de Parkinson, ce sont les neurones dopaminergiques qui sont détruits, ce qui entraîne une dérégulation de la sécrétion d'acétylcholine (Langmead, Watson et al. 2008). Le sous-type des récepteurs muscariniques humains hM1 est localisé au niveau du système nerveux central et connu pour son implication dans les mécanismes de mémorisation et d'apprentissage. Il constitue donc une cible potentiellement intéressante pour le traitement de ces maladies neuro-dégénératives.

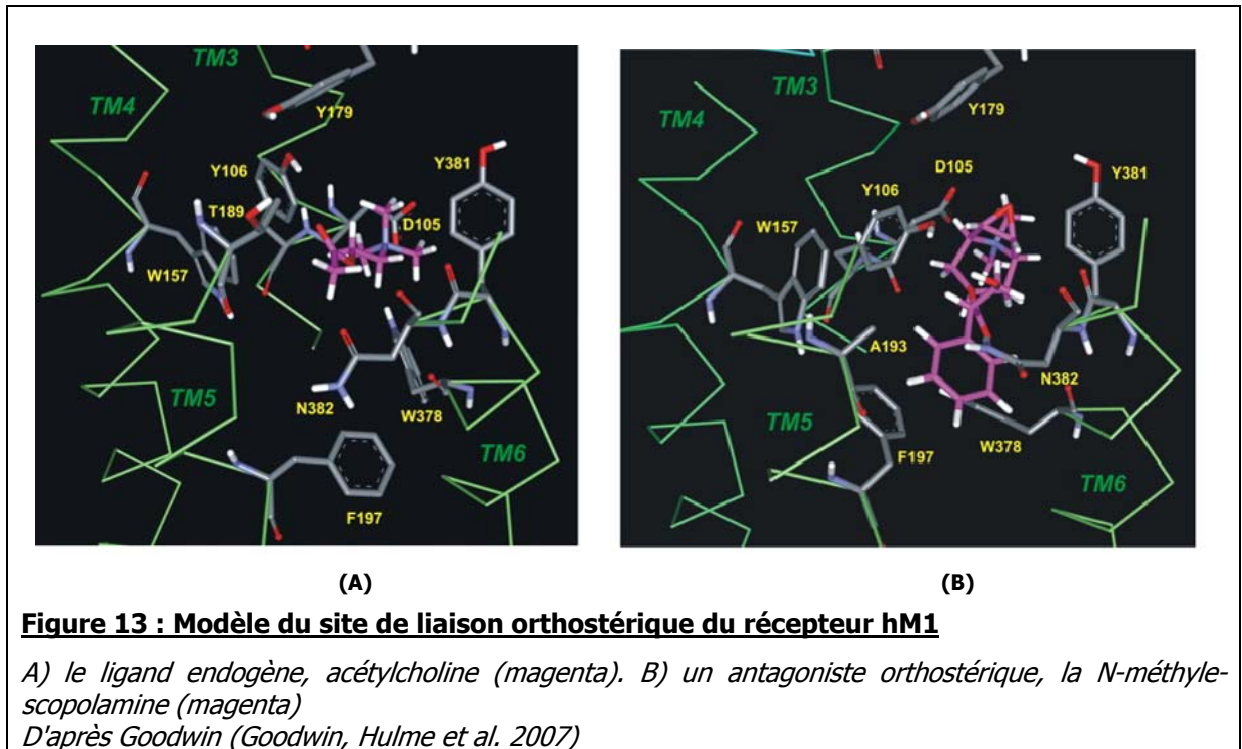
### **1.4.3. Caractérisation des sites de liaison des récepteurs muscariniques**

#### **1.4.3.1. Site orthostérique**

De nombreuses études portent sur la caractérisation des sites de liaison des récepteurs muscariniques. Il s'agit pour la plupart d'études de mutagenèse de ces récepteurs visant à déterminer les résidus importants pour l'interaction avec divers ligands. En particulier, l'équipe de Hulme a réalisé de nombreux mutants du récepteur hM<sub>1</sub> (Hulme and Lu 1998; Hulme, Lu et al. 1999; Lu and Hulme 1999; Ward, Curtis et al. 1999; Allman, Page et al. 2000; Lu and Hulme 2000; Lu, Saldanha et al. 2001; Hulme, Lu et al. 2003; Bee and Hulme 2007; Goodwin, Hulme et al. 2007) tandis que celle de Wess (Schmidt, Li et al. 2003; Han, Hamdan et al. 2005; Li, Nowak et al. 2005; Li, Han et al. 2007; Scarselli, Li et al. 2007; Li, Hamdan et al. 2008) se focalisait sur le récepteur hM<sub>3</sub>.

Les positions des résidus importants pour l'interaction récepteur / ligand orthostérique peuvent être cartographiées sur des modèles structuraux de ces récepteurs (Figure 13). Il ressort de ces études que la cavité de liaison de l'Ach sur les récepteurs muscariniques est en grande partie superposable avec son équivalent dans la

rhodopsine : la cavité de liaison du rétinale, cette cavité étant localisée profondément au cœur du faisceau d'hélices.



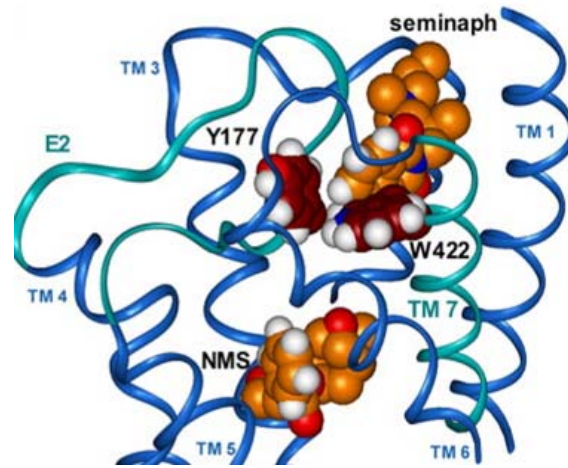
#### 1.4.3.2. Site allostérique

Les modulateurs allostériques des récepteurs muscariniques ont eux été largement étudiés par les équipes de Mohr (Mohr, Trankle et al. 2003; Trankle, Weyand et al. 2003; Huang, Prilla et al. 2005; Trankle, Dittmann et al. 2005; Grossmuller, Antony et al. 2006; Holzgrabe, De Amici et al. 2006; Jager, Schmalenbach et al. 2007) et Birdsall (Scarselli, Li et al. 2007). Ces études montrent que la liaison des ligands allostériques intervient au niveau des régions extracellulaires du récepteur et en particulier au niveau de la boucle extracellulaire E2 (Figure 14).

La conclusion de ces travaux, est qu'il existe probablement plusieurs sites de liaison des ligands allostériques.

- Le premier site, lie les modulateurs classiques tels que la gallamine, W84, ... Ce site est localisé au niveau des boucles extracellulaires E2 et E3 ainsi qu'au niveau des régions extracellulaires des hélices.
- D'autre part, il existe un certain nombre de ligands aux propriétés pharmacologiques atypiques (Duo3, tacrine). Ces résidus se lient à un ou plusieurs sites distincts du site commun.

Enfin, des travaux récents d'arrimage *in silico* (Espinoza-Fonseca and Trujillo-Ferrara 2005; Espinoza-Fonseca and Trujillo-Ferrara 2006) semblent suggérer qu'il puisse exister en réalité de multiples sites de liaisons de ligands allostériques sur le récepteur hM<sub>1</sub>. Toutefois, ces résultats ne permettent pas de conclure sur la correspondance des sites de liaison d'un sous-type de récepteur muscarinique à l'autre.



**Figure 14 : Liaison de ligands ortho et allostérique au récepteur muscarinique M2**

*Les deux ligands sont en jaune, les résidus Y177 et W422 du récepteur muscarinique M2 en rouge. D'après Jager (Jager, Schmalenbach et al. 2007)*

## 1.5. Les toxines muscariniques



**Figure 15 : Mamba vert**

Le venin du Mamba vert (Figure 15) contient de nombreuses toxines ayant des propriétés physiologiques intéressantes. On y trouve entre autres, des inhibiteurs de cholinestérase, de fasciculine (Karlsson, Mbugua et al. 1985), des bloqueurs de canaux potassiques (Harvey and Karlsson 1982) de canaux calciques (de Weille JR 1991) et des toxines ciblant les récepteurs muscariniques à l'acétylcholine (Karlsson, Risinger et al. 1991). Ces toxines sont en grande majorité de petites protéines de moins de 100 résidus résistantes

aux protéases mais n'ayant pas d'effet enzymatique.

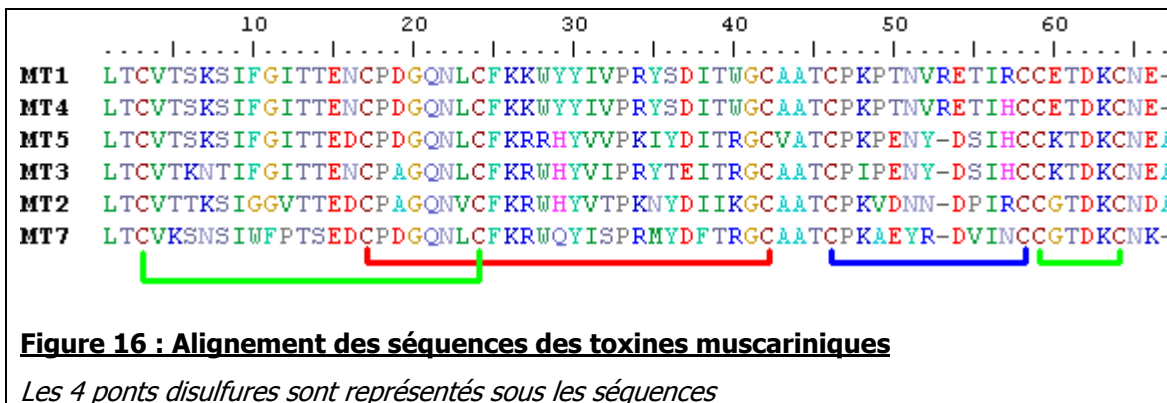
### 1.5.1. Synthèse

Les toxines muscariniques sont présentes en quantité très faible dans les venins de serpents, ce qui rend leur purification particulièrement difficile. Des systèmes d'expression en cellule ont été mis en place, mais ils ne permettent souvent que de produire des quantités limitées de toxine.

Un protocole de synthèse et de repliement des toxines muscariniques a été mis au point au CEA par G. Mourier et al (Mourier, Dutertre et al. 2003). La prise en compte des états d'oxydation de la toxine a joué un rôle majeur dans la capacité à produire une protéine correctement repliée. Au final, cette approche est basée sur une synthèse chimique en une étape en phase solide et permet d'obtenir des quantités importantes de peptides: de 3 à 5 mg de toxine MT7 par synthèse avec un rendement de l'ordre de 25% après purification.

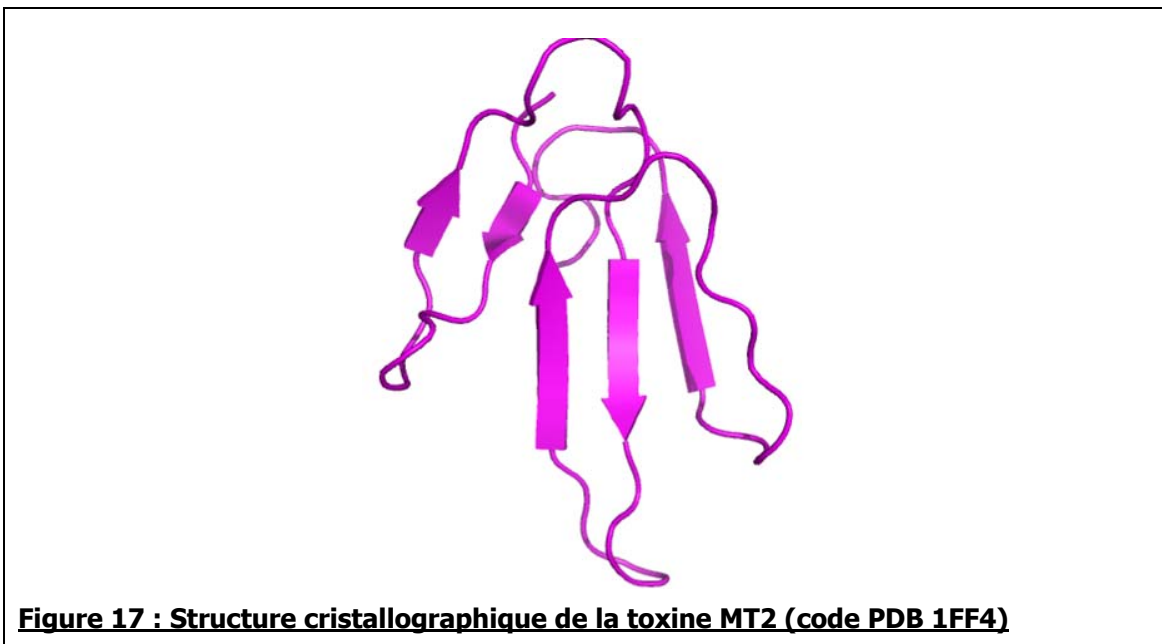
### 1.5.2. Structure

Les toxines muscariniques sont une famille de peptides de 65-66 résidus (Figure 16) à 4 ponts disulfures. On en connaît à l'heure actuelle une douzaine au total, mais toutes n'ont pas été séquencées. La structure de la toxine muscarinique MT2 (Figure 31) a été résolue par RMN (Segalas, Roumestand et al. 1995) puis par cristallisation et diffraction aux rayons X à une résolution de 1.50 Å (code PDB 1FF4). Cette structure présente un repliement à trois doigts organisés en cinq brins  $\beta$  antiparallèles interconnectés par trois boucles (I, II et III). Le cœur de la protéine est stabilisé par quatre ponts disulfures. Ce motif structural est courant dans de nombreuses familles de toxines de serpents (Galat, Gross et al. 2008), comme les  $\alpha$ -neurotoxines (Popot and Changeux 1984), les fasciculines (le Du, Marchot et al. 1992) ou les cardiotoxines.



**Figure 16 : Alignement des séquences des toxines muscariniques**

Les 4 ponts disulfures sont représentés sous les séquences



### 1.5.3. L'interaction toxine MT7 – récepteur hM1

La famille des toxines muscariniques est remarquable par le fait que, bien qu'ayant de très fortes similarités de séquences (Figure 16), elles possèdent des profils de sélectivité très différents pour les 5 sous-types de récepteurs muscariniques humains (Karlsson, Jolkkonen et al. 2000).

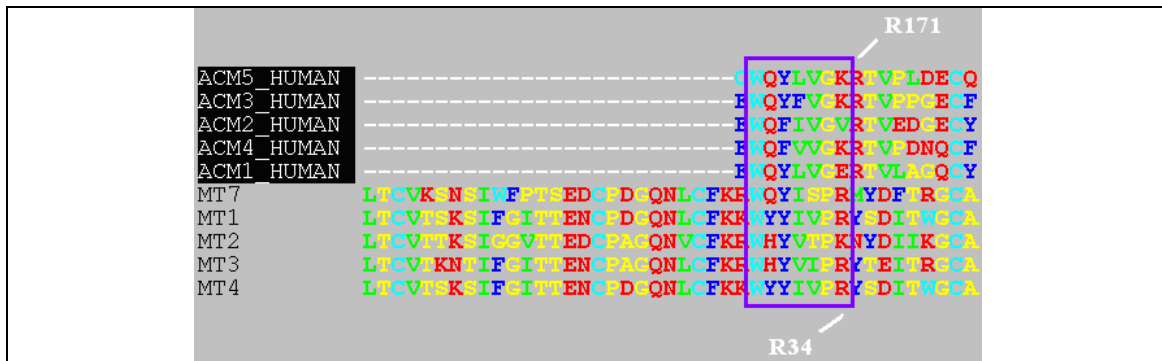
Avec une affinité subnanomolaire pour le hM1 et plus que micromolaire pour les autres récepteurs muscariniques, la MT7 est la seule toxine à être vraiment sélective d'un sous-type de récepteur muscarinique (Figure 18). La caractérisation pharmacologique de l'interaction de la MT7 sur le récepteur hM1 (Olianas, Maullu et al. 2000; Olianas, Adem et al. 2004; Fruchart-Gaillard, Mourier et al. 2006; Fruchart-Gaillard, Mourier et al. 2006) établit que cette toxine se lie au récepteur sans provoquer son activation, ce qui en fait un ligand antagoniste. D'autre part la MT7 possède la capacité de ralentir de manière non compétitive la cinétique de dissociation d'un antagoniste orthostérique : le NMS. La toxine MT7 est donc un ligand antagoniste allostérique du récepteur M1. La liaison de la MT7 avec les récepteurs muscariniques intervient sur la face extracellulaire des récepteurs, et implique ses boucles (Kukkonen, Perakyla et al. 2004; Fruchart-Gaillard, Mourier et al. 2006).

<i>Subtype</i>		<i>Affinity (Ki) nM</i>				
		<i>M1</i>	<i>M2</i>	<i>M3</i>	<i>M4</i>	<i>M5</i>
<b>M<sub>1</sub>-toxin 1</b>	MT7	0,1	> 2000	> 2000	> 2000	> 2000
	MT1	22*–49	> 1000	> 1000	29*–58	> 1000
	MT2	630	> 2000	> 2000	1900	> 2000
	MT4	62	> 1000	> 1000	87	> 1000
	MT5	180	> 1000	> 1000	540	> 1000
<b>M<sub>2</sub>-toxin</b>	nb	nb	150	nb	nb	nb
<b>M<sub>4</sub>-toxin 1</b> <b>2</b>	MT3	78–300*	> 1000	> 1000	1,4*–2	> 1000
	MT6	190	> 425	> 425	3,6	> 425
	MT $\alpha$	23	44	3	5	8
	MT $\beta$	> 1000	> 2000	140	120	350

**Figure 18 : Affinité des toxines muscariniques pour les différents sous-type de récepteurs**

*D'après (Karlsson, Jolkkonen et al. 2000)*





**Figure 19: Alignement de séquence de la boucle II des toxines et de la boucle E2 des récepteurs**

On peut remarquer une similarité de séquence troublante entre la boucle II de la toxine et la boucle E2 du récepteur (Figure 19) qui pourrait suggérer que la toxine pourrait déplacer la boucle du récepteur et prendre sa place.

### ***1.6. Prédiction de la structure des complexes protéine-protéine***

Au-delà des structures isolées, les complexes de protéine apparaissent de plus en plus comme des éléments clefs dans la compréhension du fonctionnement du vivant. Ainsi de nombreuses campagnes de détection des complexes protéine-protéine à l'échelle d'un génome ont vu le jour (Janin and Seraphin 2003). Ces études reposent sur des techniques telles que le double hybride ou le tap-tag (Rigaut, Shevchenko et al. 1999) qui peuvent être conduites à très haut débit et aboutissent à la détection de plusieurs dizaines de milliers d'interactions.

Il n'est pas envisageable de déterminer expérimentalement toutes les structures de ces complexes. En effet, les programmes d'étude structurale des complexes de protéines reposants sur la cristallographie, la RMN, ou la microscopie électronique sont efficaces mais d'un débit relativement faible. Résoudre la structure d'une protéine isolée est déjà un exercice difficile car cela implique d'arriver à produire la protéine en grande quantité et à la purifier que ce soit pour la RMN, la cristallographie ou même la microscopie électronique. Une autre étape limitante propre aux études cristallographiques est l'obtention de cristaux diffractant les rayons X à une résolution suffisante.

La détermination expérimentale de la structure d'un complexe présente plusieurs difficultés supplémentaires : les partenaires du complexe doivent généralement être produits conjointement et dans un rapport stœchiométrique précis pour s'assurer de la formation du complexe ce qui complique les étapes de production et de purification. De plus les énergies de liaison des protéines du complexe sont bien souvent insuffisantes pour leur permettre de rester associées dans les conditions de purification et de cristallisation. Enfin, chaque protéine ayant parfois plusieurs partenaires différents, le manque d'un ou de plusieurs partenaires peut rendre le complexe instable, insoluble et donc difficile à produire et à purifier. Dans le même temps, la base de données de structures de protéines PDB (Berman, Battistuz et al. 2002), voit son nombre d'entrées continuer de croître de manière exponentielle avec plus de 50.000 structures disponibles à ce jour. Cet essor est dû en partie aux programmes de génomique structurale visant à déterminer de manière quasi-automatique la structure du plus grand nombre possible de protéines d'un organisme. L'étude systématique de la levure a par exemple déjà permis

d'obtenir des cristaux pour 25 protéines et de déterminer la structure de 15 protéines isolées.

Le recours à des méthodes informatiques de prédiction de complexes basées sur les structures disponibles des protéines isolées s'impose comme une nécessité.

### **1.6.1. Nature des interfaces des complexes macromoléculaires**

#### 1.6.1.1. Nature des résidus

Les effets électrostatiques et hydrophobes sont les deux déterminants les plus importants dans la liaison entre les protéines. Au cœur d'un site d'interaction, on trouve généralement un résidu polaire qui constitue un hotspot. Plusieurs auteurs ont analysé la composition de ces hotspots. Ces analyses (Thorn and Bogan 2001; Moreira, Fernandes et al. 2007) révèlent que les trois résidus les plus fréquents sont le tryptophane (21%), l'arginine (13.3%) et la tyrosine (12.3%). Autour d'un hotspot, on trouve des résidus hydrophobes responsables de l'exclusion du solvant du site de liaison, ce qui induit une baisse locale de la constante diélectrique. Les effets électrostatiques, dépendant de la loi de Coulomb se trouvent de ce fait renforcés au niveau du hotspot.

#### 1.6.1.2. Dimension des interfaces de protéines

D'après l'analyse de 75 complexes protéiques (Lo Conte, Chothia et al. 1999), l'équipe de J. Janin a établi que, dans leur grande majorité, les complexes protéine/protéine ont des surfaces de contact dont la taille est de l'ordre de  $1600 \pm 400 \text{ \AA}^2$ . Un second groupe contenant environ le tiers des protéines étudiées rassemble des cas pour lesquels cette surface atteint des valeurs beaucoup plus importantes allant jusqu'à  $4600 \text{ \AA}^2$ . C'est en particulier le cas des protéines G et d'autres protéines impliquées dans les voies de signalisation. Cette même équipe a étudié plus précisément les interfaces présentant des surfaces de contact élevées ( $> 1800 \text{ \AA}^2$ ). La surface de contact a été analysée pour identifier si elle était composée de un ou plusieurs sites. Les interfaces composées d'un seul site de contact possèdent une surface de l'ordre de  $1600 \text{ \AA}^2$  alors que celles composées de deux sites ont une surface de l'ordre de  $2000 \text{ \AA}^2$  (Chakrabarti and Janin 2002).

### 1.6.1.3. Réarrangements structuraux

Le premier modèle proposé pour décrire le mécanisme d'interaction de deux protéines est le modèle dit clef-serrure. Proposé en 1895 par Emil Fischer, il repose sur la complémentarité de forme du site récepteur d'un l'enzyme (la serrure) et du substrat (la clef). Ce modèle historique de reconnaissance d'un ligand par un récepteur, a inspiré les premiers travaux de modélisation de complexes par assemblage de corps rigides. Il ne permet toutefois pas d'expliquer les mécanismes aujourd'hui connus des réarrangements conformationnels intervenant lors de la formation de certains complexes de protéines. En 1975, Burgen et al (Burgen, Roberts et al. 1975) proposèrent deux modèles permettant d'expliquer la liaison de ligands flexibles aux macromolécules:

- le modèle dit "d'ajustement induit" qui postule une adaptation mutuelle des deux structures : l'interaction débute par le contact limité des deux partenaires qui va entraîner des changements conformationnels permettant aux protéines de s'ajuster mutuellement.
- Le modèle de "sélection de conformation" : parmi les conformations du ligand présentes en solution, seule celle qui correspondra exactement au site d'interaction sera capable de former un complexe stable.

Des travaux de dynamique moléculaire ont montré que ces deux derniers modèles n'étaient pas incompatibles mais plutôt qu'ils se complétaient. Grunberg et al proposent un mécanisme en trois étapes : diffusion, sélection d'un conformère et ajustement induit (Grunberg, Leckner et al. 2004).

La présence dans les bases de données de structures de protéines, sous forme complexées et non complexées permet d'évaluer les réarrangements structuraux ayant lieu lors de la liaison des deux partenaires. Cela nous apprend que ces changements sont, dans leur grande majorité, limités à des mouvements de chaînes latérales optimisant la compacité des interfaces et le réseau de liaisons hydrogène (Janin and Chothia 1990). Mais il existe également des mouvements de plus grande amplitude impliquant des changements de conformation du squelette peptidique. Ces mouvements peuvent être de différentes natures: mouvements des boucles, de cisaillement, de charnière, réarrangement des structures quaternaires, etc. Une étude recense 240 types de

mouvement dans la base de données "*molecular movement database*" (Gerstein and Krebs 1998; Flores, Echols et al. 2006).

Les protéines membranaires, et en particulier celles jouant des rôles de canaux ou de transporteurs sont concernées elles aussi par ces mouvements qui peuvent être localisés au niveau des domaines cytosoliques et extracellulaires, mais également directement au cœur de la membrane (Gerstein and Echols 2004 ; Lacapere, Pebay-Peyroula et al. 2007).

Les protéines présentent donc une grande variété de mouvements, souvent essentiels à leur fonction. La capacité de prédire ces mouvements est une étape importante dans la prédiction de la structure des complexes.

### **1.6.2. Prédiction de la structure des complexes de protéine**

Les méthodes informatiques d'arrimage moléculaire ont longtemps reposé largement sur le modèle clef-serrure. Elles procédaient généralement en deux grandes étapes : une première étape de recherche et une seconde étape d'optimisation. Bien qu'utilisant des algorithmes de recherche différents, la quasi totalité des programmes de docking (RosettaDock (Gray, Moughon et al. 2003), Zdoc (Zhang, Liu et al. 2005), FTDock (Jackson, Gabb et al. 1998) utilisent une première étape de filtrage basée sur une recherche de complémentarité de forme en "corps rigides". A l'issue de cette première étape, les modèles retenus sont ensuite raffinés, en particulier afin de prendre en compte la flexibilité des protéines, pour les chaînes latérales des acides aminés et éventuellement du squelette peptidique (HADDOCK (Dominguez, Boelens et al. 2003), RosettaDock (Gray, Moughon et al. 2003).

Au cours de ces dernières années plusieurs auteurs, s'inspirant du modèle de mécanisme d'arrimage des complexes dit de "sélection de conformation" décrit plus haut, ont introduit une étape préliminaire d'exploration de l'espace conformationnel des deux partenaires par dynamique moléculaire (Ehrlich, Nilges et al. 2005; de Vries, van Dijk et al. 2007), dynamique moléculaire activée (Grunberg, Leckner et al. 2004) ou Monte-Carlo (Gray, Moughon et al. 2003). Cette approche pouvant être combinée à une approche par fragments (Rosetta) ou à une analyse par les principaux modes normaux.

Un des facteurs critiques lors des expériences de docking est la capacité à discriminer une structure proche de la réalité des structures inexacts. Pour cela, il existe différentes fonctions de score, basées soit sur des potentiels statistiques empiriques, soit sur des potentiels énergétiques. On peut distinguer deux grands types de fonctions de score :

- Les fonctions de type "physique" qui sont basées sur des lois de la physique atomique. Des termes d'énergie tel que van der Waals, électrostatique, liaisons hydrogènes, solvatation sont utilisés (Mandell, Roberts et al. 2001; Gray, Moughon et al. 2003).
- Les fonctions d'énergies basées sur des statistiques tirées des bases de données structurales. Ces bases s'étant considérablement enrichies ces dernières années, cette approche devient de plus en plus fiable. Par exemple le package 3D-dock (Gabb, Jackson et al. 1997) introduit la notion de RPScore (residue level pair potential score) basé sur des statistiques de contacts entre paires de résidus calibrées sur une base de donnée de complexes de protéines de structures connues. RosettaDock (Lyskov and Gray 2008) utilise également un potentiel empirique ajusté sur des ensembles connus de structures de protéines ou de complexes. (Kuhlman and Baker 2000)

### **1.6.3. Comparaison des méthodes de docking : le concours CAPRI**

Le concours CAPRI (**C**ritical **A**ssessment of **P**redicted **I**nteractions) (Janin, Henrick et al. 2003; Janin 2005; Janin 2005; Janin 2005; Janin 2007; Janin and Wodak 2007) propose aux groupes travaillant sur la prédiction de la structure de complexes d'évaluer les performances de leurs outils. Les cibles proposées aux participants sont des complexes dont la structure vient d'être déterminée et n'est pas encore rendue publique. Cette structure native constituera la référence pour évaluer leurs prédictions.

Les données de départ fournies peuvent être des combinaisons des coordonnées atomiques des formes liées de chaque protéine extraites de celle du complexe, celles des formes non liées déterminées individuellement, ou même des modèles par homologie. A ce jour, 15 cycles du concours ont été effectués avec un total de 36 cibles soumises depuis 2001. Les cibles proposées lors des derniers cycles étaient en majorités des problèmes de type non-lié/non-lié, représentant de ce fait un exercice de prédiction

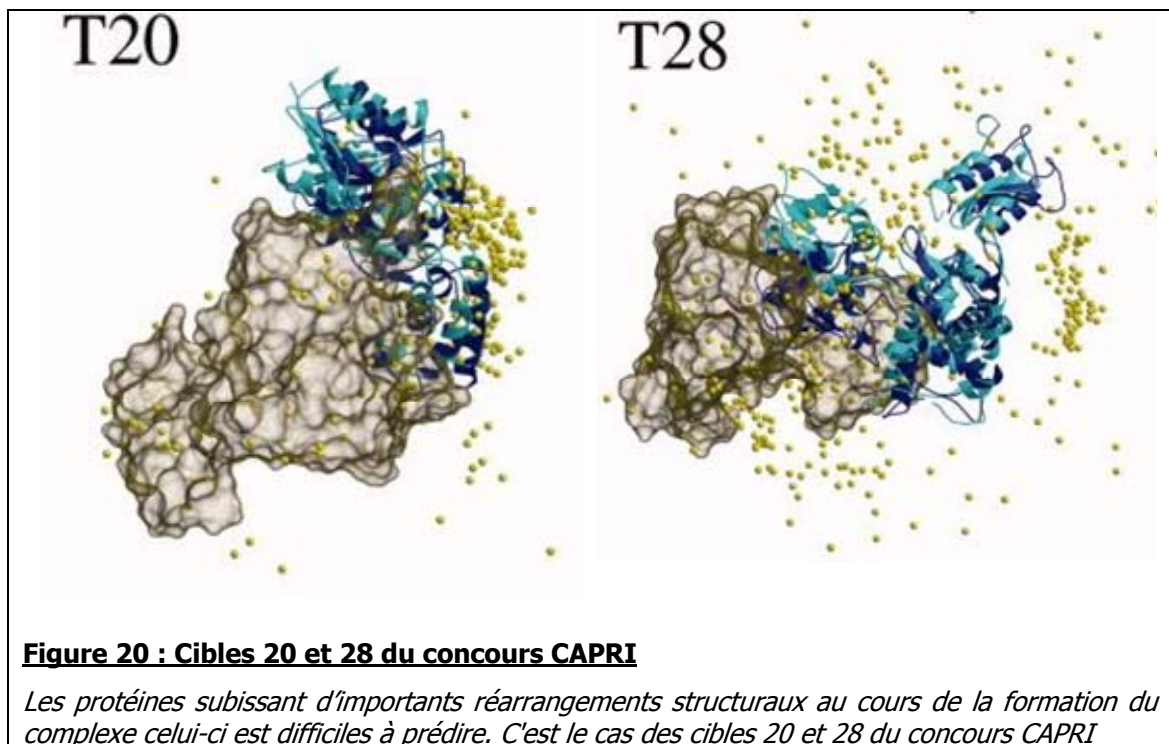
difficile, mais plus réaliste vis à vis des problématiques de la biologie que les problèmes lié/lié des débuts.

Les critères d'évaluation des modèles sont :

- $F_{nat}$  : proportion de contacts natifs prédits / nombre de contacts du complexe natif
- $L_{rms}$  : RMSD sur les atomes du squelette du ligand en superposant les récepteurs prédit et natif
- $I_{rms}$  : RMSD sur les atomes du squelette des résidus à l'interface entre modèle et complexe natif

Pour avoir une référence en terme de RMSD, Chotia et al (C. Chothia 1986) ont établi qu'en moyenne deux déterminations indépendantes de la structure cristallographique d'une même protéine conduisent à des écarts moyen de 0.5 Å. Quant aux structures déterminées par RMN, les valeurs des écarts entre les modèles peuvent atteindre plus de 1.5Å.

Un certain nombre d'enseignements ont pu être tirés de ces concours. Tout d'abord, il apparaît que la prise en compte de la flexibilité des protéines est un enjeu majeur pour les méthodes de prédiction. Autant la flexibilité des chaînes latérales est relativement simple à traiter, autant celle du squelette peptidique reste beaucoup plus difficile à prédire. Ceci apparaît de manière très nette lorsqu'on considère les cibles CAPRI dont les squelettes présentent des mouvements importants. C'est le cas par exemple de la cible T20 dont le site de liaison est composé de boucles mobiles ou de la cible T28 dont un domaine effectue une rotation importante (Figure 20).



Deux approches différentes ont été proposées par les compétiteurs de CAPRI pour prendre la flexibilité des protéines en compte. La première, représentée par les groupes de Bonvin, ou Zacharias consiste à échantillonner des ensembles de conformères pour chaque partenaire du complexe par dynamique moléculaire (Bastard, Prevost et al. 2006; Bonvin 2006). Des paires de conformères de chaque ensemble sont arimées les unes contre les autres. D'un autre côté, les groupes de Wang ou Baker introduisent la flexibilité du squelette durant le processus même du docking en utilisant un échantillonnage par Monte-Carlo d'angles de rotations du squelette à des positions choisies (Wang, Bradley et al. 2007). La dynamique moléculaire ne se révèle efficace que pour les changements conformationnels de faible amplitude et n'impliquant pas le



passage de barrières énergétiques importantes. En revanche, la méthode des angles de rotation peut produire des changements conformationnels importants, mais nécessite une connaissance préalable des régions jouant le rôle de charnière dans la structure.

Globalement, les résultats des prédictions sont bons quand au moins l'une des sous unité du complexe ne subit pas de réarrangement important de son squelette. Dans ce cas, les résidus de l'interface sont identifiés de manière fiable ainsi qu'une proportion importante de contacts natifs (> 30%). Seules les méthodes les plus performantes parviennent à prédire correctement le positionnement des chaînes latérales (Janin and Wodak 2007). Un problème majeur qui reste à résoudre à l'heure actuelle est la capacité des fonctions de score à discriminer certaines interactions spécifiques des molécules d'eau avec les interfaces des complexes. Néanmoins quelques travaux ont proposé des solutions à ce problème. Van Dijk et Bonvin par exemple ont implémenté dans HADDOCK une méthode d'arrimage des structures en état solvate (van Dijk and Bonvin 2006).

La tendance se dégageant des dernières évolutions de cette discipline est l'intégration d'information de nature "non structurale". Il a par exemple été proposé d'utiliser des données cristallographiques en combinaison avec des données de microscopie électronique, des données de couplages dipolaires obtenues en RMN (Tjandra and Bax 1997) ou des mesures de diffraction aux petits angles (SAXS) (Gabel, Simon et al. 2006). Une autre ressource utile peut être d'utiliser l'information évolutive contenue dans les séquences de protéines homologues (Dominguez, Boelens et al. 2003; Zhang, Liu et al. 2005; de Vries, van Dijk et al. 2006; Kanamori, Murakami et al. 2007; Madaoui and Guerois 2008). Ces données peuvent être utilisées soit pour limiter l'espace de recherche à des régions spécifiques de la surface de la protéine, pour appliquer des contraintes de distance (Dominguez, Boelens et al. 2003) ou filtrer à posteriori les solutions proposées par le calcul.

### 1.6.4. Méthode des cycles de doubles mutants

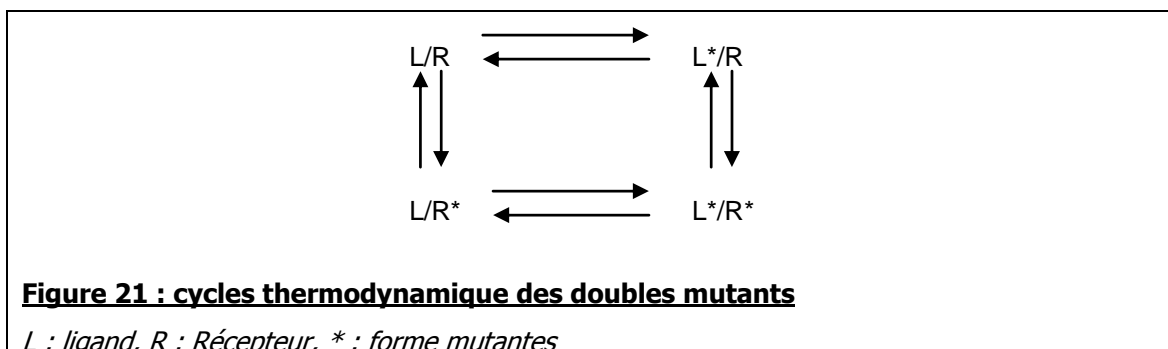
#### 1.6.4.1. Principe

Les expériences de balayage alanine (alanine scanning) consistent en la substitution systématique des résidus d'une protéine par des résidus alanine, grâce à des méthodes de mutagenèse dirigée. Ce type de d'expérience consiste à remplacer la chaîne latérale du résidu muté par un groupement méthyle relativement neutre d'un point de vue biochimique. On peut donc faire l'hypothèse que cela n'induit pas de changements conformationnels importants dans la structure de la protéine mutante. Il s'agit d'une hypothèse forte souvent difficile à vérifier. La méthode de balayage alanine permet donc d'identifier les résidus importants de la protéine pour l'interaction, et donc vraisemblablement situés au niveau de l'interface de liaison du ligand. En revanche, cette technique n'apporte pas d'information concernant les partenaires des résidus sur la structure du ligand.

La méthode de cycles de doubles mutants permet, quant à elle, de déterminer des couples de résidus en interaction dans la structure d'un complexe protéine/protéine. Ces expériences reposent sur l'utilisation conjointe de mutants alanine du récepteur et du ligand et sur la détermination pharmacologique des affinités des quatre combinaisons possibles de récepteur et de ligands sauvages et mutés.

La liaison de deux protéines s'accompagne toujours d'une libération d'énergie. La mutation alanine si elle entraîne une variation d'affinité des partenaires, va donc s'accompagner d'une variation de l'énergie libre de liaison du complexe que l'on notera  $\Delta\Delta G$ .

On peut ainsi former un cycle thermodynamique faisant intervenir les quatre  $\Delta G$ .



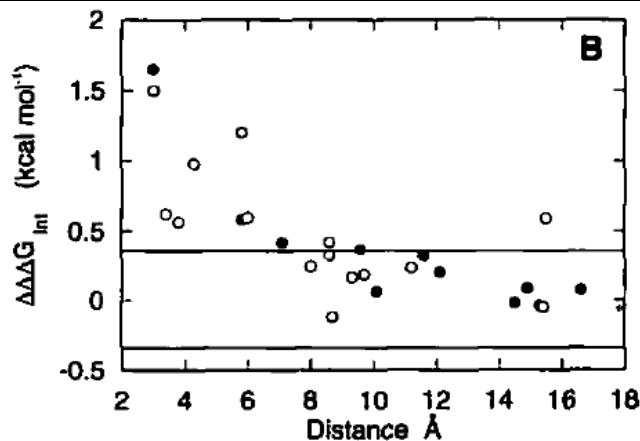
L'énergie de couplage  $\Delta\Delta G$  est définie comme :

$$\Delta\Delta G = \Delta\Delta G_{LR \rightarrow L^*R^*} - \Delta\Delta G_{LR \rightarrow L^*R} - \Delta\Delta G_{LR \rightarrow LR^*}$$

On peut alors distinguer plusieurs cas de figure :

- Si  $\Delta\Delta G = 0$ . Alors,  $\Delta\Delta G_{LR \rightarrow L^*R^*} = \Delta\Delta G_{LR \rightarrow L^*R} + \Delta\Delta G_{LR \rightarrow LR^*}$ . Alors les effets des deux mutations sont purement additifs. On peut donc supposer qu'il n'y a pas d'interaction entre les deux résidus mutés de chaque partenaire.
- Si  $\Delta\Delta G < 0$  ou  $> 0$ , cela signifie que lorsqu'on met en présence les deux mutants, on obtient un effet différent de la somme des deux mutations individuelles.

Des expériences de cycles de doubles mutants ont été conduites par Schreiber et Fersht sur le système barnase / barstar dont la structure du complexe était connue ((Schreiber and Fersht 1995), Figure 22). Ils ont ainsi démontré qu'aucune des paires de résidus ayant des énergies de couplage significativement différentes de zéro n'est distante dans la structure du complexe de plus de 8 Å.



**Figure 22 :  $\Delta\Delta G$  en fonction de la distance des résidus dans le complexe Barnase/Bastar**

Valeurs de distance entre les paires de résidus en fonction des  $\Delta\Delta G$  mesurés. Les résidus ayant des  $\Delta\Delta G$  de plus de  $0.45 \text{ kcal.mol}^{-1}$  sont distants de moins de  $8\text{Å}$   
D'après Schreiber (Schreiber and Fersht 1995) et (Vaughan, Buckle et al. 1999).

Cependant, il faut toutefois noter qu'aucune relation directe entre énergie de couplage et distance entre résidus ne peut être établie (Vaughan, Buckle et al. 1999). Ceci s'explique par l'importance jouée par la nature des résidus concernés ainsi que par leur contexte structural au sein du complexe.

### **1.6.5. Validation de l'utilisation des données de doubles mutant pour le docking**

Les expériences de cycles de doubles mutants fournissent donc des informations sur les paires de résidus en interaction dans la structure d'un complexe protéique. Cette information peut être exploitée afin de guider des expériences d'arrimage *in silico* de protéines. Elles peuvent être utilisées de différentes manières. Une première solution est d'effectuer l'étape de recherche en aveugle et d'utiliser les données expérimentales pour trier les solutions proposées. Cette approche suppose que l'on a la capacité d'explorer de manière suffisamment exhaustive l'espace conformationnel pour pouvoir être certain de ne pas passer à côté de la bonne solution. Une approche plus efficace consiste à introduire les proximités identifiées par les cycles de doubles mutants dans le calcul de docking sous la forme de contraintes de distance. Ceci permet de restreindre l'exploration de l'espace des solutions en amont de l'étape de classification.

Les structure de trois complexes protéine/protéine ont déjà été prédites à partir de données thermodynamiques au laboratoire. Des approches légèrement différentes ont été utilisé pour modéliser ces trois complexes :

1. Dans le premier exemple, le complexe entre le récepteur de l'acétylcholine et l' $\alpha$ -cobratoxine, des contraintes de distances ont été dérivées de l'analyse de cycles de doubles mutants et les complexes sélectionnés sur la base de l'énergie d'interaction et de l'analyse des violations des contraintes de distance.
2. Dans le second exemple, le complexe entre le canal potassium Kv1.1 et la BgK, des contraintes de distance ont été dérivées de l'analyse de cycles de doubles mutants et les complexes sélectionnés en comparant les valeurs expérimentales des énergies de couplage et celles recalculées sur la base des modèles proposés. Dans ce cas les complexes ont d'abord été réalisés uniquement sur la base de l'analyse des violations des contraintes de distance. Ensuite, le calcul en retour des énergies de couplage et la sélection des modèles donnant un bon accord avec les valeurs expérimentales, a permis d'identifier plusieurs liaisons hydrogène à l'interface.

3. Dans le troisième exemple, le complexe entre uPAR et uPA, des données thermodynamiques issues de mutations en alanine ont été utilisées. Des complexes ont été générés en introduisant des contraintes de distance dérivées de l'analogie entre uPA et le peptide inhibiteur. Les complexes ont été triés en comparant les valeurs des variations des énergies de liaison calculées et expérimentales.

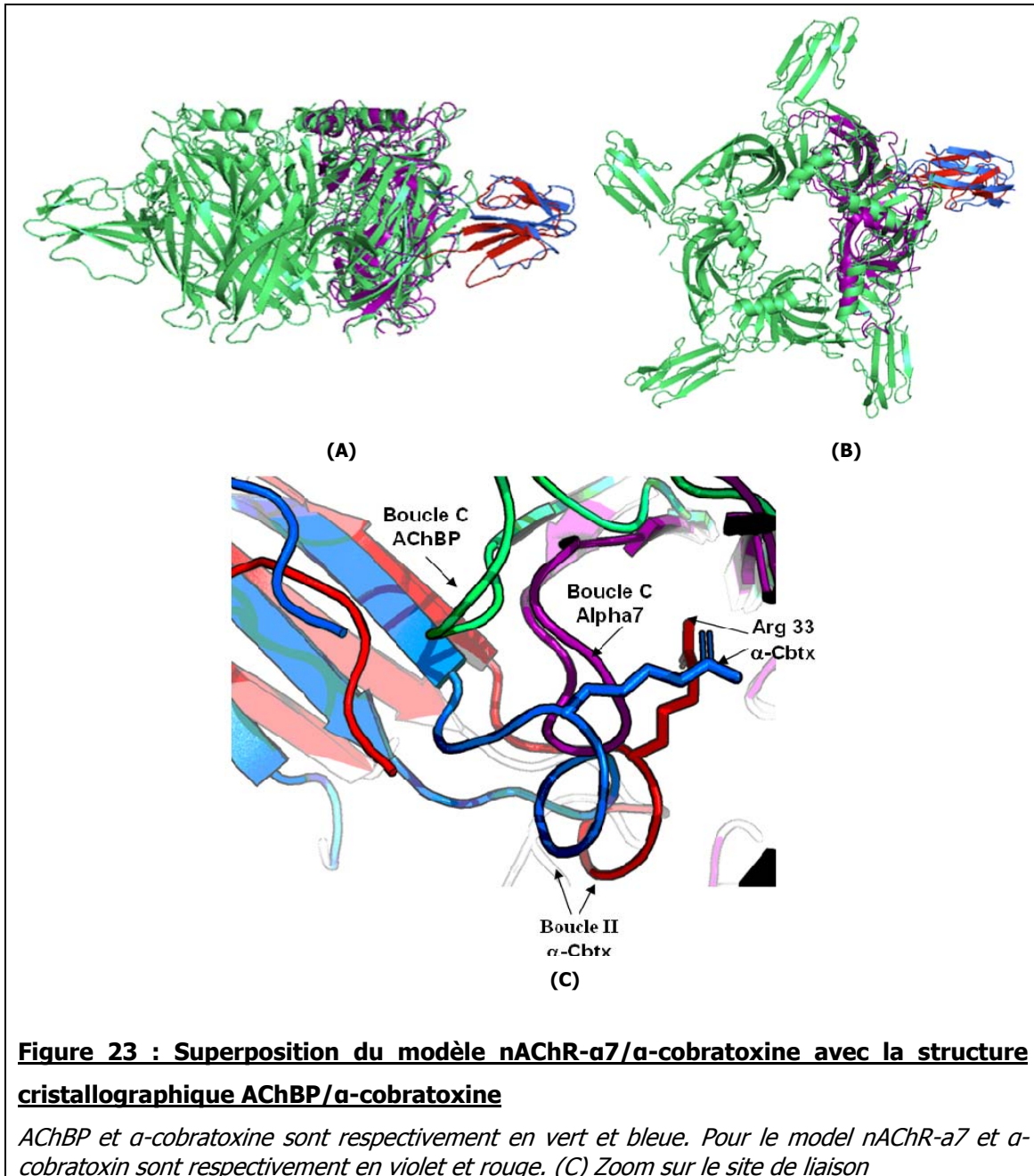
Pour deux de ces complexes, la structure cristallographique a été publiée ultérieurement à ces travaux de modélisation (Bourne, Talley et al. 2005 ) (pdb code 1VI5) ; (Huai, Mazar et al. 2006; Schneider, Ader et al. 2008) (code PDB 2FD6). Cette approche peut être validée en comparant les modèles aux structures cristallographiques. Pour le complexe BgK/Kv11 le modèle peut être comparé aux données récentes obtenues par RMN du solide (Schneider, Ader et al. 2008). Pour comparer ces modèles aux structures cristallines nous utilisons les critères proposés par CAPRI et présentés précédemment (cf § 1.6.3) : Fnat, L-rms et I\_rms. Il faut noter que dans le concours CAPRI, pour chaque cible, chaque groupe peut présenter plusieurs solutions. Dans ces travaux de modélisation de complexes à partir de données thermodynamiques un seul modèle a été proposé.

### 1.6.5.1. Evaluation du modèle Ach-7/ -cobratoxine

L' -cobratoxine est une neurotoxine, présente dans le venin des serpents. C'est un antagoniste du récepteur nicotinique à l'acétylcholine (AChR). L'AChBP est une protéine fonctionnellement et structuralement proche du domaine N-terminal de la sous unité de l'AChR et sa structure tridimensionnelle est connue (D'Adamo, Liu et al. 1998 ; Kashuba, Kvasa et al. 2001). La structure de l' -cobratoxine est également connue. Un modèle de Ach-7 a été obtenu par modélisation par homologie à partir de la structure cristalline de l'AChBP (PdB code : 1I9B) en utilisant Modeller (Sali, Potterton et al. 1995). Le pourcentage d'identité de séquence entre AchBP et la région extracellulaire de Ach-7 est de 24%. La structure cristalline -cobratoxin est disponible sur la PDB (PdB code : 2CTX). Dix contraintes ambiguës dérivées de l'analyse des cycles de doubles mutants ont été utilisées : R33-P192, R33-F185, R33-E191, R33-K143, F65-S184, F65-E191, K35-E160, K35-S184, R36-P194, F29-W53. La plupart de ces contraintes concernent la boucle centrale de l' -cobratoxine, dont deux impliquent la région C-terminale. La flexibilité a été introduite à deux niveaux.

- D'une part, différentes conformations de la boucle F du récepteur (résidus 151-170) ont été générées par dynamique moléculaire afin de sélectionner des conformations qui permettent l'accessibilité au site de liaison de la toxine.
- D'autre part durant la deuxième étape de la procédure d'amarrage, seuls les squelettes peptidiques sont maintenus dans leurs positions d'origine alors que les chaînes latérales sont libres. Après un dynamique de corps rigide avec des rayons de van der Waals réduit, dix structures sont sélectionnées et raffinées en libérant les chaînes latérales.

Toutes les structures de complexe obtenues sont similaires. Après superposition des récepteurs, l'écart entre les différentes positions de la toxine est de 0.65Å (Fruchart-Gaillard, Gilquin et al. 2002). La meilleure structure, en termes d'énergie d'interaction de van der Waals, a été sélectionnée et renumérotée pour être comparée (Figure 23A,B) à la structure cristallographique du complexe AChBP/ -cobratoxin (Bourne, Talley et al. 2005).



La superposition des éléments de structure secondaire des deux complexes donne un écart de 2.3 Å de rmsd entre le modèle et la structure cristallographique. La Figure 23 montre que les deux structures se superposent, indiquant que la position et l'orientation de la toxine a été correctement prédite. Lorsque les deux récepteurs sont superposés, l'écart entre les C des toxines (L-rms) est égal à 4.4 Å. Les déviations majeures sont localisées dans les extrémités des boucles de la toxine et dans la boucle C du récepteur, une région peu conservée entre AChR-7 et AChBP (Figure 23C). Une flexibilité importante de cette région est observée dans plusieurs structures de complexe de ligand avec AChBP (Celie, Kasheverov et al. 2005; Hansen, Sulzenbacher et al.

2005). Par contre, la boucle F du récepteur du modèle se superpose bien avec celle de la structure cristallographique, indiquant l'efficacité de la procédure de dynamique moléculaire appliquée en amont de la procédure d'amarrage.

L'interface est similaire dans le modèle et dans la structure cristalline. L'écart entre les résidus de l'interface ( $I_{rms}$ ) est égal à 3.6 Å. Le pourcentage de contacts natifs retrouvés dans le modèle est égal à 44%. Les valeurs individuelles de  $F_{nat}$  pour les résidus du « hot spot » R33, F65, K25 et R36 sont respectivement de 63%, 33 %, 100 % et 75%. Ces valeurs de  $F_{nat}$  sont satisfaisantes puisque le seuil généralement appliqué dans CAPRI pour admettre un accord avec la structure cristalline est de 30%.

Dans le modèle, comme dans la structure cristalline, les extrémités des boucles et le C-terminal de la toxine sont en contact avec le récepteur (résidus 7-9, 25-37, 49-52 et 65-69 dans le modèle et 6-9, 25-37, 51, 65-68 dans la structure cristalline). Il est à noter que bien qu'aucune contrainte impliquant la boucle I n'a été introduite, le même segment de cette boucle est en contact avec le récepteur AChBP ou AChR- 7. Dans le modèle comme dans la structure cristallographique, la toxine se lie à l'interface entre deux sous unités (Figure 23A,B). Les mêmes régions du récepteur sont en contact avec le ligand c'est-à-dire la boucle A (résidus 91 dans le modèle et la structure cristallographique), la boucle B (résidus 143, 146, 147 et 152 dans le modèle et 147-148 structure cristallographique) et la boucle C (résidus 183-193 et 195 dans le modèle et résidus 183-193 dans structure cristallographique) pour la sous-unité 1, et en contact avec la boucle I (résidus 30, 32-36 dans le modèle et résidu 34 dans la structure cristallographique), boucle D (résidus 53-55 dans le modèle et résidus 53, 55 et 57 dans la structure cristallographique), boucle E (résidus 117 dans le modèle et résidus 107, 115 et 117 dans la structure cristallographique) et boucle F (résidus 159, 162, 165-169, 174 et 177 dans le modèle et résidus 161-163 et 166 dans la structure cristallographique) pour la seconde sous unité. Cette analyse révèle que cette approche a bien mis en évidence les contacts les plus importants pour la liaison de la toxine au récepteur.





### 1.6.5.2. Evaluation du modèle du canal potassique Kv1.1/BgK

Les canaux potassiques Kv1 sont protéines tétramériques assurant le transport passif des ions potassium à travers la membrane. La BgK, une toxine d'anémone de mer bloque le flux de potassium à travers le canal Kv1.1 en se liant à sa région P. Bien qu'à l'époque aucune structure du canal Kv1.1 n'était disponible, la structure du canal bactérien KcsA qui possède deux hélices transmembranaires avait été déterminée. Cette structure a été utilisée pour modéliser par homologie la structure des régions S5-S6 de Kv1.1. (Gilquin, Racape et al. 2002).

Le modèle de S5-S6 Kv1.1 a été construit par modélisation par homologie sur la base de la structure du canal bactérien KcsA. Six contraintes de distance ont été utilisées (BgK(Ser-23)-Kv1.1(Tyr-379), BgK(Phe-6)-Kv1.1(Tyr-379), BgK(Tyr-26)-Kv1.1(Ser-357), BgK(Tyr-26)-Kv1.1(Asp-361), BgK(Asn-19)-Kv1.1(Ser-357), BgK-(Tyr-26)-Kv1.1(Tyr-379) (Gilquin, Racape et al. 2002). Le calcul en retour des énergies de couplage a montré que l'accord était bien meilleur lorsque on observait les interactions suivantes BgK(Lys-25NZ)-Kv1.1(Tyr-375O) and BgK(Phe-6)-Kv1.1(Asp-361) (Gilquin, Braud et al. 2005). Nous avons pu ainsi identifier deux liaisons hydrogène à l'interface (Figure 24C). Les seules données structurales disponibles aujourd'hui concernant l'interaction d'une toxine en interaction avec un canal potassique sont fournies par des expériences d'RMN du solide qui concerne le canal KcsA chimérique et la kaliotoxine (Lange, Giller et al. 2006). Le modèle issu de ces données RMN montre l'interaction de l'atome NZ de la lysine avec les carbonyles du pore et pointe le rôle de la tyrosine Tyr78 (Figure 24A). Nous retrouvons ces deux caractéristiques dans le modèle de S5-S6 Kv1.1/BgK que nous avons obtenu (Figure 24B). Ainsi comme suggéré par Dauplais et al (Dauplais, Lecoq et al. 1997) les données de mutagenèse et la topologie de la BgK et des toxines de scorpion (Charybdotoxine, kaliotoxine) conduisent à un mode de liaison analogue.



### 1.6.5.3. Evaluation du modèle uPAR/uPA.GFD

uPA est l'activateur du plasminogène de type urokinase. uPAr son récepteur joue un rôle important dans la malignité des cancers humains. La structure du complexe uPAR lié à un peptide antagoniste a été déterminée au laboratoire (Llinas, Le Du et al. 2005). La structure d'uPA déterminée par RMN est disponible.

Le récepteur uPAR est composé de trois domaines à trois doigts qui forment une cavité dans laquelle est logé le peptide. Trois résidus hydrophobes de ce peptide (Phe5, Leu9 and Trp10) forment un motif qui se retrouve dans uPA (Ile28, Phe25 et Tyr24) bien que la structure secondaire des deux ligands soit différente. Vingt contraintes de distance ambiguës entre uPAR and uPA ont été dérivées de cette analogie. Après une procédure de d'arrimage moléculaire classique, le classement a été réalisé en comparant les valeurs expérimentales et calculées des variations d'énergies de liaison pour des mutants alanines sur 11 positions de uPA. Pour quatre complexes un accord raisonnable est obtenu. Sur 10 des 11 positions, les valeurs de  $\Delta G$  calculées et expérimentales sont soit toutes les deux inférieures à 0.8 kcal/mol ou supérieures à 0.9 kcal/mol.

La Figure 25 montre la superposition du modèle uPAR/uPA avec la structure cristalline. L'écart entre les deux structures est en moyenne égal à  $3.2 \pm 0.5$  Å sur les carbones et le  $L_{rms}$  est égal à  $5.5 \pm 1$  Å. Cette valeur élevée du  $L_{rms}$  est due à un mouvement de fermeture d'uPAR lorsqu'il se lie à uPA qui n'est pas observé lorsque qu'il se lie au peptide. La cavité de liaison de uPA est plus petite que celle du peptide. Ce mouvement charnière n'a pas été pris en compte dans notre procédure de l'époque. Néanmoins un des modèles est proche de la structure cristalline avec un  $L_{rms}$  égal à 3.5 Å. Enfin le  $I_{rms}$  est égal à  $2.8 \pm 0.3$  Å, et le pourcentage de contacts natifs  $F_{nat}$  est égal à  $33 \pm 2$  %. En conclusion les résidus les plus importants pour l'interaction uPAR /uPA sont correctement localisés.



**Figure 25 : Superposition du modèle et de la structure cristallographique uPAR/uPA.**

*Pour la structure cristallographique uPAR et uPA sont respectivement en cyan et orange. Pour le modèle uPA est en rouge et les trois domaines D1, D2 et D3 de uPAR sont représentés en rose, magenta et bleu, respectivement.*

### 1.6.6. Conclusion

Ces trois exemples montrent que la modélisation de complexes protéine/protéine à partir de données thermodynamiques conduit à des structures en assez bon accord avec l'expérience, en particulier pour les résidus majeurs de l'interaction. Par contre, ils montrent les limites de la procédure utilisée. L'absence d'échantillonnage des conformères du récepteur et du ligand avant l'application de la procédure d'arrimage ne permet pas de tenir compte complètement de la flexibilité des partenaires. Enfin il semble que la relaxation du complexe par dynamique moléculaire doit être réalisée afin d'optimiser le modèle et de valider sa stabilité.



---

# Chapitre 2: Étude structurale des partenaires

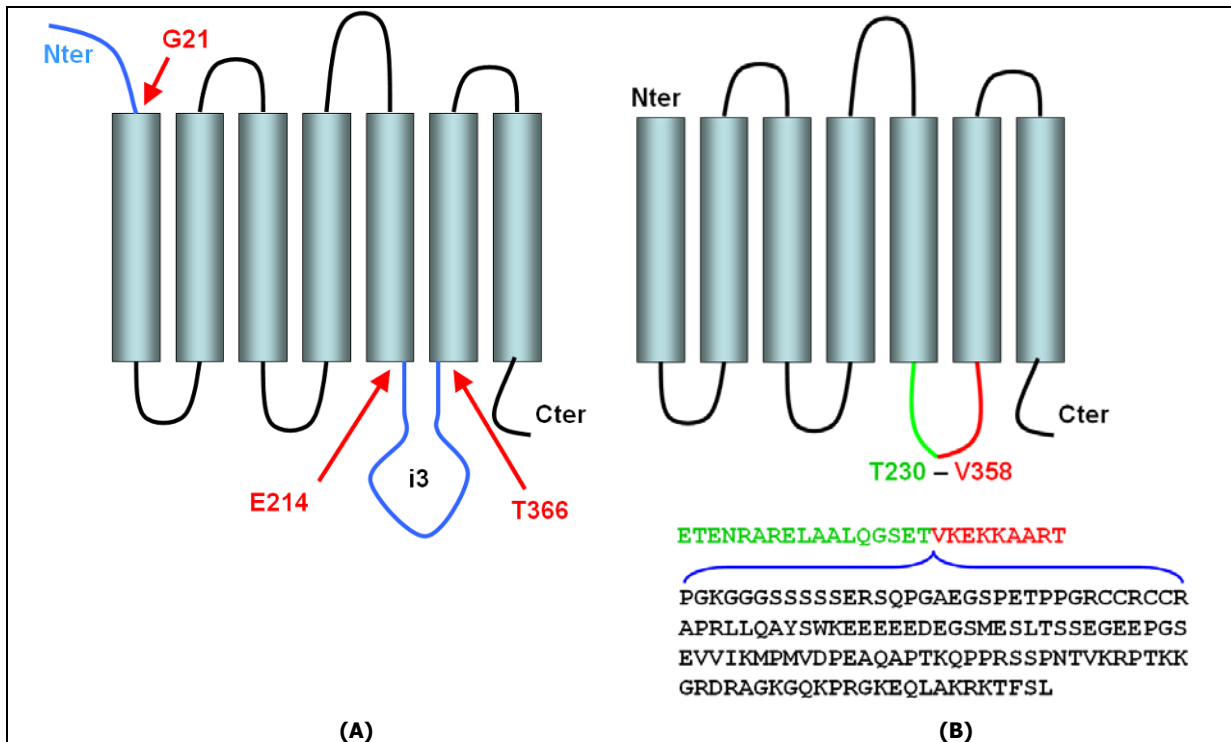
---

## ***2.1. Le récepteur hM1***

### **2.1.1. Séquences peptidique**

Les récepteurs muscariniques possèdent un court segment N terminal de 21 acides aminés avant la première hélice transmembranaire (Figure 26A). Il a été montré (Weill, Galzi et al. 1999) que l'absence de ce segment n'a aucune influence sur la liaison de ligands au récepteur. Une autre particularité des récepteurs muscariniques est la présence d'un grand domaine intracellulaire I3 (Figure 26A). Ce domaine d'une longueur de près de 180 acides aminés est situé entre les hélices transmembranaires 5 et 6. Des récepteurs muscariniques recombinants dans lesquels cette partie intracellulaire a été remplacée par une boucle plus courte (suppression des résidus 231 à 357) ont été construits et caractérisés par le groupe de B. Ilien (Weill, Galzi et al. 1999). Enfin, les sites de glycosylation ont également été mutés. Ces récepteurs recombinants possèdent toujours la capacité de lier leurs ligands et d'activer les protéines G. La séquence peptidique que nous utiliserons pour modéliser le hM1 est une forme avec le segment N-ter tronqué des résidus 1 à 20 ainsi que la forme courte de la boucle I3 décrite par Weill et al (Weill, Galzi et al. 1999) (Figure 26B).





**Figure 26 : Modification des boucles du récepteur M1**

*Organisation schématique du M1 sauvage (A) et après modification (B). la partie Nter a été supprimée jusqu'au résidu 20. Les résidus 231 à 357 de la boucle intracellulaire I3 sont également supprimés.*

### 2.1.2. Modélisation du cœur d'hélices

La rhodopsine était la seule structure de RCPG disponible au début de ce travail. Elle a donc été utilisée comme base pour construire un modèle du faisceau d'hélice du récepteur muscarinique par modélisation comparative. En dépit de valeurs d'identité de séquence relativement faibles entre les deux protéines (24% dans les hélices), leur alignement est peu ambiguë en raison de la présence de zones de forte conservation au sein des hélices (Baldwin, Schertler et al. 1997).

### 2.1.3. Prédiction des boucles

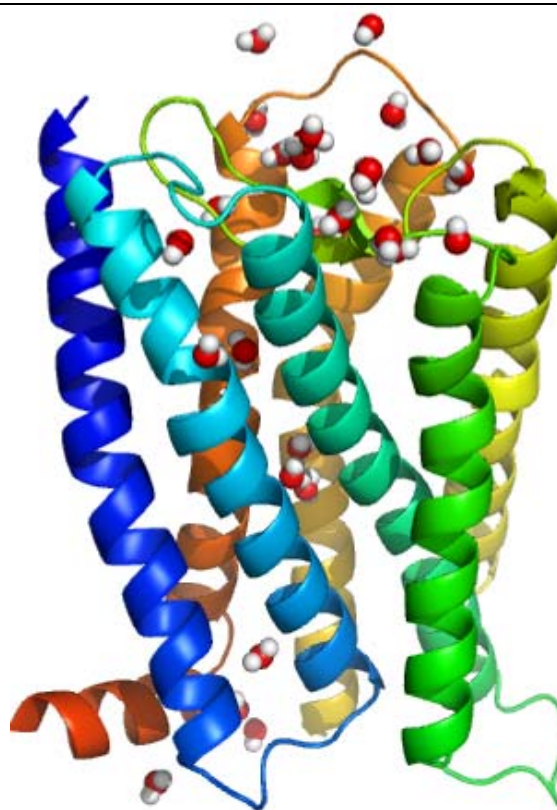
La boucle E1 du récepteur hM1 a été modélisée par comparaison avec celle de la rhodopsine (Figure 27) avec laquelle elle possède une similarité raisonnable (30%). La présence du pont disulfure entre le résidu 178 de la boucle E2 et le résidu 98 appartenant à une hélice transmembranaire contraint la conformation de la base de la boucle dans une conformation similaire à celle de la rhodopsine. Ces résidus sont donc

également modélisés par homologie. En revanche, la structure de la rhodopsine ne constitue pas un modèle convenable pour prédire les conformations de l'extrémité de la boucle E2 ainsi que de la boucle E3. L'extrémité de la boucle E2 (169-177) a donc été prédite entièrement de novo avec l'aide du programme RAPPER (Furnham, Dore et al. 2006). Pour la boucle E3 (391-396) le programme RAPPER a identifié un fragment dans la PDB (code 1KQF) afin de guider la prédiction. Les 38 molécules d'eau internes de la structure cristallographique de la rhodopsine ont été également conservées dans le modèle de hM1 (Figure 28).

Le modèle final a été évalué par le serveur RAMPAGE (Lovell, Davis et al. 2003). On observe Figure 29 que 98% des résidus se trouvent dans les régions autorisées du diagramme de Ramachandran (Ramachandran, Ramakrishnan et al. 1963). Ce modèle ne présente donc aucun défaut majeur dans sa géométrie.

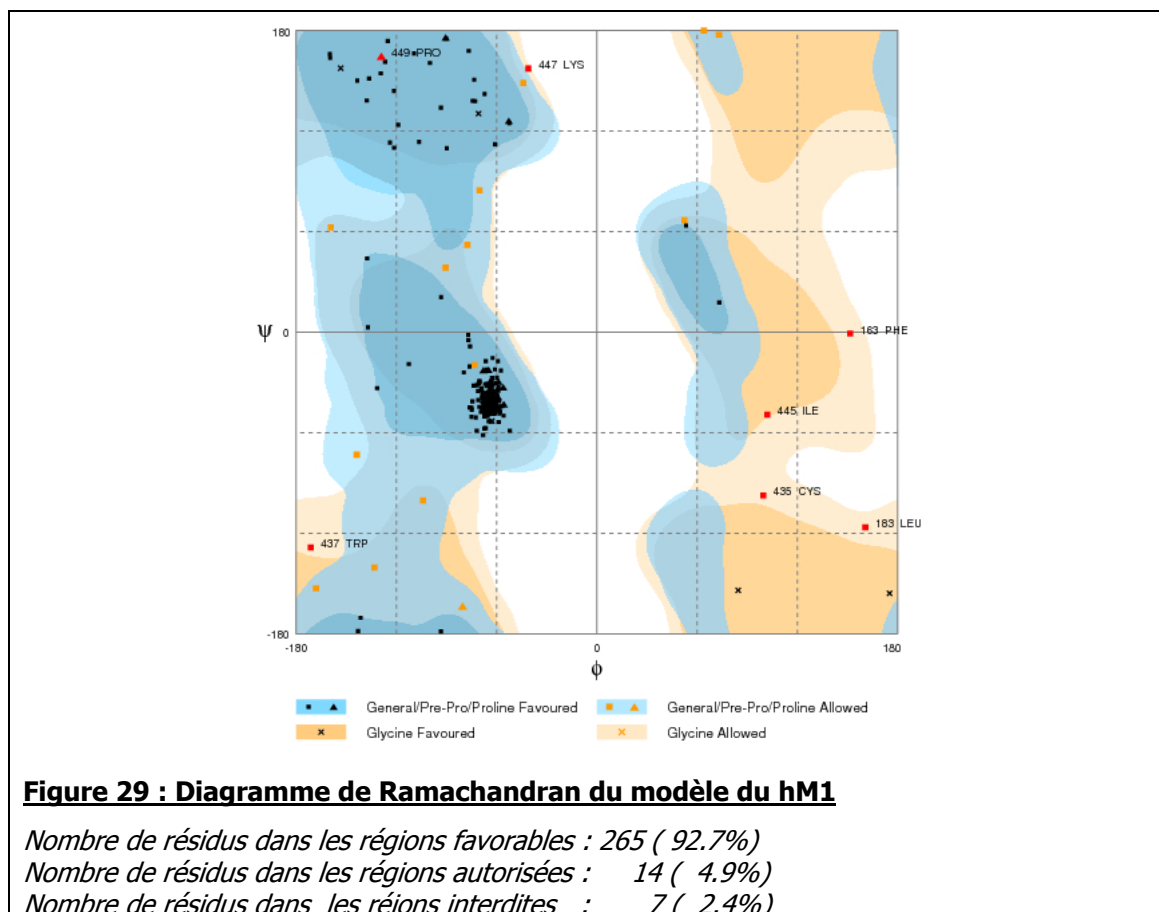
hM1            TTYLLMGHWALGTLACD:  
Rhodopsine   TS--LHGYFVFGPTGCM:

**Figure 27: Alignement des séquences des boucles extracellulaires E1 de la rhodopsine et du hM1**



**Figure 28 : Modèle du récepteur hM1**

*Le modèle est construit par homologie sur la base de la structure de la rhodopsine.*



## 2.2. La toxine MT7

*Les toxines modifiées sont produites par G. Mourier. Les structures cristallographiques des toxines MT7 modifiées ont été déterminées par E. Stura et R. Menez (CEA, SIMOPRO).*

La cristallisation de la toxine MT7 sauvage n'a toutefois pas permis d'obtenir de cristaux de bonne qualité diffractant à des résolutions inférieures à 3 Å. En revanche, deux formes modifiées de la toxine ont pu être cristallisées et leur structures déterminées :

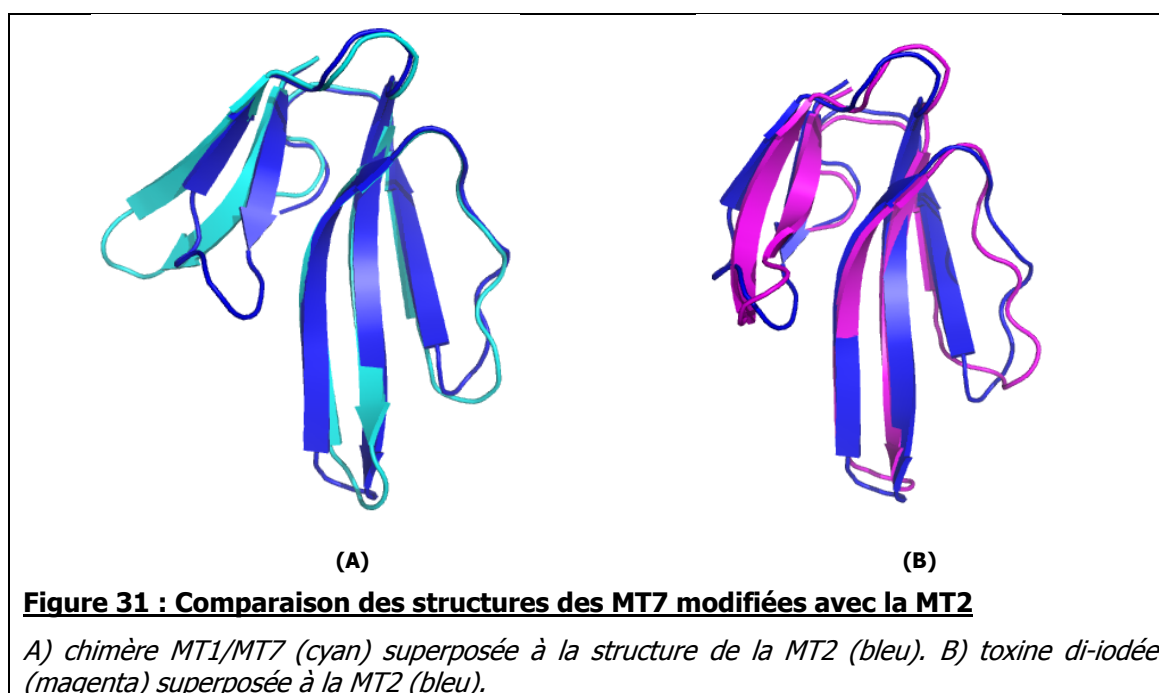
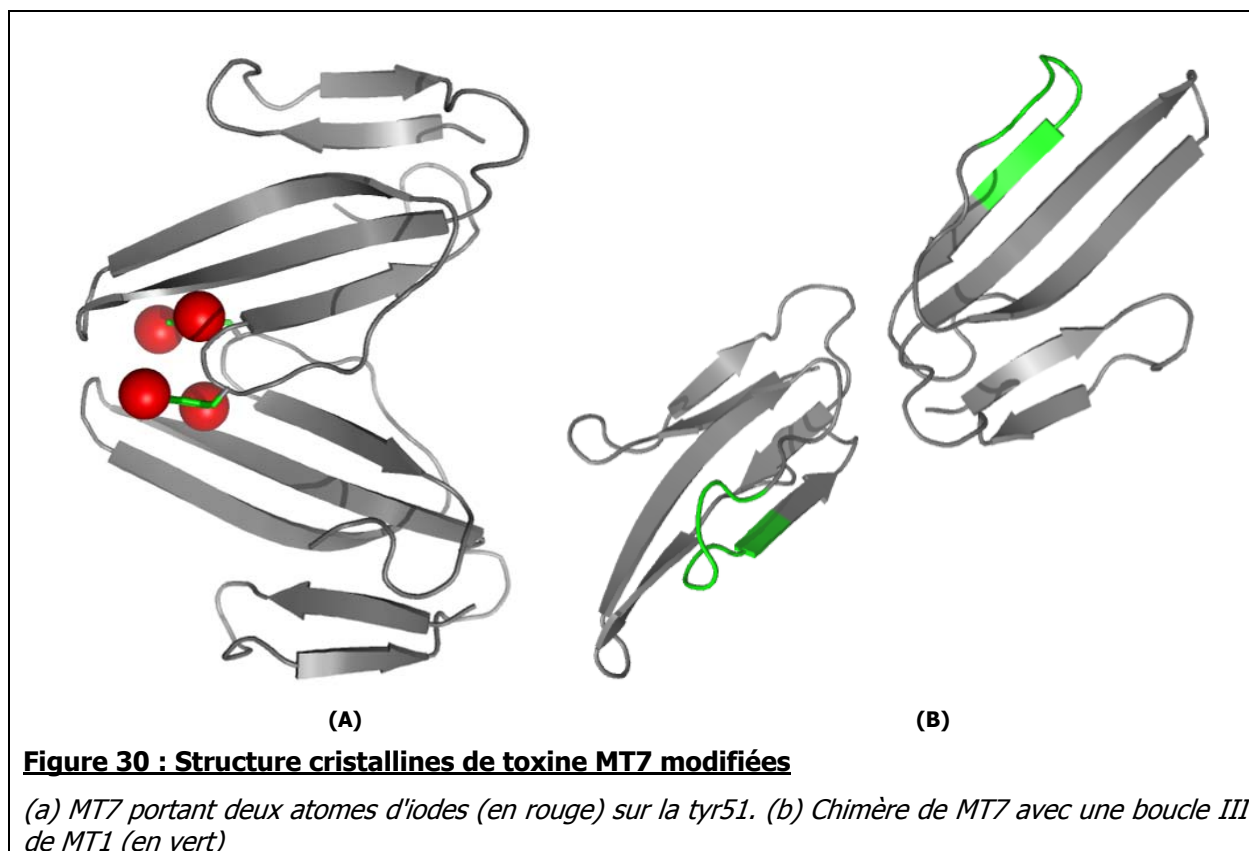
- Une forme portant deux atomes d'iode sur la boucle III au niveau de la tyrosine 51 et diffractant à une résolution de 1.42Å (Figure 30A)
- Une chimère MT7/MT1 constituée de la boucle I de la toxine MT1 et des boucles II et III natives de la toxine MT7. La structure de cette chimère a pu être déterminée avec une résolution de 1.25 Å (Figure 30B)

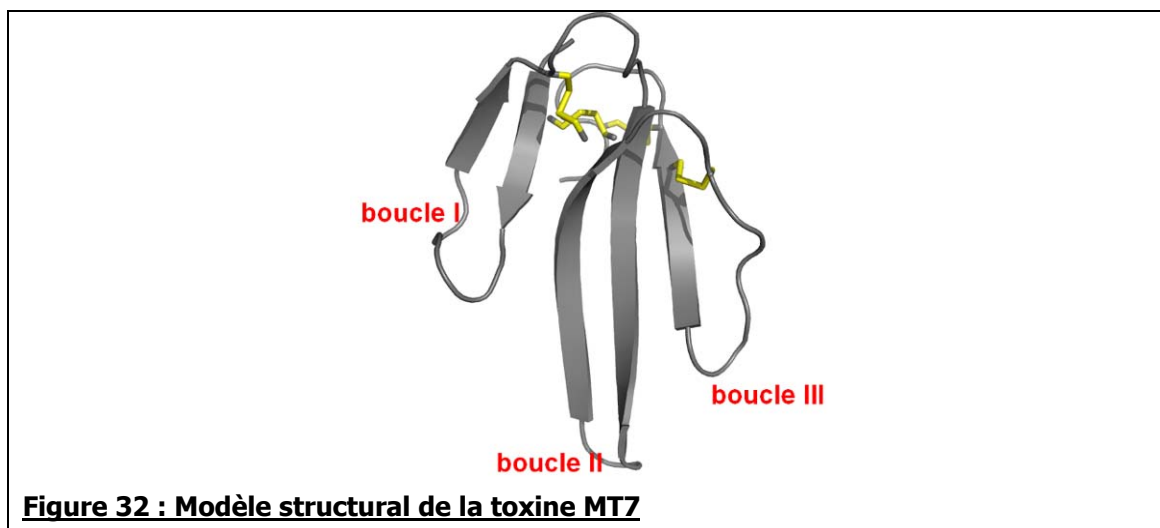
L'encombrement stérique des atomes d'iode est important et la structure de la MT7-di-iodée présente des contacts cristallins au niveau de la boucle III. Cependant, la comparaison des boucles II et III des structures de la toxine di-iodée et de la chimère démontre que les changements conformationnels dus à la greffe de l'atome d'iode sur la Tyr51 sont en fait très limités puisque l'écart entre les squelettes des toxines à l'exception de la boucle I (résidus 1-4 et 17-65) est seulement de 0.93Å.

On observe pour la structure de la chimère MT7/MT1 une conformation de la boucle I qui montre une disruption de l'interface entre les brins des boucles I et II. La chimère diverge nettement (écart de 3.5 Å) sur la conformation de la boucle I de la MT7-di-iodée et de la MT2 (Figure 31) qui présente une boucle I positionnée comme dans l'ensemble des protéines adoptant un repliement à trois doigts (Gilquin, Bourgoïn et al. 2003).

Un modèle de la structure de la toxine a donc été obtenu en fusionnant avec MODELLER v8.2 les deux structures cristallographiques. Nous avons utilisé d'une part les coordonnées des boucles I de la toxine di-iodée (résidus 5 à 16), et d'autre part, les boucles II et III de la chimère (résidus 1 à 4 and 17 à 65). Ceci nous permet au final de

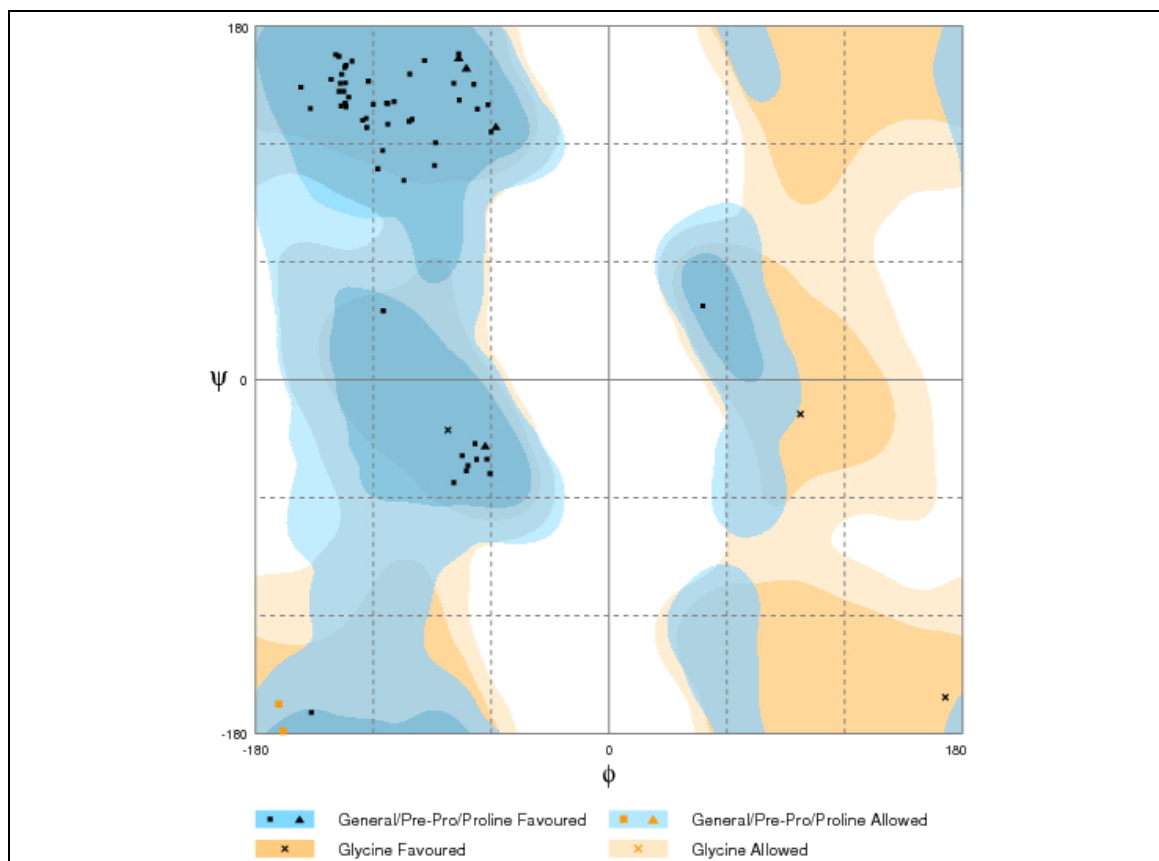
disposer d'un modèle fiable de notre ligand et de nous affranchir d'éventuels artefacts structuraux dus à la présence d'iode ou à la greffe de la boucle d'une autre toxine.





**Figure 32 : Modèle structural de la toxine MT7**

Le modèle structural (Figure 32) de la toxine MT7 présente donc un repliement à trois doigts constitué d'un cœur hydrophobe maintenu par 4 ponts disulfures et de 3 boucles issues de ce cœur. L'analyse de la géométrie de cette structure (Figure 33) sur le server RAMPAGE (Lovell, Davis et al. 2003) démontre que tous les résidus de la protéine sont localisés dans les régions autorisées (allowed) du diagramme de Ramachandran (Ramachandran, Ramakrishnan et al. 1963), et que 63 résidus sont dans les régions les plus favorables (favored). Ce modèle possède donc les propriétés d'une structure de qualité cristallographique.



**Figure 33 : Diagramme de Ramachandran de la structure de la toxine MT7**

*Nombre de résidus dans les régions favorables : 63 ( 98.4%)*

*Nombre de résidus dans les régions autorisées : 2 ( 3.0%)*

*Nombre de résidus dans les régions interdites : 0 ( 2.4%)*





---

# **Chapitre 3: Prédiction de la structure du complexe hM1/MT7**

---

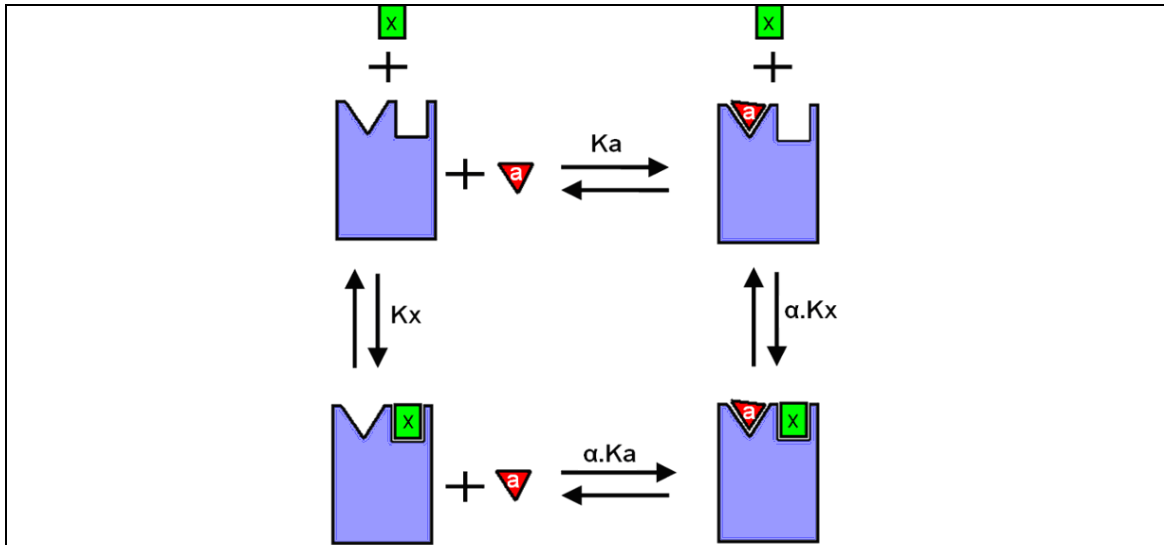
### **3.1. Mesures thermodynamiques hM1/MT7**

*Les expériences de mutagenèse et les mesures thermodynamiques ont été réalisées par le groupe de Denis Servent (CEA SIMOPRO).*

#### **3.1.1. Modèle du complexe ternaire**

Proposé initialement pour décrire le comportement d'un RCPG liant à la fois son ligand orthostérique et une protéine G (De Lean, Stadel et al. 1980), le modèle du complexe ternaire (TCM) peut en fait être généralisé à tout récepteur liant simultanément son ligand orthostérique et un autre ligand dans un second site topographiquement différent. Dans ce modèle, on ne considère qu'un seul état d'activation du récepteur auquel peuvent simultanément se lier un ligand orthostérique A et un ligand allostérique X avec des affinités respectives pour le récepteur libre  $K_A$  et  $K_X$ . La conséquence de la liaison d'un des ligands au récepteur est d'altérer les cinétiques d'association et de dissociation de l'autre ligand. La différence d'affinité du ligand orthostérique pour le récepteur libre ou lié au ligand allostérique définit la coopérativité du système noté  $\alpha$ . Un facteur  $\alpha > 1$ ,  $\alpha < 1$  ou  $\alpha = 1$  indique respectivement une coopérativité positive, négative ou neutre. Ainsi, le modèle du complexe ternaire, permet de modéliser correctement l'action de la plupart des modulateurs allostériques.

La NMS est un antagoniste du récepteur muscarinique. Elle se lie au site orthostérique et n'entre donc pas en compétition avec la toxine qui est un ligand allostérique. Fruchart-Gaillard et al ont montré que l'utilisation du modèle du TCM pour traiter les résultats d'expérience de compétition à l'équilibre était valable dans ce contexte (Fruchart-Gaillard, Mourier et al. 2006).



**Figure 34 : Modèle du complexe ternaire**

*Les ligands allostériques  $X$  et orthostérique se lient simultanément sur le récepteur. Leur coopérativité est mesurée par le coefficient  $\alpha$ .*

### 3.1.2. Alanine scanning des partenaires

*Toutes les mesures de liaison du système hM1/MT7 décrite dans ce document ont été réalisées par l'équipe de Denis Servent (CEA/SIMOPRO).*

#### 3.1.2.1. Modifications de la toxine

Les résidus modifiés de la toxine ont été remplacés par des résidus alanine, afin d'étudier l'effet de leur chaîne latérale. Les toxines modifiées sont produites par synthèse chimique, de la même manière que la toxine sauvage (Mourier, Dutertre et al. 2003). Afin d'évaluer l'effet de ces modifications sur la structure globale de la toxine, le spectre de dichroïsme circulaire de chaque mutant a été comparé au spectre de la toxine MT7 sauvage. Tous ces spectres présentent une signature typique d'un feuillet  $\beta$  et indiquent que les diverses modifications en alanine ne semblent pas perturber la structure secondaire de la toxine.

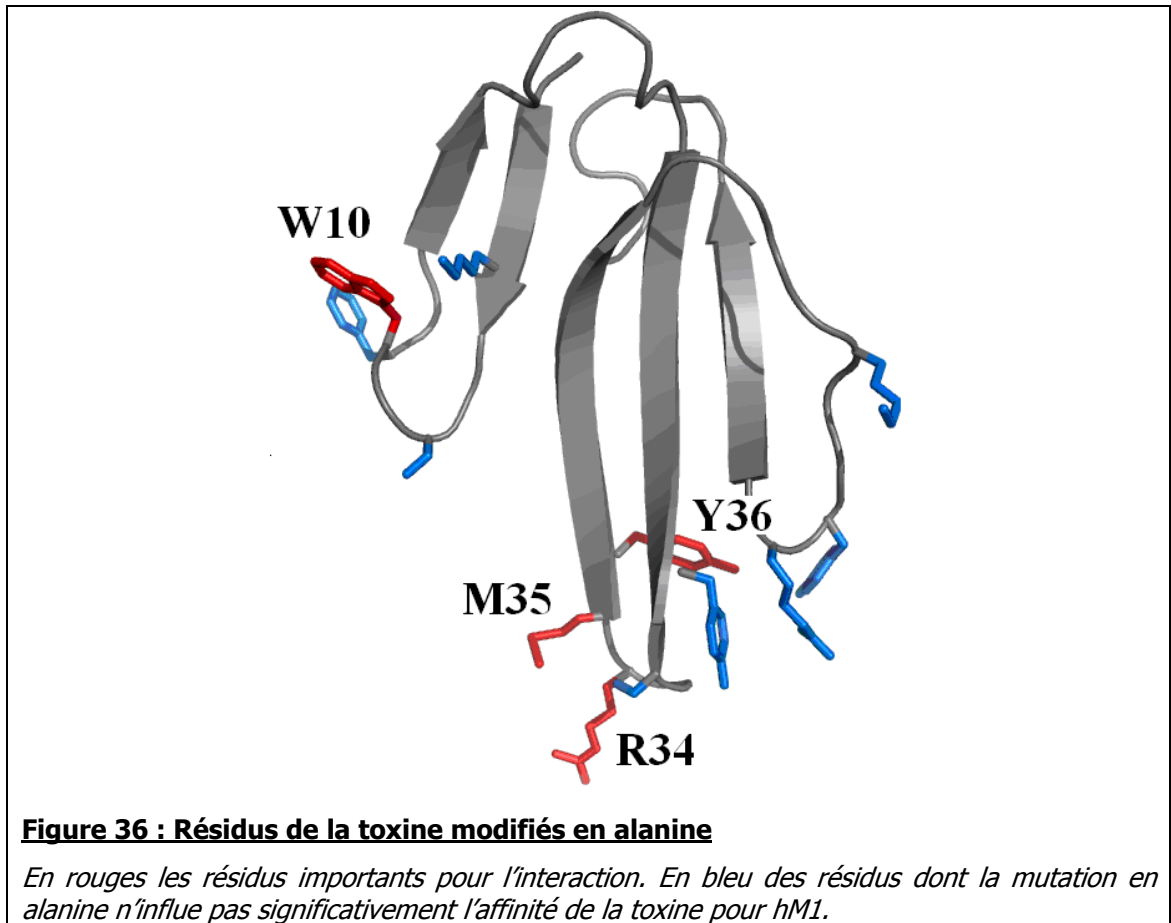
Compte tenu des données disponibles dans la littérature, les positions modifiées ont été choisies en priorité au niveau des extrémités des boucles de la toxine, sur les deux faces. Les modifications concernent d'une part des résidus très conservés dans la famille tels que Ser8, Tyr30 et Lys48, qui pourraient être responsables de l'affinité des toxines pour les récepteurs muscariniques. D'autre part, des résidus spécifiques de la toxine MT7 tels que Lys5, Trp10, Phe11, Ser32, Met35 et Tyr51 ont également été modifiés pour tenter d'expliquer la forte sélectivité de la MT7 pour le hM1. Au final, 12 positions ont été modifiées (Figure 35) : Lys5, Ser8, Trp10 et Phe11 sur la boucle I, Tyr30, Ser32, Arg34, Met35 et Tyr36 sur la boucle II et Lys48, Tyr51 et Arg52 sur la boucle III de la toxine.

Le seuil de variation minimale l'énergie de liaison pour considérer qu'une modification induit un changement significatif de est fixé à 1kcal.mol<sup>-1</sup> d'après un traitement statistique des résultats par ANOVA. Les résultats obtenus (Figure 35) démontrent que les mutations aux positions 5, 8, 11, 30, 32, 48, 51 et 52 n'affectent pas significativement la liaison de la toxine sur le récepteur. Un second groupe de modifications W10A, M35A et Y36A induisent un effet significatif mais inférieur à un facteur 10 sur l'affinité de la toxine. Le résidu ayant de loin l'effet le plus marqué est l'arginine 34 dont la modification provoque une réduction d'affinité de près de 200 fois. Ce résultat indique clairement que cette chaîne latérale, centrale dans la structure de la

toxine joue un rôle majeur dans la liaison de la toxine au récepteur. Les résidus modifiés sont représentés en bâtonnet sur la Figure 36. On peut constater que les extrémités des 3 boucles de la toxine sont impliquées dans la liaison au récepteur. Le cœur de l'interaction est l'arginine 34 qui est conservée chez les toxines muscariniques ainsi que la Tyr36 également conservée. Plus en périphérie, et notamment au niveau de la boucle I, on trouve des résidus impliqués dans l'interaction, mais spécifique de la MT7.

<b>Mutation</b>	<b>région</b>	<b>G</b>
K5A	cœur	0,6
S8A	cœur	0,2
<b>W10A</b>	<b>boucle I</b>	<b>1,1</b>
F11A	boucle I	0,5
Y30A	boucle II	0,5
S32A	boucle II	0,5
<b>R34A</b>	<b>boucle II</b>	<b>3,1</b>
<b>M35A</b>	<b>boucle II</b>	<b>1,1</b>
<b>Y36A</b>	<b>boucle II</b>	<b>1,3</b>
K48A	boucle III	0,2
Y51	boucle III	0,6
R52	boucle III	0,5

**Figure 35 : Variation d'énergies libres associées à des mutations de la toxine**  
*Les mutations entraînant des variations d'affinité significatives sont en gras*  
*D'après (Marquer 2008)*





### 3.1.2.2. Mutagenèse du récepteur

De la même manière que pour la toxine, des expériences de mutagenèse ont été réalisées sur le récepteur afin d'identifier les régions importantes pour l'interaction.

La toxine constitue, en regards des molécules chimiques de petite taille couramment étudiées, un ligand de haut poids moléculaire pour les RCPG. De plus, étant donné l'importance de la surface constituant la zone d'interaction déterminée par alanine scanning sur la toxine, ainsi que ses propriétés allostériques, il est probable que sa liaison se fasse en de multiples points au niveau de la partie extracellulaire du récepteur. 22 positions ont été mutées (Figure 39), au niveau des boucles externes du récepteur, et de la cavité apicale. La boucle extracellulaire E2 dont l'importance est clairement établie dans la littérature a été étudiée en priorité.

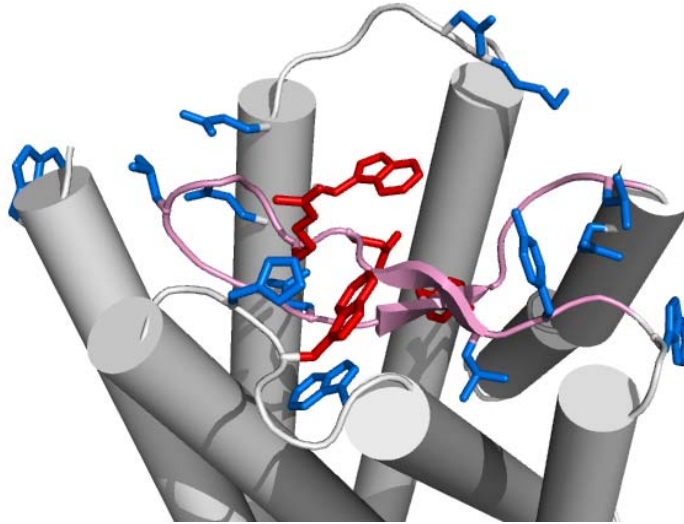
Les résidus ont été remplacés soit par des alanines, soit par leur équivalent dans la séquence du récepteur hM3, pour lequel la toxine montre une affinité supérieure au micromolaire. Les boucles externes E2 de hM1 et hM3 diffèrent par huit résidus (Figure 37). Sept de ces huit résidus ont été remplacés par ceux correspondants dans la séquence du récepteur hM3. Il s'agit des mutations L167F, E170K, L174P, A175P, Q177E, Y179F et Q185E.

La cartographie des effets des mutants sur le modèle du récepteur hM1 (Figure 38) permet de constater que les résidus importants pour l'interaction avec la toxine sont centrés au niveau des régions extracellulaires du récepteur, sur la boucle E2 et en haut des hélices transmembranaires.



```
          170          180  
hM1 QYLVGERTVLAGQCYIQFLSQPI  
hM3 QYFVGKRTVPPPGECFIQFLSEPT
```

**Figure 37: Séquences peptidiques des boucles E2 du hM1 et du hM3**



**Figure 38 : Mutagenèse du récepteur hM1**

*En rouges les résidus déterminés comme importants pour l'interaction. En bleu les résidus dont les mutants alanines sont sans effet sur l'interaction. En violet la boucle E2*

<b>Mutation</b>	<b>région</b>	<b>G</b>
W23A	N-ter	-0,31
H90A	boucle E1	0,2
<b>W91A</b>	<b>boucle E1</b>	<b>1,64</b>
W101A	TM3 / coeur	0,93
W164A	boucle E2	1,22
Y166A	boucle E2	-0,31
L167A	boucle E2	0.00
E170A	boucle E2	0,62
<b>E170K</b>	<b>boucle E2</b>	<b>1,96</b>
<b>R171A</b>	<b>boucle E2</b>	<b>1,65</b>
L174A	boucle E2	0,2
Q177E	boucle E2	0,06
<b>Y179A</b>	<b>boucle E2</b>	<b>2,33</b>
<b>Y179F</b>	<b>boucle E2</b>	<b>1,77</b>
L183A	boucle E2	0,28
Q185A	boucle E2	0,28
Q185E	boucle E2	-0,42
K392A	Boucle E2	0,35
D393A	Boucle E2	0,42
E397A	TM7	0,16
<b>W400A</b>	<b>TM7</b>	<b>2,03</b>
E401A	TM7	0,28

**Figure 39 : Variation d'énergies libres associées à des mutations du récepteur**

*Les mutations entraînant des variations d'affinité significatives sont en gras  
D'après (Marquer 2008)*

### 3.1.3. Cycles de doubles mutants sur le système hM1/MT7

Afin de déterminer les régions des structures des deux partenaires impliquées dans l'interaction, des expériences de cycles de doubles mutants ont été réalisées. Ces expériences permettent de déterminer des couples de résidus en interaction dans la structure du complexe. Ces données sont ensuite traduites en contraintes de distance lors des expériences de simulation d'arrimage.

#### 3.1.3.1. Choix des couples de résidus

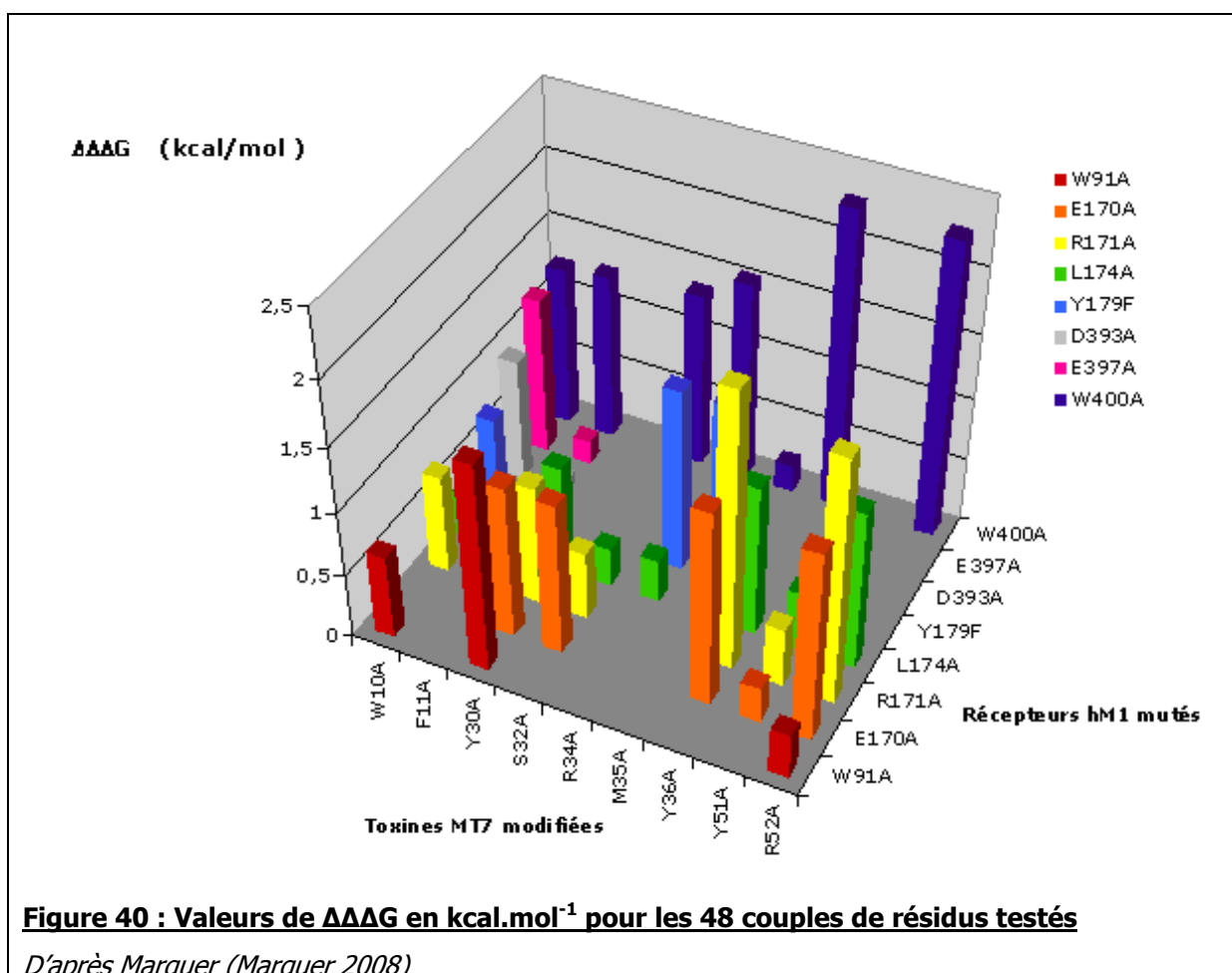
Lorsqu'une dizaine de proximités ont été disponibles les premiers modèles du complexe ont pu être construits. Sur cette base, de nouvelles mutations ont été proposées pour trancher entre les différentes hypothèses suggérées par la modélisation. Ainsi, de nombreux cycles de mutagenèse / docking ont permis de faire converger les modèles en écartant au fur et à mesure les ambiguïtés.

Le choix des premières positions mutées a été fait sur la base des résultats des expériences d'alanine scanning. Les positions à muter ont été choisies de manière à couvrir les trois boucles de la toxine, ainsi que les trois boucles extracellulaires et la cavité apicale du récepteur. Un certain nombre de positions dont la participation au site de liaison était jugée improbable ont été également choisies afin de disposer en nombre suffisant de contrôles négatifs, indispensable au traitement statistique des données.

3.1.3.2. Résultats des expériences de cycles de doubles mutants

Au total, 48 couples de résidus toxine/récepteur ont été testés correspondant à 8 positions coté récepteur et 9 positions coté toxine (Figure 40). La valeur de  $\Delta\Delta G$  considérée comme seuil de significativité pour conclure à une proximité entre résidus est fixée à  $0.7 \text{ kcal.mol}^{-1}$  ce qui correspond à un seuil deux fois plus stringent que celui de  $0.35 \text{ kcal.mol}^{-1}$  utilisé par Fersht (Schreiber and Fersht 1995). On identifie ainsi au total 23 paires de résidus en interaction dans la structure du complexe hM1/MT7.

Parmi les proximités identifiées, onze concernent la boucle E2 du récepteur, huit le haut des hélices et une au niveau de la boucle E3. Ceci correspond au niveau de la toxine à cinq contact sur la boucle I, quatorze sur la boucle II et quatre boucle III.



### ***3.2. Exploration de la flexibilité conformationnelle***

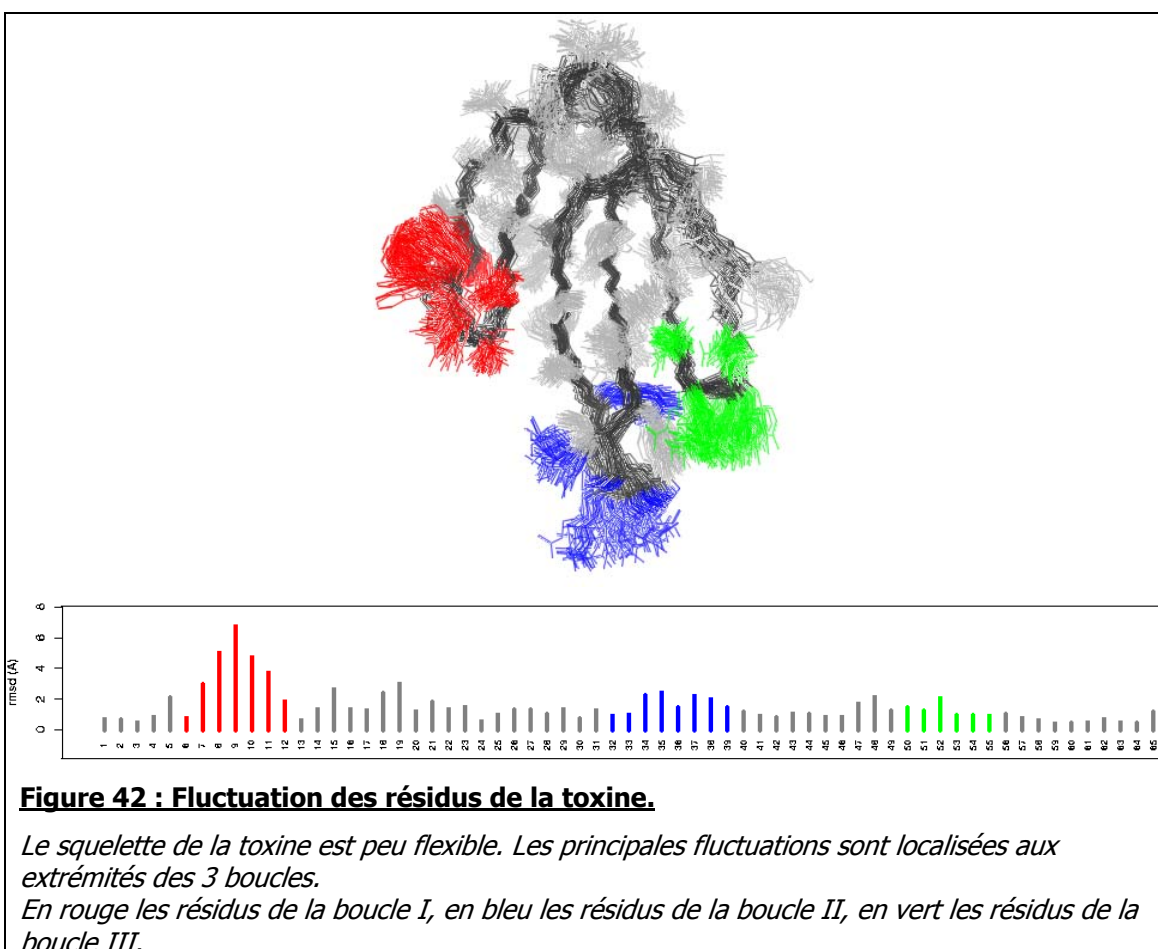
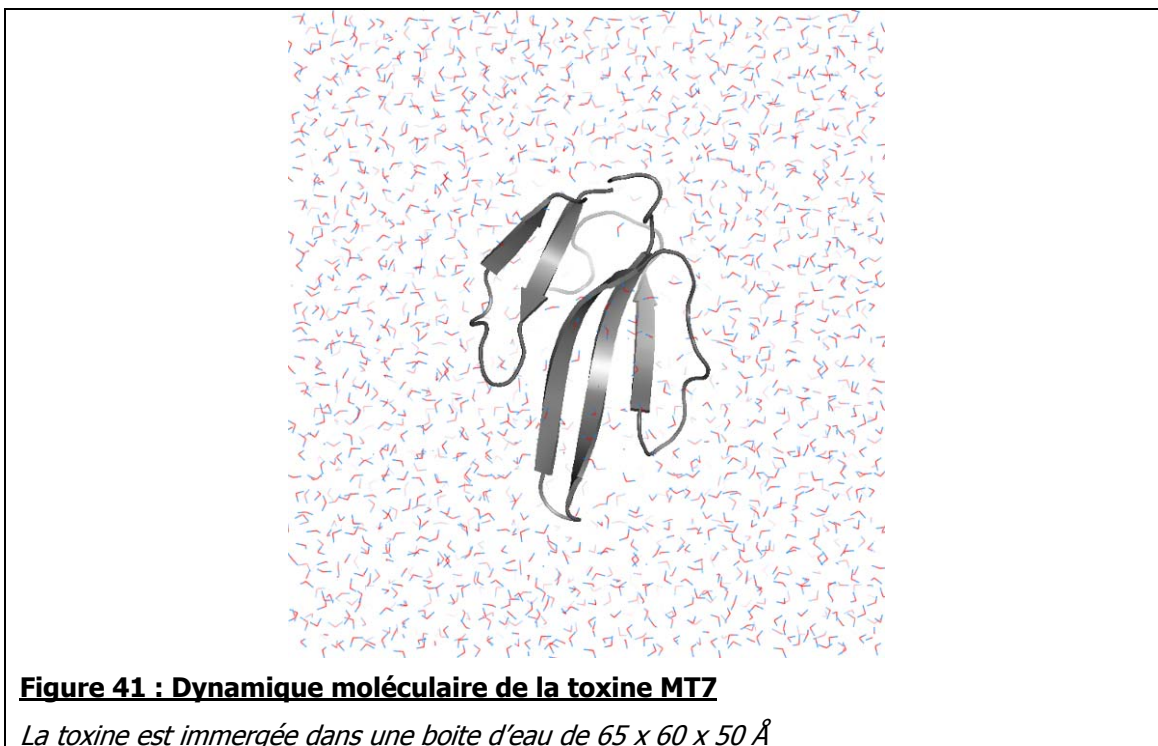
Il est maintenant largement admis que la prise en compte des modifications conformationnelles intervenant entre les formes libres et liées des protéines d'un complexe sont une des clefs du succès des expériences de docking moléculaire. Ceci a pu être vérifié à plusieurs reprises sur des cibles du concours CAPRI (cf § 1.6.3). De nombreux travaux récents s'intéressent à l'étude de l'impact de ces changements conformationnels (Grunberg, Leckner et al. 2004; Ehrlich, Nilges et al. 2005) sur la prédiction de la structure des complexes de protéines. En particulier, nous avons conclu de l'analyse des modèles construits précédemment au laboratoire (cf § 1.6.5) qu'un échantillonnage des conformations des partenaires avant arrimage aurait pu améliorer sensiblement les prédictions.

#### **3.2.1. Dynamique moléculaire de la toxine**

Pour échantillonner des conformations de la toxine, nous avons utilisé la dynamique moléculaire avec le programme CHARMM dans le champ de force CHARMM22. Les mouvements de la toxine ont été simulés pendant une durée de 2,5 nanosecondes, dans un solvant explicite (Figure 41). On observe globalement peu de déformation de la structure de la toxine au cours de cette simulation avec des écarts moyens sur le squelette carboné le long de la trajectoire de 1.4 Å (Figure 42). La région du squelette qui apparaît comme la plus flexible est l'extrémité de la boucle I avec un maximum au niveau du résidu 9. Ce résultat est en bon accord avec l'étude par RMN d'une toxine très similaire, la MT2. L'impossibilité de mesurer des effets NOEs au niveau de la boucle I de la toxine MT2 suggérant des échanges conformationnels à des temps longs pour ces résidus, et donc probablement des mouvements de grande amplitude.

Comme on pouvait s'y attendre les mouvements des chaînes latérales sont principalement localisés au niveau des extrémités des 3 boucles de la toxine. En particulier, les résidus W10 et R34 (Figure 42). Afin d'obtenir des échantillons de conformations de la toxine aussi différents que possibles les uns des autres, nous avons appliqué une méthode statistique d'agrégation des données (clustering) sur 1000 coordonnées extraites de la trajectoire de dynamique moléculaire. Un ensemble de 10

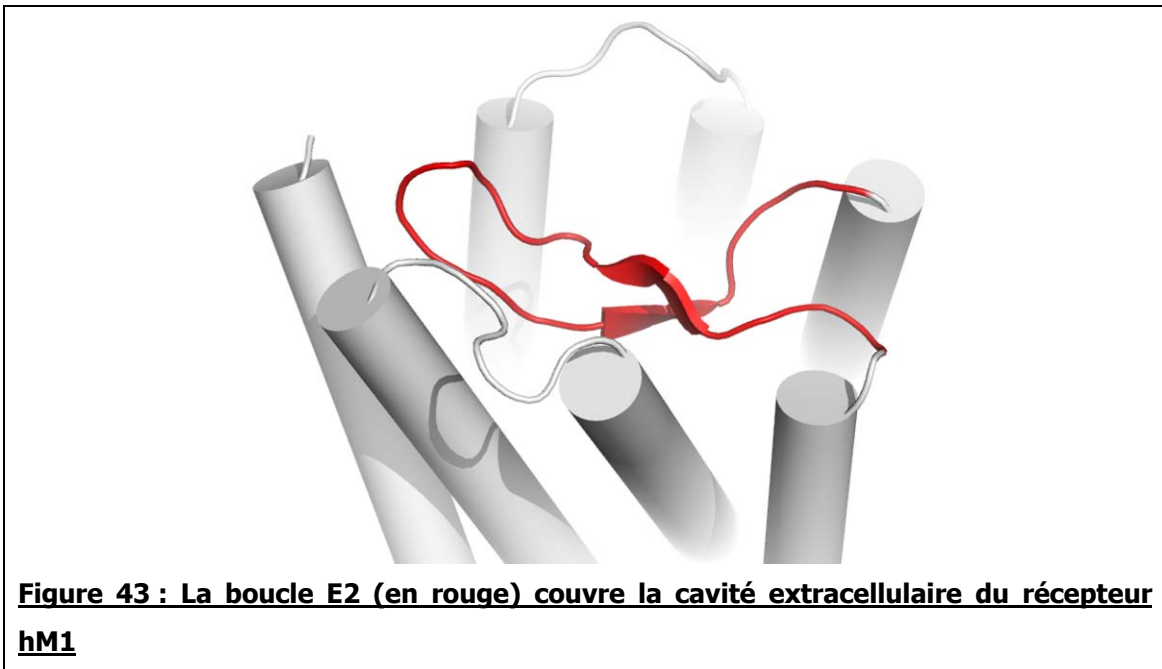
groupes a été obtenu par la méthode des nuées dynamiques (kmeans). Par la suite, les 10 structures les plus proches du noyau de chaque groupe seront conservées.



## Dynamique moléculaire du récepteur

### 3.2.1.1. La boucle extracellulaire E2

Nous ne disposons que d'un modèle du récepteur hM1, construit par modélisation comparative d'après la structure de la rhodopsine. Dans ce modèle la boucle extracellulaire E2 est dans une conformation étendue, dans la cavité apicale du récepteur (Figure 43).



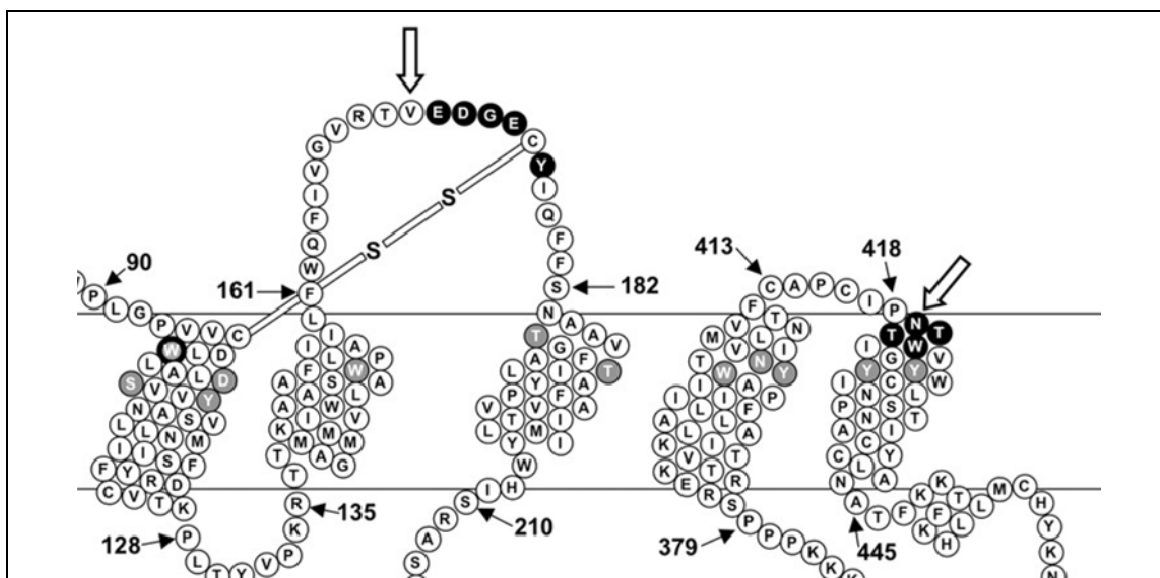
Depuis la parution de la première structure de la rhodopsine en 2000 (Palczewski, Kumasaka et al. 2000), l'apparente incompatibilité entre la conformation de cette boucle et la possibilité d'accès au site de liaison d'un ligand de haut poids moléculaire sur un RCPG est un sujet largement débattu dans la littérature (Oliveira, Hulsen et al. 2004; Klco, Wiegand et al. 2005). Plusieurs équipes ont tenté de répondre à cette question. Il apparaît par exemple d'après l'enfouissement des résidus de la boucle E2 du récepteur dopaminergique D2 que la boucle participe directement au site de liaison et adopte chez ce récepteur une position étendue dans la cavité, similaire à celle de la rhodopsine (Shi and Javitch 2004).

D'autres expériences ont permis d'explorer par mutagenèse les possibilités de changement conformationnels de cette boucle. En particulier sur le récepteur muscarinique M2, il a été réalisé une modification de la boucle en introduisant un pont

disulfure entre l'extrémité de la boucle E2 et le segment transmembranaire TM7 (Figure 44). La présence de ce pont disulfure non natif (réversible en présence de DTT) réduit fortement la flexibilité de la boucle (Avlani, Gregory et al. 2007). Les auteurs observent que lorsque ces cystéines sont réduites, l'activité du récepteur est normale, alors que lorsque le pont est oxydé, on inhibe la liaison des ligands orthostériques et allostériques du récepteur. Ceci indique donc :

- La réalisation de cette liaison implique que la boucle E2 du hM1 peut adopter une conformation étendue au dessus de la cavité, semblable à la boucle E2 de la rhodopsine.
- Cette boucle a la capacité d'adopter une conformation plus ouverte autorisant l'accès des ligands, tout comme celle du récepteur  $\beta_2$ -adrénergique.

La boucle E2 du hM1 est donc capable de réaliser des mouvements de grande amplitude entre des conformations ouvertes et fermées, et module ainsi l'accès au site de liaison.



**Figure 44 : Modification de la boucle E2 du récepteur hM2**

*Le récepteur M2 est représenté en diagramme serpent. Le pont natif de la boucle e2 est représenté. Les flèches blanches indiquent les résidus mutés en cystéine  
D'après Avlani et al (Avlani, Gregory et al. 2007)*

### 3.2.1.2. La méthode PEDC

Nous avons donc souhaité apporter un soin particulier au traitement de la flexibilité de la boucle extracellulaire E2 du récepteur. Par la dynamique moléculaire, on peut espérer échantillonner des changements conformationnels jusqu'à 2 ou 3 Å



d'écart par rapport à la structure native. Cette capacité d'exploration est insuffisante pour simuler les mouvements nécessaires pour dégager l'accès du site à un ligand de haut poids moléculaire tel que notre toxine. Nous avons donc utilisé une méthode basée sur de la dynamique moléculaire activée PEDC. Le principe de la méthode PEDC (Guilbert, Perahia et al. 1995; Gilquin, Guilbert et al. 2000) est d'explorer les chemins conformationnels du paysage énergétique d'une protéine en s'éloignant progressivement de la conformation de départ. Il est possible de diriger l'évolution du système vers une structure finale connue ou bien de le laisser évoluer librement en pénalisant les conformations proches de celle de départ. La recherche conformationnelle est assurée par une simulation de dynamique moléculaire classique à laquelle on ajoute un potentiel supplémentaire (Équation 1) dépendant du RMSD par rapport à la structure de départ. Ce potentiel va donc pénaliser les conformations proches de la structure native permettant au système de franchir plus facilement les barrières énergétiques

La contrainte de RMSD est exprimée par une fonction harmonique classique :

$$V_{dist}^J = \frac{1}{2} k_{dist} (d_R^J - d_0^J)^2,$$

**Équation 1 : Potentiel PEDC**

$d_R^J$  est le RMSD entre la structure J et la structure de référence R.  
k est la constante de force, exprimée en kcal. Mol<sup>-1</sup>

Le seuil de distance de contrainte va être augmenté progressivement au cours du temps pour d'atteindre des valeurs d'écart par rapport à la structure de départ de plus en plus importantes. Des contraintes sont également ajoutées afin de ne pas tenir compte des composantes rotationnelles et translationnelles du mouvement de la protéine dans les calculs. Il est intéressant d'observer que par cette approche, la variation conformationnelle très importante dans la boucle ne s'accompagne d'aucune variation brutale de l'énergie interne du système qui reste pratiquement constante au cours du temps. Ceci tend à prouver que la structure évolue au cours de l'exploration le long de chemins préférentiels de basse énergie.

Cette méthode a été appliquée et validée sur le lysozyme et a permis de mettre en évidence le chemin de repliement de la protéine, ceci en accord avec les données expérimentales disponibles (Gilquin, Guilbert et al. 2000).

### 3.2.1.3. Dynamique moléculaire du récepteur

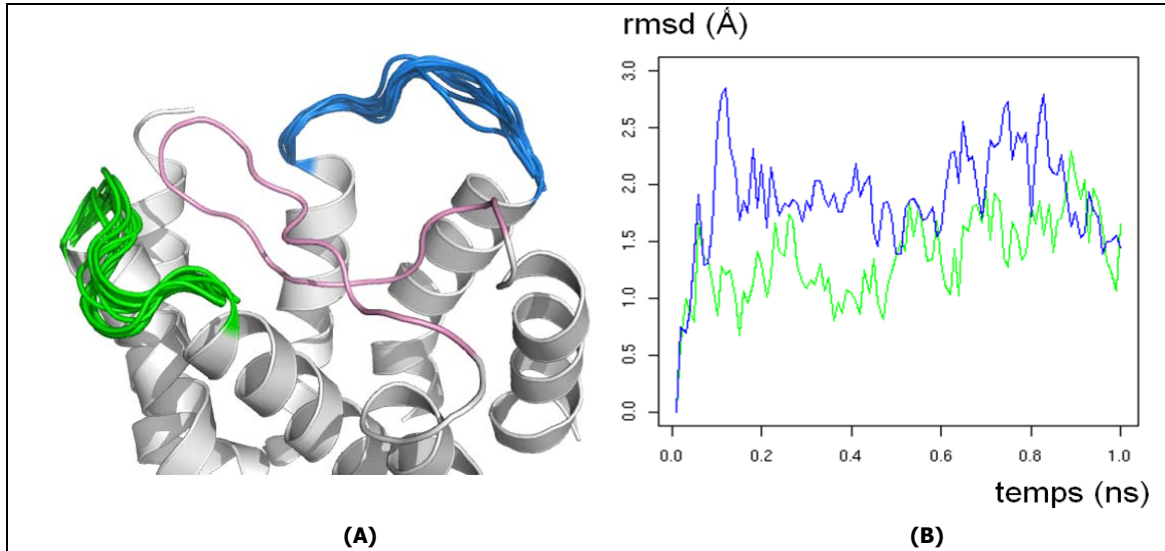
La boucle E2 du récepteur est contrainte dans la cavité de liaison par le pont disulfure établi par les cystéines 98 (hélice 4) et 178. Pour simuler l'ouverture de la cavité, nous avons fait l'hypothèse que la boucle effectuait un mouvement de rotation autour d'une charnière, dont le pont disulfure est l'un des pivots. Nous avons donc utilisé le résidu de la boucle situé en face de la cystéine 178, la glycine 169 comme second pivot (Figure 46).

Le modèle du récepteur hM1, a donc été immergé dans une boîte de molécules d'eau explicites. Afin d'éviter une attaque des hélices transmembranaires par le solvant, celle-ci sont maintenues pendant toute la durée de la simulation par des contraintes harmoniques. Enfin, le potentiel PEDC est appliqué aux résidus 169 à 177 de la boucle E2. Ce système a été soumis à une dynamique moléculaire de 1 nanoseconde. Le coefficient du potentiel PEDC est constant au cours de la simulation, mais toutes les 10 picosecondes, on incrémente le seuil de distance (exprimée en RMSD) à partir duquel la contrainte est appliquée de 0.1 Å (Figure 47, courbe rouge). Ainsi, à l'issue de la simulation, la boucle a effectué un mouvement d'une amplitude totale de 10Å (Figure 47, courbe verte).

Avec des écarts globalement inférieurs à 2.5 Å par rapport à la structure native, les conformations des boucles extracellulaires E1 et E3 du récepteur hM1 fluctuent peu durant la simulation de dynamique moléculaire (Figure 45). Ce résultat n'est pas surprenant compte tenu de la courte durée de cette simulation.

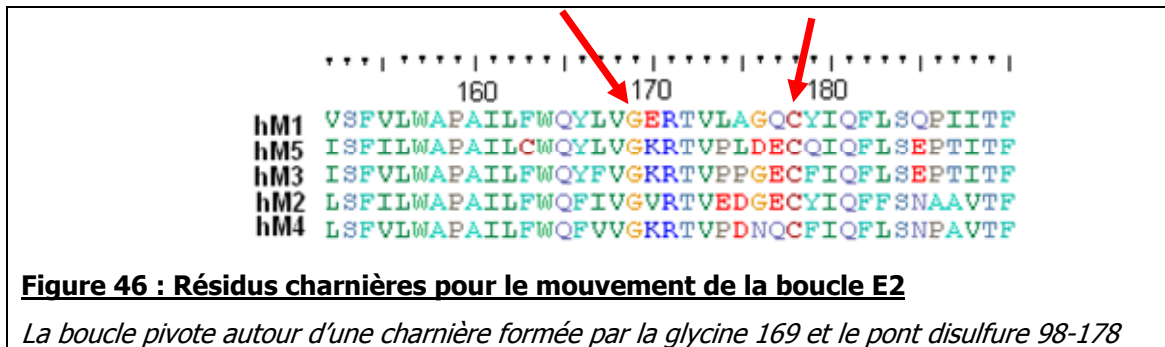
On peut mesurer la déformation de la boucle, en calculant les écarts entre des coordonnées au cours de la simulation et sa conformation initiale. Après ajustement des coordonnées des carbones- des résidus 169 à 177 (Figure 47, bleu), on constate que ces valeurs restent faibles (< 2.5Å) en regard de l'amplitude totale du mouvement qui est de 10Å. Ceci démontre que l'on assiste bien à une ouverture de la boucle et pas à une disruption complète de la structure. On observe tout de même sur cette courbe qu'une transition intervient vers 0.7ns avec une accélération de la variation du rmsd. L'analyse des structures nous montre qu'on assiste effectivement à partir de ce moment à une déformation de la boucle. Les dernières conformations (Figure 48B) devenant différentes de la forme native de la boucle.

On peut donc en conclure que la procédure PEDC nous permet bien de simuler un mouvement de charnière d'ouverture de la boucle, et non pas à un dépliement de la structure, au moins jusqu'à 7Å



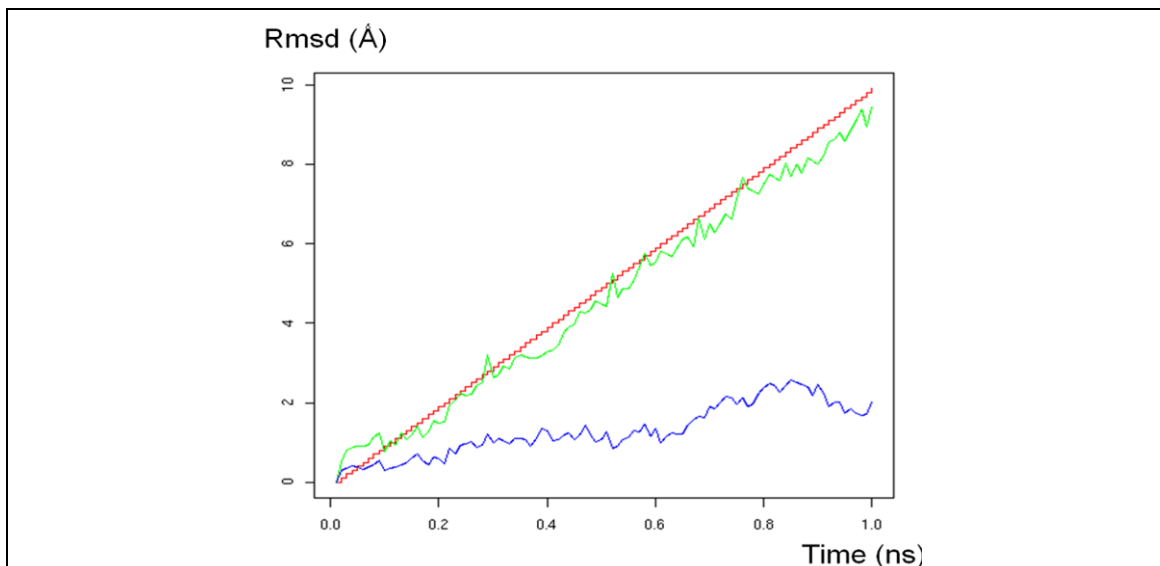
**Figure 45 : Flexibilité des boucles extracellulaires E1 et E3 du hM1**

La boucle E1 est représentée en vert et la boucle E3 en bleu  
 (A) superposition de 10 coordonnées des boucles E1 et E2  
 (B) évolution des conformations au court du temps



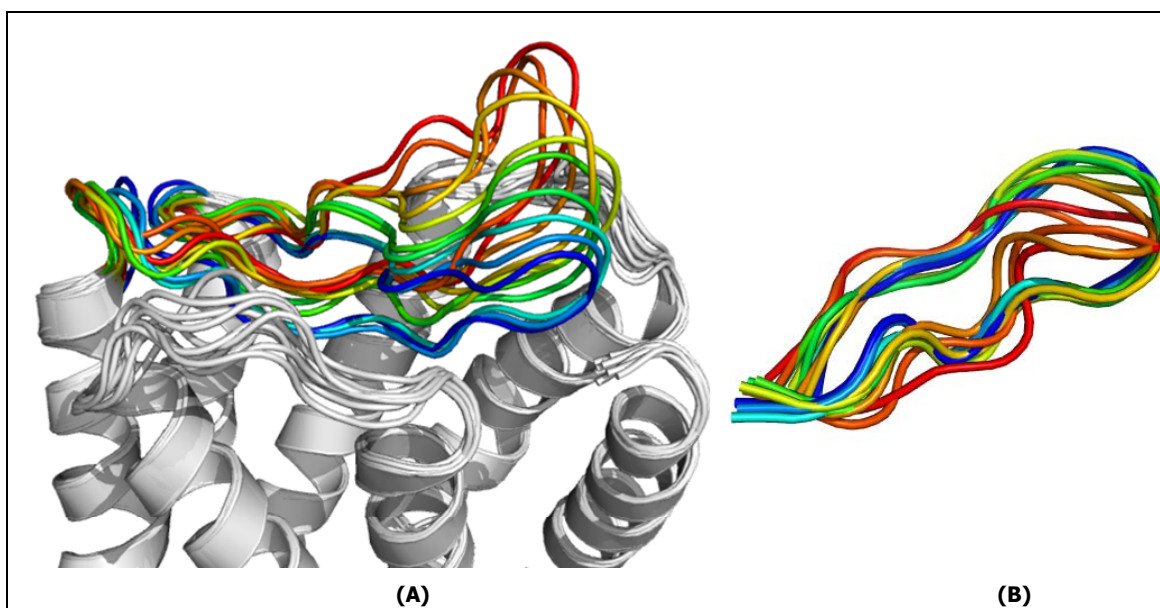
**Figure 46 : Résidus charnières pour le mouvement de la boucle E2**

La boucle pivote autour d'une charnière formée par la glycine 169 et le pont disulfure 98-178



**Figure 47 : Déformation de la boucle E2 au cours de la dynamique**

*L'intensité de la contrainte est tracée en rouge, l'écart avec la conformation de départ de la boucle E2 est en vert et la fluctuation des conformations après superposition en bleu*



**Figure 48 : Dynamique activée de la boucle E2 du récepteur**

*A : mouvement global de la boucle E2, du bleu vers le rouge  
B : superposition de conformations des boucles E2 extraites de la trajectoire  
La couleur des boucles, des plus froides vers les plus chaudes correspond au temps.*

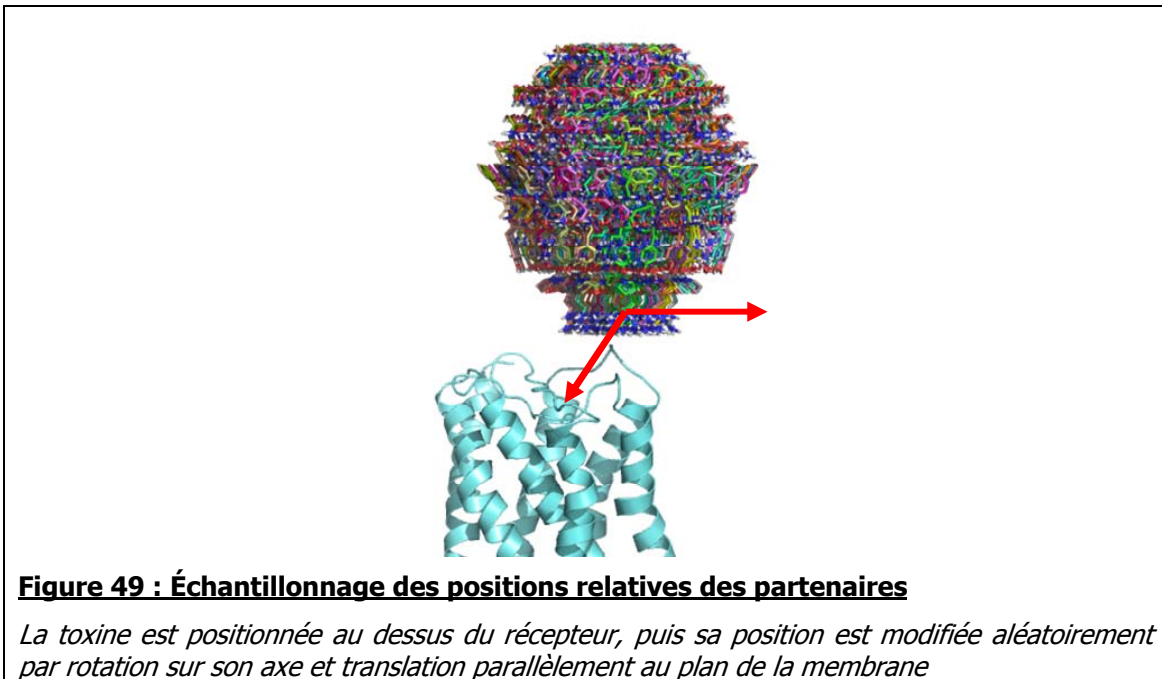
### 3.3. Procédure d'arrimage moléculaire des partenaires

#### 3.3.1.1. Dynamique moléculaire

La procédure de docking que nous avons adoptée repose sur l'utilisation de dynamique moléculaire sous contraintes. Les calculs sont effectués dans le vide avec le programme de dynamique moléculaire Charmm. La constante diélectrique est fixée à  $\epsilon = r$  (option RDIE). Afin d'alléger le temps de calcul nécessaire, nous utilisons également le champ de force à atomes unifiés Charmm19.

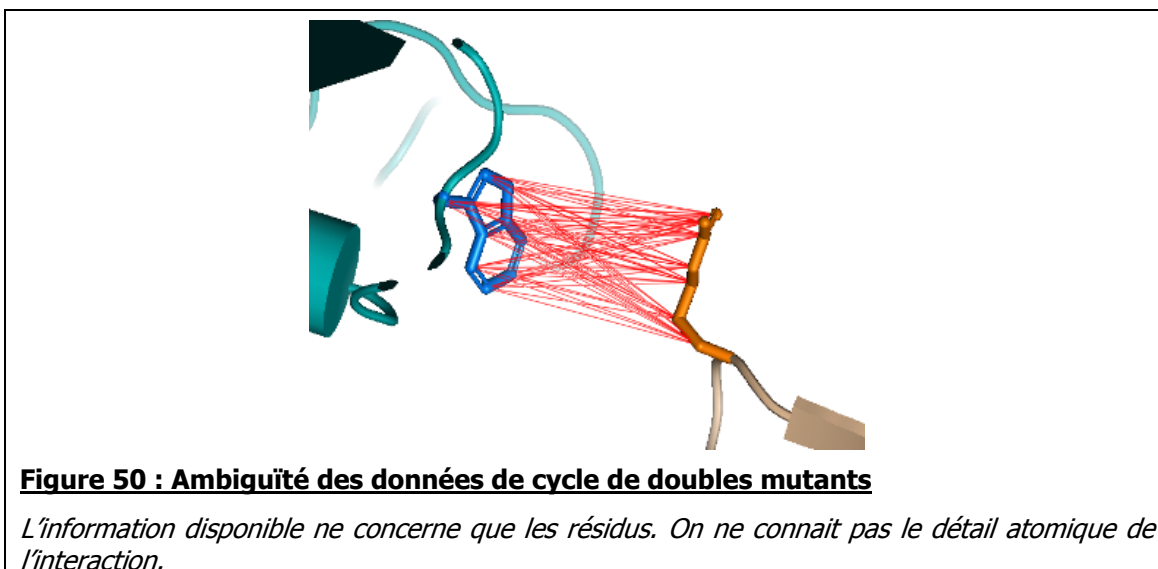
#### 3.3.1.2. Échantillonnage des positions de départ

Pour chaque répétition de la simulation d'arrimage, les positions relatives du récepteur et de la toxine sont échantillonnées (Figure 49). La toxine est placée verticalement, au dessus de la membrane, les extrémités des boucles dirigées vers le récepteur. Ensuite, on effectue une rotation aléatoire de la toxine autour d'un axe perpendiculaire à la membrane, puis une translation de 0 à 5Å dans une direction aléatoire dans le plan de la membrane.



3.3.1.3. Introduction des données expérimentales

L'information fournie par les cycles de doubles mutants ne concerne que les résidus et ne donne pas de détails à l'échelle atomique (Figure 50), il est donc nécessaire d'introduire de l'ambiguïté dans le traitement des contraintes.



La RMN des protéines est également confrontée au même type d'ambiguïté dans l'attribution des effets nucléaires Overhauser (NOE). Des contraintes ambiguës de distance sont utilisées pour attribuer les dérivées les effets NOE et déterminer de la structure des protéines en solution. Dans le cas de la RMN l'intensité de l'effet NOE est une somme pondérée de la contribution de chaque noyau à l'interaction.

Dans le cas des cycles de doubles mutants en revanche, on ne dispose d'aucun moyen de corréler la mesure de  $G$  à la distance de la paire de résidus. En effet, cette mesure va dépendre de la nature de chaque résidu muté, de leur contexte structural, etc. On ne dispose finalement que d'une information qualitative binaire sur la proximité ou non des deux résidus. Nous avons choisi d'utiliser un critère simple pour tenir compte de l'ambiguïté des données: la distance entre les deux atomes les plus proches de la paire de résidus, à l'exclusion des protons (Équation 2).

La valeur du potentiel que nous appliquons sur la paire d'atome la plus proche de chaque paire de résidu sous contrainte prend donc la forme suivante :

$$\left\{ \begin{array}{l} \text{Si } d > d_{\text{cible}}: E_{\text{cons}} = \frac{1}{2} \cdot K \cdot (d - d_{\text{cible}})^2 \\ \text{Si } d < d_{\text{cible}}: E_{\text{cons}} = 0 \\ \text{Avec } E_{\text{cons}}: \text{potentiel de contrainte} \end{array} \right.$$

**Équation 2 : Traitement des contraintes de distance ambiguë**

*d* : distance de la paire d'atomes les plus proches

*d<sub>cible</sub>* : distance en dessous de laquelle la contrainte est nulle

*K* : Coefficient arbitraire

La valeur du paramètre  $d_{\text{cible}}$  (Équation 2) appliquée pour le docking est de 5Å. Il s'agit en fait d'un compromis, entre la distance minimale entre deux atomes dans la structure d'une protéine, de l'ordre de 2Å et la valeur seuil de 8Å entre deux résidus prédits comme proches par les doubles mutants (Schreiber and Fersht 1995). En effet, appliquer uniformément des contraintes de 8Å de distance ne permet pas d'obtenir des modèles aux interfaces suffisamment compactes, tandis que l'utilisation de contraintes plus fortes conduit à des complexes ayant des énergies de van der Waals inacceptables.

3.3.1.4. Utilisation de la dynamique moléculaire en 4D

Échantillonner de manière exhaustive les possibilités d'association d'un complexe protéine/protéine, se révèle extrêmement coûteux en termes de puissance de calcul. Classiquement, les méthodes de docking procèdent donc à une première étape visant à réduire l'espace de recherche en utilisant des méthodes d'exploration rapide. C'est par exemple le cas de Rosetta qui procède à une première étape de docking en corps rigide avec une représentation simplifiée des chaînes latérales (Gray, Moughon et al. 2003). Dans le cas des expériences de docking par dynamique moléculaire sous contrainte que nous réalisons, l'espace de recherche conformationnel se trouve fortement réduit grâce aux données expérimentales. Ceci nous permet de débiter nos calculs directement avec une représentation tout atome et flexible des protéines.

Un autre facteur limitant la recherche conformationnelle, est la rugosité du paysage énergétique des protéines, constitué de nombreuses " vallées " dans lesquelles les algorithmes de recherche sont susceptibles de se trouver piégés. Pour limiter ce

phénomène, il est classique d'utiliser des champs de force simplifiés ayant un nombre de paramètres plus faibles afin de lisser le paysage énergétique.

Pour notre part, nous avons plutôt utilisé les avantages que nous offre la dynamique moléculaire. D'une part, nous pouvons contrôler la température de la simulation. En augmentant cette température, on fournit de l'énergie cinétique au système et on permet ainsi le passage de barrières énergétiques. Ce levier doit cependant être manié avec délicatesse, car en abuser conduit inévitablement à la rupture d'interactions légitimes au sein de la structure des protéines, ce qui se traduit par sa dénaturation, ou à une divergence des calculs.

Une autre possibilité basée sur l'utilisation de dynamique moléculaire en 4D a été introduite par Torda et Van Gunsteren (van Schaik, Berendsen et al. 1993). Cette méthode consiste à projeter les coordonnées tridimensionnelles des protéines dans un espace à 4 dimensions. En augmentant ainsi le nombre de degrés de liberté du système, on obtient une hyper-surface en 4D dans laquelle les barrières énergétiques sont abaissées. On garde cependant une représentation où l'on peut faire l'hypothèse que les minimums locaux restent superposables avec ceux de la 3D. Les coordonnées des protéines dans la 4<sup>ème</sup> dimension restant strictement contraintes, l'espace de recherche ne s'en trouve pas augmenté. On peut donc contrôler de manière dynamique au cours de la simulation, la distance entre les deux partenaires dans la 4<sup>ème</sup> dimension fictive. Ceci permet d'agir très finement sur le degré de lissage du paysage énergétique que l'on souhaite obtenir.

#### 3.3.1.5. Conclusion

La procédure de docking que nous utilisons consiste donc en une procédure de recuit simulé, au cours duquel, la température du système est augmentée progressivement jusqu'à 500K et les protéines séparées dans une quatrième dimension fictive. On contrôle le décalage des deux protéines dans la 4<sup>ème</sup> dimension en appliquant des contraintes harmoniques de distance imposant un écart jusqu'à 1 Å durant la simulation. Ensuite, après la phase d'échantillonnage, le système est progressivement refroidi tandis que les deux protéines sont ramenées lentement dans le même plan de la 4<sup>ème</sup> dimension. Durant, cette simulation, les données expérimentales sont introduites sous la forme de contraintes de distance ambiguës.



### 3.4. Topologie des contraintes

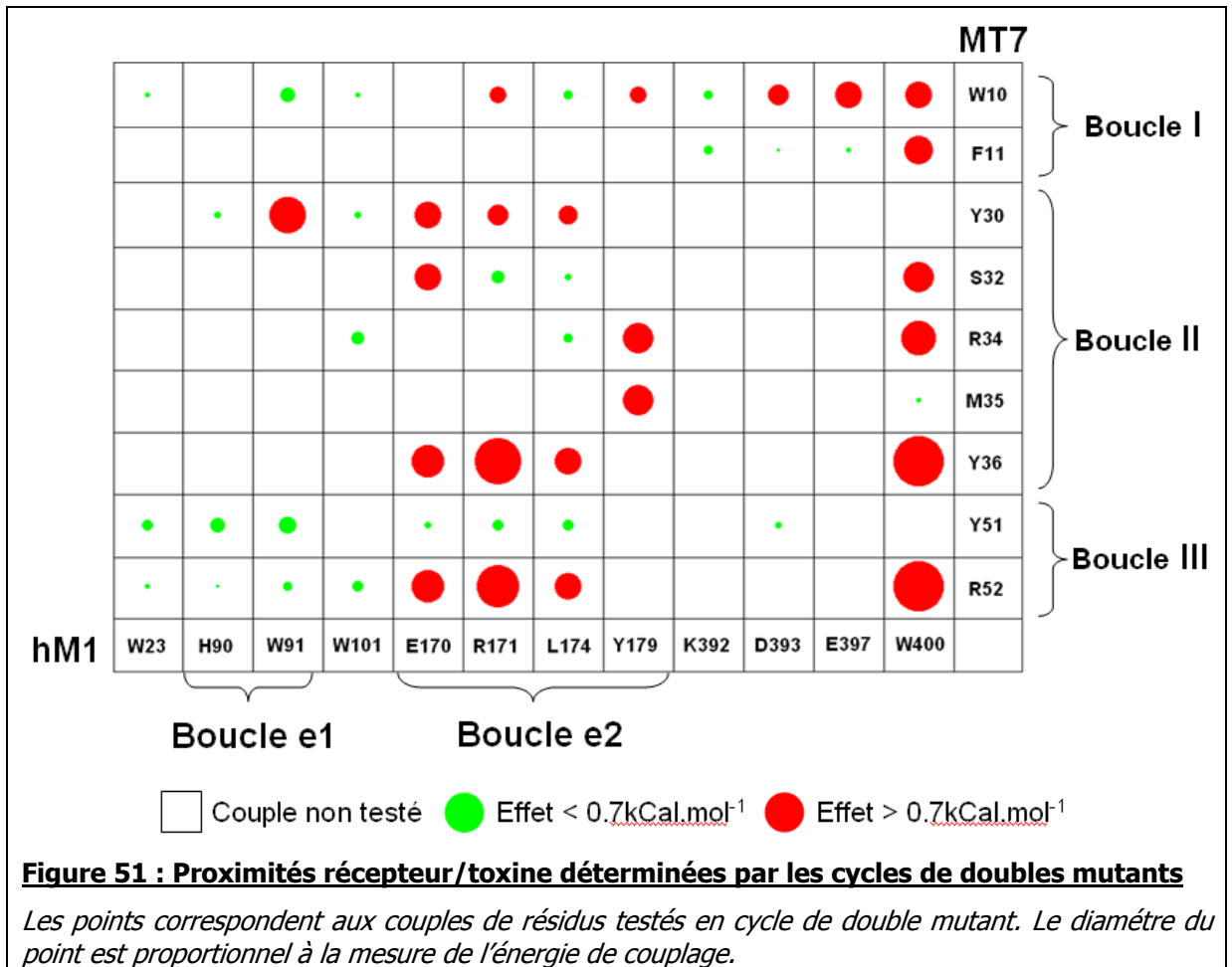
La topologie des contraintes déduite des expériences de cycles de doubles mutants (

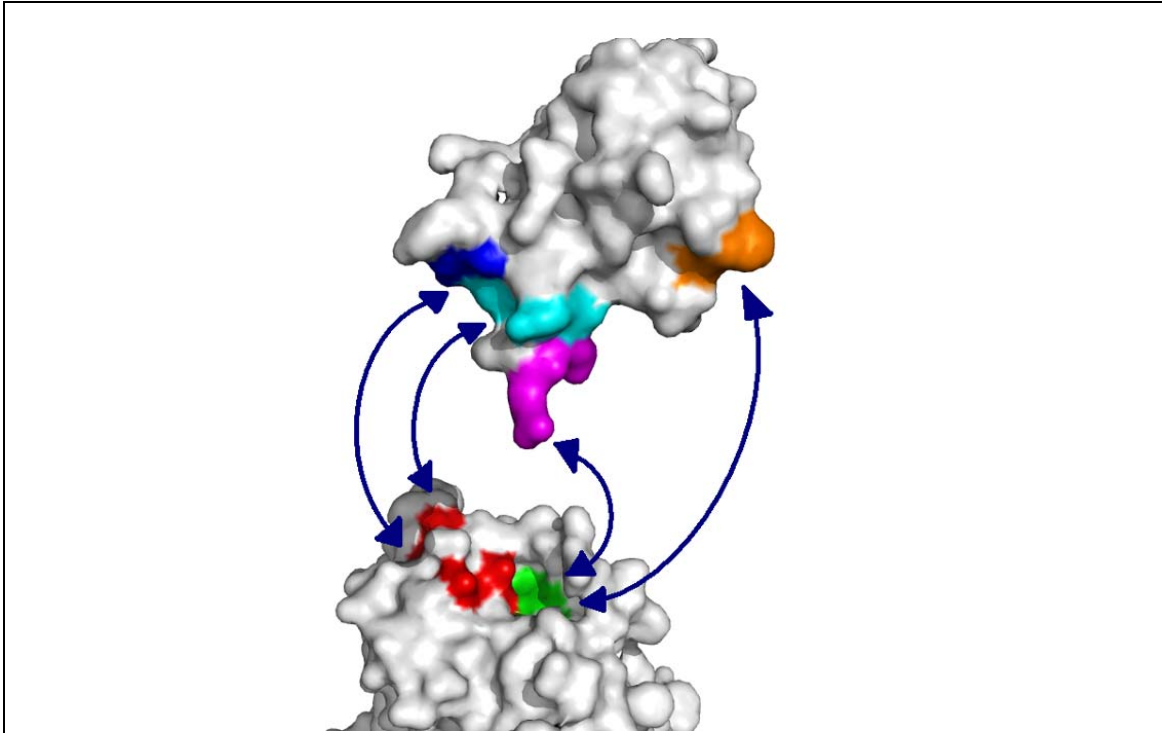
□ Couple non testé ● Effet < 0.7kCal.mol<sup>-1</sup> ● Effet > 0.7kCal.mol<sup>-1</sup>

Figure 51) est cohérente puisque des groupes de résidus proches dans la structure de la toxine sont en contacts avec les mêmes groupes de résidus proches dans celle du récepteur (Figure 52). Par exemple,

- le groupe boucle I de la toxine (W10 et F11) est localisé dans la cavité du récepteur (Y179 et W400).
- Le haut de la boucle II de la toxine (S32, Y36) est en contact avec la boucle E2 du récepteur (E170, R171, L174)
- Le bas de boucle II de la toxine (R34, M35) est situé dans la cavité du récepteur (Y179 et W400)
- R52, seul résidu identifié pour la boucle III de la toxine est en contact avec des résidus de la boucle E2 du récepteur (E170, R171, L174)

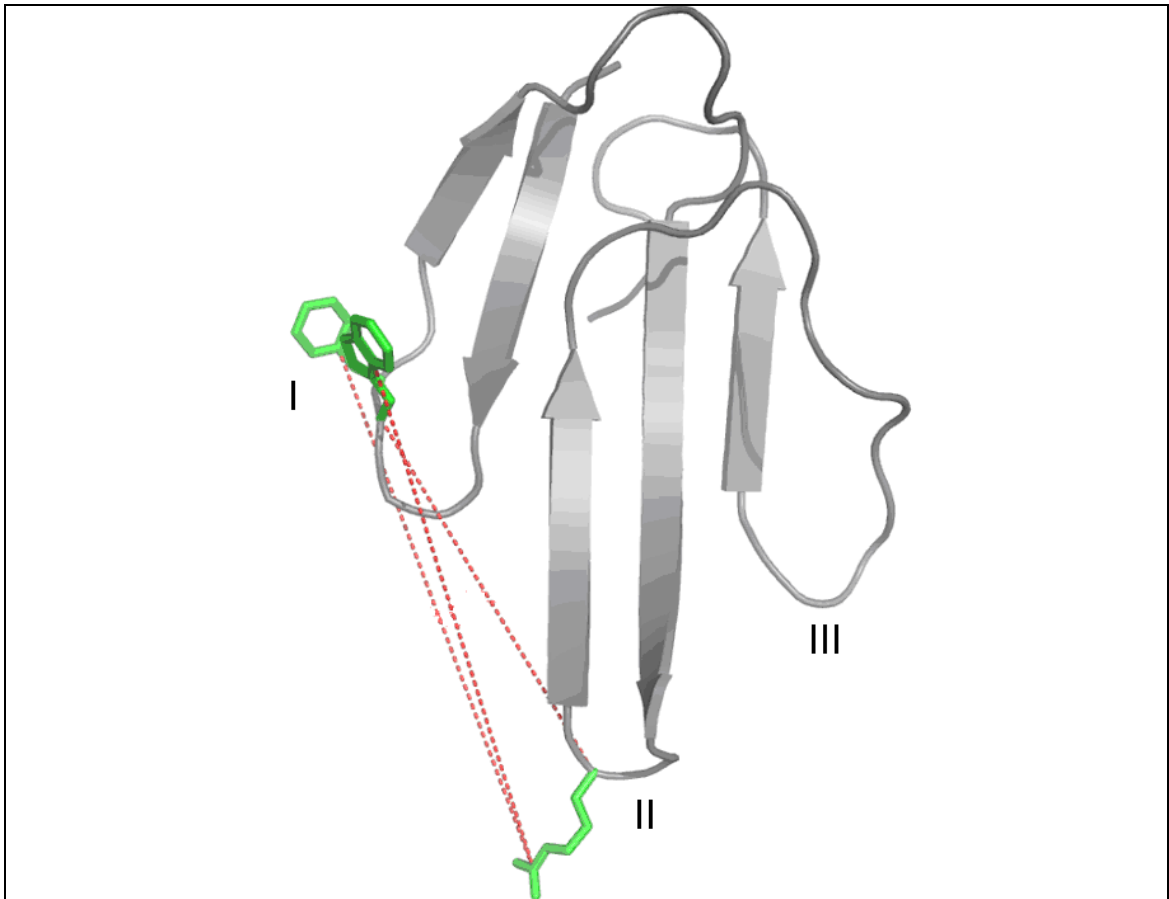
Mais en parallèle, on observe que les mêmes partenaires au niveau de la cavité du récepteur sont partagés par les résidus de la boucle I et II de la toxine. Or les résidus concernés coté toxine (Figure 53) ont leurs carbones distants de plus de 25 Å et leurs atomes les plus proches à 16 Å.





**Figure 52 : Groupes de résidus en interactions**

*Pour le récepteur : En rouge, les résidus boucle E2, en vert les résidus de la cavité de liaison*  
*Pour la toxine : en orange les résidus boucle I, en magenta le bas de la boucle II en cyan les M35 et Y36 également sur la boucle II et en bleu R52 sur la boucle III*



**Figure 53 : Incohérence dans les contraintes**

*Résidus du bas de la boucle II et de la boucle I sont distants de plus de 20 et partagent pourtant les mêmes points de contact coté récepteur*

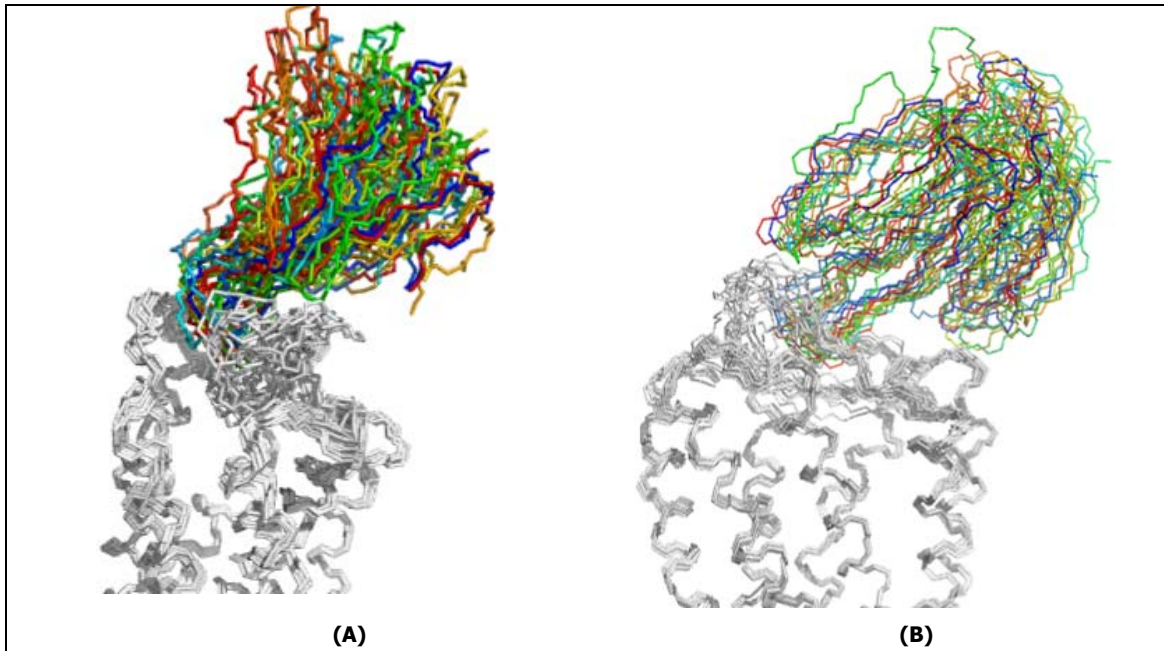


### 3.5. Premiers modèles

Ce paradoxe apparent dans la topologie des contraintes est à l'origine d'échecs systématiques dans la construction de modèles du complexe hM1-MT7 en utilisant les 23 contraintes. Aucun modèle hM1-MT7 satisfaisant l'ensemble des contraintes ancrées sur les trois boucles de la toxine et des données structurales n'a pu être construit. Nous avons donc adopté une stratégie visant à déterminer un sous ensemble de contraintes qui pourraient être satisfaites simultanément dans un modèle hM1 MT7. Les 23 contraintes disponibles ont donc été séparées en trois groupes, en fonction de la boucle de la toxine sur laquelle elles sont ancrées. Les effets très importants mesurés pour la mutation de l'arginine 34 localisée au bout de la boucle II ainsi que sa position centrale dans la structure de la toxine indiquent clairement que la boucle II de la toxine est au cœur de l'interaction avec le récepteur.

10.000 essais d'arrimage (correspondant au total à 1 $\mu$ s de dynamique moléculaire) ont donc été réalisés avec les contraintes boucle I+II (17 contraintes) et boucle II+III (19 contraintes). Les simulations basées sur l'utilisation des contraintes impliquant les boucles I et II conduisent toutes à des modèles incorrects. De nombreuses contraintes restent insatisfaites dans les structures prédites. D'autre part, les modèles présentant le moins de violations présentent caractéristiques physiques inacceptables : en particulier des énergies d'interaction de van der Waals faibles, voir positives. En revanche, l'utilisation des contraintes impliquant les boucle II et boucle III permet d'aboutir à des modèles où la quasi-totalité des contraintes sont satisfaites et qui présentent des énergies d'interaction négatives.

Nous obtenons au final 17 modèles corrects satisfaisants la totalité des contraintes entrées. On peut constater une variabilité dans la détermination du positionnement du haut de la toxine. Ceci s'explique par l'absence de données expérimentales pour cette région. Toutefois, les résidus établissant les contacts avec le récepteur fluctuent très peu. Ainsi, bien que les contraintes de distances introduites dans les simulations soient peu précises, la redondance des données permet d'obtenir des modèles avec une bonne définition spatiale.

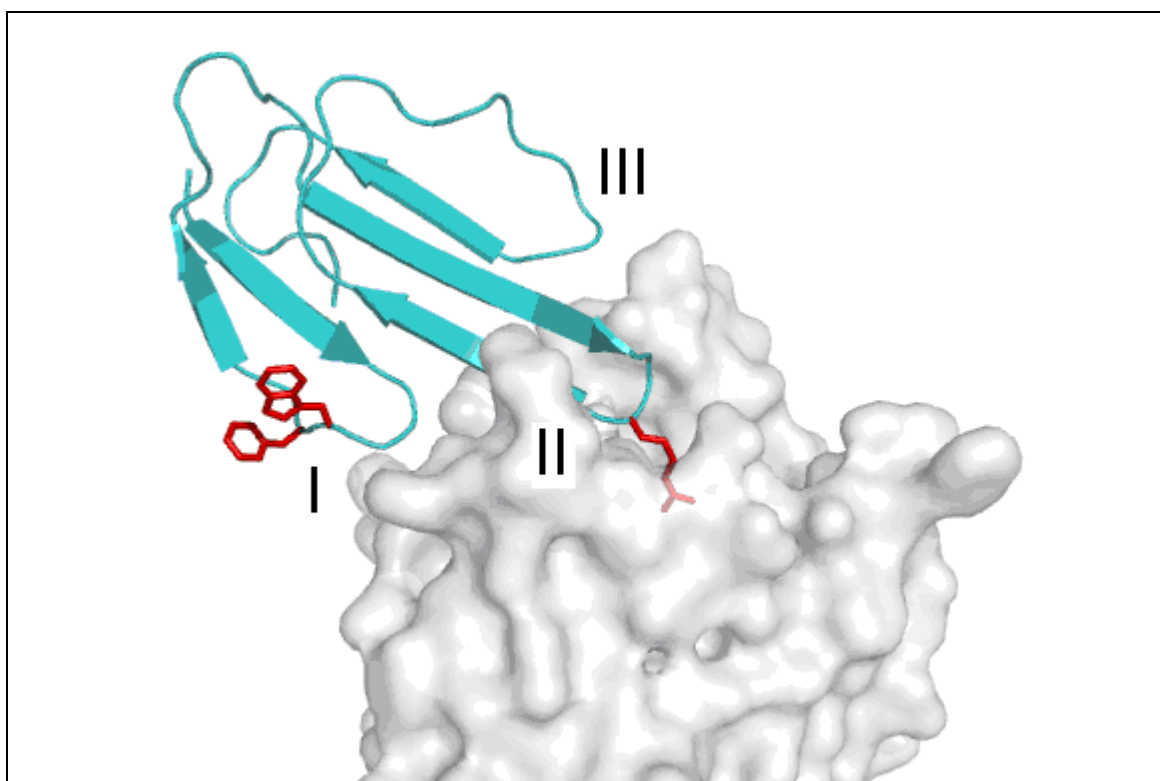


**Figure 54 : Variabilité dans le positionnement de la toxine parmi les solutions**

*La position du haut de la toxine n'est pas totalement déterminée mais l'interface qui est bien couverte par les données expérimentales fluctue très peu.*

3.5.1.1. Vers un dimère de récepteur

Il apparaît donc que seules les données expérimentales concernant les points de contacts entre les boucles II et III de la toxine et le récepteur peuvent être prises en compte dans des modèles structuraux d'une toxine MT7 arrimée sur un récepteur hM1. Ces modèles convergent tous vers un même mode d'interaction de la toxine sur le récepteur. Par contre, aucun modèle satisfaisant tenant compte des informations disponibles au niveau de la boucle I de la toxine n'a pu être produit. En revanche, on remarque sur les modèles boucle II + III que la boucle I se trouve au-dessus de la membrane, dans une position très décalée latéralement par rapport à la partie extracellulaire du récepteur (Figure 55). Ceci suggère que le cœur de l'interaction serait localisé au niveau du résidu R34 mais que la boucle I de la toxine pourrait en fait interagir avec le second récepteur d'un dimère de hM1.



**Figure 55 : Positionnement de la toxine sur le récepteur**

*La boucle I de la toxine est décalée par rapport au récepteur et se trouve au-dessus de la membrane. En rouge, les résidus W10 et F11 (boucle I) et R34 (boucle II)*





### ***3.6. Démonstration expérimentale de l'existence de dimère de hM1***

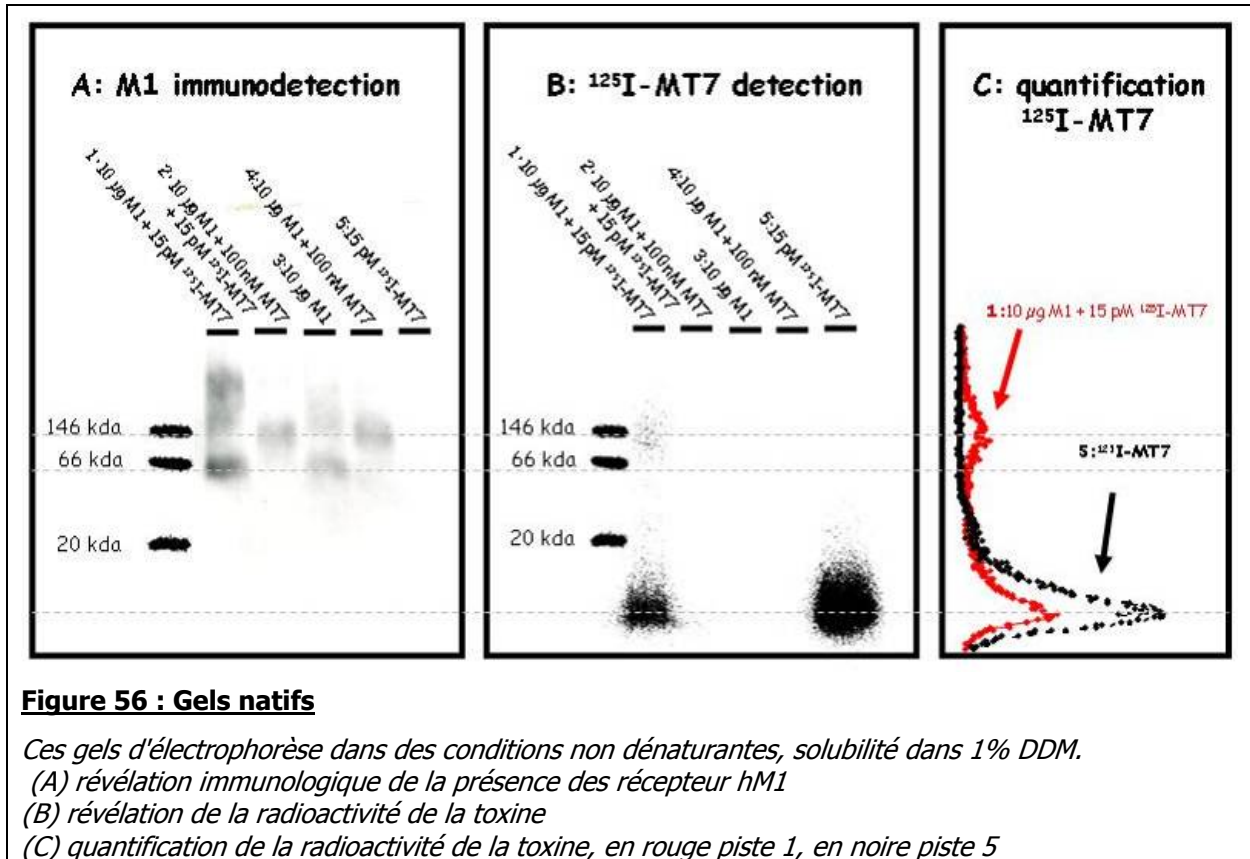
#### **3.6.1. Expériences de gels natif**

*Ces expériences ont été réalisées par Carole Fruchard-Gaillard (CEA SIMOPRO).*

Dans ces expériences les membranes de cellules CHO-hM1 ont été incubées en absence ou en présence de différents ligands puis solubilisées par une solution à 1% de DDM. Enfin, la présence de récepteur muscarinique est révélée par western blot avec un anticorps dirigé contre le récepteur M1

En absence de toxine, la présence de dimères préexistants de récepteurs a été observée par ce type d'expérience. Les conditions de détergents sont donc choisies de telle manière que ces dimères soient déstabilisés. Dans ces conditions, en absence de toxine, on révèle par détection immunologique des récepteurs monomériques (Figure 56, A, ligne 3). En présence de toxine, on observe que le hM1 migre sur le gel à une distance correspondant au poids moléculaire d'un dimère (Figure 56, A, ligne 2). Ceci semble indiquer que la toxine protège les dimères de hM1 de la déstabilisation par le détergent.

Sur le même gel, on peut révéler la présence de toxine radiomarquée à l'iode par -imager. En absence de récepteur, la toxine MT7 chaude migre à une distance correspondant aux faibles poids moléculaires (Figure 56, B, ligne 5). En présence de récepteurs, on révèle sa présence sur le gel à une distance correspondant à un poids moléculaire double de celui du récepteur hM1 (Figure 56, B, ligne 1). Ceci indique que la toxine possède la capacité de se lier sur un dimère de récepteurs hM1.



### 3.6.2. Transfert de fluorescence par photo-blanchiment de l'accepteur

Ce travail a été réalisé par Catherine Marquer (CEA, SIMOPRO) en collaboration avec Emmanuelle Girard (CNRS, Gif-sur-Yvette). Les lignées cellulaires exprimant les récepteurs modifiés ont été fournies par Brigitte Ilien (ESBS, Illkirch).

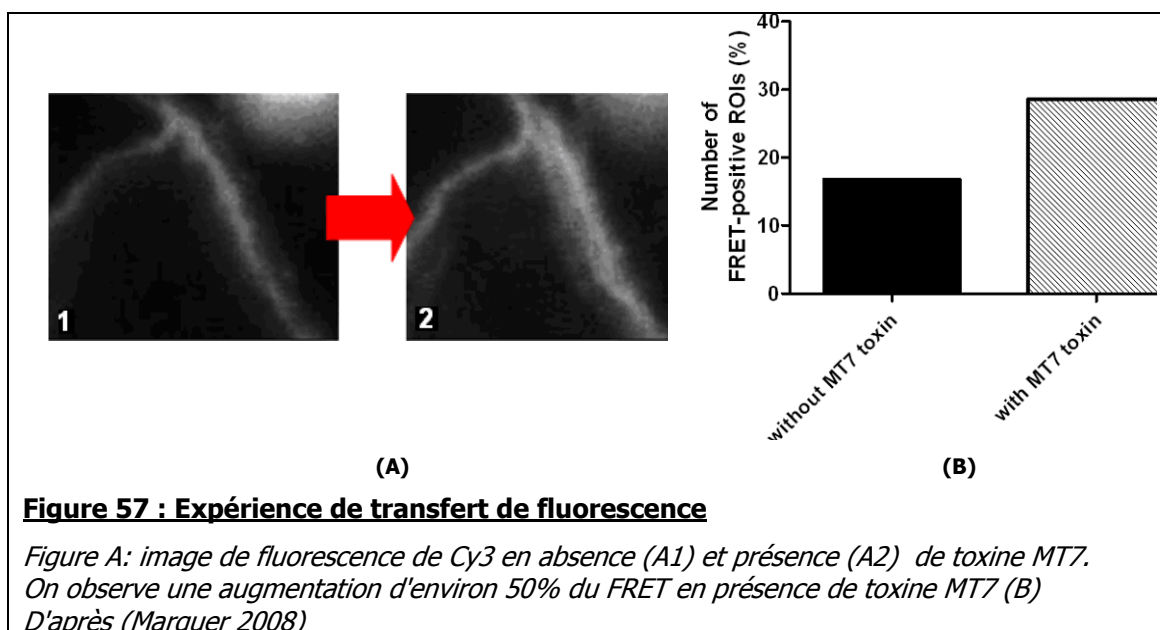
Des cellules co-exprimant deux modifications du récepteur hM1 on été utilisées

- l'un fusionné à l'eGFP ainsi (excitation ~ 450 nm, émission à ~ 530 nm)
- l'autre comportant un épitope FLAG, auquel se lie un anticorps couplé au fluorophore Cy3 (excitation ~ 540 nm, émission à 575nm)

Les deux fluorophores sont choisis de manière à ce que la longueur d'onde d'émission de fluorescence du premier, le donneur, (eGFP) corresponde à celle d'excitation du second, l'accepteur (Cy3). Si les deux fluorophores sont proches spatialement, lorsqu'on excite le donneur à sa longueur d'onde d'émission, on observe un

transfert d'énergie vers l'accepteur : c'est le transfert de fluorescence (FRET). Dans ce cas, on a donc une diminution de la fluorescence mesurable à la longueur d'onde d'émission du donneur et une augmentation de celle de l'accepteur. Le photoblanchiment consiste à détruire sélectivement un fluorophore en l'irradiant de façon prolongée à sa longueur d'onde d'excitation. Donc, s'il y a FRET, on s'attend à mesurer une augmentation du signal d'émission du donneur lorsque l'on détruit l'accepteur.

Les mesures réalisées démontrent, en absence de toxine, une augmentation de l'émission de fluorescence de l'eGFP après destruction de Cy3 (Figure 57). Cette expérience démontre donc une proximité des récepteurs qui indique la présence, in cellulo, de dimères préexistants de récepteurs. D'autre part on observe un signal de fluorescence du donneur beaucoup plus important en présence de MT7. Ceci indique donc que l'on observe plus de dimère de récepteurs en présence de toxine. Ce qui confirme donc la capacité de la toxine à stabiliser la forme dimérique du hM1.

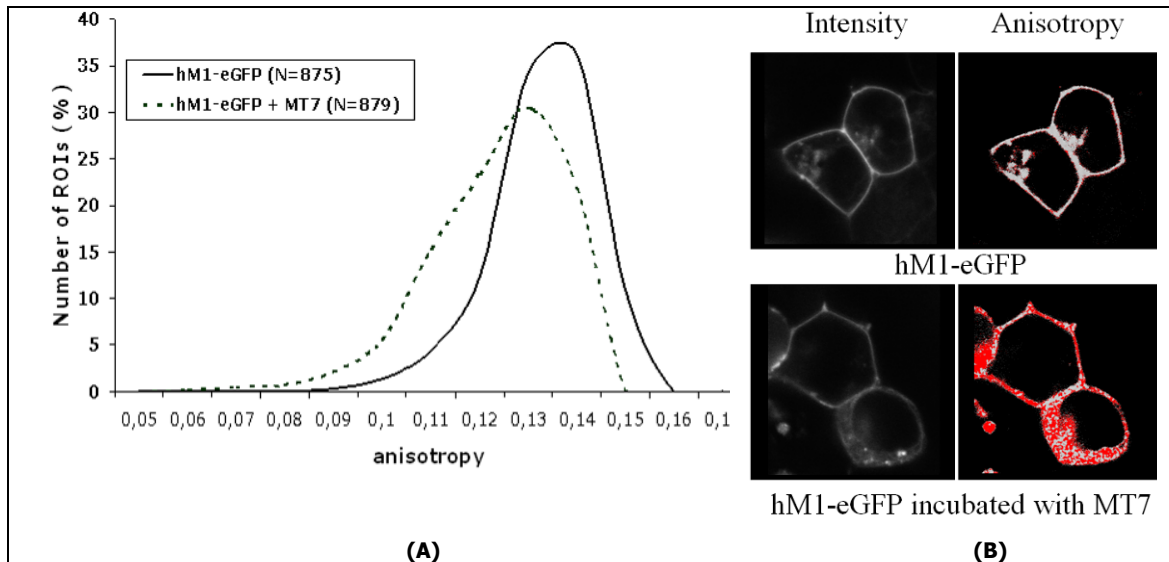


### 3.6.3. Expérience d'anisotropie de fluorescence

*Cette étude a été réalisée par Catherine Marquer (CEA, SIMOPRO) en collaboration avec Olivier Grandjean (INRA, Versailles).*

Lorsqu'on illumine une population de fluorophores avec un rayonnement polarisé, on excite de manière sélective ceux dont l'orientation est compatible avec la polarisation du faisceau incident. On peut alors mesurer l'intensité de la fluorescence émise par le fluorophore, dans une direction parallèle au faisceau incident, et dans une direction perpendiculaire et ainsi calculer la perte de polarisation du signal de fluorescence: l'anisotropie de fluorescence. Un certain nombre d'évènements peuvent intervenir pendant la durée de vie de l'état excité du fluorophore et modifier l'orientation du rayonnement émis. L'eGFP est une molécule de grande taille, dont le temps de rotation (24ns) est supérieur à la durée de vie de son état excité (2.6ns). Il n'y a donc pas d'augmentation de l'anisotropie de fluorescence de cette molécule due au mouvement brownien. En revanche, les spectres d'excitation et d'émission de l'eGFP se recouvrent partiellement. Un transfert de fluorescence entre 2 molécules d'eGFP proches dans l'espace est donc possible. Ce phénomène est appelé homo-FRET. La fluorescence émise par l'eGFP accepteur sera dans ce cas dirigée dans une direction différente de celle du donneur, ce qui entraîne la dépolariation du signal de fluorescence.

Les mesures réalisées (Figure 58) démontrent une diminution significative de l'anisotropie de fluorescence des molécules d'eGFP fusionnées aux récepteurs en présence de la toxine MT7. Ceci indique donc qu'il existe une augmentation de l'homo-FRET entre les molécules d'eGFP en présence de la toxine et donc que la toxine stabilise la forme dimérique du hM1.



**Figure 58 : Expérience d'anisotropie de fluorescence**

La figure A présente la distribution des valeurs d'anisotropie pour des cellules incubées ou non en présence de MT7. La figure B représente les intensités de fluorescence et d'anisotropie pour des cellules incubées ou non en présence de MT7. Les pixels rouges correspondent aux plus basses valeurs d'anisotropie.

### 3.6.4. Conclusion

Ces résultats confirment l'existence de dimères constitutifs de récepteurs hM1 démontrée par ailleurs par Goin et al. (Goin and Nathanson 2006). La toxine MT7 possède la capacité de se lier sur ces dimères de récepteurs et ainsi de les protéger de l'action du détergent. Enfin, la toxine démontre une capacité à stabiliser la forme dimérique du récepteur M1.

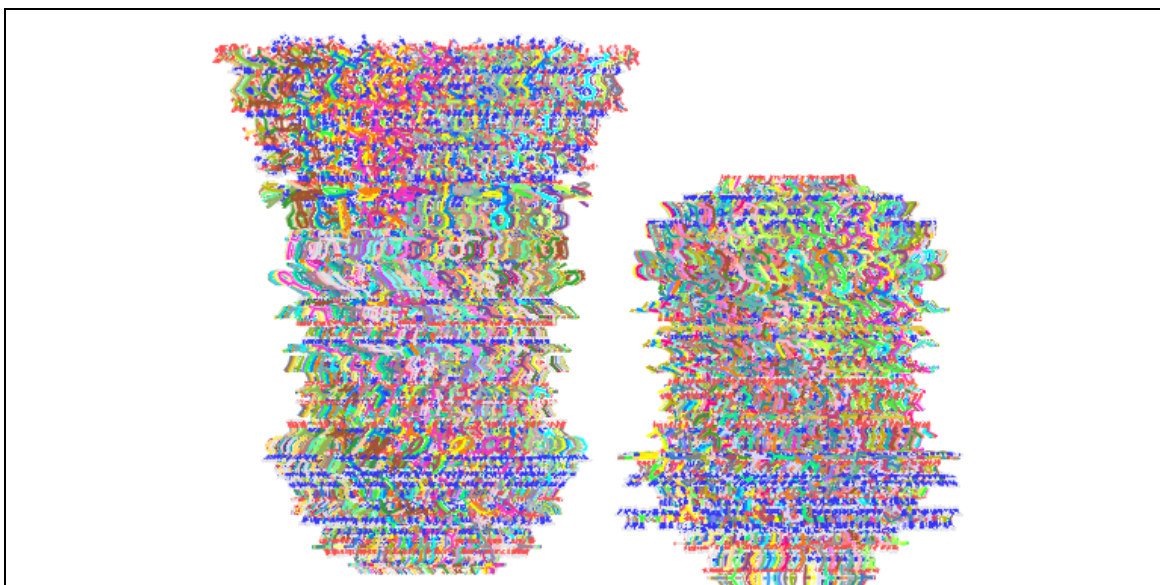
### **3.7. Modèles trimériques du complexe dimère de hM1 / MT7**

#### **3.7.1. Construction d'un modèle de dimère de hM1**

Nous avons souhaité prédire l'assemblage trimérique d'une toxine MT7 sur un dimère de hM1. Toutefois, aucune donnée n'est disponible actuellement concernant les régions de ce récepteur potentiellement impliquées dans l'interface de dimérisation. Nous avons donc utilisé les contraintes de distance toxine/récepteur pour prédire l'arrangement du dimère. Le premier modèle, satisfaisant les données expérimentales des boucles II+III de la toxine a été utilisé comme base. Un second récepteur est ajouté au système, parallèlement au premier, dans le plan de la membrane. Les boucles intracellulaires des deux récepteurs sont éliminées. Les positions de départ des deux récepteurs sont tirées aléatoirement (Figure 59) en effectuant des rotations autour d'un axe perpendiculaire à la membrane.

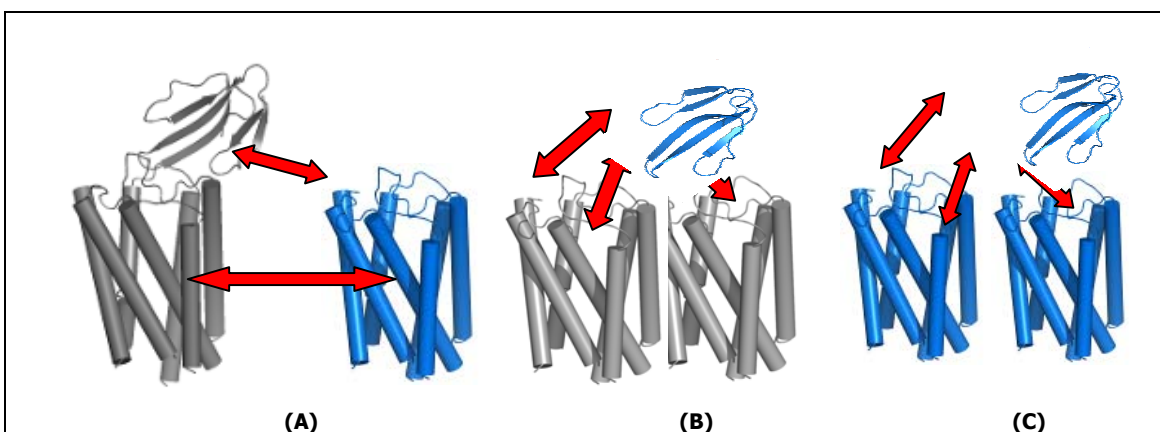
Pour cette simulation, on utilise une procédure d'arrimage moléculaire semblable à la construction du premier modèle, mais on procède cette fois en 3 étapes :

- Le récepteur B est guidé vers le complexe récepteur A/toxine en utilisant les contraintes disponibles boucle I (Figure 60 A). Pour empêcher l'axe du récepteur de "basculer", une contrainte supplémentaire, avec une constante de force importante (50 kcal.mol<sup>-1</sup>) est appliqué entre les centres de gravité des deux récepteurs. Les chaînes latérales sont flexibles.
- La seconde étape (Figure 60 B) consiste à raffiner le positionnement de la toxine. Pour cela, on maintient fixe le dimère de récepteur et on introduit toutes les contraintes de distances disponibles. Les chaînes latérales et la boucle extracellulaires E2 des récepteurs sont flexibles.
- Enfin, un ultime raffinement (Figure 60 C) est effectué en relâchant les 3 partenaires et en utilisant l'ensemble des contraintes. Les chaînes latérales et la boucle extracellulaires E2 des récepteurs sont flexibles.



**Figure 59 : Échantillonnage des positions des partenaires**

*On effectue une rotation d'un angle aléatoire de chaque partenaire autour d'un axe perpendiculaire à la membrane et passant par leur centre de gravité*



**Figure 60 : Étapes de construction d'un modèle trimérique**

*Lors de la première étape, un second récepteur M1 est rapproché du complexe en utilisant les données boucle I de la toxine. Lors de la seconde étape, le dimère de récepteur est maintenu fixe et la position de la toxine raffinée en introduisant toutes les contraintes. Enfin, les chaînes latérales des 3 protéines sont relâchées*

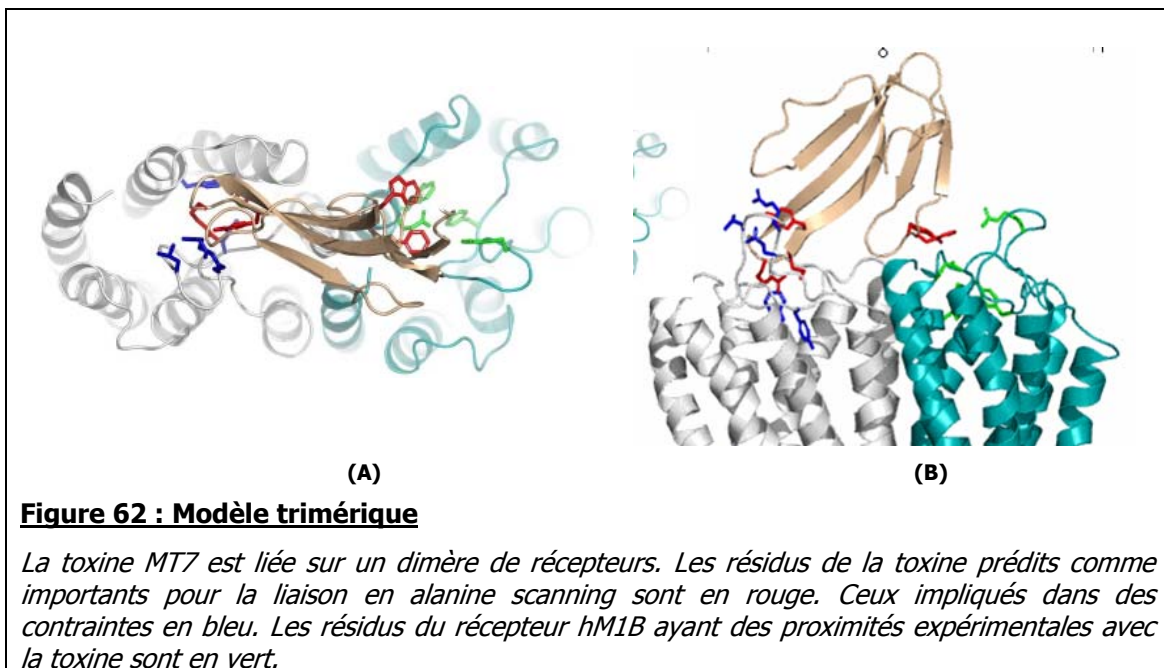
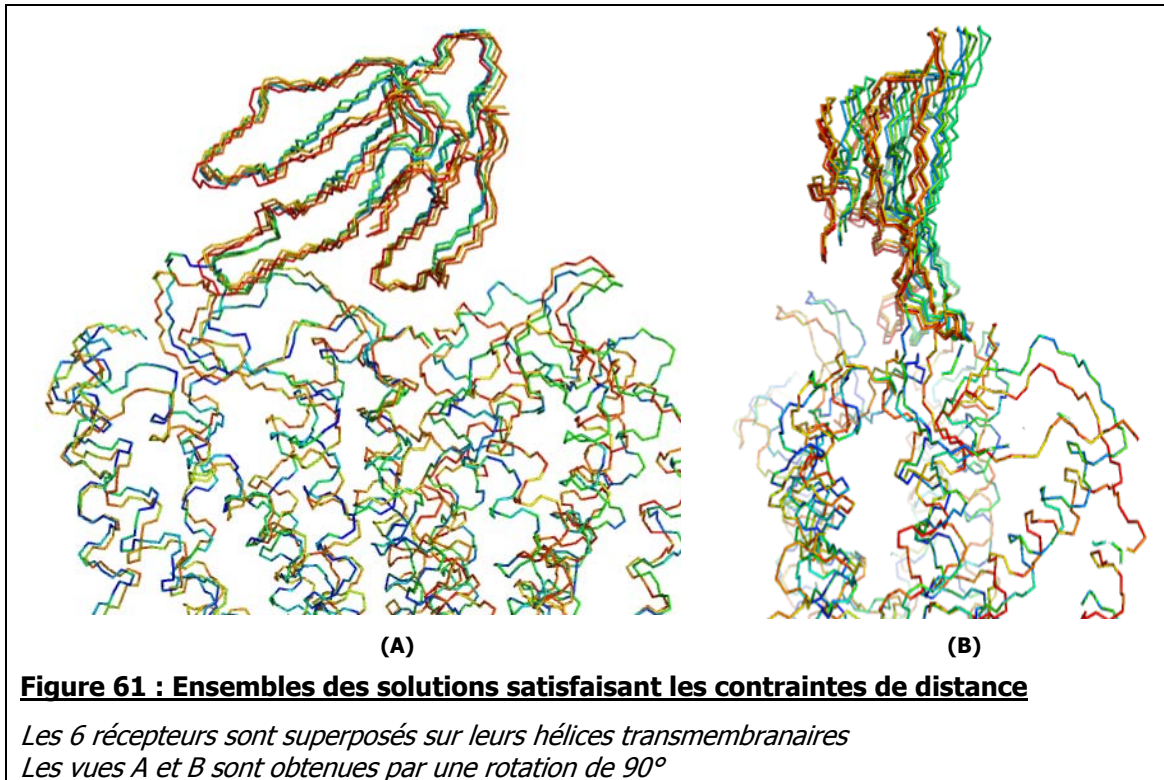
*En bleu : protéines libres, en gris protéine fixes. Les flèches représentent les contraintes appliquées*

A l'issue de 10.000 simulations, on obtient 12 solutions (Figure 61) dans lesquelles l'ensemble des 21 contraintes de distance (à l'exclusion de R52/W400 et W91/Y30) sont satisfaites, c'est-à-dire que les résidus sous contraintes sont distants de moins de 8Å. Le RMS moyen sur les carbones des hélices de l'ensemble des modèles en accord avec les données est, après superposition, de 2.8 Å. Le rms moyen sur les



toxines est alors de 2.5Å. Ceci indique que les calculs convergents vers un unique mode d'assemblage des récepteurs.

On constate que comparativement à la première étape, l'incertitude dans le positionnement de la toxine est moins grande grâce aux 6 contraintes de distance supplémentaire qui ont été utilisées. Parmi ces 6 modèles, le meilleur en terme de surface d'interaction toxine/récepteur (Figure 62) a été optimisé par la suite.





---

# **Chapitre 4: Validation et optimisation du modèle**

---

## ***4.1. Introduction***

Le besoin de puissance de calcul lors de l'étape de prédiction des modèles modèle du complexe hM1/MT7 nous a contraints à utiliser des approximations lors des calculs de mécanique moléculaire. En particulier, la dynamique moléculaire a été réalisée dans le vide en utilisant le champ de force à atomes unifiés charmm19. D'autre part, la présence de contraintes de distance lors de la simulation d'arrimage ne nous permet pas de conclure que l'assemblage prédit restera stable en leur absence. La simulation de ce modèle en dynamique moléculaire libre poursuit donc plusieurs objectifs :

- Étudier la stabilité du modèle, en absence des contraintes de distance
- Relaxer les tensions induites par les contraintes lors de l'arrimage
- Optimiser les interactions entre les 3 partenaires

Pour réaliser ces objectifs, il est nécessaire de prendre en compte de manière rigoureuse l'environnement des protéines. Le RCPG étant des protéines membranaires, le système sera donc simulé dans un champ de force tout atome (charmm22), et en présence de lipides et de molécules d'eau explicites.

### **4.1.1. Dynamique moléculaire des RCPG**

#### **4.1.1.1. Travaux méthodologiques**

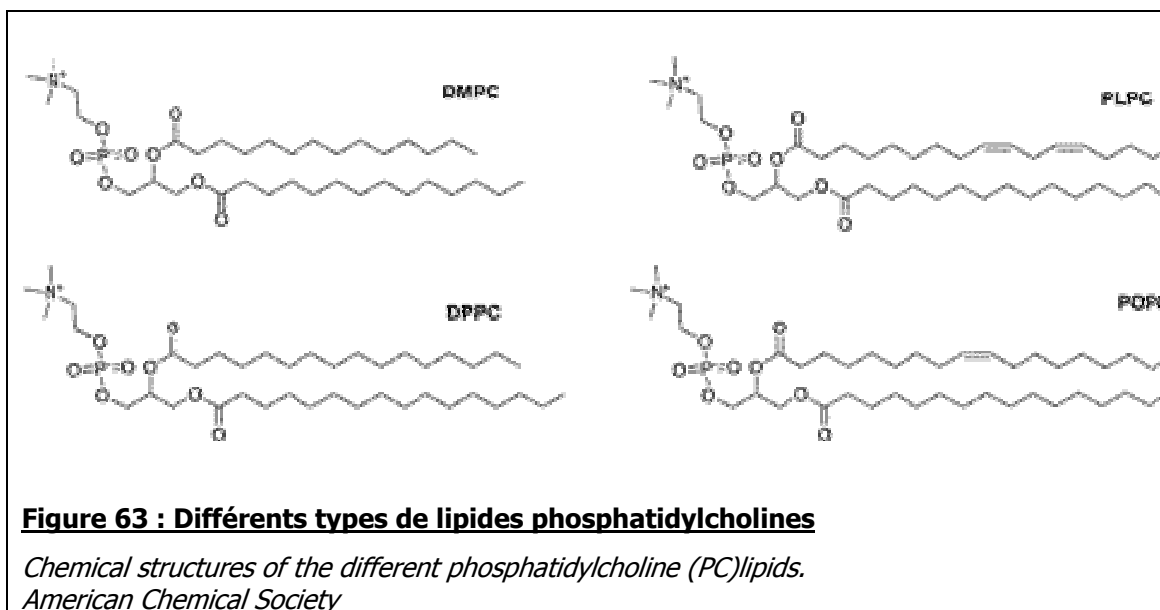
A la suite de la publication de la structure de la rhodopsine, de nombreux travaux de dynamique moléculaire de RCPG ont été réalisés. En premier lieu, certains auteurs, et en particulier l'équipe de Perez, se sont attachés à mettre en place et à valider un cadre méthodologique de simulation des RCPG, en environnement membranaire explicite.

L'influence du traitement des interactions à longue portée ainsi que des ensemble thermodynamique ont fait l'objet d'un premier travail (Cordomi, Edholm et al. 2007). En réalisant cinq simulations de la dynamique de la rhodopsine, sur une durée de 16ns, les auteurs ont exploré les effets du traitement des interactions électrostatiques

par la méthode PME ou par des troncatures. Il apparaît que l'utilisation de troncatures permettant de gagner un facteur 2 par rapport à PME en vitesse de calcul, au prix d'artefacts, en particulier une dérive dans l'épaisseur de la bicouche.

L'influence de l'utilisation des ensemble thermodynamiques NVT (volume et température constants) et NPT (pression et température constante) a également été étudiée. Il apparaît que l'utilisation de NPT en combinaison avec PME constitue les paramètres de simulation les plus réalistes puisqu'elle autorise des réarrangements des lipides et un traitement rigoureux de l'électrostatique. L'utilisation de l'ensemble NVT contraint les paramètres de la membrane, qui ne peut plus se réorganiser après équilibration. Cet ensemble semble cependant donner des résultats comparables à NPT en ce qui concerne l'évolution des structures des protéines et en particulier des passages transmembranaires. Les auteurs concluent donc que l'utilisation de troncature en combinaison avec l'ensemble NVT constitue une alternative raisonnable à la méthode rigoureuse NPT + PME.

Les mêmes auteurs ont également étudié l'influence de l'espèce de lipide choisie pour construire la bicouche lipidique sur l'évolution de simulation de rhodopsine (Cordomi and Perez 2007). Quatre types de lipides classiquement utilisés dans des calculs de dynamique moléculaire ont été testés : DMPC, DPPC, PLPC, et POPC (Figure 63). Ils observent que la surface de contact entre lipide et rhodopsine semble favorisée pour les espèces possédant les chaînes aliphatiques les plus longues et des insaturations. En particulier, les bicouches de lipides insaturés DMPC et DPPC démontrent une faible capacité à s'adapter à la structure du récepteur. Dans le cas du DMPC, la faible longueur des chaînes aliphatiques contraint les lipides à adopter une conformation très étendue peu réaliste pour accommoder la surface de la protéine. Les auteurs démontrent également que les lipides jouent un rôle important dans la modulation de l'inclinaison des hélices au niveau des coudes à proline.



On sait que la durée de vie du contact d'un lipide à la surface d'une protéine est de l'ordre de plusieurs dizaines de nanosecondes, alors que celle d'une molécule d'eau n'est que de quelques dizaines de picosecondes. L'organisation de la membrane lipidique intervient donc à des échelles de temps de très loin supérieures à celle de l'eau. La question de la difficulté d'équilibration des simulations de membranes de lipides est aujourd'hui largement reconnue. Il est admis que des temps de plusieurs nanoseconde, voir dizaine de nanoseconde sont nécessaire pour équilibrer ce type de simulation (Faraldo-Gomez, Forrest et al. 2004). Grossfield et al (Grossfield, Feller et al. 2007) ont adressé le problème particulier de la convergence des calculs de dynamique moléculaire d'un RCPG. Il s'agit de répondre à la question : "*la simulation a-t-elle duré assez longtemps pour que l'on puisse avoir confiance en ses prédictions ?*". Les auteurs ont réalisé 26 simulations de 100 ns de la rhodopsine dans un environnement lipidique. Leurs résultats les amènent à conclure que les fluctuations des protéines ne convergent pas pour une simulation unique dont l'échelle de temps est de l'ordre de 100 ns. D'autre part, alors qu'on pouvait s'attendre à ce que les régions extra-membranaires s'ordonnent à des échelles de temps comparables à celles des protéines solubles, les auteurs observent qu'elles sont également largement sous-échantillonnées.

Ces résultat invitent donc à la plus grand prudence quand à l'interprétation des résultats de dynamiques moléculaires de protéines membranaires. On prendra en particulier soin d'équilibrer le système pendant un temps de plusieurs dizaines de nanosecondes au moins.

4.1.1.2. Activation de la rhodopsine

Le rétinale est un cofacteur lié covalamment à la rhodopsine et dont l'isomérisation par la lumière entraîne l'activation du récepteur. La nature exacte de l'événement précoce qui va à lui seul entraîner tous les changements structuraux aboutissant finalement à l'activation du récepteur est donc parfaitement caractérisée. De nombreux travaux sur la rhodopsine ont visés à étudier par la dynamique moléculaire les événements de l'évolution de la structure du récepteur après isomérisation du rétinale.

Plusieurs auteurs ont tenté de répondre à cette question en se basant sur des simulations longues (jusqu'à 150 ns). (Crozier, Stevens et al. 2003 ; Huber, Botelho et al. 2004 ; Lemaitre, Yeagle et al. 2005; Crozier, Stevens et al. 2007 ). Cependant, on sait que si l'isomérisation du rétinale est un événement très rapide, qui intervient en moins de 200 fs, les changements structuraux qu'il entraîne dans la structure de la rhodopsine interviennent à une échelle de temps de l'ordre de la microseconde à la milliseconde (Okada and Palczewski 2001). Des événements se produisant à ces échelles de temps ne peuvent pas être capturés par les méthodes classiques de dynamique moléculaire. On pourra par ces méthodes tout au plus envisager d'étudier les événements précoces à l'origine de l'activation.

Pour circonvenir ces limitations Kong et al (Kong and Karplus 2007) sont partis de l'hypothèse que, bien que les mouvements de grande amplitude ne sont pas accessibles par la dynamique moléculaire, ils restent cependant encodés dans la structure de la rhodopsine et sont donc déterminés par son comportement à l'équilibre. Ils ont donc réalisé, plutôt qu'une unique simulation longue de la rhodopsine, plusieurs simulations indépendantes (8 pour un total de 22ns) et analysé leurs trajectoires en s'intéressant aux corrélations d'énergies d'interaction des paires de résidus. Cette étude leur permet d'identifier un réseau de couplages, notamment au niveau des hélices VI et VII qui semblent impliqués directement dans la transduction du signal d'activation, ceci en accord avec le rôle connu des motifs peptidiques D(E)RY et NPxxY.

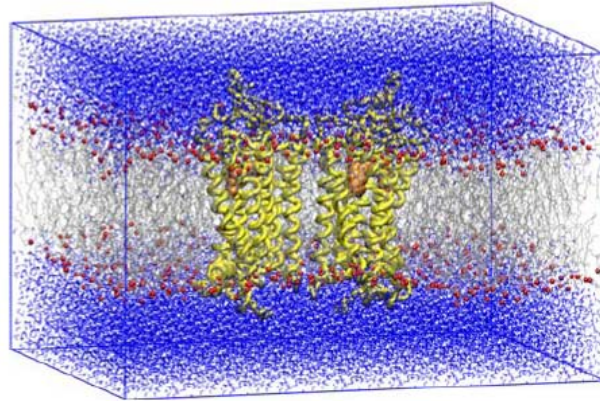
#### 4.1.1.3. Modèles

La dynamique moléculaire a également été utilisée assez largement pour étudier et/ou valider des modèles de RCPG. En particulier, Espinoza-Fonseca et al (Espinoza-Fonseca, Pedretti et al. 2008) ont démontré qu'il est possible de simuler de manière stable un modèle de récepteur muscarinique hM1 dans une membrane de POPC pendant 10 nanoseconde. Des coordonnées extraites de cette trajectoire leur ont ensuite permis de réaliser des expériences d'arrimage de l'acétylcholine. Les propriétés de ces modèles sont en bon accord avec les données expérimentales concernant le site et le mode de liaison de l'acétylcholine.

Des travaux de dynamique moléculaire ont été appliqués à l'étude de dimère de RCPG. Periole et al (Periole, Huber et al. 2007) ont démontré par des simulations de dynamique à gros grain la tendance de monomères de rhodopsine à s'auto assembler. Ils notent également, la formation d'irrégularités de la bicouche lipidique au voisinage des hélices II, IV et VII. Ces déformations locales de la membrane joueraient un rôle important dans la détermination des sites d'oligomérisation des RCPG.

Enfin d'autre auteurs (Filizola, Wang et al. 2006) ont simulé la dynamique d'un modèle tout atome d'un dimère de rhodopsine. Un modèle du dimère basé sur les arrangements de récepteur observé par des expériences d'AFM (Fotiadis D 2006) a été inséré dans une bicouche de POPC et simulé pendant une durée de 45ns (Figure 64). De manière intéressante, les auteurs observent que les deux monomères initialement relativement distants l'un de l'autre dans le modèle de microscopie de force atomique démontrent une tendance à se rapprocher lors de la simulation. Ceci suggère que les dimères de rhodopsine pourraient posséder une interface plus compacte que celle observée par AFM.





**Figure 64 : Modèle de dimère de rhodopsine dans une bicouche lipidique**

*Les dimères de rhodopsine modélisés d'après les données AFM ont été insérés dans une bicouche lipidique et simulés par dynamique moléculaire. d'après Filizola (Filizola, Wang et al. 2006)*

#### 4.1.2. Effets électrostatiques et protéines

Les interactions électrostatiques entre les charges des biomolécules gouvernent une grande partie de la physique des protéines. Elles sont entre autres à la base de la formation des liaisons hydrogène dont l'importance est cruciale au fonctionnement des macromolécules. De ce fait, le traitement fiable des interactions électrostatiques constitue depuis toujours un des enjeux majeurs de la biophysique.

Les lois physiques gouvernant les interactions électrostatiques sont parfaitement connues et caractérisées, tant par la mécanique newtonienne, que par la mécanique quantique. Dans une approximation newtonienne, on décrit l'interaction entre deux charges par la loi de Coulomb. Ce type de calcul devient particulièrement lourd dans le cas des macromolécules biologiques en raison de la taille des systèmes, puisque le nombre d'interaction évolue selon le carré du nombre d'atomes. De plus si on souhaite traiter de manière correcte les interactions entre les protéines, le solvant et les ions libres, on augmente encore de plusieurs ordres de grandeur la taille du système. Enfin, pour que les mesures des interactions avec le solvant aient un sens, il est nécessaire de les moyenniser sur une simulation de longue durée. Ceci rend le coût computationnel nécessaire pour le traitement exhaustif des interactions électrostatiques inaccessible en l'état actuel des moyens de calculs. Le recours à des approximations est donc indispensable.

#### 4.1.2.1. Seuils de distance sphériques

La méthode la plus simple pour limiter le coût des calculs des interactions électrostatiques à longue portée, est de les ignorer purement et simplement au-delà d'un seuil de distance pour lequel on considère que leur contribution n'est plus significative. Plus la distance seuil choisie sera grande, plus on aura un traitement réaliste des interactions et plus le temps de calcul nécessaire augmente. Afin d'assurer la continuité des potentiels, il est nécessaire d'introduire des fonctions de lissage masquant la coupure brutale de l'interaction électrostatique. C'est ce rôle qui est assuré par exemple par les fonctions SHIFT et SWITCH implémentées dans Charmm. Les méthodes de troncature ont été largement exploitées par les programmes de dynamique moléculaire. Cependant, il est aujourd'hui admis que ces approximations entraînent inévitablement des artefacts dans les simulations de dynamique moléculaire. Leur description précise et leurs conséquences réelles restent toutefois un sujet de débats.

#### 4.1.2.2. Particle Mesh Ewald

Lors des calculs de dynamiques moléculaires, pour éviter l'apparition d'artefacts résultant des conditions aux limites, on simule en fait un système périodique : On crée des images de la boîte principale dans toutes les directions de l'espace. Ceci permet d'assurer la continuité au niveau des limites de la boîte de simulation, au prix de l'introduction d'une périodicité artificielle dans le système. Il convient alors d'évaluer les interactions interatomiques, non plus dans la seule boîte principale, mais également entre des atomes pouvant appartenir aux images.

Pour les potentiels à courte portée dont l'énergie décroît très rapidement avec la distance, il est raisonnable de se limiter à calculer l'interaction entre deux atomes  $i$  et  $j$  en considérant uniquement leurs représentations les plus proches (qui ne sont donc pas nécessairement situés dans l'image principale). En revanche pour les forces à longue portée, telles que celles résultant des interactions électrostatiques, l'énergie décroît très lentement. En conséquence, des charges situées au delà de leur représentation la plus proche du système périodique peuvent encore contribuer de manière significative au potentiel total. Il est donc nécessaire de tenir compte de l'ensemble des images périodiques du système pour calculer les potentiels à longue portée.

La méthode de sommation d'Ewald a été développée dans le cadre de l'étude de l'énergie électrostatique dans les cristaux ioniques (Ewald 1921). Les systèmes périodiques simulés par la dynamique moléculaire peuvent, dans une certaine mesure être assimilés à des systèmes cristallins. La méthode de la sommation d'Ewald peut donc être utilisée pour résoudre la difficulté posée par la prise en compte des charges présentes dans les différentes images périodiques d'une simulation de dynamique moléculaire. Cette méthode exprime l'équation du potentiel électrostatique sous la forme d'une somme de deux termes :

$$\phi(\mathbf{r}) \stackrel{\text{def}}{=} \phi_{sr}(\mathbf{r}) + \phi_{lr}(\mathbf{r})$$

**Équation 3 : Sommation d'Ewald**

Le premier  $\phi_{sr}(\mathbf{r})$  représente la contribution des interactions à faible portée, calculées en écrantant les charges par une distribution (généralement gaussienne) de charge signe opposé. Ce terme est alors calculable rapidement dans l'espace réel. Le second terme,  $\phi_{lr}(\mathbf{r})$ , permet de calculer les interactions à longue distance dans l'espace de Fourier. L'utilisation des sommes de Fourier suppose que le système étudié soit d'une périodicité infinie.

Particle Mesh Ewald (PME) (York, Darden et al. 1993) est une méthode permettant le calcul rapide des sommations d'Ewald. Les termes à courte portée sont calculés en utilisant des troncatures classiques basées des seuils de distance sphériques. Pour le calcul des interactions à longue portée, les charges sont interpolées par les points d'une grille et la somme approchée par des convolutions utilisant des transformations rapides de Fourier. Ainsi, la méthode PME est définie par deux paramètres principaux : Le seuil de distance utilisé pour la troncature des interactions de faible portée ainsi que la taille du pas de la grille sur laquelle sont interpolées les charges. L'approche particle mesh Ewald permet une plus grande précision dans le calcul des potentiels électrostatiques ( $\sim 10^{-6}$ ) lors des simulations de dynamique moléculaire.

#### 4.1.2.3. Équation de Poisson Boltzmann

La représentation explicite de chaque atome du solvant pour les calculs de mécanique moléculaire implique une charge computationnelle élevée dédiée au calcul des interactions soluté/solvant, mais aussi des interactions solvant/solvant qui n'ont très souvent aucun intérêt pour le sujet étudié. Une approche possible pour limiter le temps de calcul dépensé au traitement des interactions avec le solvant dans les calculs de mécanique moléculaire, est de considérer le solvant comme un continuum dont les propriétés peuvent être exprimées de manière analytique. Le modèle le plus simple de solvant implicite est de le considérer que le milieu possède une constante diélectrique homogène (constant **dielectric**). Une évolution de cette méthode (**radial dielectric**) propose de faire varier cette constante en fonction de la distance par rapport au soluté. Bien qu'étant très efficace sur le plan de la vitesse de calcul, ce type de modèle purement empirique ne permet pas de tenir compte correctement de l'effet de l'environnement des charges sur les interactions électrostatiques.

En pratique, il existe trois grandes familles de modèles de solvants implicites utilisés à l'heure actuelle. Ils sont basés sur : le calcul de la surface accessible au solvant, le formalisme de Born généralisé ou sur les équations de Poisson Boltzmann. Dans un article récent, Lopes et al (Lopes, Alexandrov et al. 2007), ont montré que pour la prédiction de changement de stabilité associés à des mutations ponctuelles ou le placement de chaînes latérales, les modèles de Born et ceux basés sur la surface accessible au solvant ont des performances comparables. L'équation de Poisson-Boltzmann (Équation 4) est à l'heure actuelle la méthode de référence, la plus précise pour traiter les effets électrostatiques en solution. Elle peut par exemple être appliquée aux protéines pour le calcul du potentiel électrostatique de surface (Baker, Sept et al. 2001), l'étude des pKa (Jo, Vargyas et al. 2008), ou les calculs d'énergie libre (Eriksson and Roux 2002).

L'équation PB (Équation 4) permet de décrire le potentiel électrostatique  $\phi(x)$  d'une charge situé à la position  $x$ , en fonction de son environnement. Cette équation ne peut être résolue analytiquement pour des systèmes complexes tels que les macromolécules, mais elle peut l'être efficacement par la méthode numérique des différences finies.

$$-\nabla \cdot \varepsilon(x) \nabla \phi(x) - \sum_{i=1}^m c_i q_i e^{-q_i \phi(x) - V_i(x)} = \frac{4\pi e_c^2}{kT} \rho(x)$$

**Équation 4 : Equation Poisson Boltzmann**

*$\varepsilon(x)$  est la constante diélectrique en un point donné  $x$  de l'espace,  $c_i$  la concentration en ions mobiles  $i$ ,  $q_i$  la charge de ces ions  $V_i(x)$  décrit les interactions stériques entre le soluté et les ions de l'espèce  $i$ ,  $e_c$  la charge des électrons,  $T$  la température et  $k$  la constante de Boltzmann et  $\rho(x)$  est la charge du solvant.*

Le recours aux modèles de solvation implicite, et en particulier à l'équation de Poisson Boltzmann permet donc de réaliser les calculs d'énergie libre de manière beaucoup plus rapide que les méthodes alternative (Free energy perturbation, alchemical transformation, ...) tout en gardant un degré de précision suffisant pour notre usage. La méthode MM-PBSA par exemple (Molecular dynamics Poisson Boltzmann Solvent Accessible) développée par Kollman et Case (Kollman, Massova et al. 2000) repose sur l'utilisation d'un modèle de solvation implicite pour calculer les énergies libres.

Plus récemment une approche de calcul rapide des énergies libres basée sur les équations Poisson-Boltzmann (Eriksson and Roux 2002) a été proposé par Roux et al. La variation d'énergie libre est exprimée sous la forme de la somme d'un terme d'énergie lié aux interactions non polaires  $G_{np}$  et d'un terme électrostatique  $G_{elec}$ . La contribution électrostatique peut être calculée par les équations Poisson-Boltzmann (résolues numériquement grâce au module PBEQ de CHARMM) en assignant une constante dielectrique  $\epsilon_{prot}$  à l'intérieur de la protéine. La contribution non polaire à l'énergie de liaison est approchée de manière empirique par une fraction de l'énergie de van der Waals  $E_{vdw}$ . Les paramètres  $\epsilon_{prot}$  et  $\alpha$  sont optimisés à partir des données disponibles pour 13 mutations alanine du complexe barnase/barstar. Cette approche ne tient pas compte des contributions entropiques. On admet que celles-ci s'annulent lorsque l'on fait la différence entre les énergies calculées pour les protéines natives et mutante.

### 4.1.3. Insertion du complexe dans une membrane

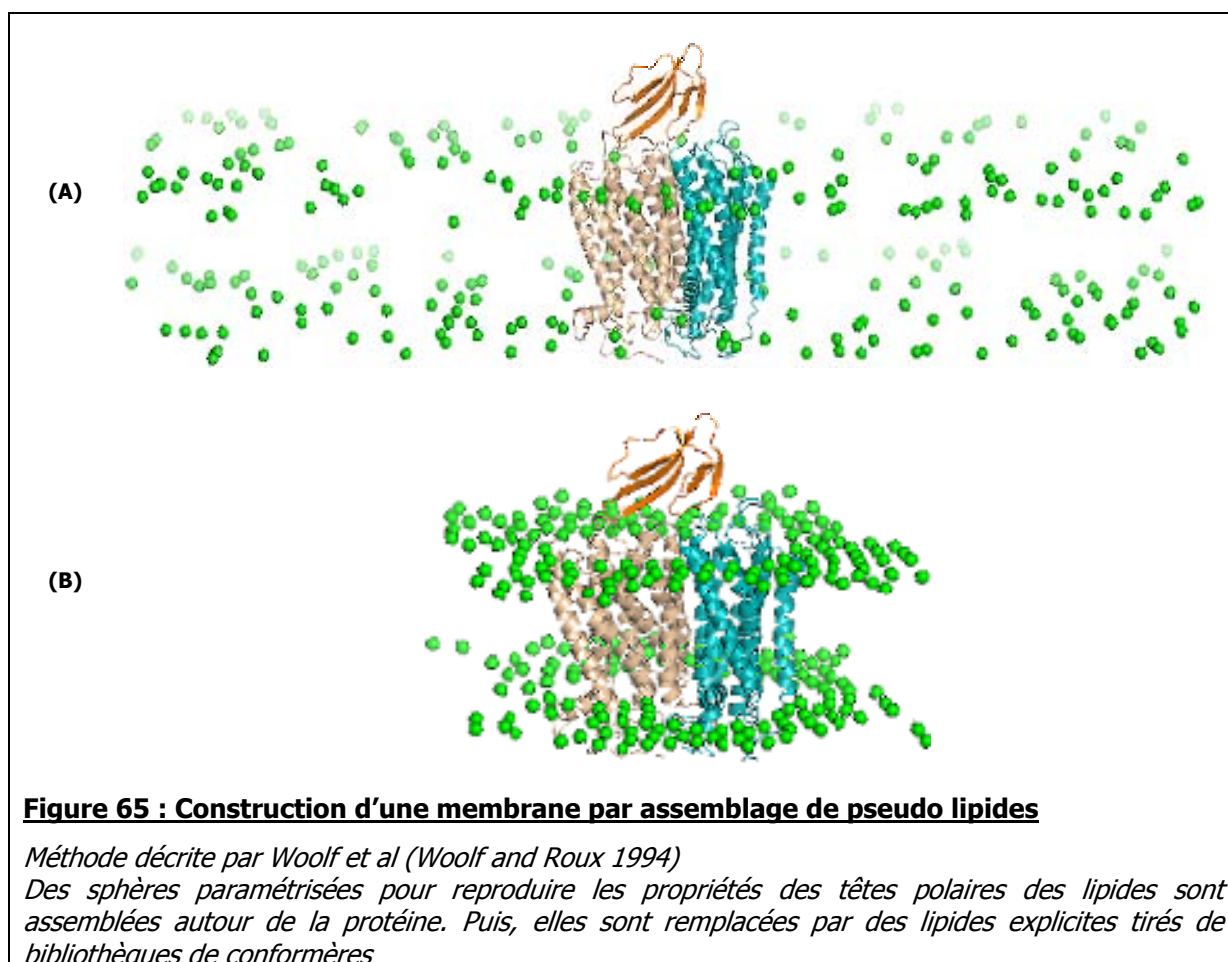
Pour réaliser la simulation de la membrane lipidique, nous avons choisi de travailler avec des lipides POPC (Palmitoyle-oleoyle-phosphatidyle-choline, Figure 63). Ces lipides, paramétrés dans le champ de force Charmm27 présentent une insaturation de la chaîne aliphatique, démontrée comme bénéfique pour l'interaction avec le récepteur (Cordomi and Perez 2007).

Avec leurs régions extracellulaires de grande taille ainsi que leurs huitième hélice positionnée parallèlement à la membrane, au niveau de la région polaire des lipides, construire et équilibrer autour d'une protéine membranaire une bicouche lipidique n'est pas une question triviale (Kandt, Ash et al. 2007).

Une des méthodes les plus simples consiste à superposer les structures de la protéine et des lipides et à supprimer les molécules de lipides ayant des atomes dont les positions recouvrent celles de la protéine. Malheureusement, ceci mène, en raison de la longueur des chaînes lipidique à un système dont la configuration de départ va présenter un vide de plusieurs angströms entre la protéine et la membrane. Cette inhomogénéité artefactuelle dans la densité des lipides due à la méthode de construction se révèle par la suite difficile à résorber, en partie car les molécules d'eau ont tendance à s'y engouffrer massivement.

Une autre méthode consiste à créer progressivement un trou dans la membrane en appliquant des contraintes excluant les lipides localisés à l'intérieur d'un cylindre virtuel. Cette méthode n'est efficace que pour les protéines dont la forme est régulière et pour lesquelles la modélisation de l'enveloppe par un cylindre n'est pas trop simplificatrice. Ce n'est malheureusement pas le cas pour les RCPG.

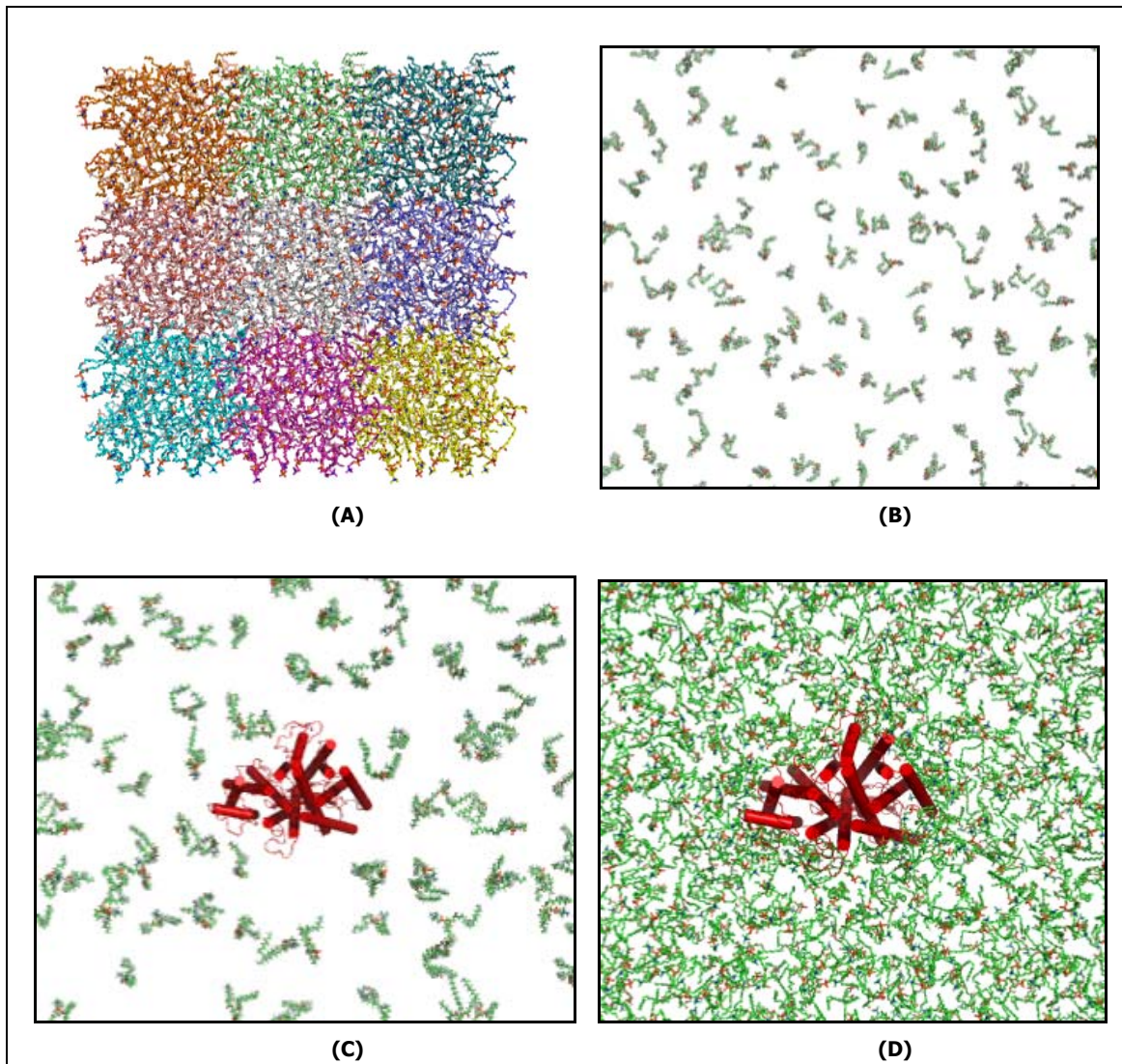
Une approche différente a été proposée par B. Roux (Woolf and Roux 1994) consiste à construire la bicouche lipidique autour de la protéine par une procédure comportant plusieurs étapes (Figure 65). Des sphères paramétrées pour reproduire les propriétés des têtes polaires des lipides et contraintes dans des plans parallèles à la future membrane sont compactées autour de la protéine. Puis, ces sphères sont remplacées par des lipides explicites, dont les conformations sont tirées au hasard dans des bibliothèques de lipides pré-équilibrés. Enfin, les clashes stériques sont minimisés par des transformations (translation et rotation) des coordonnées des lipides traités en corps rigides. Cette méthode se révèle toutefois d'un coût relativement élevé en termes de calcul, en particulier pour les systèmes comportant un grand nombre de lipides pour lesquels l'étape de minimisation peut se révéler relativement longue.



Nous avons donc utilisé une approche proposée récemment par l'équipe de Tieleman (Kandt, Ash et al. 2007), basée sur la translation des coordonnées des lipides.

- Une membrane équilibrée de 72 molécules de lipides a été dupliquée pour obtenir une bicouche lipidique d'une taille suffisante pour y insérer notre protéine. On obtient donc un système de 648 lipides mesurant 144 Å de côté (Figure 66 A)
- Dans un second temps, la membrane est expansée jusqu'à 4 fois sa taille de départ par translation des coordonnées des lipides dans le plan x/y.
- Les protéines sont placées au centre de la bicouche lipidique, et les lipides ayant des atomes distants de moins de 14 Å de la protéine sont supprimés (Figure 66 B).
- Enfin, la membrane est contractée autour des protéines (Figure 66 C), par des étapes successives de translation/minimisation des lipides. Ceux-ci sont rapprochés par contraction d'un facteur 0.9 des coordonnées X/Y, suivi de 1000 pas de minimisation. A chaque étape de contraction, on mesure les paramètres de la membrane (aire par lipide et distance P-P). On cesse les contractions lorsque ces paramètres atteignent des valeurs proches de ce qu'ils étaient dans la membrane de départ (Figure 66 D).
- Enfin, la bicouche est découpée de manière à ne garder que 5 couches de lipides entre les protéines et les bords de la boîte. Des couches d'eau sont ensuite construites de part et d'autre de la membrane, puis certaines molécules d'eau remplacées avec VMD (Humphrey, Dalke et al. 1996) par des ions Na et Cl afin d'équilibrer les charges du système.



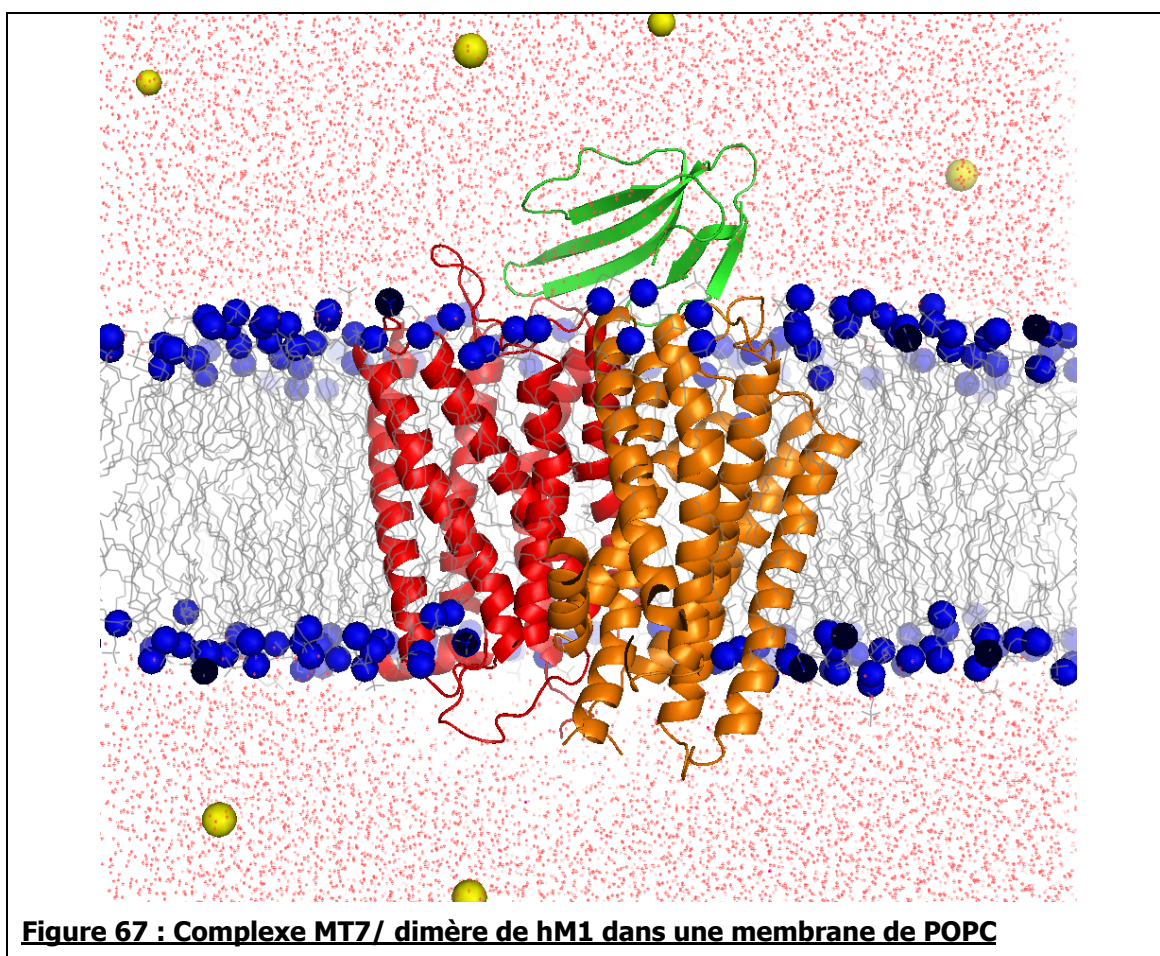


**Figure 66 : Construction de la membrane par expansion contraction**

*A : duplication d'une membrane pré-équilibrée. B : expansion de la membrane par translation des lipides. C : insertion de la protéine et suppression des lipides proches. C : re-compression de la membrane autour de la protéine. Méthode décrite par Tieleman et al*

## 4.2. Dynamique moléculaire du modèle

Parmi les 6 solutions obtenues par arrimage moléculaire (cf § 3.7.1) la meilleure en termes de surface d'interaction toxine/récepteurs est sélectionnée pour la simulation. Le système simulé (Figure 67) comporte au total presque 200 000 atomes, dont 37 000 molécules d'eau et 429 lipides.

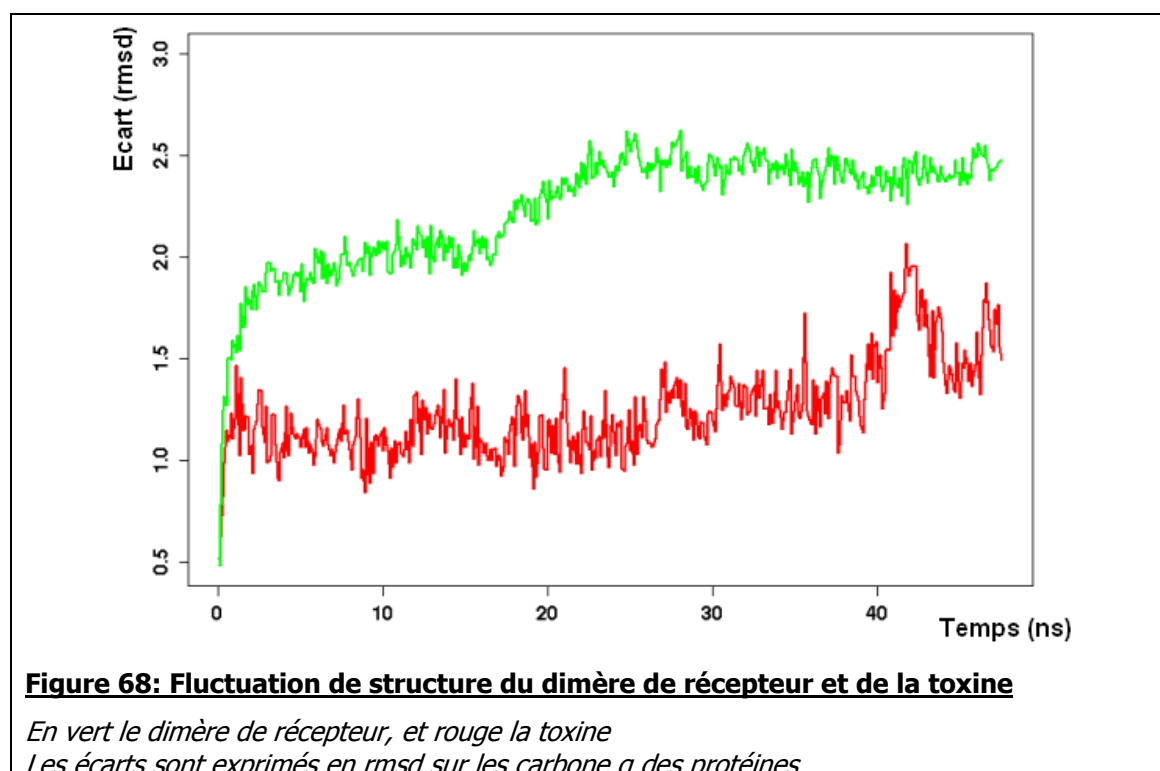


Après minimisation, chauffage progressif et équilibration, le système a été simulé pendant une durée totale de 50ns avec le programme NAMD (Phillips, Braun et al. 2005).

Le degré d'ordre dans une bicouche lipidique est un indicateur important de la qualité d'une simulation. Il est généralement évalué en calculant la surface disponible par lipide. Cependant, ce critère dont la mesure est directe pour un système constitué uniquement de lipides est en revanche difficile à utiliser dans le cadre de la simulation de RCPG (Cordomi and Perez 2007). En effet la forme irrégulière ainsi que la présence

de l'hélice N° 8, tangente à la membrane au niveau des têtes polaires des lipides rend la détermination de la surface occupée par le récepteur difficile. Nous utiliserons donc le critère de l'épaisseur de la membrane qui donne une indication sur l'extension des chaînes aliphatiques des lipides et donc sur le degré d'ordre de la membrane. Nous observons durant notre simulation une distance moyenne P-P de la membrane fluctuant autour de  $41.2 \pm 0.1 \text{ \AA}$ . Cette valeur est très proche de celle rapportée par Cordomi et Perez (Cordomi and Perez 2007) de  $41.12 \pm 0.03 \text{ \AA}$

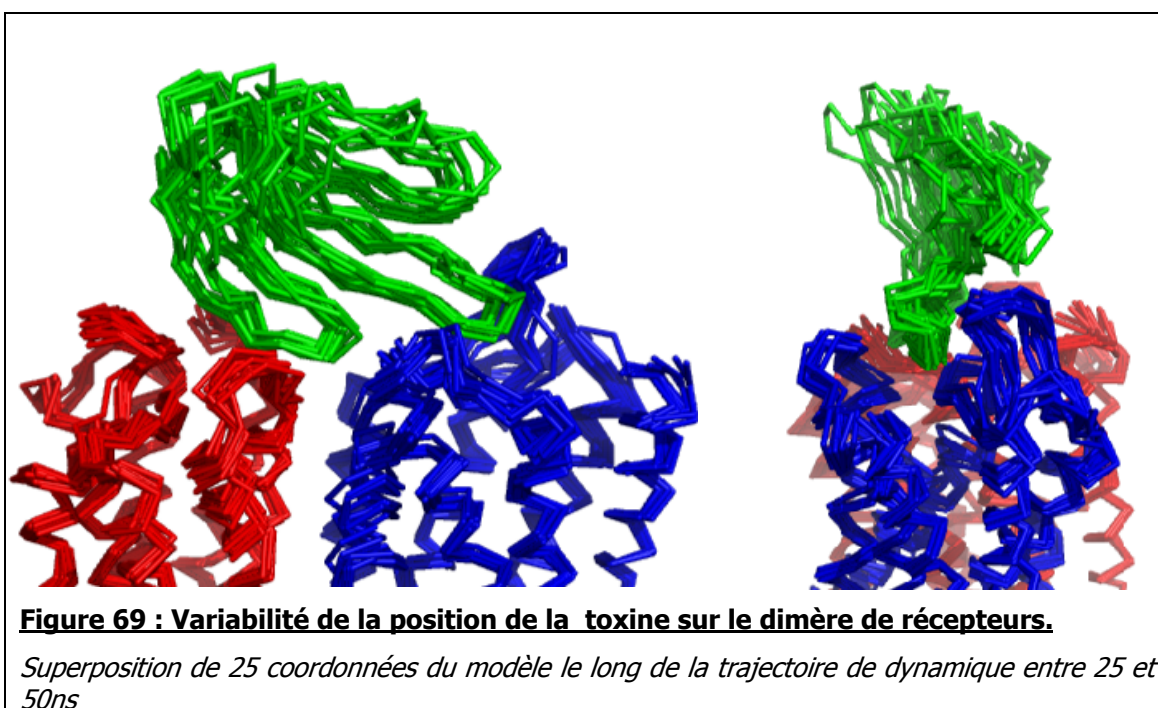
Les fluctuations des structures de la toxine et du dimère de récepteurs restent stables (Figure 68) au cours de la simulation. Les écarts sur la toxine ne dépassent pas  $2 \text{ \AA}$ . Pour le dimère de récepteurs, ils se stabilisent autour de  $2.5 \text{ \AA}$  par rapport à la structure de départ.



On constate dans les premières nanosecondes de simulation une évolution des positions relatives des 3 partenaires du complexe. Les deux récepteurs s'éloignent l'un de l'autre, la distance entre leurs centres de masse passant de  $28.5 \text{ \AA}$  à  $30 \text{ \AA}$ . Cet écartement entre les récepteurs correspond à un RMSD mesuré à  $2.5 \text{ \AA}$  en superposant les hélices des deux récepteurs. Le même type de repositionnement peut être observé côté toxine. Les distances entre les centres de masses des faisceaux d'hélices des deux

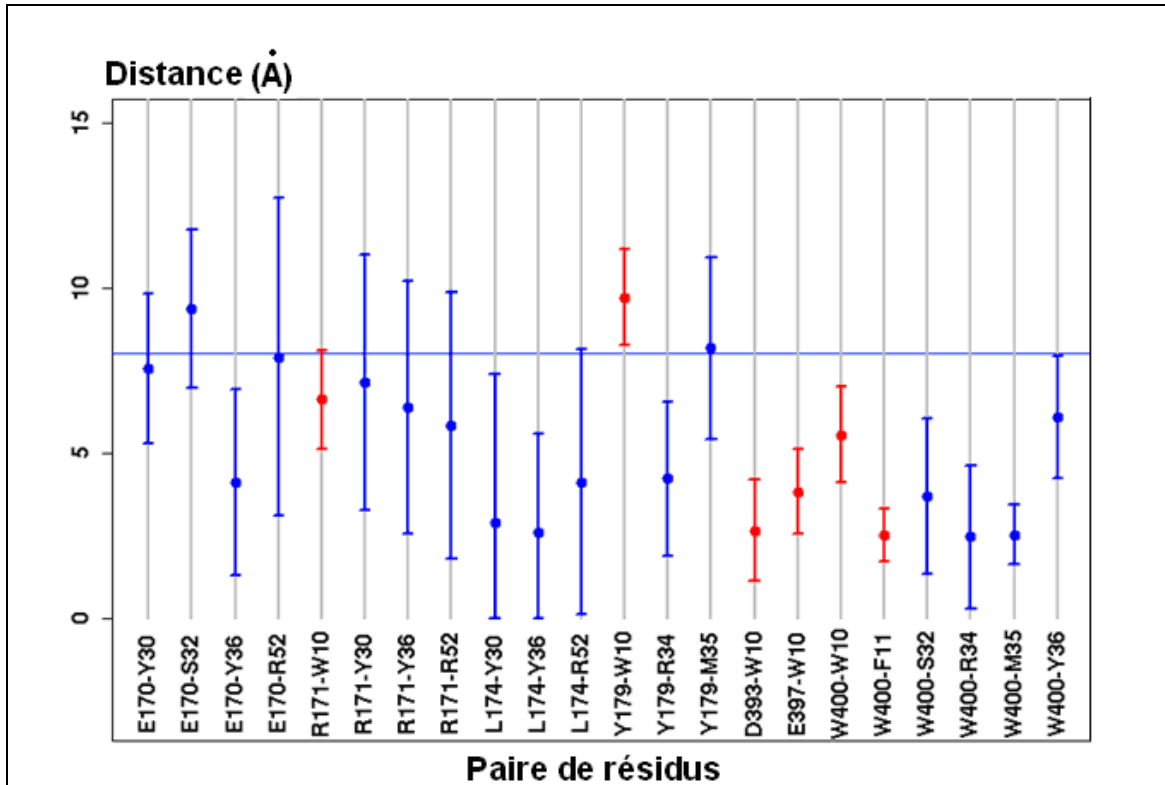
récepteurs A et B et celui de la toxine passant de 40 à 45 et 40 à 41 Å respectivement. Ces déplacements relatifs des 3 protéines semblent correspondre à une relaxation des tensions induites dans la structure du complexe par les contraintes introduites au cours de la simulation d'arrimage.

Pendant la dynamique on observe des fluctuations du plan de la toxine par rapport à un axe perpendiculaire à la membrane (Figure 69). L'amplitude de ce mouvement atteint un angle de l'ordre de 45° le long de la trajectoire de dynamique entre 25 et 50ns. Les résidus du bas des boucles I, II et III de la toxine en revanche fluctuent peu (rmsd < 2Å).



Les 21 paires de résidus sous contrainte durant l'étape d'arrimage restent proches durant la dynamique libre. En particulier, les distances de toutes ces paires restent en moyenne inférieures à 8 Å au cours du temps (Figure 70). 11 sont inférieures à 5 Å. Le couple formé par les résidus Tyr179 (hM1B) et Trp10 (MT7) présente la plus grande distance parmi les proximités dont nous disposons. La distance moyenne entre ces deux résidus fluctue autour d'une moyenne de 10Å, et ne repasse sous la barre des 8Å que très occasionnellement durant la simulation (Figure 71).

Pour la suite de cette étude, nous ne considérerons que les 25 dernières nanosecondes de simulation qui correspondent à une portion de simulation à partir de laquelle la distance entre les centres de gravité des deux monomères reste stable.



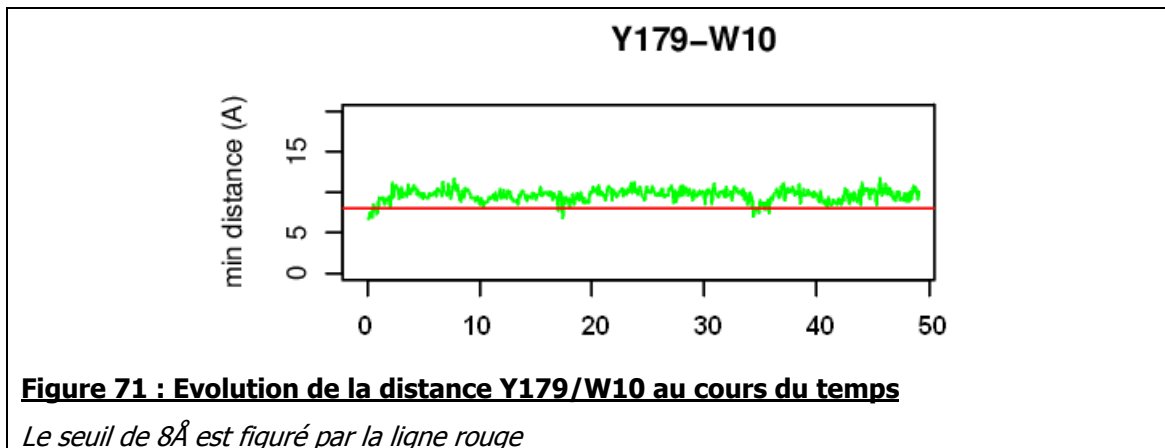
**Figure 70 : Fluctuation des distances des paires de résidus**

Les distances sont moyennées sur la période de 25 à 50 ns de la simulation.

La barre d'erreur correspond à  $\pm$  l'écart type des fluctuations

En bleu les couplages correspondant au récepteur A

En rouge les couplages correspondant au récepteur B



**Figure 71 : Evolution de la distance Y179/W10 au cours du temps**

Le seuil de 8Å est figuré par la ligne rouge

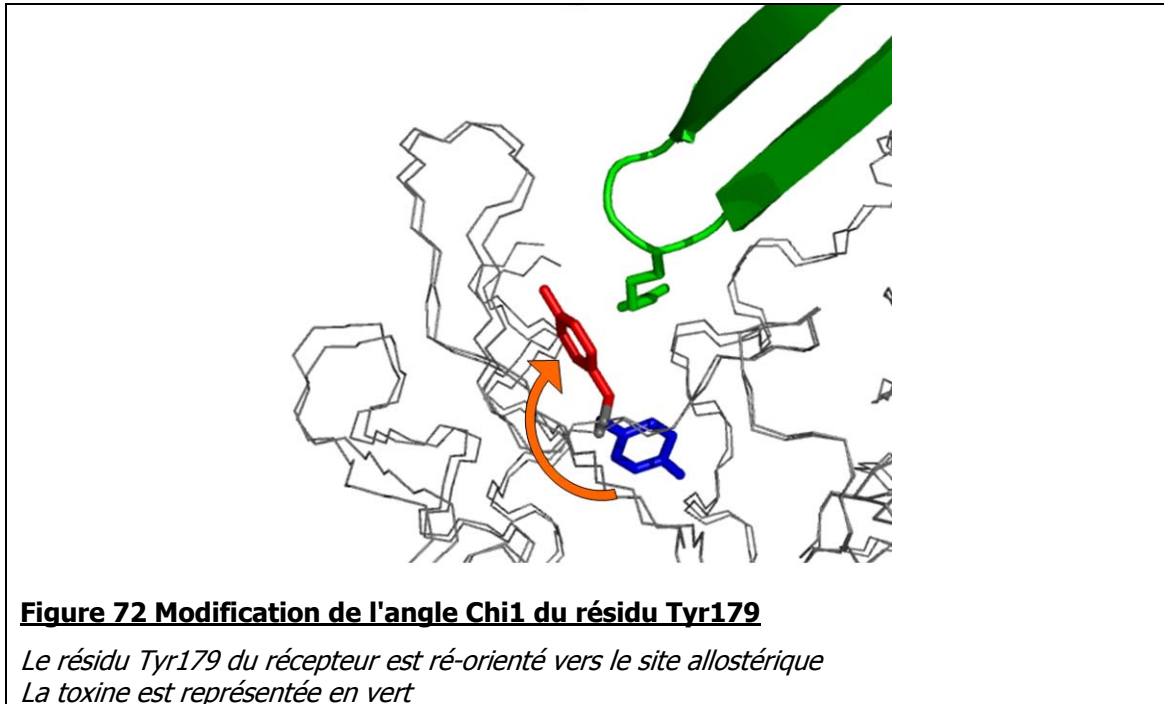
### ***4.3. Validation du modèle par recalculs d'énergie libre***

Des mesure expérimentales de variations d'énergie libre on été réalisée pour 12 mutants de la toxine MT7, et 22 coté récepteur hM1. Pour chacune de ces mutations, la variation d'énergie libre correspondante a été recalculée à partir du modèle. On utilise pour cela la méthode décrite par Eriksson et al sur le canal K<sup>+</sup> (Eriksson and Roux 2002) basée sur l'utilisation des équations Poisson Boltzman (cf 4.1.2.3). Les valeurs sont moyennées sur 25 jeux de coordonnées extraits des 25 dernières nanosecondes de la trajectoire de dynamique moléculaire.

Nous avons remarqué que la qualité des corrélations entre les valeurs expérimentales et recalculée des  $\Delta G$  est relativement différente pour les résidus de la toxine et du récepteur. Pour les mutants MT7, on observe un bon accord avec les valeurs expérimentales avec des écarts inférieurs à 1 kcal.mol<sup>-1</sup>. Côté récepteur en revanche, pour 4 mutations (W91A, W164A, Y179A, E397A) on obtient des écarts supérieurs à 1 kcal.mol<sup>-1</sup> (Figure 74).

Pour W91A et Y179A, les énergies calculées sont proches de zéro alors que les valeurs expérimentales sont de 1.64 et 2.33 kcal/mol respectivement. La non-interaction de W91 résulte du fait de l'absence de contrainte entre ce résidu et la toxine ainsi que la distance importante qui sépare les résidus (cf § 5.1.2.1). Dans le cas de Tyr179 en revanche, les contraintes de distances sont satisfaites lors du docking: des résidus de la toxine sont proches dans la structure du complexe.

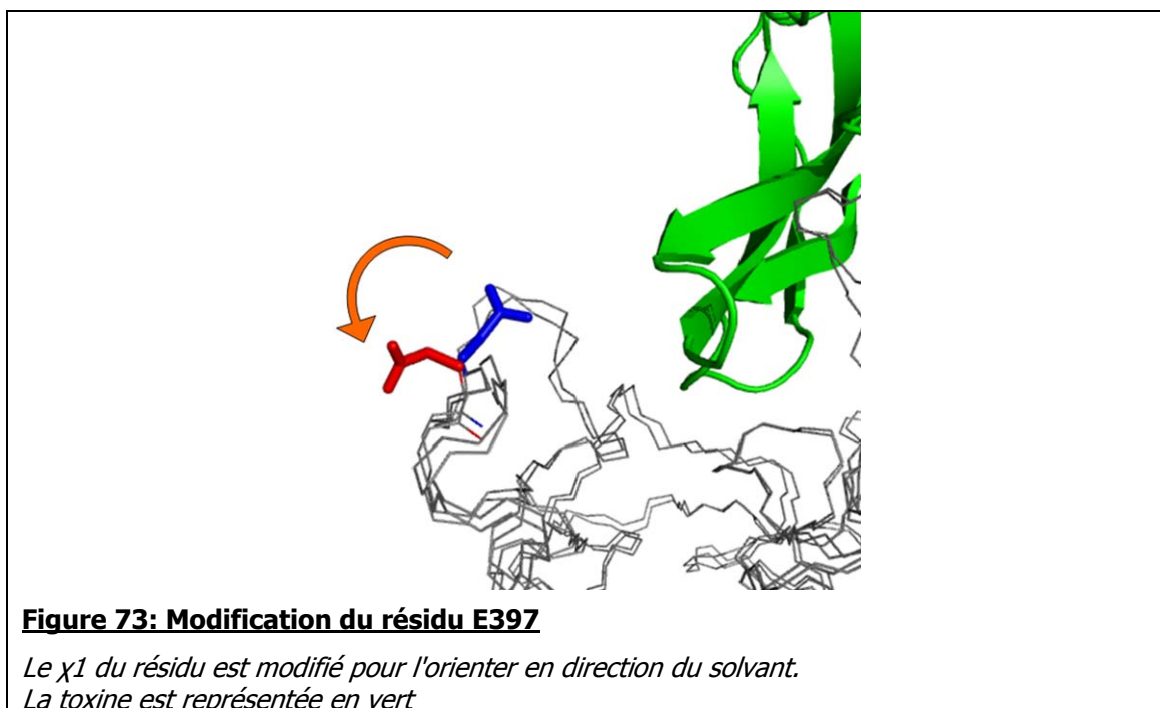
La chaîne latérale du résidu Y179 est positionnée en direction du site orthostérique dans la structure de la rhodopsine. Or, cette tyrosine établie des contacts avec plusieurs résidus de la toxine. Lors de la modélisation du récepteur, l'angle  $\chi_1$  de Tyr179, a donc été modifié afin de l'orienter en direction du site allostérique, en accord avec les travaux de Kukkonen et al (Kukkonen, Perakyla et al. 2004). Lors de la simulation d'arrimage, l'orientation de ce résidu sur le monomère A est revenue, dans une position semblable à la rhodopsine.



Après repositionnement de la chaîne latérale de Tyr179 (Figure 72) en direction de la toxine et optimisation de son interaction avec le résidu de la toxine R34 proche, la valeur de  $\Delta\Delta G$  progresse de 0.5 kcal/mol. Pour la mutation Y179F la valeur calculée d'énergie libre reste donc sous évaluée à seulement 0.3 kcal/mol par rapport à la valeur expérimentale de 1.7kcal/mol.

D'un autre côté, la valeur de  $\Delta\Delta G$  calculée pour le mutant E397A est très supérieure à la valeur mesurée expérimentalement. Cette chaîne latérale se trouve enfouie lors de la construction du complexe. Le repliement de chaînes latérales polaires lors de simulation de dynamique moléculaire dans le vide est un artefact classique.

La chaîne latérale de E397 a donc été réorientée vers le solvant (Figure 73) ce qui diminue la valeur de  $\Delta\Delta G$  de 0.4 kcal/mol.

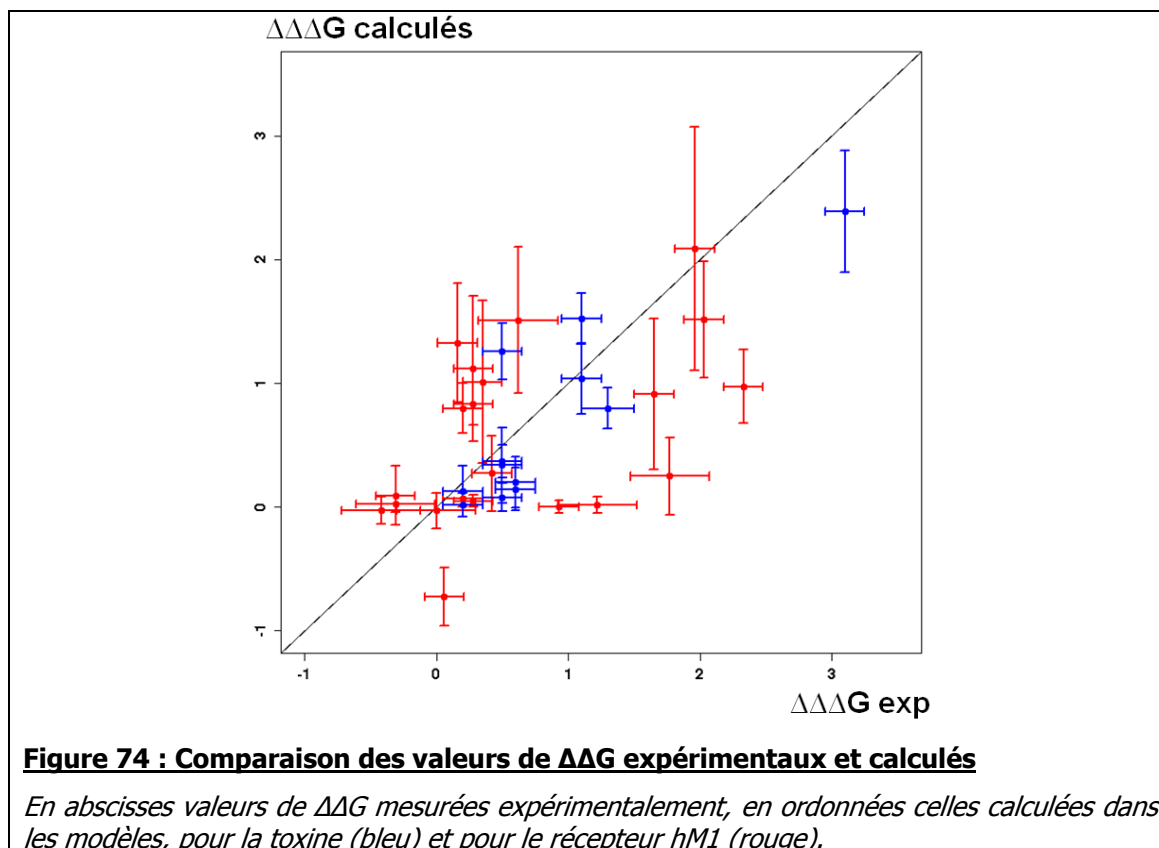




Les différences entre les valeurs expérimentales et recalculées des  $\Delta\Delta G$  semble donc être le résultat de chaînes latérales mal positionnées et non pas d'un mauvais positionnement global de la toxine.

La Figure 74 montre les valeurs des  $\Delta\Delta G$  calculés après optimisation des chaînes latérales de Tyr179 et E397, en fonctions des mesures expérimentales. Les fluctuations des valeurs d'énergie libre sont de 0.3 kcal.mol<sup>-1</sup> pour les valeurs calculées et 0.25 kcal.mol<sup>-1</sup> pour les valeurs expérimentales. La valeur absolue de l'erreur moyenne est de 0.8 kcal.mol<sup>-1</sup> et le coefficient de corrélation de Pearson 0.6. Des valeurs du même ordre de grandeur ont été obtenues par d'autres auteurs utilisant MM-PBSA pour calculer des valeurs d'énergie libre (Moreira, Fernandes et al. 2007), (Carlsson, Ander et al. 2006). Moreira et al 2006 considèrent un seuil de déviation de 1kcal.mol<sup>-1</sup> entre valeur expérimentale et calculée comme acceptable. Selon ce critère on obtient un accord satisfaisant pour 85% des positions étudiées.

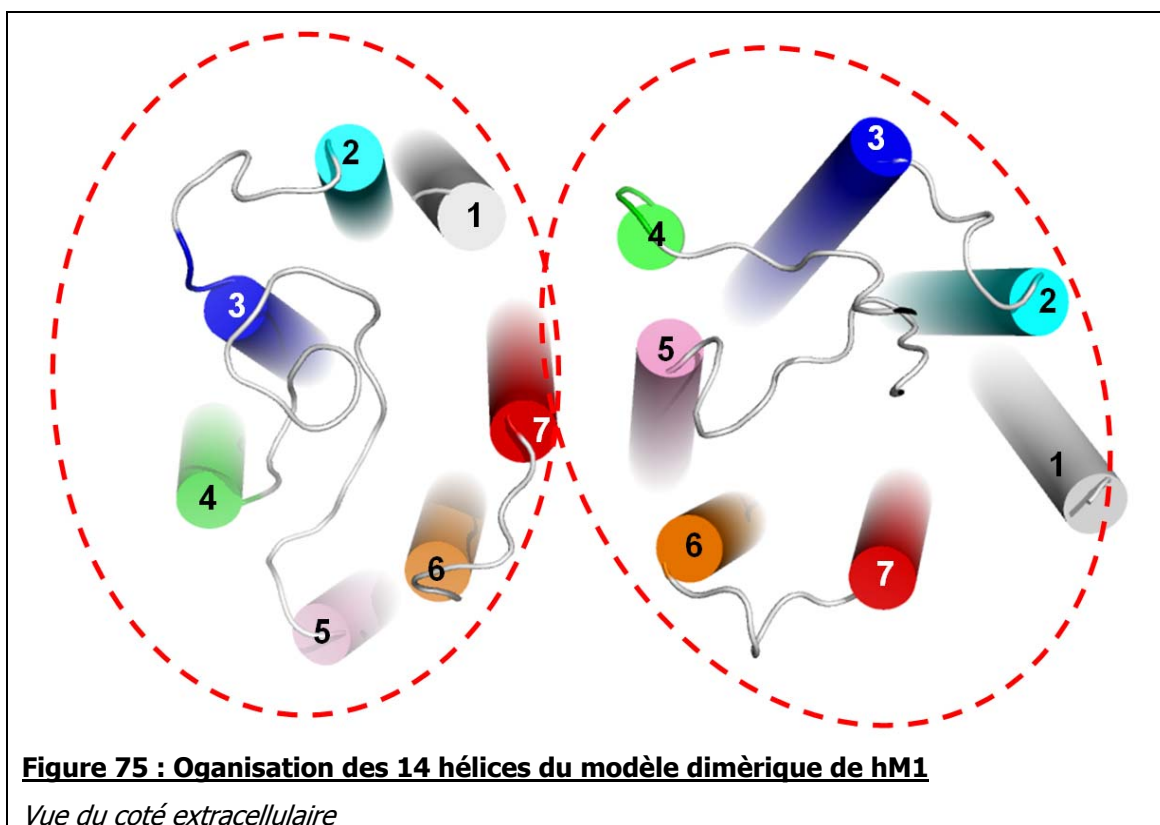
Notre modèle hM1/MT7 est donc capable de reproduire les propriétés énergétiques mesurées expérimentalement avec un bon accord.



#### 4.4. Analyse du modèle MT7/dimère de hM1

##### 4.4.1. Interface de dimérisation

L'interface de dimérisation des récepteurs est située au niveau des hélices TM4 et TM5 du côté de hM1A et TM1, TM6 and TM7 sur hM1B. La valeur totale de la surface accessible au solvant enfouie par la formation du dimère de récepteur passe de  $2234 \pm 96 \text{ \AA}^2$  à l'issue de la simulation d'arrimage à  $2609 \pm 114 \text{ \AA}^2$  après dynamique moléculaire. On assiste donc à une relaxation du modèle brut d'arrimage au cours de la simulation. Il est probable que la contrainte de  $50 \text{ kcal.mol}^{-1}$  appliquée entre les centres de masse des récepteurs pour assembler le dimère a créé des tensions dues aux effets de van de Walls.



#### 4.4.2. Site de liaison de la toxine

Le site de liaison de la toxine s'étend sur une surface totale de  $1996 \pm 141 \text{ \AA}^2$  sur les deux récepteurs hM1 assemblés en dimère. Les boucles II et III établissent des contacts avec le récepteur hM1A sur une surface de  $1331 \pm 103 \text{ \AA}^2$  et la boucle I avec un second récepteur hM1B sur une surface de  $693 \pm 114 \text{ \AA}^2$ . Ces surfaces étaient à l'issue de l'étape d'arrimage de respectivement  $1378 \pm 78 \text{ \AA}^2$  et  $502 \pm 75 \text{ \AA}^2$ , ce qui démontre que l'interface d'interaction s'est optimisée au cours de la simulation de dynamique moléculaire.

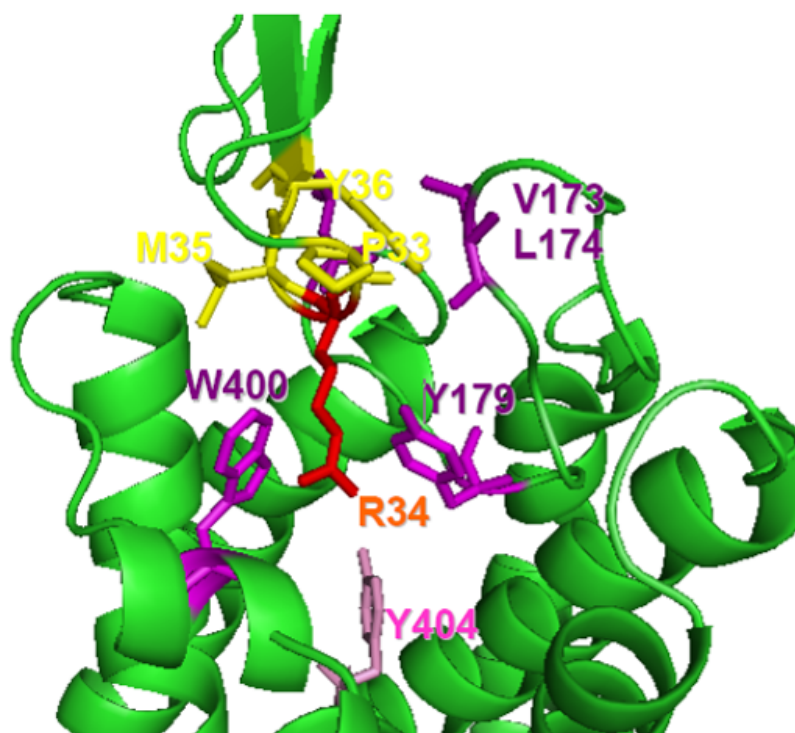
Pour définir la région de la toxine impliquée dans l'interaction avec les récepteurs, les résidus de la toxine ayant une surface enfouie de plus de  $10 \text{ \AA}^2$  par la formation du complexe sont sélectionnés (Figure 78). Ces résidus correspondent aux segments 8 à 12 de la boucle I, 28 à 38 de la boucle II et 51 à 52 boucle III. Le dernier résidu, la lysine 65, est isolé au niveau du domaine C-terminal.

Les fluctuations des coordonnées de ces résidus sont mesurées en superposant le long de la trajectoire de dynamique moléculaire (25 à 50 ns) les coordonnées des hélices des récepteurs. Nous constatons que cette interface est stable avec une déviation des coordonnées de moins de  $4 \text{ \AA}$ . Cette variation est due en majorité à la boucle III de la toxine, dont les fluctuations moyennes sont de l'ordre  $4.5 \text{ ns}$ . Les deux autres boucles de la toxine (8-12 et 28-38) ont des déplacements de carbones- inférieures à  $3 \text{ \AA}$ .

##### 4.4.2.1. Interface MT7/hM1A

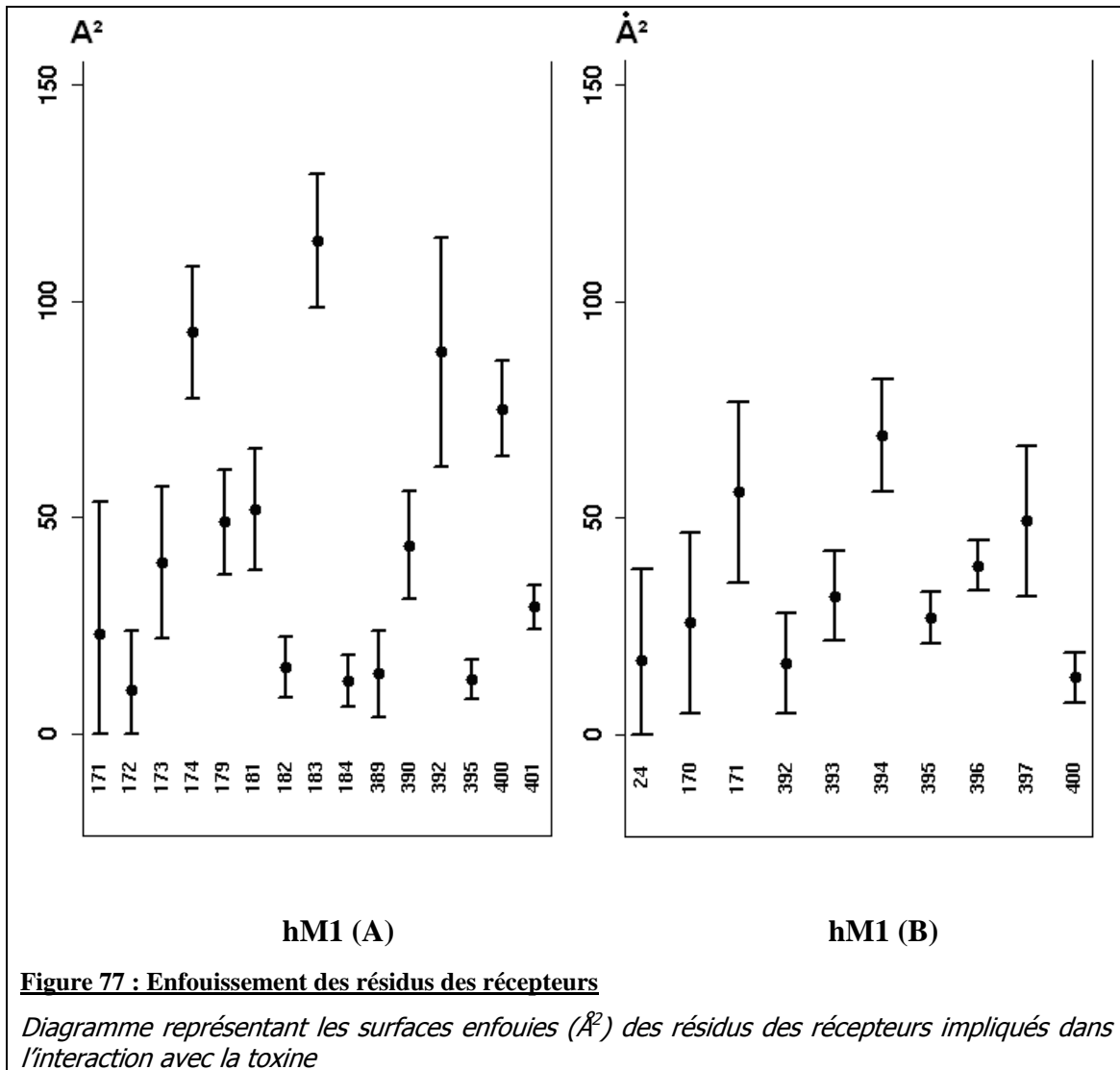
Les boucles II et III de la toxine établissent des contacts avec le protomère hM1A. Les régions du récepteur enfouies lors de la liaison de la toxine (Figure 77A) sont localisées sur la boucle extracellulaire E2, dans les parties supérieures des hélices TM5 (168-190), TM6 (388-397) et enfin TM7 (400-401). Les résidus les plus enfouis sont Leu174, Tyr179, Gln181, Leu183 Lys392 and Trp400. (Figure 76).

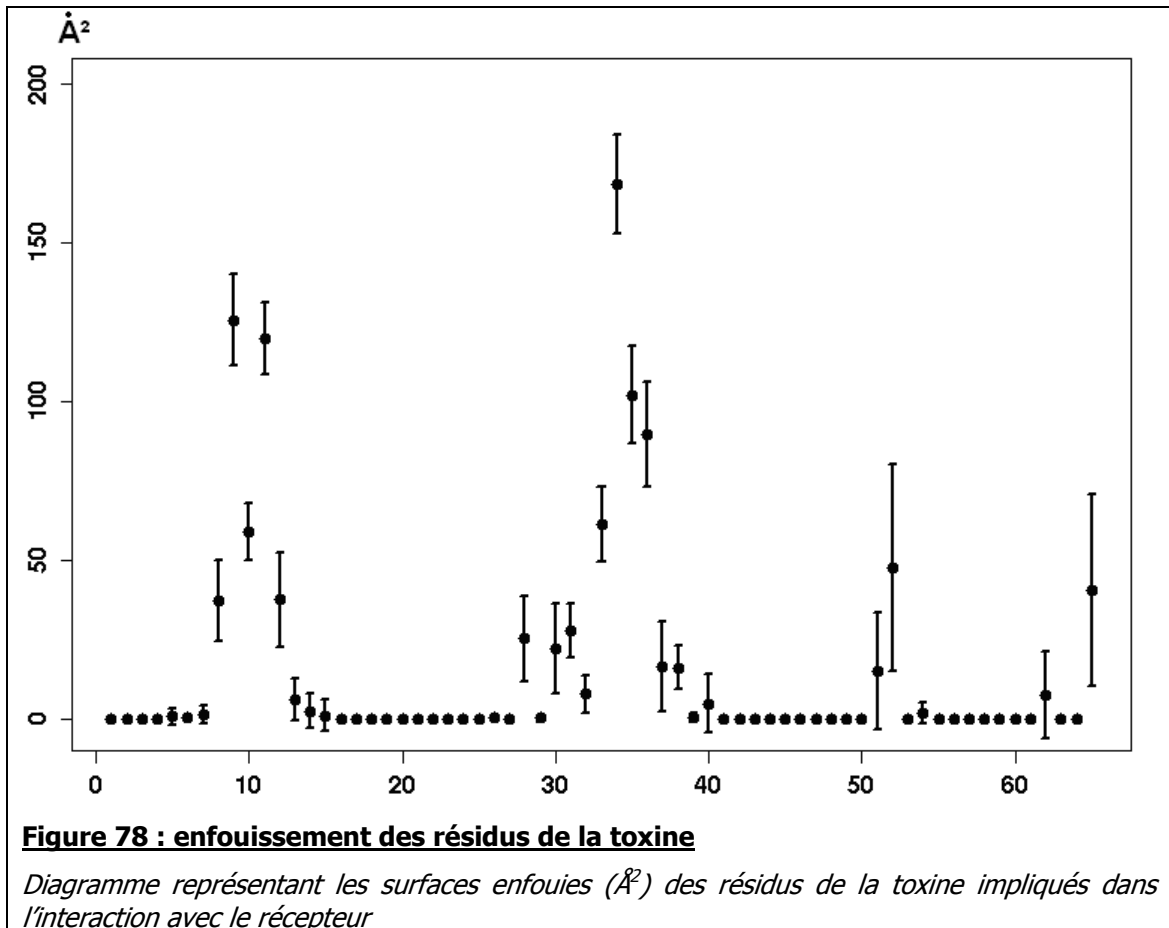
Coté toxine, l'analyse des surfaces enfouies montre que les contacts sont localisés au niveau des résidus Pro33, Arg34, Met35 et Tyr36 (Figure 76). Le plus enfoui de ces résidus est Arg34. Par ailleurs, ce résidu réalise également la contribution électrostatique la plus importante par des interactions cation- avec les résidus Trp400, Tyr179 and Tyr404, à équidistance de ces 3 résidus (Figure 76). La Tyr404 appartient à la fois au fond de la cavité de liaison des ligands allostérique et au plafond du site orthostérique (cf § 331.4.3.1). Les résidus hydrophobes Pro33, Met35 and Tyr36 participent à l'interaction principalement par des contributions à l'énergie de van der Waals. Ils forment un couvercle hydrophobe situé au dessus de la cavité de liaison occupée par R34. Ainsi, le haut de la boucle II de la toxine contribue à renforcer l'interaction électrostatique majeure établie par l'arginine 34, en la protégeant de l'accès du solvant (Figure 76).



**Figure 76: Agrandissement de l'interface de liaison de la toxine sur le récepteur hM1A**

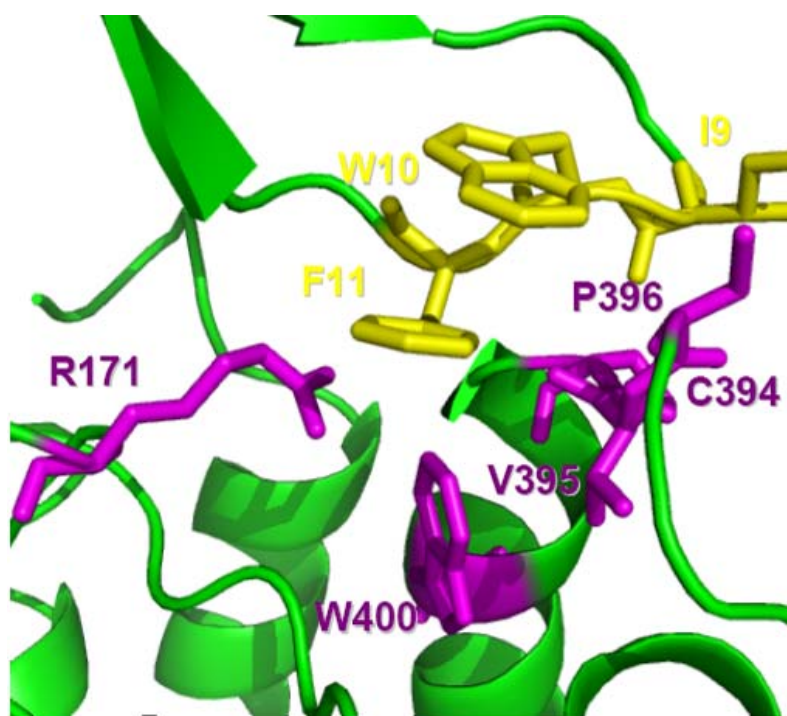
*En jaune les résidus de la boucle II de la toxine avec en rouge l'arginine 34, en violet les principaux contacts sur le récepteur.*





4.4.2.2. Interface MT7/hM1B

La boucle I de la toxine établit des contacts avec un second récepteur hM1B. Les résidus du récepteur enfouis par cette interaction (Figure 77B) sont également localisés sur la boucle E2 (Glu170, Arg171) ainsi que sur la boucle E3 (Asp393, Cys394, Val395, Pro396 and Glu397). Côté toxine (Figure 78), sont enfouis par cette interaction les résidus Ser8, Ile9, Trp10, Phe11 and Pro12. La Figure 79 montre Trp10 et Phe11, forment un groupe hydrophobe avec Cys394, Val395, Pro396 et Trp400. Par ailleurs côté toxine, Ile9, Trp10 and Phe11 ont également des contributions d'énergie de van der Waals importantes dans l'interaction.



**Figure 79 : Agrandissement de l'interface de liaison de la toxine sur le récepteur hM1B**

*En jaune les résidus de la boucle I de la toxine, en violet les principaux contacts sur le récepteur*





---

# Chapitre 5: Discussion

---

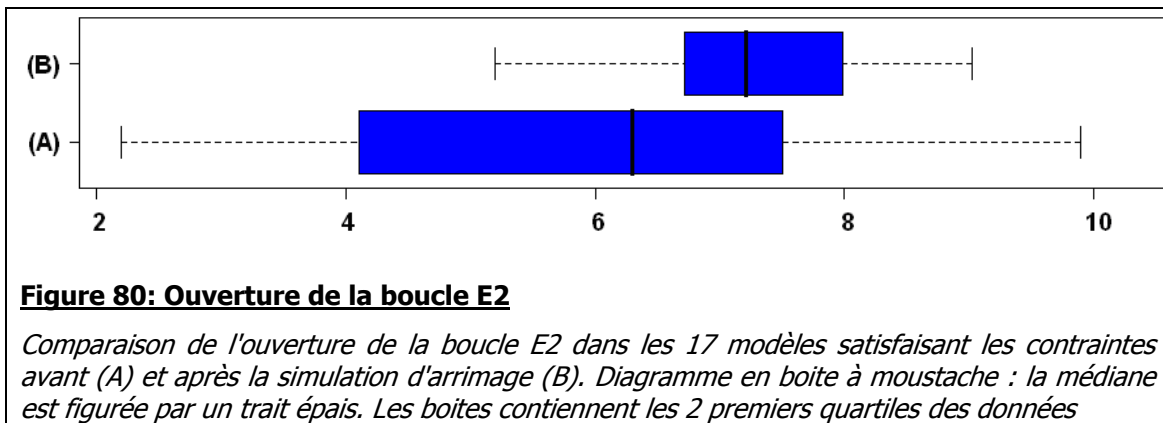
## ***5.1. Apports et limites de la dynamique moléculaire***

La dynamique moléculaire est omniprésente à toutes les étapes de notre procédure : Elle est utilisée pour échantillonner les changements conformationnels des partenaires avant arrimage, pour les assembler en complexe et enfin pour optimiser et valider le modèle. Cette rigueur dans la prise en compte de la flexibilité des protéines possède néanmoins un coût computationnel très élevé. La seule simulation de dynamique dans la membrane représente plus de 50.000 heures de calcul. L'étape d'arrimage se révèle également très lourde. Un gain de temps très important pourrait être réalisé à ce niveau. Tout d'abord en optimisant les procédures, mais également en réalisant un pré-filtrage lors de l'échantillonnage des positions de départ sur la base des données de doubles mutants.

### **5.1.1. Influence de la boucle E2**

Des conformations de la boucle extracellulaire E2 échantillonnées par dynamique moléculaire activée. Ces conformations qui décrivent un mouvement d'ouverture de la boucle de 0 à 10 Å de rmsd ont été utilisées comme point de départ pour les simulations d'arrimage sous contrainte. Les solutions acceptables obtenues lors de l'arrimage de la toxine correspondent à des conformations de départ de la boucle E2 ouverte de quatre à huit Angstroms (Figure 80B). Au cours des simulations d'arrimage, cette boucle est flexible et soumise à des contraintes de distances. On observe dans les modèles corrects du complexe que les boucles E2 ont à la fin de la simulation d'arrimage une ouverture de l'ordre de sept à huit Angstrom par rapport la position initiale (Figure 80A).

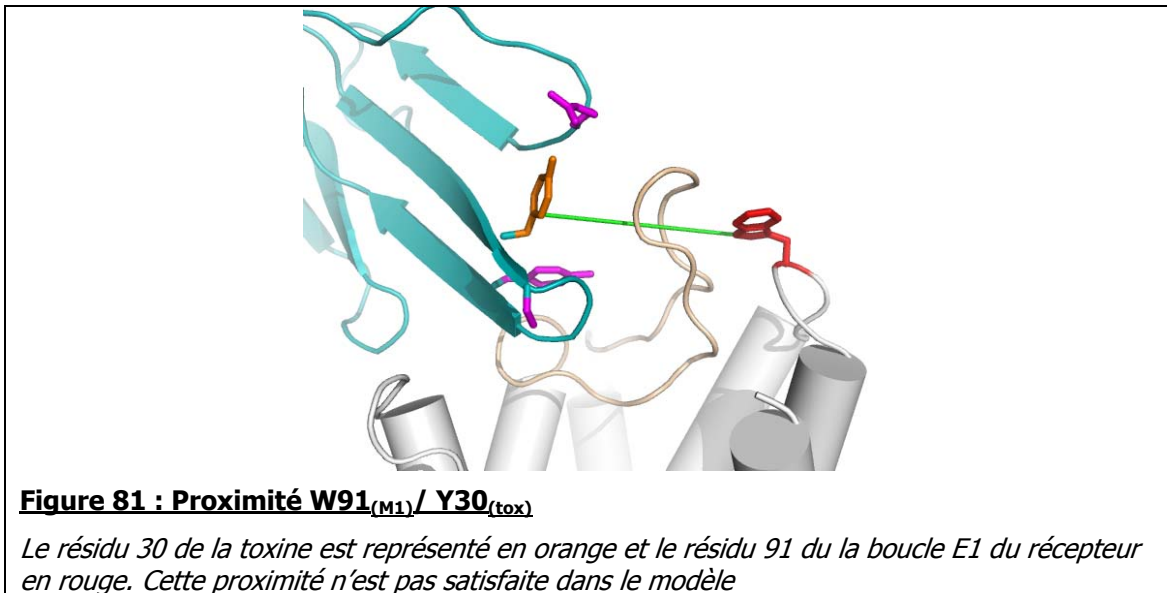
Il apparaît donc que la simulation de l'ouverture de la boucle par dynamique moléculaire activée avec le potentiel PEDC, et également sous les contraintes de données expérimentales, est un facteur déterminant dans la réussite de la prédiction de la structure du complexe hM1. Ce résultat est en accord avec les résultats de Alvani et al (Alvani, Gregory et al. 2007) : la mobilité de la boucle extracellulaire E2 est essentielle pour la liaison de ligands, par son rôle de "portail" modulant l'accessibilité du site de liaison du récepteur et par sa participation directe à l'interaction.



### 5.1.2. Données non prises en comptes

#### 5.1.2.1. W91<sub>M1</sub> / Y30<sub>tox</sub>

Le résidu Trp91 du hM1 est prédit d'après les expériences de doubles mutant comme proche du résidu Tyr30 de la toxine. Ces deux résidus sont localisé à environ 15 Å l'un de l'autre dans les modèles. De plus, la boucle E2 du récepteur vient se placer de telle manière qu'il est probable qu'elle écrante l'interaction entre ces deux résidus.



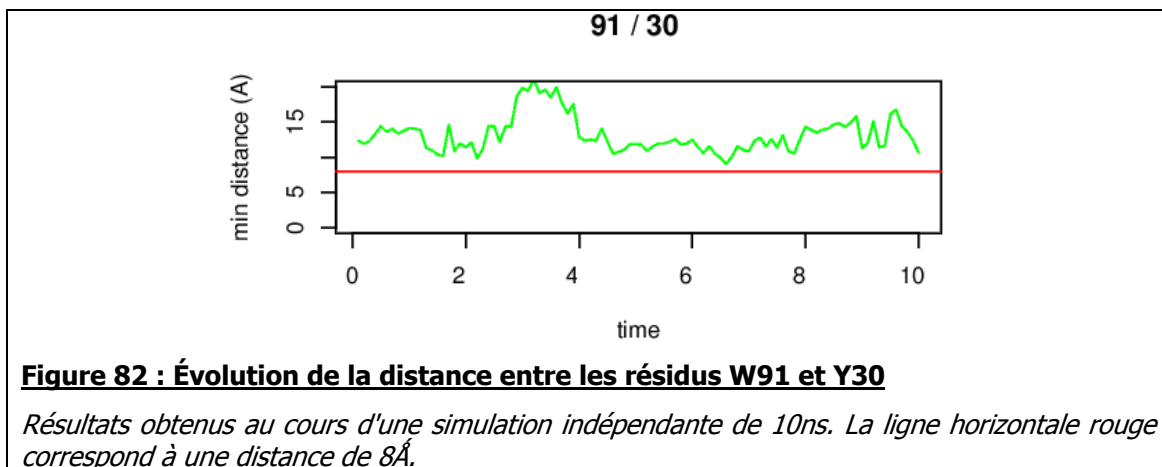
Le résidu W91 est le seul point de contact prédit au niveau de la boucle extracellulaire E1 du récepteur et d'autre part une seule proximité entre ce résidu et la toxine a été observée. Ceci ne nous permet donc pas vérifier si ce contact est cohérent par rapport au reste des données. Soit l'erreur se situe du coté des données expérimentales et la proximité Trp91/Tyr30 est fausse, soit ces résidus sont bien proches dans le complexe et ils ne sont pas correctement modélisés. Deux hypothèses principales peuvent expliquer notre incapacité à tenir compte de ces données.

- Les résidus W91 et Y30 sont situés de part et d'autre de la boucle E2 du récepteur dans les modèles. Dans une telle configuration, il est possible que la mutation de W91 entraîne un affaissement de la boucle E2. Or, le principe des cycles de doubles mutants repose sur l'hypothèse que les mutations alanine n'entraînent pas de changements structuraux. Dans le cas contraire, on risque de mesurer un G

traduisant en réalité un effet indirect et non pas une proximité entre les résidus. Dans ce cas, l'effet observé en 91 et 52 pourrait en fait traduire la perte d'interaction entre Arg52 et la boucle E2. De nouvelles mesures de cycle de doubles mutants avec les autres résidus de la toxine situé autour d'Y30 pourraient permettre de trancher cette hypothèse.

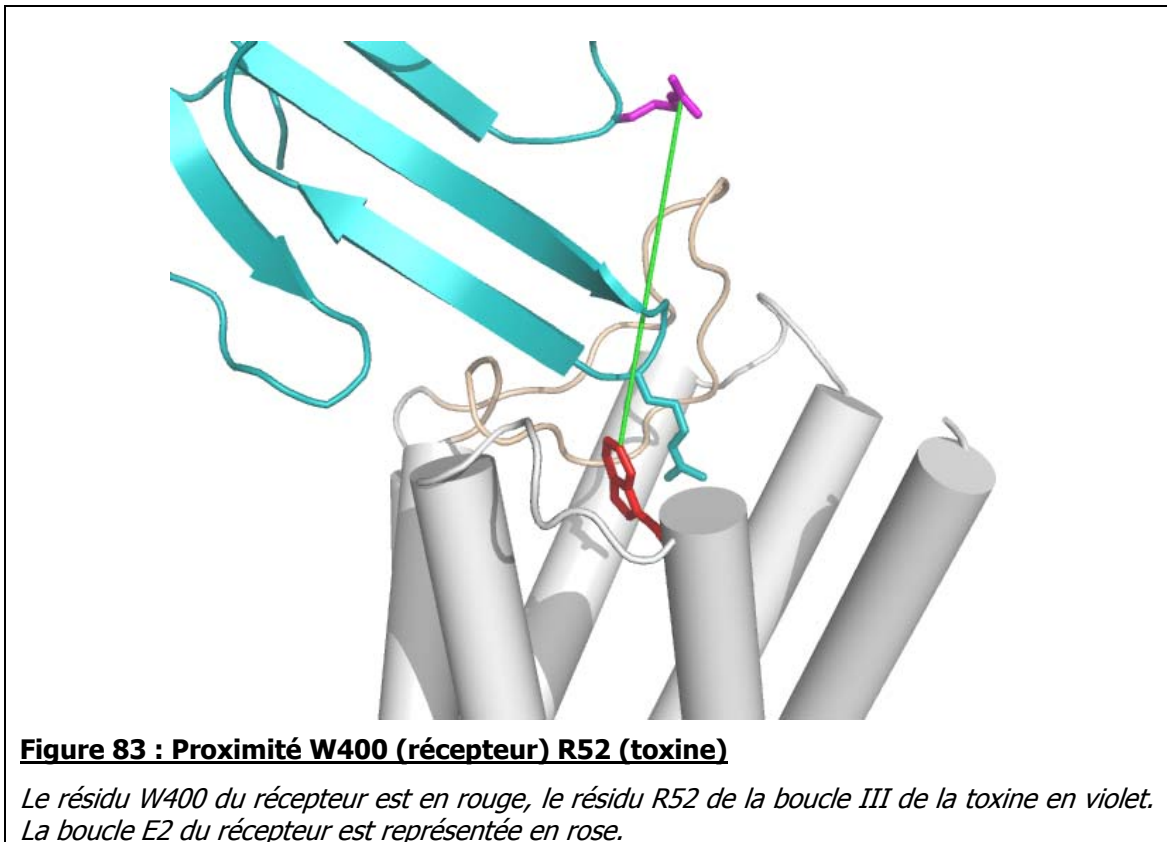
- Ce résidu est situé sur une boucle extracellulaire, modélisée d'après la rhodopsine. La conformation de cette boucle n'a pas été échantillonnée avant, ou pendant, la simulation d'arrimage. Il est donc tout à fait possible qu'elle ait été modélisée de manière incorrecte, ou qu'elle possède une flexibilité permettant d'établir des contacts avec la toxine. Un mouvement de la boucle E1 du récepteur positionnant W91 à une distance de près de 8 Å de Y30 a d'ailleurs été observé lors un calcul de dynamique moléculaire libre (Figure 82).

L'incapacité à expliquer les données expérimentales concernant le couple 91/30 pourrait traduirait donc une prise en compte encore insuffisante de la flexibilité de la boucle E1 du récepteur.



R52<sub>tox</sub> / W400<sub>M1</sub>

La seconde proximité qu'il semble impossible d'expliquer concerne les résidus W400 et R52, ces deux résidus étant distants de plus de 19Å. Cette proximité est beaucoup plus difficile à expliquer. En effet, les deux résidus ont chacun plusieurs autres partenaires détectées ce qui confirme leur implication dans la liaison. Et d'autre part plusieurs G négatifs ont été mesuré pour chacun d'eux, ce tend à démontrer la spécificité de ces mutants.



## 5.2. *Dimérisation des récepteurs muscariniques*

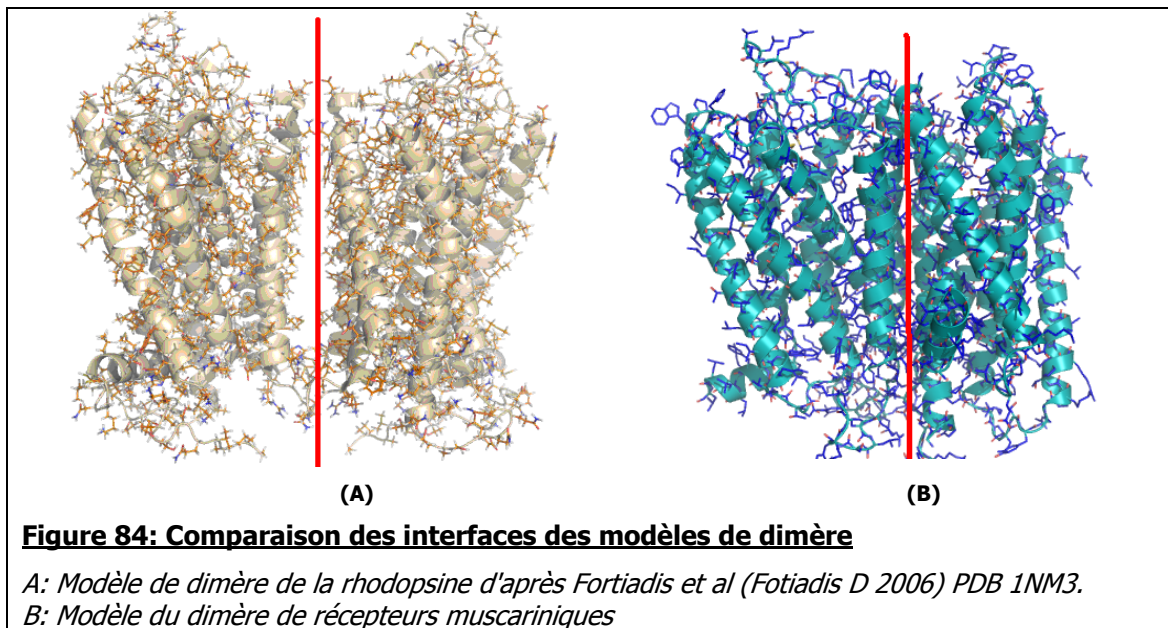
### 5.2.1. Nature de la surface d'interaction

Peu de données structurales sur la dimérisation des RCPG sont disponibles à l'heure actuelle. Le système le plus étudié est une fois encore la rhodopsine, pour laquelle on dispose d'un modèle moléculaire construit à partir des résultats de microscopie de force atomique (Fotiadis D 2006). Ce modèle décrit un assemblage peu compact des récepteurs, les hélices TM4 et TM5 se faisant face (Figure 84A). Pour ce modèle, la zone de contact au niveau des faisceaux d'hélices faible avec seulement 185 Å<sup>2</sup> de surface enfouie au niveau des hélices transmembranaire et 972 Å<sup>2</sup> pour la totalité de la protéine. La zone de contact est principalement localisée au niveau de la boucle intracellulaire I2.

Notre modèle présente une surface d'interaction beaucoup plus importante avec une 2609 Å<sup>2</sup> de surface enfouie par la formation du dimère. De nombreuses données concernant l'interface standard des protéines solubles sont disponibles dans la littérature. En revanche, les protéines membranaires étant peu représentées dans les bases de données, on ne dispose pas d'étude systématique de leurs surfaces d'interaction. On peut toutefois observer que la taille de la surface d'interaction de notre dimère est comparable aux 2370 Å<sup>2</sup> enfouis à l'interface de deux sous-unités du récepteur nicotinique (PDB 1YI5).

Pour évaluer la compacité de l'interface du dimère on compare l'enfouissement des hélices TM2 et TM3 du récepteur, positionnées au cœur du faisceau d'hélices, avec l'enfouissement des hélices à l'interface du dimère. Les valeurs d'enfouissement des hélices TM2 et TM3 sont respectivement  $1990 \pm 30 \text{ \AA}^2$  et  $1961 \pm 92 \text{ \AA}^2$ . On mesure des valeurs comparables pour les hélices impliquées dans la formation du dimère :

- hM1A: TM4A= $1472 \pm 40 \text{ \AA}^2$ , TM5A= $2035 \pm 91 \text{ \AA}^2$
- hM1B: TM1B= $1574 \pm 74 \text{ \AA}^2$ , TM6B= $2407 \pm 70 \text{ \AA}^2$ , TM7B= $1968 \pm 36 \text{ \AA}^2$



Ceci indique que la surface enfouie des hélices situées à l'interface du dimère est semblable à celle des hélices au cœur d'un monomère. Le type d'interface que nous prédisons est donc de nature très différente de celle proposée pour la rhodopsine (Fotiadis D 2006).

### 5.2.2. Mode d'assemblage du dimère de récepteurs

Très peu de données structurales concernant l'interface de dimérisation de RCPG sont disponibles pour nous permettre de comparer notre modèle. D'après ces résultats, les hélices transmembranaires TM3 et TM4 semblent impliqués dans la dimérisation des récepteur 5-HT<sub>4</sub>, (Berthouze, Rivail et al. 2007), les hélices TM4 et TM5 pour les récepteur D<sub>2</sub>-dopaminergiques (Guo, Shi et al. 2003) et C5a (Klco, Lassere et al. 2003) ou l'hélice TM6 du  $\beta_2$ AR (Hebert, Moffett et al. 1996). Ces données indiquent donc que

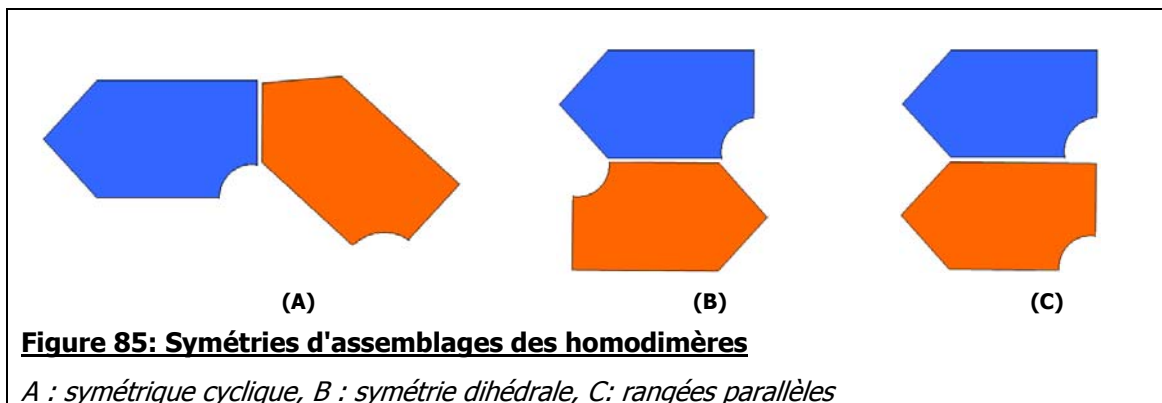


les modes d'assemblage des dimères de RCPG sont vraisemblablement très différents d'une famille de RCPG à l'autre.

Tous ces auteurs s'accordent en revanche pour décrire des assemblages présentant des symétries diédrale dans les dimères de RCPG. Cependant, il faut noter que ces résultats reposent essentiellement sur des expériences de crosslinking. Ce type d'expérience consistant à introduire une cystéine dans les hélices et à vérifier si on observe la formation d'un dimère ponté de manière covalente et résistant aux purifications. Ce type d'approche n'est donc, par principe, pas en mesure de détecter autre chose que des assemblages ayant des symétries diédrale.

Or, les protéines membranaires ne sont pas libres de s'assembler en trois dimensions, mais sont contraintes dans le plan de la membrane. il n'existe pas de raison a priori pour que les RCPG s'assemblent selon une symétrie diédrale, on peut envisager 3 cas de figures pour assembler des dimères dans la membrane :

- Une symétrie diédrale, ou la même face de chaque récepteur contribue à l'interface
- Une symétrie cyclique, autour d'un axe perpendiculaire à la membrane.
- Des rangées, si les monomères sont parfaitement parallèles.



Levy et al (Levy, Boeri Erba et al. 2008) ont observé que les homodimères présentant des symétries cyclique sont 11 fois plus rares que ceux présentant des symétries diédrale dans leur jeu de données. Cette observation trouve son origine dans les mécanismes évolutifs. En effet, dans les complexes diédraux, les protéines sont face à face tandis que dans les complexes cycliques est sont face à dos. La formation d'interface de type cyclique par mutations ponctuelles aléatoire est donc moins probable.

La propriété la plus étonnante dans le modèle du dimère que nous avons prédit est donc l'assemblage d'un homodimère de récepteurs présentant une symétrie cyclique. Le mode de dimérisation des récepteurs que nous observons n'est pas le plus probable d'un point de vue évolutif. Ceci est corroboré par l'analyse de la conservation des résidus impliqués dans les interfaces de chaque monomère. On observe en effet un taux de conservation moyen de respectivement 64% et 74% pour les résidus impliqués dans les interfaces des récepteurs hM1a et hM1b du dimère. La conservation moyenne sur l'ensemble de la protéine étant de 70%, ceci indique que les régions impliquées dans la formation du dimère ne subissent pas de pression de sélection particulière.

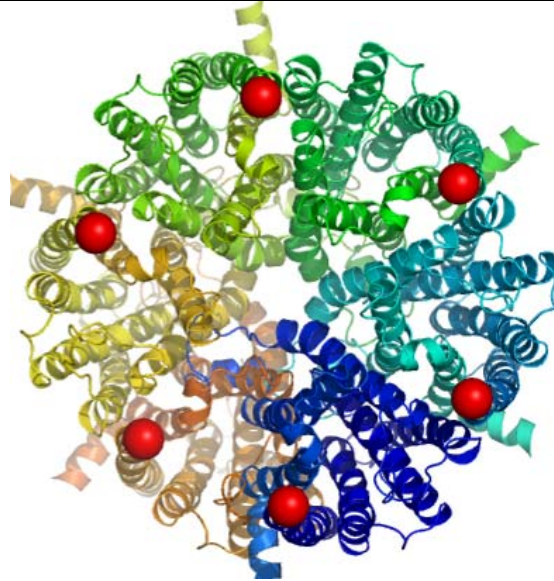
Il est donc possible que le mode de dimérisation des récepteurs que nous prédisons ne corresponde pas à un dimère natif de récepteurs. Ce mode d'assemblage particulier pourrait être induit sous la contrainte de la toxine MT7 liant les sites allostériques des deux récepteurs.

### 5.2.3. Oligomères d'ordre supérieurs

Au-delà de leur capacité à former des dimères, l'existence d'oligomères d'ordre supérieur à 2 a effectivement été démontrée pour certains RCPG. Par exemple, des expériences d'immunoprécipitation, immunopurification et immunoblotting ont montré qu'il existe, pour les récepteurs muscariniques M2, des assemblages au moins trimériques et probablement d'ordre supérieur. De même, plus récemment, l'assemblage multimérique de récepteurs  $\alpha 1b$  a été observé par le groupe de Miligan grâce à des techniques de FRET à 3 partenaires (Lopez-Gimenez, Canals et al. 2007).

L'existence d'une interface qui ne présente pas de symétrie diédrale dans le modèle de dimère pose la question de la possibilité pour ces récepteurs de s'assembler en chaîne et de former des oligomères d'ordre supérieur à 2. Compte tenu de l'interface de dimérisation que nous avons prédit, les récepteurs formant un angle d'environ  $240^\circ$ , la taille maximale d'un oligomère ne peut excéder le stade de l'hexamère (Figure 86). Il est étonnant de constater qu'on obtient, peu de chose prêt, un assemblage fermé avec un nombre entier de récepteurs.

Il serait très intéressant d'appliquer la méthode du FRET à trois partenaires de Miligan afin d'étudier les possibilités de formation d'assemblages de plus de deux récepteurs chez le hM1.



**Figure 86: Oligomère putatif de M1**

*Construction d'un oligomère de hM1 en dupliquant le modèle 6 fois et en superposant les monomère  $B_{(n+1)}$  sur les monomères  $A_{(n)}$ . Les sphères rouges représentent le N-ter des récepteurs.*

*La vue est perpendiculaire à la membrane, du côté extracellulaire.*

### ***5.3. Bases moléculaire des propriétés pharmacologiques de la toxine***

#### **5.3.1. Site de liaison**

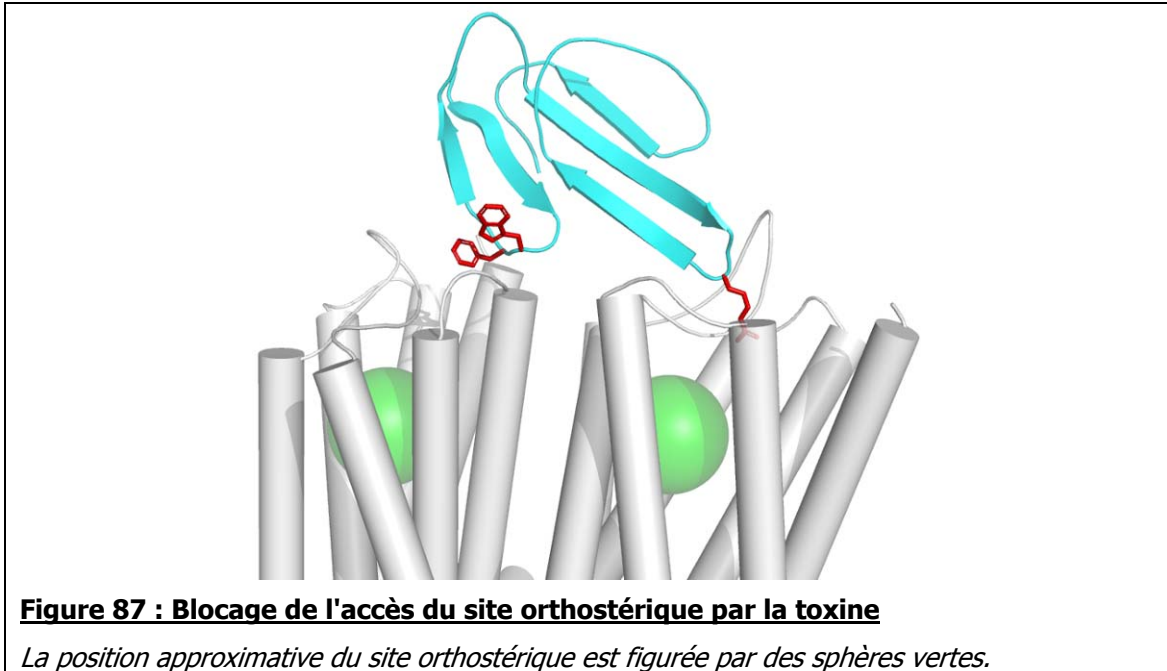
Le site de liaison de la toxine implique de nombreux résidus de la boucle extracellulaire E2 (Gly169-Tyr179) et Trp400 en haut de l'hélice transmembranaire TM7. Ces résidus sont en bon accord avec les résultats la littérature qui indiquent que les résidus de la boucle E2 sont essentiels pour la liaison de ligands allostériques (Kukkonen, Perakyla et al. 2004; Fruchart-Gaillard, Mourier et al. 2006). Les données de mutagénèse et l'analyse du modèle converge pour accorder un rôle important au résidu Tyr179. Ce résidu est équivalent au Tyr177 du récepteur M2 et a été identifié comme prépondérant pour la liaison de ligands allostériques (Mohr, Trankle et al. 2003; Trankle, Dittmann et al. 2005). Les travaux de ce groupe suggèrent également que résidu Y179 se trouve à la jonction entre le haut de la cavité orthostérique et le fond du site allostérique.

#### **5.3.2. Stoechiométrie NMS / MT7 / hM1**

Les travaux de l'équipe de D. Servent établissent, par l'analyse des résultats des expériences de liaison, que la stœchiométrie des ligands MT7 et NMS est équimolaire sans pouvoir toutefois conclure sur la stœchiométrie MT7/hM1 ou NMS/hM1.

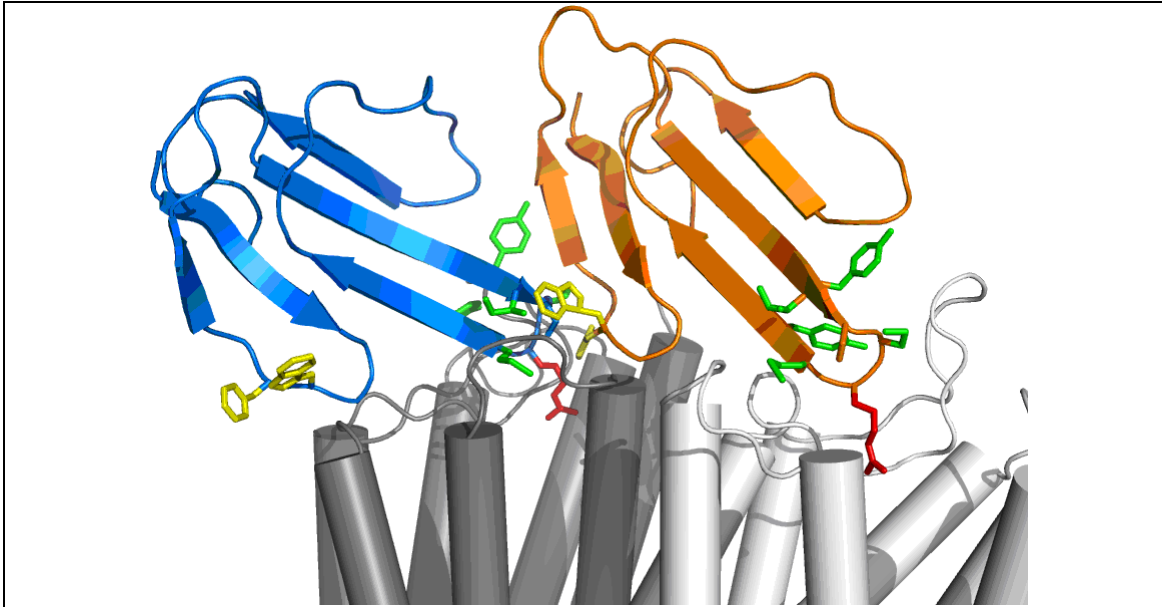
##### **5.3.2.1. Hypothèse 1 : 1 dimère de M1, 1 toxine, 1 NMS**

La première hypothèse que l'on peut formuler, est que la stœchiométrie du système NMS, MT7, hM1 est 1:1:2. Le positionnement de la toxine, la boucle II engagée profondément dans la cavité extracellulaire du récepteur, avec le résidu Arg34 affleurant le haut de la poche de liaison des ligands orthostériques. Le positionnement exact du NMS n'étant pas parfaitement connu, il est possible que l'Arg34 interfère avec le site orthostérique et empêche la liaison du NMS. Du côté de la boucle I de la toxine en revanche, les contacts hydrophobes réalisés essentiellement par les résidus Trp10 et Phe11 sont localisés plus en surface et ne semblent pas pouvoir gêner directement l'accès à la cavité orthostérique.



#### 5.3.2.2. Hypothèse 2 : 1 dimère de M1, 2 toxines, 2 NMS

On constate (Figure 88) que notre modèle n'interdit pas la possibilité de lier deux toxines sur un dimère de récepteurs. Une telle configuration ne présente pas d'incompatibilité stérique. Un modèle formé d'un trimère de récepteur liant deux toxine peut être minimisé avec des énergies de van der Waals faibles. On constate alors que les résidus hydrophobes de la boucle I de la toxine se positionnent de telle manière qu'ils complètent le "bouchon" hydrophobe formé par les résidus du haut de la boucle II (Pro33, Met35, Tyr30) et pourraient ainsi contribuer à protéger l'interaction polaire réalisée par l'Arg34.



**Figure 88 : Possibilité de formation d'oligomères d'ordre supérieur**

*La liaison d'une seconde toxine de la même manière que la première ne présente pas d'incompatibilité stérique majeure*

Enfin, nous avons vu que la possibilité de formation d'oligomères de plus de deux récepteurs ne peut être exclue. On peut donc imaginer des situations plus complexes plusieurs toxines pourraient se lier sur des assemblages de plus de deux récepteurs. La formation d'oligomères fermés (cf § 5.2.3) pourrait être compatible avec une stœchiométrie 1:1:1.

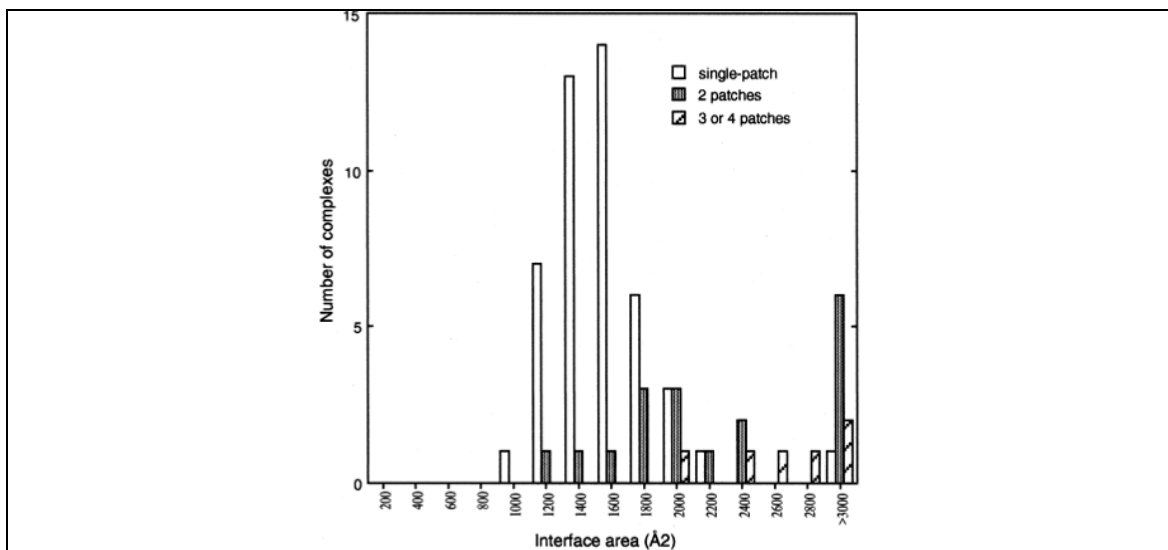
### 5.3.3. Haute affinité de liaison

La toxine MT7 présente une très haute affinité pour le récepteur hM1, de l'ordre de 29pM. Bien qu'il n'existe pas relation linéaire directe entre surface d'interaction et affinité de la liaison, on peut tout de même considérer que de manière générale, les interactions très affines correspondent à des surfaces d'interaction de grande taille (Brooijmans, Sharp et al. 2002).

La surface d'interaction mesurée lorsque la toxine est arrimée à un unique récepteur est de  $1331 \pm 103 \text{ \AA}^2$ , ce qui correspond à une valeur basse par rapport au standard des interfaces protéine/protéine de  $1600 \pm 400 \text{ \AA}^2$  (Lo Conte, Chothia et al. 1999). Il est donc peu probable qu'une interface de cette taille soit suffisante pour expliquer la haute affinité de la toxine. Si l'on considère la toxine en complexe avec un

dimère de récepteur en revanche, on atteint une valeur de  $1996 \text{ \AA}^2$  qui se positionne dans les valeurs de tête des standards et correspond mieux à l'affinité picomolaire.

Une étude portant sur la composition des interfaces de protéine (Chakrabarti and Janin 2002) montre que les protéines ayant des surfaces d'interaction de plus de  $2000 \text{ \AA}^2$  sont constituées d'au moins 2 groupes de contacts (Figure 89). L'interface de liaison de notre toxine qui établie des contacts au niveau des groupes de résidus principaux, boucle I et boucle II présente donc une topologie cohérente avec les données de la littérature.



**Figure 89: Composition des interfaces protéine/protéine en fonction de la surface d'interaction**

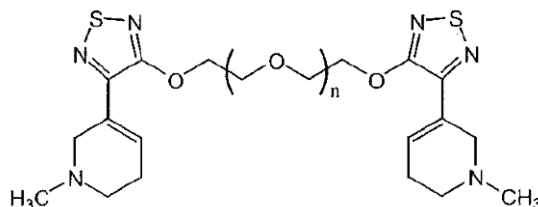
*D'après Chakrabarti (Chakrabarti and Janin 2002)*

*Les protéines possédant des interfaces de plus de  $2000 \text{ \AA}^2$  possèdent au moins 2 sites de contacts distincts*

Les travaux de (Rajeswaran, Cao et al. 2001) s'intéressent à la conception de ligands divalents du récepteur muscariniques M1. Les auteurs ont réalisé des constructions ou des agonistes du M1 (carbachol, xanoméline) sont pontés par des chaînes de longueur variables (Figure 90). Pour certaines tailles de linkers, les affinités mesurées de la construction sont jusque 100 fois plus élevées que celle du pharmacophore natif. Il est intéressant de constater que les construction démontrent une affinité maximale pour le récepteur avec 3 répétitions du motif linker. Dans cette configuration, les pharmacophores sont distants de  $15.5 \text{ \AA}$ . Les auteurs font l'hypothèse que ce ligand pourrait se lier à plusieurs sites sur un monomère de récepteurs muscariniques. Cette distance étant comparable à celle du même ordre ( $\sim 15 \text{ \AA}$ ) qui



sépare le bas des boucles I et II de la toxine, on peut faire l'hypothèse que les modes d'action des deux ligands sont en fait comparables et qu'ils lient tous les deux des dimères de récepteurs.



**Figure 90 : Ligand divalent du récepteur hM1**

*D'Après (Rajeswaran, Cao et al. 2001)*

Ces résultats indiquent donc que la capacité d'un ligand à lier simultanément des récepteurs assemblés en dimère pourrait être un facteur déterminant dans la haute affinité. Pour vérifier cette hypothèse, on pourrait introduire des mutations dans les hélices des récepteurs abolissant la dimérisation. On s'attend alors à observer une baisse d'affinité de la toxine MT7 pour ces récepteurs recombinants.

On sait déjà que le mutant W10A ne provoque pas de perte de la capacité à stabiliser le dimère pour la toxine. Toutefois, le remplacement d'un tryptophane par une alanine au niveau d'un site hydrophobe ne constitue pas une perturbation très importante. Des mutants de MT7 perturbant la nature hydrophobe de la boucle I pourraient être synthétisés et leur capacité à influencer la dimérisation de hM1 étudiée systématiquement.

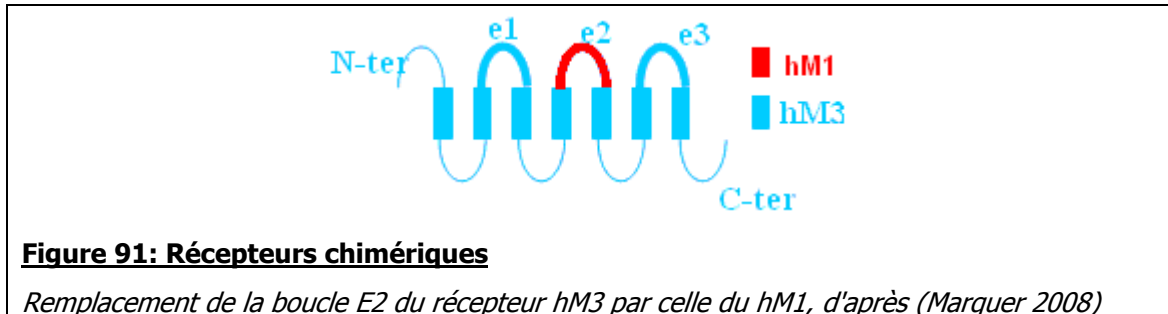
En revanche, il ne semble pas que l'induction de la dimérisation soit liée à la capacité des ligands à ponter les dimères. En effet, un certain nombre d'auteurs rapportent des effets de stabilisation de formes dimériques similaires pour de petits ligands. Par exemple les récepteurs hSSTR5 (Patel, Lange et al. 2002), les récepteurs Adenosine A2A et dopaminergique D2 (Vidi, Chemel et al. 2008) et les récepteur TRHR (Kroeger, Hanyaloglu et al. 2001) ont la propriété de dimériser sous l'action d'agonistes.

### 5.3.4. Sélectivité pour le sous-type M1 des récepteurs muscariniques

En dépit de la haute similarité de séquence entre les récepteurs hM1 et hM3, la toxine MT7 présente une affinité 100 fois supérieure pour la forme M1. La construction de récepteurs chimériques (Marquer 2008) a démontré que le remplacement de la boucle extracellulaire E2 du M3 par celle du M1 (Figure 91) permet d'atteindre des valeurs d'affinité proche de celle du M1. Notre modèle apporte un éclairage sur les bases structurales de cette sélectivité.

On trouve parmi les résidus du récepteur en contact avec la toxine un certain nombre de position peu conservées dans la famille des mAChR qui pourraient constituer des déterminants dans la sélectivité des toxines. Il s'agit des résidus Glu170, Leu174, Val395 et Glu397.

L'acide glutamique 170 du récepteur M1 est particulièrement intéressant car il correspond à une inversion de charge par rapport aux autres sous-types de récepteurs muscariniques qui possèdent des lysines à cette position (Figure 92).

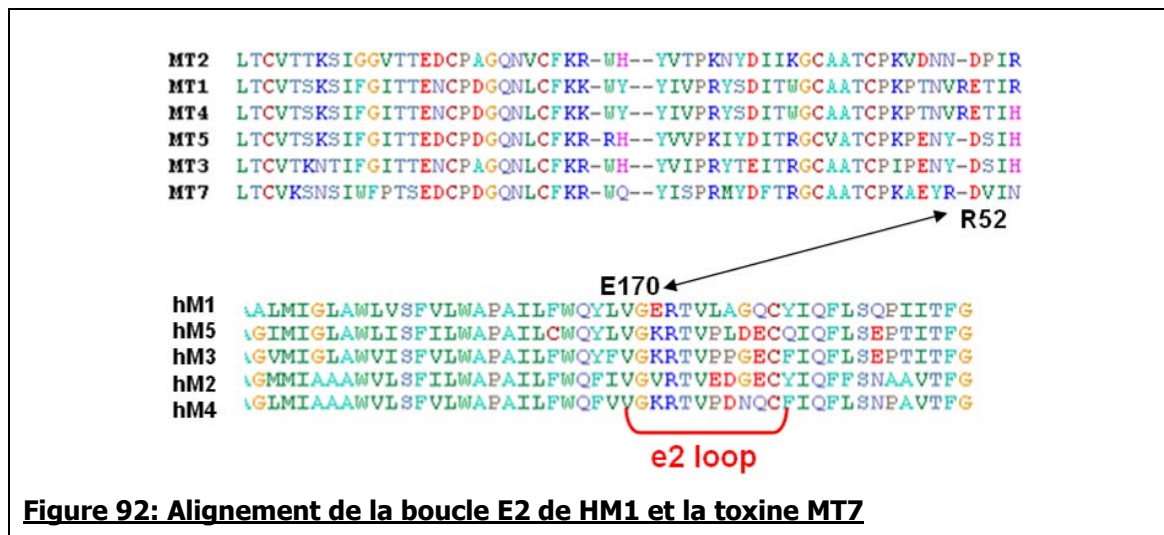


En ce qui concerne la toxine, on constate que bien qu'ils soient spécifiques de la MT7, les résidus Trp10 et Phe11 de la boucle I de la toxine ont pour équivalents des acides aminés ayant des propriétés similaires chez les autres toxines muscariniques (Phe et Gly). Il est donc peu vraisemblable que cette boucle I explique la sélectivité de la MT7. Ceci indique que la liaison à un second récepteur par cette boucle pourrait être une propriété partagée par toutes les toxines muscariniques. Cette hypothèse pourra être vérifiée en réalisant des expériences de transfert de fluorescence entre différents récepteurs muscariniques et les toxines MT1 à 6.

En revanche, la boucle III de la toxine et en particulier les résidus Tyr51 et Arg52 impliqué dans des contacts avec le récepteur selon notre modèle sont variables

dans la famille. En particuliers, le résidu polaire Arg52 correspond à une position majoritairement hydrophobe (Figure 92) dans le reste de la famille des toxines muscariniques.

On constate que les deux résidus Arg52-(tox) et Glu170-(M1) qui correspondent à des acides aminés ayant des propriétés originales par rapport aux familles de la toxine et du récepteur respectivement sont justement en interaction l'un avec l'autre dans notre modèle. L'affinité de liaison du double mutant R52A+E170A conduit à une perte d'affinité d'un facteur 60 environ, à comparer avec l'affinité 100 fois plus basse de la MT7 pour le hM3 que pour le hM1. Le mutant E170K a été réalisé. Il serait intéressant de réaliser une toxine modifiée R52E et d'effectuer l'expérience de double mutant avec le couple R52E+E170K pour quantifier la part que l'interaction de cette seule paire de résidus est susceptible de jouer dans la sélectivité de la toxine pour le hM1.





---

# Chapitre 6: Conclusion et Perspectives

---

Les récepteurs muscariniques à l'acétylcholine sont des protéines transmembranaires jouant des rôles dans de nombreux processus physiologiques. La toxine MT7 est un puissant modulateur allostérique de ces récepteurs. D'autre part, cette toxine est le seul ligand connu spécifique du sous-type 1 des récepteur muscariniques.

Nous avons souhaité étudier les bases moléculaires de l'interaction entre la toxine MT7 et le récepteur hM1. Pour cela, nous avons tout d'abord effectué un échantillonnage des structures des deux partenaires en dynamique moléculaire. Les mouvements de grande amplitude des boucles du récepteur ont été prédits par dynamique activée. Puis, nous avons, à partir des conformères échantillonnés, réalisé l'arrimage flexible de la toxine sur le récepteur en nous basant sur des données de mutagénèse et des calculs de dynamique moléculaire sous contraintes ambiguës. Ensuite, ce modèle a été optimisé par une simulation de dynamique moléculaire libre, en environnement membranaire explicite. Enfin, des calculs en retours des valeurs d'énergie libre de liaison ont été effectués afin de valider le modèle. Nous prédisons que la toxine se lie sur un dimère de récepteurs hM1.

Le modèle que nous proposons indique que la toxine serait en fait en contact simultané avec deux récepteurs, assemblés en dimère. A partir des données expérimentales, nous sommes en mesure de prédire l'assemblage de ce complexe trimérique : Les boucles II et III sont en contact avec un premier récepteur et forme le cœur de l'interaction au niveau de l'arginine 34 qui plonge dans la cavité du récepteur. La boucle I de la toxine réalise des contacts hydrophobes avec un second récepteur.

Le mode de d'interaction de la toxine permet d'atteindre une surface de contact avec le récepteur relativement importante qui pourrait expliquer la haute affinité de liaison de ce couple. D'autre part, l'analyse de ce modèle apporte des éléments de compréhension des bases moléculaires de la sélectivité de la toxine pour ce sous-type de récepteur. Il apparaît que cette sélectivité est essentiellement portée par la boucle extracellulaire E2 du récepteur.

### ***6.1. Perspectives : les bases moléculaires de l'allostérie***

L'utilisation thérapeutique de modulateurs allostérique présente de nombreux avantages (Gilchrist 2007). En particulier, leur action est subordonnée à la liaison simultanée d'un ligand orthostérique, et donc saturable. Ainsi, cette propriété limite les risques de toxicité de la molécule par surdosage. Autre avantage, les modulateurs allostériques exercent leur influence que dans les tissus dans lesquels l'agoniste naturel possède un effet physiologique (Christopoulos 2002; May and Christopoulos 2003). Il est donc particulièrement intéressant de comprendre les bases moléculaires des mécanismes allostériques des ligands de RCPG. L'existence de dimères de récepteurs ouvre de nouvelles perspectives pour la compréhension des effets allostérique de certains ligands (Milligan and Smith 2007). En particulier le mode de liaison original de la toxine MT7 prédit par notre modèle apporte un nouvel éclairage sur son mode d'action. Il serait donc très intéressant d'étudier les propriétés allostériques de ce ligand à la lumière de la structure de ce complexe.

Les effets allostériques sont, par nature, dynamiques puisqu'ils modifient les cinétiques d'association et/ou de dissociation des ligands. Or, notre modèle apporte un éclairage purement statique sur l'interaction hM1/MT7. De nombreuses questions restent sans réponses sur la manière dont la toxine, ligand allostérique, et le NMS, ligand orthostérique, interagissent.

La suite logique de ce travail serait donc de s'intéresser aux propriétés dynamiques et allostériques du récepteur muscarinique. Pour cela, nous envisageons d'utiliser des techniques de dynamique moléculaires.

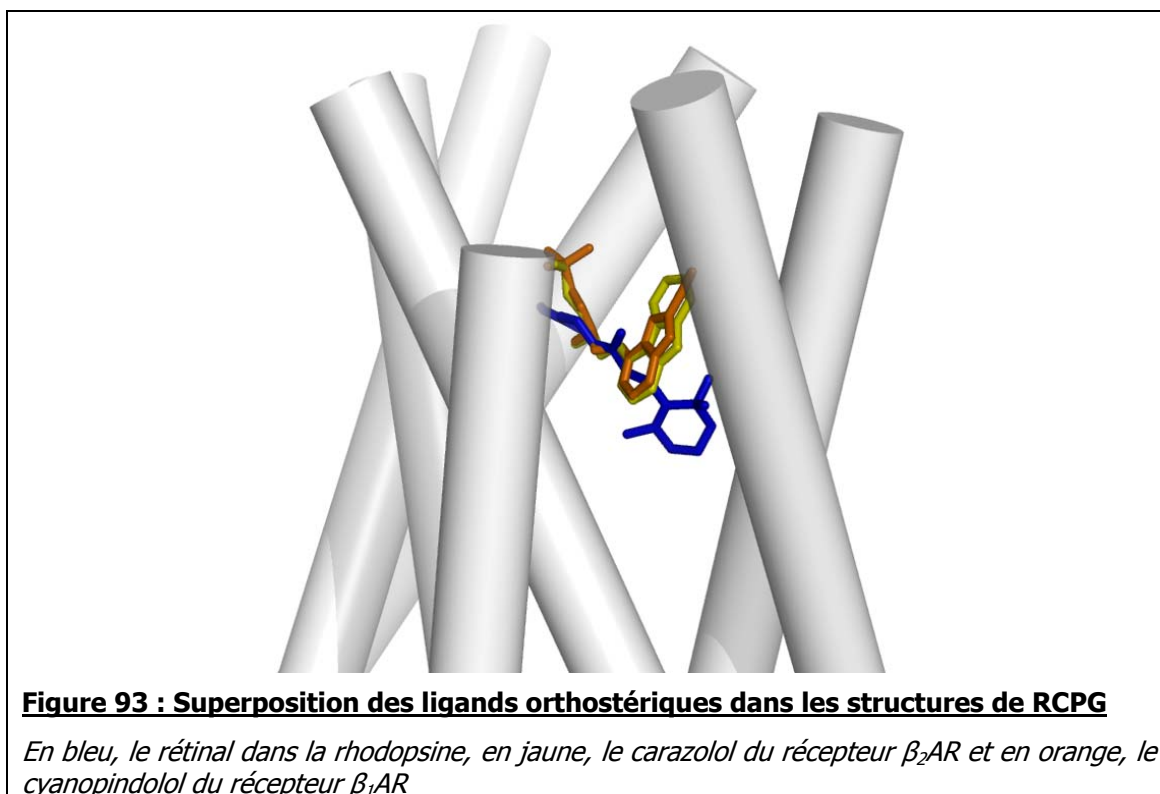
Nous disposons dans la littérature et grâce au travail accompli par de groupe de D. Servent de nombreuses données expérimentales concernant les propriétés pharmacocinétique des ligands MT7 et NMS. En plus des affinités de liaison de la toxine sur un récepteur libre de tout ligand, des expériences similaires ont été menées sur des récepteurs préalablement incubés dans de la NMS et dont les sites de liaison sont donc "occupés".

Nous disposerons de données thermodynamique pour divers cas de figure possible avec le trio NMS/MT7/hM1:

- Liaison du NMS seul
- Liaison de la toxine seule
- Liaison de la toxine sur un récepteur libre
- Liaison de la toxine sur un récepteur occupé par le NMS

Toutes ces informations seront d'un intérêt crucial pour guider et/ou valider des simulations de dynamique moléculaire.

Pour réaliser ce type de simulation, il sera nécessaire de se poser la question du positionnement du NMS au niveau du site orthostérique. On dispose de données structurales sur les positions du rétinol dans la rhodopsine ainsi que de deux antagonistes des récepteur  $\beta_1$  et  $\beta_2$  adrénergiques. Ces trois sites sont partiellement superposables mais leurs localisations précises restent cependant relativement différentes. Le cycle du rétinol étant situé nettement plus profond à l'intérieur du domaine transmembranaire que le Carazolol ou le Cyanopindolol (Figure 93).





Comme on peut le voir, le NMS est une molécule qui semble s'apparenter plus au rétinol qu'aux antagonistes de récepteurs AR. D'autre part, un certain nombre d'expériences de mutagenèse ont permis de cartographier la cavité de liaison du NMS. Ces informations permettront de guider la modélisation d'un récepteur occupé par une molécule de NMS. La position du NMS au sein du récepteur est une donnée structurale importante, qui détermine la possibilité ou non de positionner à la fois la boucle II de la toxine MT7 et le NMS sur un même récepteur.

Des modèles pourront ensuite être construits, avec un ou deux NMS, situés sur l'un et/ou l'autre des deux récepteurs hM1 formant un dimère. L'étude de la stabilité d'une simulation de dynamique moléculaire d'un modèle avec un NMS dans la cavité orthostérique et une toxine MT7 dans le site allostérique pourrait permettre d'analyser l'influence de la présence du NMS sur la localisation de MT7 et peut être d'éclairer la question de la stœchiométrie MT7/hM1/NMS.

Selon le modèle Monod-Wyman-Changeux, la base d'une interaction allostérique est l'induction de changements structuraux du récepteur par la liaison d'un ligand A, affectant le site de liaison d'un ligand B et donc ses cinétiques de liaison.

Nous avons vu que les sites allostériques et orthostériques du récepteur hM1 sont relativement proches dans la structure du récepteur. Si les changements structuraux induits par la liaison de la toxine sur le site de liaison du NMS sont essentiellement des réarrangements locaux par des mouvements de chaînes latérales, on peut étudier de tels mouvements par des simulations de dynamique moléculaire.

Si en revanche il s'agit de changements conformationnels comparables à ceux responsables de l'activation du récepteur, de grande amplitude et impliquant la réorganisation du faisceau d'hélice, ces mouvements ne seront pas observables à l'échelle de temps de la dynamique moléculaire et devront être traités par d'autres approches (Lacapere, Pebay-Peyroula et al. 2007).

La méthodologie mise en place par Kong et Karplus (Kong and Karplus 2007; Kong and Karplus 2008) est basée sur l'étude de corrélation d'énergie d'interactions calculées sur une dynamique moléculaire en environnement membranaire explicite. Cette approche permet d'obtenir des informations sur les résidus impliqués dans les transitions induites par l'isomérisation du rétinol. Elle pourrait également être appliquée pour obtenir un aperçu des perturbations causées par la présence du NMS.

L'étude par l'analyse vibrationnelle des principaux modes normaux pourrait nous permettre d'obtenir un aperçu macroscopique des mouvements coordonnés de grande amplitude de la structure tel que les changements d'angle des hélices. Cette approche a par exemple permis de reproduire les mouvements allostériques d'ouverture du pore du récepteur nicotinique (Taly, Delarue et al. 2005; Taly, Corringer et al. 2006).

Enfin, on peut également envisager de réaliser une analyse en composante principale d'une dynamique moléculaire du système en environnement membranaire explicite d'une durée suffisamment longue. Une telle approche permettrait de profiter de l'avantage que nous procure la description explicite de la bicouche lipidique et de son interaction avec les protéines (Debret, Valadie et al. 2008) tout en appréhendant les mouvements de grande amplitude.



---

# Chapitre 7: Index des figures

---

FIGURE 1 : LA SIGNALISATION CELLULAIRE .....	12
FIGURE 2 : CASCADE DE SIGNALISATIONS DECLENCHEES PAR LES RCPG .....	14
FIGURE 3 : ORGANISATION D'UN RCPG.....	15
FIGURE 4: STRUCTURE DE LA RHODOPSINE BOVINE.....	17
FIGURE 5 : DOMAINES EXTRACELLULAIRES DE LA RHODOPSINE ET DU RECEPTEUR $\alpha_2$ ADRENERGIQUE .....	19
FIGURE 6 : POSITIONNEMENT DES LIGANDS ORTHOSTERIQUE DANS LES STRUCTURES DE RCPG .....	20
FIGURE 7 : MODELISATION DE RANGEES DE RHODOPSINES A PARTIR DE DONNEES AFM.....	23
FIGURE 8 : EVOLUTION DE LA BASE DE DONNEES <i>PROTEIN DATA BANK</i> .....	25
FIGURE 9 : SYNAPSE CHOLINERGIQUE .....	29
FIGURE 10 : AMANITE TUE-MOUCHE .....	29
FIGURE 11 : ALIGNEMENT DES SEQUENCES DES 5 SOUS-TYPES DE RECEPTEURS MUSCARINIQUES .....	31
FIGURE 12 : IMPLICATION DES RECEPTEURS MUSCARINIQUES DANS LES PATHOLOGIES HUMAINES .....	32
FIGURE 13 : MODELE DU SITE DE LIAISON ORTHOSTERIQUE DU RECEPTEUR hM1 .....	34
FIGURE 14 : LIAISON DE LIGANDS ORTHO ET ALLOSTERIQUE AU RECEPTEUR MUSCARINIQUE M2.....	36
FIGURE 15 : MAMBA VERT.....	37
FIGURE 16 : ALIGNEMENT DES SEQUENCES DES TOXINES MUSCARINIQUES .....	38
FIGURE 17 : STRUCTURE CRISTALLOGRAPHIQUE DE LA TOXINE MT2 (CODE PDB 1FF4).....	38
FIGURE 18 : AFFINITE DES TOXINES MUSCARINIQUES POUR LES DIFFERENTS SOUS-TYPE DE RECEPTEURS..	39
FIGURE 19: ALIGNEMENT DE SEQUENCE DE LA BOUCLE II DES TOXINES ET DE LA BOUCLE E2 DES RECEPTEURS.....	40
FIGURE 20 : CIBLES 20 ET 28 DU CONCOURS CAPRI.....	47
FIGURE 21 : CYCLES THERMODYNAMIQUE DES DOUBLES MUTANTS .....	49
FIGURE 22 : $\Delta G$ EN FONCTION DE LA DISTANCE DES RESIDUS DANS LE COMPLEXE BARNASE/BASTAR	50
FIGURE 23 : SUPERPOSITION DU MODELE nAChR- $\alpha 7$ / $\alpha 10$ -COBRATOXINE AVEC LA STRUCTURE CRISTALLOGRAPHIQUE AChBP/ $\alpha 10$ -COBRATOXINE.....	54
FIGURE 24 MODELISATION DU CANAL POTASSIQUE.....	58
FIGURE 25 : SUPERPOSITION DU MODELE ET DE LA STRUCTURE CRISTALLOGRAPHIQUE UPAR/uPA.....	60
FIGURE 26 : MODIFICATION DES BOUCLES DU RECEPTEUR M1 .....	64
FIGURE 27: ALIGNEMENT DES SEQUENCES DES BOUCLES EXTRACELLULAIRES E1 DE LA RHODOPSINE ET DU hM1.....	65
FIGURE 28 : MODELE DU RECEPTEUR hM1 .....	65
FIGURE 29 : DIAGRAMME DE RAMACHANDRAN DU MODELE DU hM1 .....	66
FIGURE 30 : STRUCTURE CRISTALLINES DE TOXINE MT7 MODIFIEES.....	68
FIGURE 31 : COMPARAISON DES STRUCTURES DES MT7 MODIFIEES AVEC LA MT2.....	68
FIGURE 32 : MODELE STRUCTURAL DE LA TOXINE MT7 .....	69
FIGURE 33 : DIAGRAMME DE RAMACHANDRAN DE LA STRUCTURE DE LA TOXINE MT7.....	70

FIGURE 34 : MODELE DU COMPLEXE TERNAIRE .....	74
FIGURE 35 : VARIATION D'ENERGIES LIBRES ASSOCIEES A DES MUTATIONS DE LA TOXINE .....	76
FIGURE 36 : RESIDUS DE LA TOXINE MODIFIES EN ALANINE.....	77
FIGURE 37: SEQUENCES PEPTIDIQUES DES BOUCLES E2 DU hM1 ET DU hM3 .....	80
FIGURE 38 : MUTAGENESE DU RECEPTEUR hM1 .....	80
FIGURE 39 : VARIATION D'ENERGIES LIBRES ASSOCIEES A DES MUTATIONS DU RECEPTEUR.....	81
FIGURE 40 : VALEURS DE $\Delta G$ EN KCAL.MOL <sup>-1</sup> POUR LES 48 COUPLES DE RESIDUS TESTES .....	83
FIGURE 41 : DYNAMIQUE MOLECULAIRE DE LA TOXINE MT7.....	85
FIGURE 42 : FLUCTUATION DES RESIDUS DE LA TOXINE.....	85
FIGURE 43 : LA BOUCLE E2 COUVRE LA CAVITE EXTRACELLULAIRE DU RECEPTEUR hM1 .....	86
FIGURE 44 : MODIFICATION DE LA BOUCLE E2 DU RECEPTEUR hM2 .....	87
FIGURE 45 : FLEXIBILITE DES BOUCLES EXTRACELLULAIRES E1 ET E3 DU hM1.....	90
FIGURE 46 : RESIDUS CHARNIERES POUR LE MOUVEMENT DE LA BOUCLE E2 .....	90
FIGURE 47 : DEFORMATION DE LA BOUCLE E2 AU COURS DE LA DYNAMIQUE .....	91
FIGURE 48 : DYNAMIQUE ACTIVEE DE LA BOUCLE E2 DU RECEPTEUR .....	91
FIGURE 49 : ÉCHANTILLONNAGE DES POSITIONS RELATIVES DES PARTENAIRES .....	92
FIGURE 50 : AMBIGUÏTE DES DONNEES DE CYCLE DE DOUBLES MUTANTS .....	93
FIGURE 51 : PROXIMITES RECEPTEUR/TOXINE DETERMINEES PAR LES CYCLES DE DOUBLES MUTANTS.....	97
FIGURE 52 : GROUPES DE RESIDUS EN INTERACTIONS .....	98
FIGURE 53 : INCOHERENCE DANS LES CONTRAINTES .....	99
FIGURE 54 : VARIABILITE DANS LE POSITIONNEMENT DE LA TOXINE PARMIS LES SOLUTIONS .....	102
FIGURE 55 : POSITIONNEMENT DE LA TOXINE SUR LE RECEPTEUR .....	103
FIGURE 56 : GELS NATIFS.....	106
FIGURE 57 : EXPERIENCE DE TRANSFERT DE FLUORESCENCE.....	107
FIGURE 58 : EXPERIENCE D'ANISOTROPIE DE FLUORESCENCE.....	109
FIGURE 59 : ÉCHANTILLONNAGE DES POSITIONS DES PARTENAIRES .....	111
FIGURE 60 : ÉTAPES DE CONSTRUCTION D'UN MODELE TRIMERIQUE.....	111
FIGURE 61 : ENSEMBLES DES SOLUTIONS SATISFAISANT LES CONTRAINTES DE DISTANCE .....	112
FIGURE 62 : MODELE TRIMERIQUE.....	112
FIGURE 63 : DIFFERENTS TYPES DE LIPIDES PHOSPHATIDYLCHOLINES.....	117
FIGURE 64 : MODELE DE DIMERE DE RHODOPSINE DANS UNE BICOUCHE LIPIDIQUE.....	120
FIGURE 65 : CONSTRUCTION D'UNE MEMBRANE PAR ASSEMBLAGE DE PSEUDO LIPIDES .....	126
FIGURE 66 : CONSTRUCTION DE LA MEMBRANE PAR EXPANSION CONTRACTION .....	128
FIGURE 67 : COMPLEXE MT7/ DIMERE DE hM1 DANS UNE MEMBRANE DE POPC .....	129
FIGURE 68: FLUCTUATION DE STRUCTURE DU DIMERE DE RECEPTEUR ET DE LA TOXINE .....	130
FIGURE 69 : REPOSITIONNEMENT DE LA TOXINE SUR LE DIMERE DE RECEPTEURS. ....	131
FIGURE 70 : FLUCTUATION DES DISTANCES DES PAIRES DE RESIDUS.....	132
FIGURE 71 : ÉVOLUTION DE LA DISTANCE Y179/W10 AU COURS DU TEMPS .....	132
FIGURE 72 MODIFICATION DE L'ANGLE CHI1 DU RESIDU TYR179.....	134
FIGURE 73: MODIFICATION DU RESIDU E397 .....	135

FIGURE 74 : COMPARAISON DES VALEURS DE $G$ EXPERIMENTAUX ET CALCULES .....	136
FIGURE 75 : ORGANISATION DES 14 HELICES DU MODELE DIMERIQUE DE HM1 .....	137
FIGURE 76: AGRANDISSEMENT DE L'INTERFACE DE LIAISON DE LA TOXINE SUR LE RECEPTEUR HM1A....	139
FIGURE 77 : ENFOUISSEMENT DES RESIDUS DES RECEPTEURS.....	140
FIGURE 78 : ENFOUISSEMENT DES RESIDUS DE LA TOXINE .....	141
FIGURE 79 : AGRANDISSEMENT DE L'INTERFACE DE LIAISON DE LA TOXINE SUR LE RECEPTEUR HM1B ...	142
FIGURE 80: OUVERTURE DE LA BOUCLE E2 .....	146
FIGURE 81 : PROXIMITE $W91_{(M1)}/Y30_{(TOX)}$ .....	147
FIGURE 82 : ÉVOLUTION DE LA DISTANCE ENTRE LES RESIDUS W91 ET Y30 .....	148
FIGURE 83 : PROXIMITE W400 (RECEPTEUR) R52 (TOXINE) .....	149
FIGURE 84: COMPARAISON DES INTERFACES DES MODELES DE DIMERE.....	151
FIGURE 85: SYMETRIES D'ASSEMBLAGES DES HOMODIMERES .....	152
FIGURE 86: OLIGOMERE PUTATIF DE M1 .....	155
FIGURE 87 : BLOCAGE DE L'ACCES DU SITE ORTHOSTERIQUE PAR LA TOXINE .....	157
FIGURE 88 : POSSIBILITE DE FORMATION D'OLIGOMERES D'ORDRE SUPERIEUR .....	158
FIGURE 89: COMPOSITION DES INTERFACES PROTEINE/PROTEINE EN FONCTION DE LA SURFACE D'INTERACTION .....	159
FIGURE 90 : LIGAND DIVALENT DU RECEPTEUR HM1 .....	160
FIGURE 91: RECEPTEURS CHIMERIQUES.....	161
FIGURE 92: ALIGNEMENT DE LA BOUCLE E2 DE HM1 ET LA TOXINE MT7 .....	162
FIGURE 93 : SUPERPOSITION DES LIGANDS ORTHOSTERIQUES DANS LES STRUCTURES DE RCPG.....	167



---

## Chapitre 8: Bibliographie

---

- Abrams, P., K. E. Andersson, et al. (2006). "Muscarinic receptors: their distribution and function in body systems, and the implications for treating overactive bladder." Br J Pharmacol.
- Allman, K., K. M. Page, et al. (2000). "Scanning mutagenesis identifies amino acid side chains in transmembrane domain 5 of the M(1) muscarinic receptor that participate in binding the acetyl methyl group of acetylcholine." Mol Pharmacol **58**(1): 175-84.
- Avlani, V. A., K. J. Gregory, et al. (2007). "Critical role for the second extracellular loop in the binding of both orthosteric and allosteric G protein-coupled receptor ligands." J Biol Chem **282**(35): 25677-86.
- Baker, N. A., D. Sept, et al. (2001). "Electrostatics of nanosystems: application to microtubules and the ribosome." Proc Natl Acad Sci U S A **98**(18): 10037-41.
- Baldwin, J. (1993). "The probable arrangement of the helices in G protein-coupled receptors." EMBO J. **12**(4): 1693-1703.
- Baldwin, J. M., G. F. Schertler, et al. (1997). "An alpha-carbon template for the transmembrane helices in the rhodopsin family of G-protein-coupled receptors." J Mol Biol **272**(1): 144-64.
- Barth, P., J. Schonbrun, et al. (2007). "Toward high-resolution prediction and design of transmembrane helical protein structures." Proc Natl Acad Sci U S A **104**(40): 15682-7.
- Bastard, K., C. Prevost, et al. (2006). "Accounting for loop flexibility during protein-protein docking." Proteins **62**(4): 956-69.
- Bee, M. S. and E. C. Hulme (2007). "Functional analysis of transmembrane domain 2 of the M1 muscarinic acetylcholine receptor." J Biol Chem **282**(44): 32471-9.
- Berman, H. M., T. Battistuz, et al. (2002). "The Protein Data Bank." Acta Crystallogr D Biol Crystallogr **58**(Pt 6 No 1): 899-907.
- Berthouze, M., L. Rivail, et al. (2007). "Two transmembrane Cys residues are involved in 5-HT4 receptor dimerization." Biochem Biophys Res Commun **356**(3): 642-7.
- Bissantz, C., P. Bernard, et al. (2003). "Protein-based virtual screening of chemical databases. II. Are homology models of G-Protein Coupled Receptors suitable targets?" Proteins **50**(1): 5-25.



- Bockaert J, P. J. (1999). "Molecular tinkering of G protein-coupled receptors: an evolutionary success." EMBO J **18**: 1723–1729.
- Bonvin, A. M. (2006). "Flexible protein-protein docking." Curr Opin Struct Biol **16**(2): 194-200.
- Bourne, Y., T. T. Talley, et al. (2005). "Crystal structure of a Cbtx-AChBP complex reveals essential interactions between snake alpha-neurotoxins and nicotinic receptors." Embo J **24**(8): 1512-22.
- Brooijmans, N., K. A. Sharp, et al. (2002). "Stability of macromolecular complexes." Proteins **48**(4): 645-53.
- Burgen, A. S., G. C. Roberts, et al. (1975). "Binding of flexible ligands to macromolecules." Nature **253**(5494): 753-5.
- C. Chothia, A. M. L. (1986). "The relation between the divergence of sequence and structure in proteins." EMBO Rep **5**(4): 823–826.
- Carlsson, J., M. Ander, et al. (2006). "Continuum solvation models in the linear interaction energy method." J Phys Chem B **110**(24): 12034-41.
- Celie, P. H., I. E. Kasheverov, et al. (2005). "Crystal structure of nicotinic acetylcholine receptor homolog AChBP in complex with an alpha-conotoxin PnIA variant." Nat Struct Mol Biol **12**(7): 582-8.
- Chabre, M., R. Cone, et al. (2003). "Biophysics (communication arising): Is rhodopsin dimeric in native retinal rods?" **426**(6962): 30-31.
- Chabre, M. and M. leMaire (2005). "Monomeric G-Protein-Coupled Receptor as a Functional Unit." Biochemistry **44**(27): 9395-9403.
- Chakrabarti, P. and J. Janin (2002). "Dissecting protein-protein recognition sites." Proteins **47**(3): 334-43.
- Chance, M. R., A. Fiser, et al. (2004). "High-throughput computational and experimental techniques in structural genomics." Genome Res **14**(10B): 2145-54.
- Changeux JP, E. S. (2005). "Allosteric mechanisms of signal transduction." Science **308**(5727): 1424 - 1428.
- Cherezov, V., D. M. Rosenbaum, et al. (2007). "High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor." Science **318**(5854): 1258-65.
- Chothia, C. and A. M. Lesk (1986). "The relation between the divergence of sequence and structure in proteins." Embo J **5**(4): 823-6.
- Christopoulos, A. (2002). "Allosteric binding sites on cell-surface receptors: novel targets for drug discovery." Nat Rev Drug Discov **1**(3): 198-210.

- Conn PM, M. J., McMillian M, Stern J, Rogers D, Hamby M, Penna A, Grant E. (1982). "Gonadotropin-releasing hormone stimulation of luteinizing hormone release: A ligand-receptor-effector model." Endocr Rev. **79**(23): 7307-7311.
- Cordomi, A., O. Edholm, et al. (2007). "Effect of different treatments of long-range interactions and sampling conditions in molecular dynamic simulations of rhodopsin embedded in a dipalmitoyl phosphatidylcholine bilayer." J Comput Chem **28**(6): 1017-30.
- Cordomi, A. and J. J. Perez (2007). "Molecular dynamics simulations of rhodopsin in different one-component lipid bilayers." J Phys Chem B **111**(25): 7052-63.
- Costanzi, S. (2008). "On the applicability of GPCR homology models to computer-aided drug discovery: a comparison between in silico and crystal structures of the beta2-adrenergic receptor." J Med Chem **51**(10): 2907-14.
- Crozier, P. S., M. J. Stevens, et al. (2003). "Molecular dynamics simulation of dark-adapted rhodopsin in an explicit membrane bilayer: coupling between local retinal and larger scale conformational change." J Mol Biol **333**(3): 493-514.
- Crozier, P. S., M. J. Stevens, et al. (2007). "How a small change in retinal leads to G-protein activation: initial events suggested by molecular dynamics calculations." Proteins **66**(3): 559-74.
- D. Fotiadis, Y. L., S. Filipek, D.A. Saperstein, A. Engel and K. Palczewski (2003). "Atomic-force microscopy: rhodopsin dimers in native disc membranes." Nature **421**(6919): 127.
- D'Adamo, M. C., Z. Liu, et al. (1998). "Episodic ataxia type-1 mutations in the hKv1.1 cytoplasmic pore region alter the gating properties of the channel." Embo J **17**(5): 1200-7.
- Dauplais, M., A. Lecoq, et al. (1997). "On the convergent evolution of animal toxins. Conservation of a diad of functional residues in potassium channel-blocking toxins with unrelated structures." J Biol Chem **272**(7): 4302-9.
- de Graaf, C., N. Foata, et al. (2008). "Molecular modeling of the second extracellular loop of G-protein coupled receptors and its implication on structure-based virtual screening." Proteins **71**(2): 599-620.
- De Lean, A., J. M. Stadel, et al. (1980). "A ternary complex model explains the agonist-specific binding properties of the adenylate cyclase-coupled beta-adrenergic receptor." J Biol Chem **255**(15): 7108-17.
- de Vries, S. J., A. D. van Dijk, et al. (2006). "WHISCY: what information does surface conservation yield? Application to data-driven docking." Proteins **63**(3): 479-89.
- de Vries, S. J., A. D. van Dijk, et al. (2007). "HADDOCK versus HADDOCK: new features and performance of HADDOCK2.0 on the CAPRI targets." Proteins **69**(4): 726-33.

- de Weille JR, S. H., Maes P, Tartar A, Lazdunski M (1991). "Calciseptine, a peptide isolated from black mamba venom, is a specific blocker of the L-type calcium channel." PNAS **88**(6): 2437-2440.
- Debret, G., H. Valadie, et al. (2008). "New insights of membrane environment effects on MscL channel mechanics from theoretical approaches." Proteins **71**(3): 1183-96.
- Dohlman HG, T. J., Caron MG, Lefkowitz RJ. (1991). "Model systems for the study of seven-transmembrane-segment receptors." Annu Rev Biochem. **60**: 653-88.
- Dominguez, C., R. Boelens, et al. (2003). "HADDOCK: a protein-protein docking approach based on biochemical or biophysical information." J Am Chem Soc **125**(7): 1731-7.
- Edrington, T. C., M. P. Bennett, et al. (2008). "Calorimetric Studies of Bovine Rod Outer Segment Disk Membranes Support a Monomeric Unit for Both Rhodopsin and Opsin." Biophys J **95**: 2859-2866.
- Ehrlich, L. P., M. Nilges, et al. (2005). "The impact of protein flexibility on protein-protein docking." Proteins **58**(1): 126-33.
- Eriksson, M. A. and B. Roux (2002). "Modeling the structure of agitoxin in complex with the Shaker K<sup>+</sup> channel: a computational approach based on experimental distance restraints extracted from thermodynamic mutant cycles." Biophys J **83**(5): 2595-609.
- Espinoza-Fonseca, L. M., A. Pedretti, et al. (2008). "Structure and dynamics of the full-length M1 muscarinic acetylcholine receptor studied by molecular dynamics simulations." Arch Biochem Biophys **469**(1): 142-50.
- Espinoza-Fonseca, L. M. and J. G. Trujillo-Ferrara (2005). "Identification of multiple allosteric sites on the M1 muscarinic acetylcholine receptor." FEBS Lett **579**(30): 6726-32.
- Espinoza-Fonseca, L. M. and J. G. Trujillo-Ferrara (2006). "The existence of a second allosteric site on the M1 muscarinic acetylcholine receptor and its implications for drug design." Bioorg Med Chem Lett **16**(5): 1217-20.
- Ewald, P. P. (1921). "Die Berechnung optischer und elektrostatischer Gitterpotentiale." Ann. Phys **363**(3): 253-287.
- Faraldo-Gomez, J. D., L. R. Forrest, et al. (2004). "Conformational sampling and dynamics of membrane proteins from 10-nanosecond computer simulations." Proteins **57**(4): 783-91.
- Filizola, M., S. X. Wang, et al. (2006). "Dynamic models of G-protein coupled receptor dimers: indications of asymmetry in the rhodopsin dimer from molecular dynamics simulations in a POPC bilayer." J Comput Aided Mol Des **20**(7-8): 405-16.

- Flores, S., N. Echols, et al. (2006). "The Database of Macromolecular Motions: new features added at the decade mark." Nucleic Acids Res **34**(Database issue): D296-301.
- Fotiadis D, J. B., Philippsen A, Müller DJ, Palczewski K, Engel A. (2006). "Structure of the rhodopsin dimer: a working model for G-protein-coupled receptors." Curr Opin Struct Biol. **16**(2): 252-9.
- Fredriksson R, L. M., Lundin LG, Schiöth HB. (2003). "The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints." Mol Pharmacol. **63**: 1256-1272.
- Fruchart-Gaillard, C., B. Gilquin, et al. (2002). "Experimentally based model of a complex between a snake toxin and the alpha 7 nicotinic receptor." Proc Natl Acad Sci U S A **99**(5): 3216-21.
- Fruchart-Gaillard, C., G. Mourier, et al. (2006). "How three-finger-fold toxins interact with various cholinergic receptors." J Mol Neurosci **30**(1-2): 7-8.
- Fruchart-Gaillard, C., G. Mourier, et al. (2006). "Identification of various allosteric interaction sites on M1 muscarinic receptor using 125I-Met35-oxidized muscarinic toxin 7." Mol Pharmacol **69**(5): 1641-51.
- Furnham, N., A. S. Dore, et al. (2006). "Knowledge-based real-space explorations for low-resolution structure determination." Structure **14**(8): 1313-20.
- Gabb, H. A., R. M. Jackson, et al. (1997). "Modelling protein docking using shape complementarity, electrostatics and biochemical information." J Mol Biol **272**(1): 106-20.
- Gabel, F., B. Simon, et al. (2006). "A target function for quaternary structural refinement from small angle scattering and NMR orientational restraints." Eur Biophys J **35**(4): 313-27.
- Galat, A., G. Gross, et al. (2008). "Conserved structural determinants in three-fingered protein domains." Febs J **275**(12): 3207-25.
- Gerstein, M. and N. Echols (2004). "Exploring the range of protein flexibility, from a structural proteomics perspective." Curr Opin Chem Biol **8**(1): 14-9.
- Gerstein, M. and W. Krebs (1998). "A database of macromolecular motions." Nucleic Acids Res **26**(18): 4280-90.
- Gilchrist, A. (2007). "Modulating G-protein-coupled receptors: from traditional pharmacology to allosterics." Trends Pharmacol Sci **28**(8): 431-7.
- Gilman, A. G. (1994). "G proteins and regulation of adenylyl cyclase." Nobel Lecture.
- Gilquin, B., M. Bourgoïn, et al. (2003). "Motions and structural variability within toxins: implication for their use as scaffolds for protein engineering." Protein Sci **12**(2): 266-77.

- Gilquin, B., S. Braud, et al. (2005). "A variable residue in the pore of Kv1 channels is critical for the high affinity of blockers from sea anemones and scorpions." J Biol Chem **280**(29): 27093-102.
- Gilquin, B., C. Guilbert, et al. (2000). "Unfolding of hen egg lysozyme by molecular dynamics simulations at 300K: Insight into the role of the interdomain interface." Proteins: Structure, Function, and Genetics **41**(1): 58-74.
- Gilquin, B., J. Racape, et al. (2002). "Structure of the BgK-Kv1.1 complex based on distance restraints identified by double mutant cycles. Molecular basis for convergent evolution of Kv1 channel blockers." J Biol Chem **277**(40): 37406-13.
- Goin, J. C. and N. M. Nathanson (2006). "Quantitative analysis of muscarinic acetylcholine receptor homo- and heterodimerization in live cells: regulation of receptor down-regulation by heterodimerization." J Biol Chem **281**(9): 5416-25.
- Goodwin, J. A., E. C. Hulme, et al. (2007). "Roof and floor of the muscarinic binding pocket: variations in the binding modes of orthosteric ligands." Mol Pharmacol **72**(6): 1484-96.
- Gray, J. J., S. Moughon, et al. (2003). "Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations." J Mol Biol **331**(1): 281-99.
- Gray, J. J., S. E. Moughon, et al. (2003). "Protein-protein docking predictions for the CAPRI experiment." Proteins **52**(1): 118-22.
- Grossfield, A., S. E. Feller, et al. (2007). "Convergence of molecular dynamics simulations of membrane proteins." Proteins **67**(1): 31-40.
- Grossmuller, M., J. Antony, et al. (2006). "Allosteric site in M2 acetylcholine receptors: evidence for a major conformational change upon binding of an orthosteric agonist instead of an antagonist." Naunyn Schmiedebergs Arch Pharmacol **372**(4): 267-76.
- Grunberg, R., J. Leckner, et al. (2004). "Complementarity of structure ensembles in protein-protein binding." Structure **12**(12): 2125-36.
- Guilbert, C., D. Perahia, et al. (1995). "A method to explore transition paths in macromolecules. Applications to hemoglobin and phosphoglycerate kinase." Comp Phys Comm(91): 263-273.
- Guo, W., L. Shi, et al. (2003). "The fourth transmembrane segment forms the interface of the dopamine D2 receptor homodimer." J Biol Chem **278**(7): 4385-8.
- Gurevich, V. V. and E. V. Gurevich (2008). "GPCR monomers and oligomers: it takes all kinds." Trends Neurosci **31**(2): 74-81.
- Gurevich, V. V. and E. V. Gurevich (2008). "How and why do GPCRs dimerize?" Trends Pharmacol Sci **29**(5): 234-40.

- Han, S. J., F. F. Hamdan, et al. (2005). "Pronounced conformational changes following agonist activation of the M(3) muscarinic acetylcholine receptor." J Biol Chem **280**(26): 24870-9.
- Hansen, S. B., G. Sulzenbacher, et al. (2005). "Structures of Aplysia AChBP complexes with nicotinic agonists and antagonists reveal distinctive binding interfaces and conformations." Embo J **24**(20): 3635-46.
- Harvey, A. L. and E. Karlsson (1982). "Protease inhibitor homologues from mamba venoms: facilitation of acetylcholine release and interactions with prejunctional blocking toxins." Br J Pharmacol.
- Hebert, T., S. Moffett, et al. (1996). "A Peptide Derived from a beta 2-Adrenergic Receptor Transmembrane Domain Inhibits Both Receptor Dimerization and Activation." J. Biol. Chem. **271**(27): 16384-16392.
- Holzgrabe, U., M. De Amici, et al. (2006). "Allosteric modulators and selective agonists of muscarinic receptors." J Mol Neurosci **30**(1-2): 165-8.
- Huai, Q., A. P. Mazar, et al. (2006). "Structure of human urokinase plasminogen activator in complex with its receptor." Science **311**(5761): 656-9.
- Huang, X. P., S. Prilla, et al. (2005). "Critical amino acid residues of the common allosteric site on the M2 muscarinic acetylcholine receptor: more similarities than differences between the structurally divergent agents gallamine and bis(ammonio)alkane-type hexamethylene-bis-[dimethyl-(3-phthalimidopropyl)ammonium]dibromide." Mol Pharmacol **68**(3): 769-78.
- Huang, Y., J. Cai, et al. (2004). "Classifying G-protein coupled receptors with bagging classification tree." Comput Biol Chem. **28**(4): 275-280.
- Huber, T., A. V. Botelho, et al. (2004). "Membrane model for the G-protein-coupled receptor rhodopsin: hydrophobic interface and dynamical structure." Biophys J **86**(4): 2078-100.
- Hulme, E. C. and Z. L. Lu (1998). "Scanning mutagenesis of transmembrane domain 3 of the M1 muscarinic acetylcholine receptor." J Physiol Paris **92**(3-4): 269-74.
- Hulme, E. C., Z. L. Lu, et al. (2003). "Scanning mutagenesis studies of the M1 muscarinic acetylcholine receptor." Receptors Channels **9**(4): 215-28.
- Hulme, E. C., Z. L. Lu, et al. (1999). "The conformational switch in 7-transmembrane receptors: the muscarinic receptor paradigm." Eur J Pharmacol **375**(1-3): 247-60.
- Humphrey, W., A. Dalke, et al. (1996). "VMD: visual molecular dynamics." J Mol Graph **14**(1): 33-8, 27-8.
- Iwata, N., M. Higuchi, et al. (2005). "Metabolism of amyloid-beta peptide and Alzheimer's disease." Pharmacol Ther. **100**: 51.

- Jackson, R. M., H. A. Gabb, et al. (1998). "Rapid refinement of protein interfaces incorporating solvation: application to the docking problem." J Mol Biol **276**(1): 265-85.
- Jager, D., C. Schmalenbach, et al. (2007). "Allosteric small molecules unveil a role of an extracellular E2/transmembrane helix 7 junction for G protein-coupled receptor activation." J Biol Chem **282**(48): 34968-76.
- Janin, J. (2005). "Assessing predictions of protein-protein interaction: the CAPRI experiment." Protein Sci **14**(2): 278-83.
- Janin, J. (2005). "Sailing the route from Gaeta, Italy, to CAPRI." Proteins **60**(2): 149.
- Janin, J. (2005). "The targets of CAPRI rounds 3-5." Proteins **60**(2): 170-5.
- Janin, J. (2007). "The targets of CAPRI rounds 6-12." Proteins **69**(4): 699-703.
- Janin, J. and C. Chothia (1990). "The structure of protein-protein recognition sites." J Biol Chem **265**(27): 16027-30.
- Janin, J., K. Henrick, et al. (2003). "CAPRI: a Critical Assessment of PRedicted Interactions." Proteins **52**(1): 2-9.
- Janin, J. and B. Seraphin (2003). "Genome-wide studies of protein-protein interaction." Curr Opin Struct Biol **13**(3): 383-8.
- Janin, J. and S. Wodak (2007). "The third CAPRI assessment meeting Toronto, Canada." Structure **15**(7): 755-9.
- Jo, S., M. Vargyas, et al. (2008). "PBEQ-Solver for online visualization of electrostatic potential of biomolecules." Nucleic Acids Res **36**(Web Server issue): W270-5.
- Kanamori, E., Y. Murakami, et al. (2007). "Docking of protein molecular surfaces with evolutionary trace analysis." Proteins **69**(4): 832-8.
- Kandt, C., W. L. Ash, et al. (2007). "Setting up and running molecular dynamics simulations of membrane proteins." Methods **41**(4): 475-88.
- Karlsson, E., M. Jolkkonen, et al. (2000). "Snake toxins with high selectivity for subtypes of muscarinic acetylcholine receptors." Biochimie **82**(9-10): 793-806.
- Karlsson, E., P. M. Mbugua, et al. (1985). "Anticholinesterase toxins.." Pharmacol Ther.
- Karlsson, E., C. Risinger, et al. (1991). "Amino acid sequence of a snake venom toxin that binds to the muscarinic acetylcholine receptor." Toxicon **29**(4-5): 521-6.
- Kashuba, V. I., S. M. Kvasha, et al. (2001). "Initial isolation and analysis of the human Kv1.7 (KCNA7) gene, a member of the voltage-gated potassium channel gene family." Gene **268**(1-2): 115-22.

- Kim, J. Y. and P. N. Devreotes (1994). "Random chimeragenesis of G-protein-coupled receptors. Mapping the affinity of the cAMP chemoattractant receptors in Dictyostelium." J Biol Chem. **269**(46): 28724-28731.
- Klco, J. M., T. B. Lassere, et al. (2003). "C5a receptor oligomerization. I. Disulfide trapping reveals oligomers and potential contact surfaces in a G protein-coupled receptor." J Biol Chem **278**(37): 35345-53.
- Klco, J. M., C. B. Wiegand, et al. (2005). "Essential role for the second extracellular loop in C5a receptor activation." Nat Struct Mol Biol **12**(4): 320-6.
- Kobilka, B. and G. F. Schertler (2008). "New G-protein-coupled receptor crystal structures: insights and limitations." Trends Pharmacol Sci. **29**(2): 79-83.
- Kollman, P. A., I. Massova, et al. (2000). "Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models." Acc Chem Res **33**(12): 889-97.
- Kong, Y. and M. Karplus (2007). "The signaling pathway of rhodopsin." Structure **15**(5): 611-23.
- Kong, Y. and M. Karplus (2008). "Signaling pathways of PDZ2 domain: A molecular dynamics interaction correlation analysis." Proteins.
- Kroeger, K. M., A. C. Hanyaloglu, et al. (2001). "Constitutive and agonist-dependent homo-oligomerization of the thyrotropin-releasing hormone receptor. Detection in living cells using bioluminescence resonance energy transfer." J Biol Chem **276**(16): 12736-43.
- Kuhlman, B. and D. Baker (2000). "Native protein sequences are close to optimal for their structures." Proc Natl Acad Sci U S A **97**(19): 10383-8.
- Kukkonen, A., M. Perakyla, et al. (2004). "Muscarinic Toxin 7 Selectivity Is Dictated by Extracellular Receptor Loops." J. Biol. Chem. **279**(49): 50923-50929.
- Lacapere, J. J., E. Pebay-Peyroula, et al. (2007). "Determining membrane protein structures: still a challenge!" Trends Biochem Sci **32**(6): 259-70.
- Lange, A., K. Giller, et al. (2006). "Toxin-induced conformational changes in a potassium channel revealed by solid-state NMR." Nature **440**(7086): 959-62.
- Langmead, C. J., J. Watson, et al. (2008). "Muscarinic acetylcholine receptors as CNS drug targets." Pharmacol Ther. **117**(2): 232-43.
- le Du, M. H., P. Marchot, et al. (1992). "1.9-A resolution structure of fasciculin 1, an anti-acetylcholinesterase toxin from green mamba snake venom." J Biol Chem **267**(31): 22122-30.
- Lefkowitz (1978). "Identification and regulation of alpha- and beta-adrenergic receptors." Fed Proc. **12**(1): 1-13.



- Lefkowitz (2007). "Seven transmembrane receptors: something old, something new." Acta Physiologica **190**: 9-19.
- Lemaitre, V., P. Yeagle, et al. (2005). "Molecular dynamics simulations of retinal in rhodopsin: from the dark-adapted state towards lumirhodopsin." Biochemistry **44**(38): 12667-80.
- Levoye, A., J. Dam, et al. (2006). "Do orphan G-protein-coupled receptors have ligand-independent functions? New insights from receptor heterodimers." EMBO Rep **7**(11): 1094-1098.
- Levy, E. D., E. Boeri Erba, et al. (2008). "Assembly reflects evolution of protein complexes." Nature **453**(7199): 1262-5.
- Li, B., N. M. Nowak, et al. (2005). "Random mutagenesis of the M3 muscarinic acetylcholine receptor expressed in yeast: identification of second-site mutations that restore function to a coupling-deficient mutant M3 receptor." J Biol Chem **280**(7): 5664-75.
- Li, J. H., F. F. Hamdan, et al. (2008). "Ligand-specific changes in M3 muscarinic acetylcholine receptor structure detected by a disulfide scanning strategy." Biochemistry **47**(9): 2776-88.
- Li, J. H., S. J. Han, et al. (2007). "Distinct structural changes in a G protein-coupled receptor caused by different classes of agonist ligands." J Biol Chem **282**(36): 26284-93.
- Li, Y. and W. A. Goddard (2008). "Prediction of structure of G-protein coupled receptors and of bound ligands, with applications for drug design." Pac Symp Biocomput: 344-53.
- Liang, Y., D. Fotiadis, et al. (2003). "Organization of the G protein-coupled receptors rhodopsin and opsin in native membranes." J Biol Chem. **278**(24): 21655-21662.
- Llinas, P., M. H. Le Du, et al. (2005). "Crystal structure of the human urokinase plasminogen activator receptor bound to an antagonist peptide." Embo J **24**(9): 1655-63.
- Lo Conte, L., C. Chothia, et al. (1999). "The atomic structure of protein-protein recognition sites." J Mol Biol **285**(5): 2177-98.
- Lopes, A., A. Alexandrov, et al. (2007). "Computational sidechain placement and protein mutagenesis with implicit solvent models." Proteins **67**(4): 853-67.
- Lopez-Gimenez, J., M. Canals, et al. (2007). "The  $\alpha$ 1b-Adrenoceptor Exists as a Higher-Order Oligomer: Effective Oligomerization Is Required for Receptor Maturation, Surface Delivery, and Function 10.1124/mol.106.033035." Mol Pharmacol **71**(4): 1015-1029.
- Lovell, S. C., I. W. Davis, et al. (2003). "Structure validation by Calpha geometry: phi,psi and Cbeta deviation." Proteins **50**(3): 437-50.

- Lu, Z. L. and E. C. Hulme (1999). "The functional topography of transmembrane domain 3 of the M1 muscarinic acetylcholine receptor, revealed by scanning mutagenesis." *J Biol Chem* **274**(11): 7309-15.
- Lu, Z. L. and E. C. Hulme (2000). "A network of conserved intramolecular contacts defines the off-state of the transmembrane switch mechanism in a seven-transmembrane receptor." *J Biol Chem* **275**(8): 5682-6.
- Lu, Z. L., J. W. Saldanha, et al. (2001). "Transmembrane domains 4 and 7 of the M(1) muscarinic acetylcholine receptor are critical for ligand binding and the receptor activation switch." *J Biol Chem* **276**(36): 34098-104.
- Lyskov, S. and J. J. Gray (2008). "The RosettaDock server for local protein-protein docking." *Nucleic Acids Res* **36**(Web Server issue): W233-8.
- Madaoui, H. and R. Guerois (2008). "Coevolution at protein complex interfaces can be detected by the complementarity trace with important impact for predictive docking." *Proc Natl Acad Sci U S A* **105**(22): 7708-13.
- Mandell, J. G., V. A. Roberts, et al. (2001). "Protein docking using continuum electrostatics and geometric fit." *Protein Eng* **14**(2): 105-13.
- Marquer, C. (2008). "Caractérisation pharmacologique et structurale de l'interaction de la toxine muscarinique MT7 sur le récepteur muscarinique de l'acétylcholine hM1." *Thèse de doctorat Université Paris VI*.
- May, L. T. and A. Christopoulos (2003). "Allosteric modulators of G-protein-coupled receptors." *Curr Opin Pharmacol* **3**(5): 551-6.
- Milligan, G. (2007). "G-protein-coupled receptors: an update." *Acta Physiologica*.
- Milligan, G. (2004). "G protein-coupled receptor dimerization: function and ligand pharmacology." *Mol Pharmacol*. **66**: 1-7.
- Milligan, G. and N. J. Smith (2007). "Allosteric modulation of heterodimeric G-protein-coupled receptors." *Trends Pharmacol Sci* **28**(12): 615-20.
- Mohr, K., C. Trankle, et al. (2003). "Structure/activity relationships of M2 muscarinic allosteric modulators." *Receptors Channels* **9**(4): 229-40.
- Monod, J., J. Wyman, et al. (1965). "On the nature of allosteric transitions : a plausible model." *J Mol Biol*. **12**: 88-118.
- Moreira, I. S., P. A. Fernandes, et al. (2007). "Computational alanine scanning mutagenesis--an improved methodological approach." *J Comput Chem* **28**(3): 644-54.
- Moreira, I. S., P. A. Fernandes, et al. (2007). "Hot spots--a review of the protein-protein interface determinant amino-acid residues." *Proteins* **68**(4): 803-12.

- Mourier, G., S. Dutertre, et al. (2003). "Chemical synthesis of MT1 and MT7 muscarinic toxins: critical role of Arg-34 in their interaction with M1 muscarinic receptor." Mol Pharmacol **63**(1): 26-35.
- Okada, T., O. P. Ernst, et al. (2001). "Activation of rhodopsin: new insights from structural and biochemical studies." Trends Biochem Sci. **26**(5): 318-324.
- Okada, T., Y. Fujiyoshi, et al. (2002). "Functional role of internal water molecules in rhodopsin revealed by X-ray crystallography." Proc Natl Acad Sci U S A **99**(9): 5982-7.
- Okada, T. and K. Palczewski (2001). "Crystal structure of rhodopsin: implications for vision and beyond." Curr Opin Struct Biol. **11**(4): 420-426.
- Okada, T., M. Sugihara, et al. (2004). "The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure." J Mol Biol. **342**(2): 571-83.
- Olianas, M. C., A. Adem, et al. (2004). "Action of the muscarinic toxin MT7 on agonist-bound muscarinic M1 receptors." Eur J Pharmacol **487**(1-3): 65-72.
- Olianas, M. C., C. Maullu, et al. (2000). "Inhibition of acetylcholine muscarinic M(1) receptor function by the M(1)-selective ligand muscarinic toxin 7 (MT-7)." Br J Pharmacol **131**(3): 447-52.
- Oliveira, L., T. Hulsen, et al. (2004). "Heavier-than-air flying machines are impossible." FEBS Lett **564**(3): 269-73.
- Palczewski, K., T. Kumasaka, et al. (2000). "Crystal structure of rhodopsin: A G protein-coupled receptor." Science. **130**(6): 865
- Park, J. H., P. Scheerer, et al. (2008). "Crystal structure of the ligand-free G-protein-coupled receptor opsin." Nature **454**(7201): 183-7.
- Patel, R. C., D. C. Lange, et al. (2002). "Photobleaching fluorescence resonance energy transfer reveals ligand-induced oligomer formation of human somatostatin receptor subtypes." Methods **27**(4): 340-8.
- Pebay-Peyroula, E., G. Rummel, et al. (1997). "X-ray structure of bacteriorhodopsin at 2.5 angstroms from microcrystals grown in lipidic cubic phases." Science **277**(5332): 1676-81.
- Periole, X., T. Huber, et al. (2007). "G protein-coupled receptors self-assemble in dynamics simulations of model bilayers." J Am Chem Soc **129**(33): 10126-32.
- Phillips, J. C., R. Braun, et al. (2005). "Scalable molecular dynamics with NAMD." J Comput Chem **26**(16): 1781-802.
- Plakidou-Dymock, S., D. Dymock, et al. (1998). "A higher plant seven-transmembrane receptor that influences sensitivity to cytokinins." Curr Biol. **11**(7): 535.
- Popot, J. L. and J. P. Changeux (1984). "Nicotinic receptor of acetylcholine: structure of an oligomeric integral membrane protein." Physiol Rev **64**(4): 1162-239.

- Rajeswaran, W. G., Y. Cao, et al. (2001). "Design, Synthesis, and Biological Characterization of Bivalent 1-Methyl-1,2,5,6-tetrahydropyridyl-1,2,5-thiadiazole Derivatives as Selective Muscarinic Agonists." J. Med. Chem. **44**(26): 4563-4576.
- Ramachandran, G. N., C. Ramakrishnan, et al. (1963). "Stereochemistry of polypeptide chain configurations." J Mol Biol **7**: 95-9.
- Raman, P., V. Cherezov, et al. (2006). "The Membrane Protein Data Bank." Cell Mol Life Sci **63**(1): 36-51.
- Rasmussen, S., H. Choi, et al. (2007). "Crystal structure of the human beta2 adrenergic G-protein-coupled receptor." Nature **450**: 383-387.
- Rigaut, G., A. Shevchenko, et al. (1999). "A generic protein purification method for protein complex characterization and proteome exploration." Nat Biotechnol **17**(10): 1030-2.
- Rosenbaum, D. M., V. Cherezov, et al. (2007). "GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function." Science **318**(5854): 1266-73.
- Sali, A., L. Potterton, et al. (1995). "Evaluation of comparative protein modeling by MODELLER." Proteins **23**(3): 318-26.
- Salom, D., D. T. Lodowski, et al. (2006). "Crystal structure of a photoactivated deprotonated intermediate of rhodopsin." Proc Natl Acad Sci U S A **103**(44): 16123-8.
- Scarselli, M., B. Li, et al. (2007). "Multiple residues in the second extracellular loop are critical for M3 muscarinic acetylcholine receptor activation." J Biol Chem **282**(10): 7385-96.
- Schertler, G. F. (2005). "Structure of rhodopsin and the metarhodopsin I photointermediate." Curr Opin Struct Biol **15**(4): 408-15.
- Schmidt, C., B. Li, et al. (2003). "Random mutagenesis of the M3 muscarinic acetylcholine receptor expressed in yeast. Identification of point mutations that "silence" a constitutively active mutant M3 receptor and greatly impair receptor/G protein coupling." J Biol Chem **278**(32): 30248-60.
- Schneider, R., C. Ader, et al. (2008). "Solid-state NMR spectroscopy applied to a chimeric potassium channel in lipid bilayers." J Am Chem Soc **130**(23): 7427-35.
- Schreiber, G. and A. R. Fersht (1995). "Energetics of protein-protein interactions: Analysis of the Barnase-Barstar interface by single mutations and double mutant cycles." Journal of Molecular Biology **248**(2): 478-486.
- Schwartz, T. W., T. M. Frimurer, et al. (2006). "Molecular mechanism of 7TM receptor activation--a global toggle switch model." Annu Rev Pharmacol Toxicol **46**: 481-519.

- Segalas, I., C. Roumestand, et al. (1995). "Solution structure of a green mamba toxin that activates muscarinic acetylcholine receptors, as studied by nuclear magnetic resonance and molecular modeling." Biochemistry **34**(4): 1248-60.
- Shi, L. and J. A. Javitch (2004). "The second extracellular loop of the dopamine D2 receptor lines the binding-site crevice." Proc Natl Acad Sci U S A **101**(2): 440-5.
- Standfuss, J., G. Xie, et al. (2007). "Crystal structure of a thermally stable rhodopsin mutant." J Mol Biol **372**(5): 1179-88.
- Sun, Y., D. McGarrigle, et al. (2007). "When a G protein-coupled receptor does not couple to a G protein." Mol Biosyst. **3**: 849 - 854.
- Takeda, S., S. Kadowaki, et al. (2002). "Identification of G protein-coupled receptor genes from the human genome sequence." FEBS Lett **520**(1-3): 97-101.
- Taly, A., P. J. Corringer, et al. (2006). "Implications of the quaternary twist allosteric model for the physiology and pathology of nicotinic acetylcholine receptors." Proc Natl Acad Sci U S A **103**(45): 16965-70.
- Taly, A., M. Delarue, et al. (2005). "Normal mode analysis suggests a quaternary twist model for the nicotinic receptor gating mechanism." Biophys J **88**(6): 3954-65.
- Terrillon, S. and M. Bouvier (2004). "Roles of G-protein-coupled receptor dimerization." EMBO Rep. **5**(1):30-4
- Thorn, K. and A. Bogan (2001). "ASEdb : a database of alanine mutations and their effects on the free energy of binding in protein interactions." Bioinformatics **17**(3): 284-285.
- Tjandra, N. and A. Bax (1997). "Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium." Science **278**(5340): 1111-4.
- Trankle, C., A. Dittmann, et al. (2005). "Atypical muscarinic allosteric modulation: cooperativity between modulators and their atypical binding topology in muscarinic M2 and M2/M5 chimeric receptors." Mol Pharmacol **68**(6): 1597-610.
- Trankle, C., O. Weyand, et al. (2003). "Interactions of orthosteric and allosteric ligands with [3H]dimethyl-W84 at the common allosteric site of muscarinic M2 receptors." Mol Pharmacol **64**(1): 180-90.
- Vaidehi, N., W. B. Floriano, et al. (2002). "Prediction of structure and function of G protein-coupled receptors." Proc Natl Acad Sci U S A **99**(20): 12622-7.
- van Dijk, A. D. and A. M. Bonvin (2006). "Solvated docking: introducing water into the modelling of biomolecular complexes." Bioinformatics **22**(19): 2340-7.
- van Schaik, R. C., H. J. Berendsen, et al. (1993). "A structure refinement method based on molecular dynamics in four spatial dimensions." J Mol Biol **234**(3): 751-62.

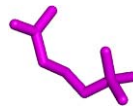
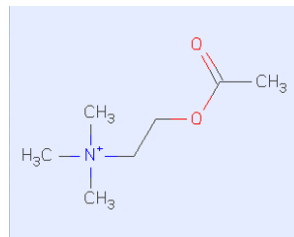
- Vaughan, C. K., A. M. Buckle, et al. (1999). "Structural response to mutation at a protein-protein interface." J Mol Biol **286**(5): 1487-506.
- Venter, J. C., M. D. Adams, et al. (2001). "The sequence of the human genome." Science **291**(5507): 1304-51.
- Vidi, P.-A., B. R. Chemel, et al. (2008). "Ligand-Dependant Oligomerization of Dopamine D2 and Adenosine A2A Receptors in Living Neuronal Cells" Mol Pharmacol: **74**(3):544-51
- Volpicelli, L. and A. Levey (2004). "Muscarinic acetylcholine receptor subtypes in cerebral cortex and hippocampus." Prog Brain Res **145**: 59-66.
- Wang, C., P. Bradley, et al. (2007). "Protein-protein docking with backbone flexibility." J Mol Biol **373**(2): 503-19.
- Ward, S. D., C. A. Curtis, et al. (1999). "Alanine-scanning mutagenesis of transmembrane domain 6 of the M(1) muscarinic acetylcholine receptor suggests that Tyr381 plays key roles in receptor function." Mol Pharmacol **56**(5): 1031-41.
- Weill, C., J. L. Galzi, et al. (1999). "Functional characterization and potential applications for enhanced green fluorescent protein- and epitope-fused human M1 muscarinic receptors." J Neurochem **73**(2): 791-801.
- Wess, J., A. Duttaroy, et al. (2003). "M1-M5 muscarinic receptor knockout mice as novel tools to study the physiological roles of the muscarinic cholinergic system." Receptors Channels **9**(4): 279-90.
- Wess, J., R. M. Eglén, et al. (2007). "Muscarinic acetylcholine receptors: mutant mice provide new insights for drug development." Nat Rev Drug Discov **6**(9): 721-33.
- Woolf, T. B. and B. Roux (1994). "Molecular dynamics simulation of the gramicidin channel in a phospholipid bilayer." Proc Natl Acad Sci U S A **91**(24): 11631-5.
- Yair Ben-Chaim, B. C., Nathan Dascal, Francisco Bezanilla, Itzhak Parnas & Hanna Parnas (2006). "Movement of 'gating charge' is coupled to ligand binding in a G-protein-coupled receptor." Nature **444**: 106-109.
- York, D. M., T. A. Darden, et al. (1993). "Molecular dynamics simulation of HIV-1 protease in a crystalline environment and in solution." Biochemistry **32**(6): 1443-53.
- Zhang, C., S. Liu, et al. (2005). "Docking prediction using biological information, ZDOCK sampling technique, and clustering guided by the DFIRE statistical energy function." Proteins **60**(2): 314-8.

---

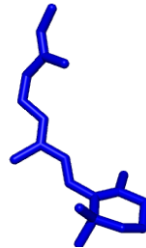
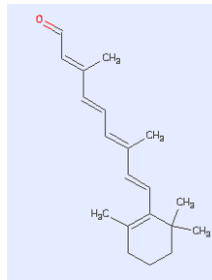
# Chapitre 9: Annexes

---

9.1.1. ligands orthostériques des RCPG

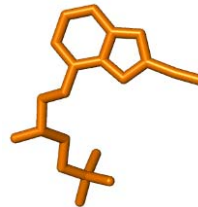
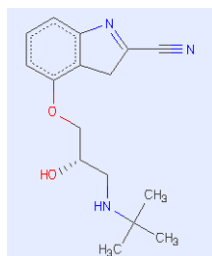


(A) Acétylcholine



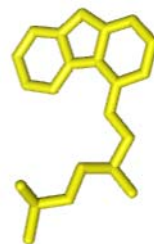
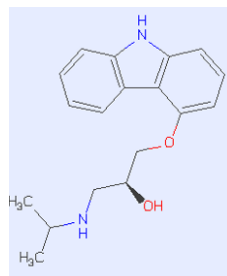
(B) Rétinal

PDB : 1U19



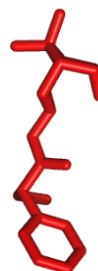
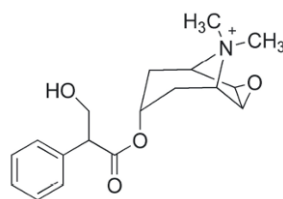
(C) Cyanopindolol

PDB : 2VT4



(B) Carazolol

PDB : 2R1H



(C) N-méthyl-scopolamine





### 9.1.3. Simulation d'arrimage

Les simulations d'arrimage sont réalisées avec le programme Charmm, version 32b1.

Ci-dessous, un fichier type de simulation d'arrimage :

```
* docking sous contrainte ambigues
*

!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!
!!! liste des variables
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!

!! maxtry      :   nombre maximal de structure de depart testee avant
plantage
!! distmax    :   rmax dans NOE
!! distmin    :   rmin dans NOE
!! radius     :   diametre de la sphere dans laquelle la toxine est
translatee
!! seed       :   seed ...
!! consfile   :   fichier de contrainte
!! prefix     :   prefix des fichiers de sortie
!! temp       :   temperature de la dyna pendant docking
!! repeat     :   nombre de repetition avec chaque couple tox/hM1
!! complex    :   fichier pdb du complexe de depart

!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!
!!! chargement des parametres
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!

read rtf card name $TOPPAR/toph19.inp
read para card name $TOPPAR/param19.inp
read psf card name @wd/hm1-tox.psf
read coor pdb name @wd/@complex.pdb resi

coor copy comp
random uniform iseed @seed

!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!
!!!definitions
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!

define backbone sele .bygroup. (TYPE ca .OR. TYPE n .OR. TYPE c ) end
define sidechain sele .NOT. backbone end

define rec sele segid A1 .or. segid A2 end
define lig sele segid B end

define loop sele rec .and. resid 169:177  end

!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!
!!! recherche position de depart
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!
```

```

set try = 0
set diam
calc diam = @radius * 2
label tryagain
incr try by 1

coor copy main ! recharge les coor de depart

!!!!!!!!!!!! translation

calc xtrans = ?rand * @diam - @radius
calc ztrans = ?rand * @diam - @radius
coor trans xdir @xtrans ydir 0 zdir @ztrans sele lig end

!!!!!!!!!!!! rotation

calc yrot = @seed
coor stat sele lig end
coor rotate xdir 0 ydir 1 zdir 0 PHI @yrot xcen ?xave ycen ?yave zcen
?zave sele lig end

!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!
!!! debut du docking
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!

!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!! contraintes !!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!

shake bonh param tol 1.0e-6
cons fix sele rec .and. ( backbone .or. type CB ) .and. .not. loop end
cons harm clear          !!! contrainte harmonique distances
cons harm bestfit force 100.0 comp sele lig .and. ( backbone .or. type
CB ) end
set distmax 4.0 ! distance max pour atomes des residus sous
contrainte
set distmin 0.0

!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!! parametrage 4D !!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!

skip elec
fourd k4dinitial 250.0 skbond skangle skdihedral skconstraints
cons fix4 sele all end !!! cons => fixe les atomes dans la 4D
dimension
scalar fdim set 0.0 sele rec end !!! fixe hM1 dans la 4eme dim
scalar fdim set 1.0 sele lig end

!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!! docking !!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
SET fo = 1.0
STREAM @wd/@consfile
mini abnr nstep 250
DYNAMICS START NSTEP 2000 TIMESTEP 0.001 NPRINT 100 iprfrq 1000
IASVEL 1 FIRSTT @temp FINALT @temp TSTRUCT @temp ieqfrq 10 ISEED
@seed

SET fo = 2.0
STREAM @wd/@consfile
mini abnr nstep 250
DYNAMICS START NSTEP 2000 TIMESTEP 0.001 NPRINT 100 iprfrq 1000
IASVEL 1 FIRSTT @temp FINALT @temp TSTRUCT @temp ieqfrq 10 ISEED
@seed

SET fo = 4.0

```

```
STREAM @wd/@consfile
mini abnr nstep 250
DYNAMICS START NSTEP 10000 TIMESTEP 0.001 NPRINT 100 iprfrq 1000
IASVEL 1 FIRSTT @temp FINALT @temp TSTRUCT @temp ieqfrq 10 ISEED
@seed
mini abnr nstep 250
DYNAMICS START NSTEP 10000 TIMESTEP 0.001 NPRINT 100 iprfrq 1000
IASVEL 1 FIRSTT @temp FINALT @temp TSTRUCT @temp ieqfrq 10 ISEED
@seed

set 4DIM = 1
calc step = @4dim / ( ( @temp - 300 ) / 10 )
label loop_freeze
    set oldtemp = @temp
    decr temp by 10
    decr 4dim by @step
    scalar fdim set @4dim sele lig end
    dynamics start nstep 1500 timestep 0.001 nprint 100 iprfrq
1000 iasvel 1 firstt @oldtemp finalt @temp tstruct @oldtemp IEQFRQ
10 iseed @seed
    mini abnr nstep 250
if temp gt 300 goto loop_freeze

!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!! Ecriture des coordonnees !!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
SKIP EXCL ALL
ENERGY
SET restr = ?noe
SKIP ALL EXCL VDW ELEC NOE
INTERACTION SELE rec END SELE lig END

open write card unit 1 name @wd/@prefix_@complex_@repeat.pdb
write coor pdb unit 1
* PARAM TEMP @temp DISTMAX @distmax DISTMIN @distmin
* MODEL @prefix_@complex_@repeat SEED @seed
* CONSOLE @consfile
* ENERGY ?ener ?elec ?vdw @restr
*
    close unit 1
```

### 9.1.4. Simulation de dynamique moléculaire activée

Les simulations de dynamique moléculaire activée avec utilisation du potentiel PEDC sont réalisées avec une version modifiée de Charmm 32b1.

Ci-dessous, un fichier type de simulation PEDC pour Charmm :

```
* run pedc
*
read rtf card name $TOPPAR/top_all22_prot.inp
read para card name $TOPPAR/par_all22_prot.inp
read psf card name hml.psf
read coor card name equ-all.crd
coor copy comp

NBONDED nbxmod 5 atom cdiel shift vatom vdistance vswitch -
cutnb 14.0 ctofnb 12.0 ctonnb 10.0 eps 1.0 e14fac 1.0 wmin 1.5

!=====
!set up periodic images
!=====

set 7 85
set 8 70
set 9 55

read image card name waterbox.img
image byres xcen 0.0 ycen 0.0 zcen 0.0 sele all end

shake bonh tole 1.0e-6 param
define hml sele segid A1 .or. segid A2 end
define helix sele
                resid 23:51 -
                .or. resid 60:87 -
                .or. resid 96:127 -
                .or. resid 140:161 -
                .or. resid 187:211 -
                .or. resid 360:389 -
                .or. resid 398:421 end
define bb sele type CA .or. type C .or. type N .or. type O end
define e2 sele resid 169:177 end
cons harm clear
cons harm force 5.0 comp sele hml .and. helix .and. bb .and. .not. e2
end

! target coordinates
! the coordinate file is identical to hpgk-cdp-bipg-mdequi-pedc.crd
! but one of the coordinates was changed very slightly
! this coordinate set SHOULD NOT BE IDENTICAL TO ORIGINAL
! COORDINATES
! The coordinates should be in charmm card format (not pdb)
pedc read ref unit 12 name equ-all.crd
close unit 12

!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!

set counter 0
set rmsd 0.0

label loop
incr rmsd by 0.1

pedc forc opti 1 fcgra 0.0 frota 0.0 frms 50.0 disref @rmsd -
```

```
sele e2 .and. bb end
! opti 0 : pot harmonique
! opti 1 : pot tends vers 0
! opti 2 : pot tends vers inf
! opti 4-5 : square well
open read formatted unit 30 name pedc-@counter.res
incr counter by 1
open write formatted unit 31 name pedc-@counter.res
open write uniform unit 33 name pedc-@counter.cor

dyna verlet restart iseed 3141516 -
      iunwri 31 iunrea 30 isvfrq 1000 -
      nstep 10000 timestep 0.001 -
      nprint 10 iprfrq 100 -
      iuncrd 33 nsavc 1000 -
      inbfrq 10 imgfrq 50

close unit 33
close unit 31
close unit 30

open write unit 20 card name pedc-@counter.pdb
write coor pdb unit 20
* MRMS @counter
*
goto loop
```

### 9.1.5. Simulations de dynamique moléculaire dans la membrane

Les simulations de dynamique moléculaire en environnement membranaire sont réalisées avec le programme NAMD sur la machine TERA-10 du Centre Calcul Recherche et Technologie.

Ci-dessous, un fichier d'entrée type de simulation NPT pour NAMD :

```
#####
#####
#forcefield
paratypecharmm      on
parameters          par_all27_prot_lipid.prm

#####
#####
#molecules

structure           start.psf
source              option.conf

#####
#####
#temp & pressure coupling

set temperature     310

langevin            on
langevinTemp        $temperature

useFlexibleCell     yes
useConstantArea     yes
LangevinPiston      on
LangevinPistonTarget 1.013
langevinDamping      1
LangevinPistonPeriod 100
LangevinPistonDecay 50
LangevinPistonTemp  $temperature
LangevinHydrogen    off

#####
#####
#integrator

timestep            1
nonbondedFreq       1
fullElectFrequency  2
stepspercycle       10

#####
#####
#approximations

rigidBonds          all
rigidTolerance       0.00000001
cutoff              12
switching           on
switchdist          10
pairlistdist        14
margin              3
exclude             scaled1-4
1-4scaling          1.0 # 1.0 for Charmm, 0.833333 for Amber
PME                 on
```

```
cellBasisVector1 120 0 0
cellBasisVector2 0 120 0
cellBasisVector3 0 0 130
PMEGridSizeX    160
PMEGridSizeY    160
PMEGridSizeZ    160

wrapWater        on
wrapAll          on                ;# wrap other molecules too

#####
#####
#####
#####
#output

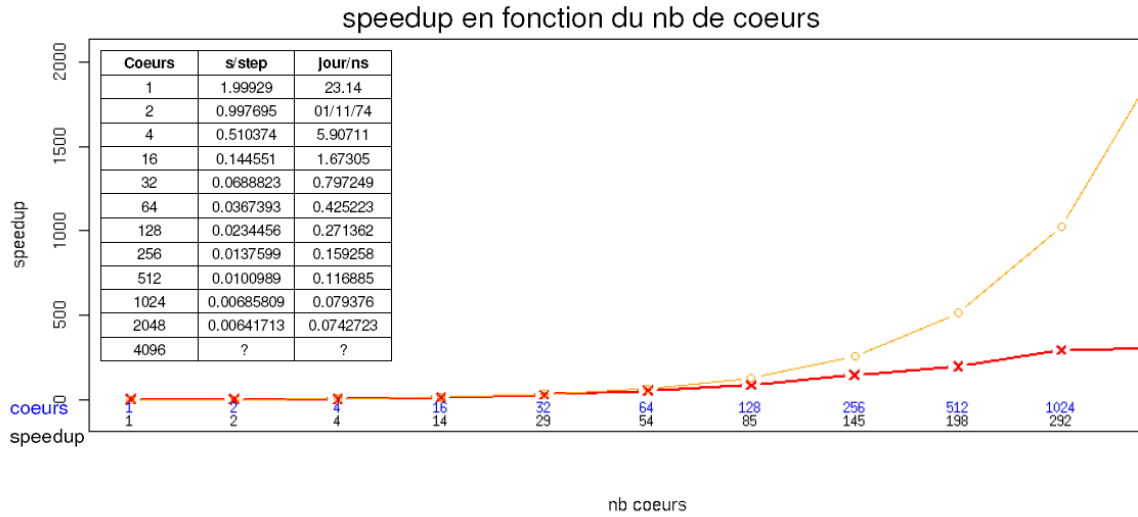
outputEnergies  1000
DCDfreq         1000
restartfreq     1000
binaryrestart   no
binaryoutput    yes
outputTiming    100

#####

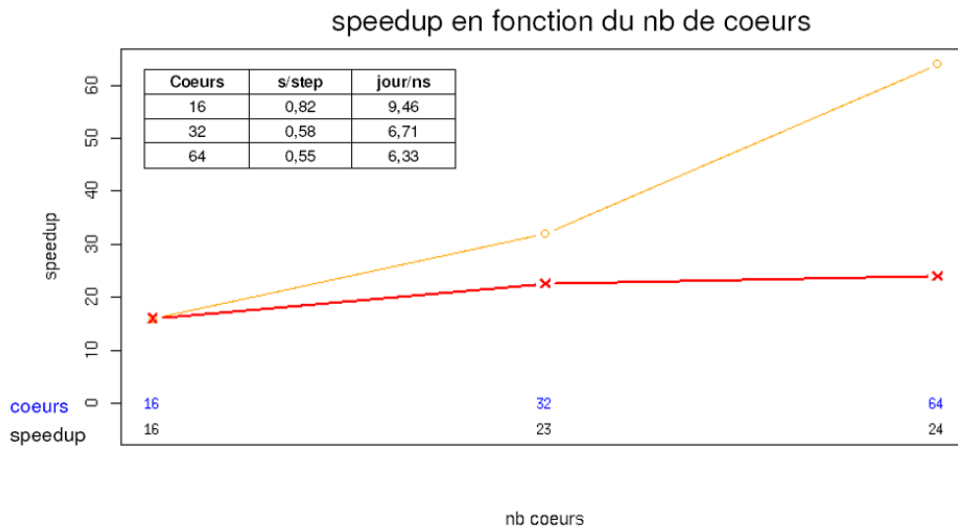
run 1000000
```



### 9.1.6. Benchmark des codes de dynamique moléculaire sur la machine TERA10



o courbe théorique  
x mesure platine



o courbe théorique  
x mesure platine