Allosteric modulation of pentameric ligand gated ion channels : from the jiggling of atoms to neuropharmacological strategies

Nicolas Martin

Université de Strasbourg

UNIVERSITÉ DE STRASBOURG

EDSC École Doctorale des Sciences Chimiques

ÉCOLE DOCTORALE DES SCIENCES CHIMIQUES

Institut de Sciences et d'Ingénierie Supramoleculaires

THÈSE présentée par :

Nicolas MARTIN

soutenue le : **20 Décembre 2017**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Chimie

ALLOSTERIC MODULATION OF PENTAMERIC LIGAND GATED ION CHANNELS

FROM THE JIGGLING OF ATOMS TO NEUROPHARMACOLOGICAL STRATEGIES

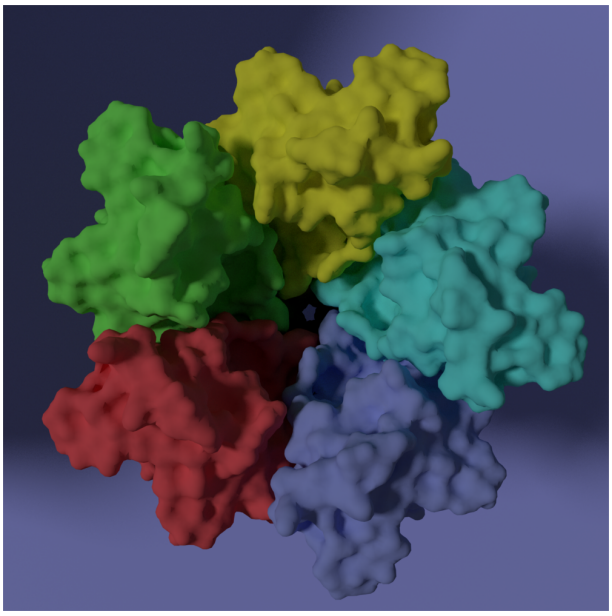
THÈSE dirigée par :
M. CECCHINI Marco Dr., Université de Strasbourg

RAPPORTEURS :
M. TAREK Mounir Dr., CNRS-Université de Lorraine
M. BAADEN Marc Dr., CNRS-Université Paris Diderot

AUTRES MEMBRES DU JURY :
Mme DEJAEGERE Annick Pr., Université de Strasbourg

ALLOSTERIC MODULATION OF PENTAMERIC LIGAND GATED ION CHANNELS: FROM THE JIGGLING OF ATOMS TO NEUROPHARMACOLOGICAL STRATEGIES

NICOLAS MARTIN



Thesis Director: Dr. Marco Cecchini
March 9, 2018

À mon grand-père

PUBLICATIONS

Some ideas and figures have appeared previously in the following publications:

- NE. Martin, S. Malik, N. Calimet, J-P. Changeux and M. Cecchini . "Un-Gating and Allosteric Modulation of a Pentameric Ligand-Gated Ion Channel captured by Molecular Dynamics" *PLoS Computational Biology*, **2017**.
- AH. Cerdan, NE. Martin and M. Cecchini. "The physiologically Active state of the Glycine Receptor captured by Molecular Dynamics". *Submitted to Structure*, **2017**.

RÉSUMÉ DE THÈSE

Régulation allostérique des récepteurs pentameriques canaux : de l'agitation atomique aux stratégies pharmacologiques.

1. INTRODUCTION

Les *pentameric ligand gated ion channels* (pLGICs), aussi connus sous le nom de cys-loop, sont des récepteurs neuronaux impliqués dans la neurotransmission rapide. On les retrouve aussi bien dans le cerveau humain que chez des procaryotes avec une importante similarité de structure, qui s'explique par leur rôle crucial dans le maintien de l'homéostasie. Parmi ces récepteurs on peut citer les nAchR, $GABA_AR$, GlyR ou $5HT_3R$. Chacune des cinq sous unités est composée d'un domaine intracellulaire (ICD) (absent chez les procaryotes), d'un domaine transmembranaire (TMD) et d'un domaine extracellulaire (ECD).
Lorsqu'ils ne fonctionnent pas correctement ils peuvent être impliqués dans des maladies telles qu'Alzheimer et Parkinson, ou encore dans certains types d'épilepsie et de schizophrénie. Ils sont activés par des neurotransmetteurs endogènes qui par définition se lient au site orthostérique mais aussi par des ligands allostériques. Ils sont la cible de nombreux médicaments déjà commercialisés tels que le STROMECTOL® (ivermectine) ou le REMINYL® (galantamine). Cependant le mécanisme d'action détaillé de ces récepteurs est encore inconnu et sa compréhension pourrait permettre des avancées majeures dans la mise au point de nouveaux traitements.
Avec l'avènement de techniques de visualisation avancées, permettant d'observer les structures des proteins à l'échelle atomique, l'étude des structures de protéines et par conséquent de leur méchanisme de fonctionnement s'est grandement développé. En particulier, Les études cristallographiques se sont montrées très efficaces pour décrire les structures de ces protéines mais ne donnent que peu d'informations sur leur mécanisme d'action ou leur dynamique. L'utilisation de techniques de dynamique moléculaire (MD) en addition à ces structures haute résolution permet de comprendre les mécanismes d'actions de ces protéines à l'échelle atomique. Cependant ces méthodes sont grandement dépendant des structures produites expérimentalement rendant la collaboration entre différents domaines d'expertise essentielle.
Ce n'est que très récemment que la première structure hétéromérique d'un pLGICs humain a été élucidée[1] et c'est notamment pour cette raison que la structure d'un récepteur homologue issue d'un eucaryote, i.e., GluCl de *Caenorhabditis Elegans*[2] a été utilisée dans cette étude et dans de nombreuses autres études computationnelle sur les pLGICs[3–7]. Pour comprendre le mécanisme d'action à l'échelle atomique de cette famille de récepteurs ainsi que leur modulation allostérique nous avons étudié quatre systèmes par dynamique moléculaire classique ; i) GluCl lié à l'ivermectine (IVM) et au L-Glutamate (L-Glu) ii) GluCl en absence

d'IVM iii) GluCl en absence d'IVM et en présence de L-Glu et iv) GluCl apo. Dans un second temps nous avons mis au point un protocole de calcul d'énergie libre pour étudier la modulation allostérique des pLGICs et appliquer ce dernier à un petit jeu de molécules choisies pour leur activité supposée.

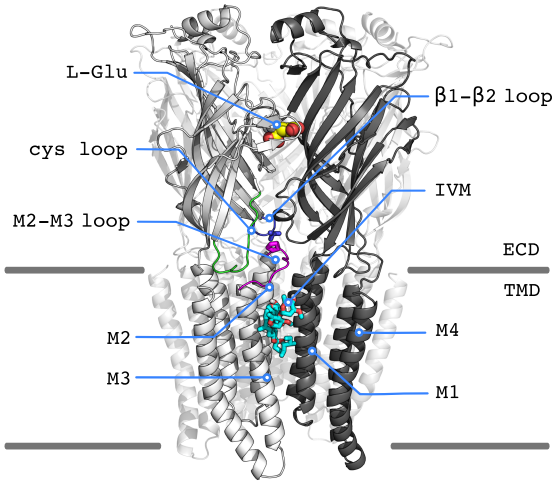


Figure 1: représentation de la structure active de GluCl (PDBID 3RIF) en complexe avec l'ivermectine et le L-glutamate.

2. RÉSULTATS ET DISCUSSIONS

2.1 Capturer la transition entre l'état ouvert et l'état fermé

Deux simulations de GluCl actif (PDBID 3RIF) auquel nous avons retiré l'IVM ont été réalisées pour un total de 5 microsecondes. La capacité de l'IVM à moduler l'activité ou à activer GluCl est encore sujette à débat[8]. Cependant nos simulations ont montré que sa déliaison était suffisante pour conduire à la relaxation du récepteur dans son état fermé en moins d'une microseconde. En effet, quand l'IVM est retirée de son site de liaison dans la structure active plusieurs observations majeures sont faites.
L'angle de twist évolue de 14 (valeur compatible avec un état actif) à environ 23 degrés (valeur compatible avec un état de repos) pour les deux simulations concernées (cf. Fig 2). L'importance de cette isomérisation quaternaire est aussi suggérée par les structures cristallographiques et avait été proposée par Taly et al. en 2005 grâce à l'analyse de modes normaux[9].
Le diamètre du pore décroît de façon significative pour atteindre des valeurs correspondant avec un état fermé (cf. Fig 2) alors qu'il reste stable quand l'IVM est liée. Le point de constriction alors observé se situe au même endroit que celui observé sur la structure apo cristallographique ou produite par MD, i.e., en position 9′. Ce rétrécissement s'accompagne d'une chute significative du flux de molécules d'eau au travers du pore jusqu'à être nul.
Lorsque que l'IVM est retirée de son site actif on observe la déliaison spontanée de quatre

(ou cinq dans le cas de la simulation B cf. Fig 2) des cinq L-Glu liés initialement au récepteur. Lorsque que l'IVM reste liée cependant, le nombre de L-Glu liés au récepteur ne décroît jamais au dessous de trois. Ceci suggère que le ligand allostérique a un effet direct sur l'affinité de liaison du ligand orthostérique.
Enfin, si on s'intéresse à l'écart quadratique moyen (RMSD) à la structure apo, qui caractérise le degré de convergence de la structure active vers la structure apo, pour les deux domaines (ECD et TMD) les conclusions suivantes peuvent être faites: i) la transition de twisting fait converger le TMD de la structure active vers la structure apo ii) une fois cette même transition complétée et tous les ligands déliés, l'ECD dévie grandement (environ 3 Å) de celle de la structure apo (données non présentées ici). Cette observation à l'échelle des sous unités mais aussi d'autres observations plus détaillées qui pour des raisons de concision ne sont pas présentées ici (cf Ref[10]) nous amènent à conclure que nous avons pu produire par MD un état qui s'approche par de nombreux aspects à l'état de repos produit par cristallographie X-Ray, rendant la transition observée parfaitement adaptée à l'étude du mécanisme d'ouverture/fermeture des pLGICs et illustrant les capacités prédictives de la MD.

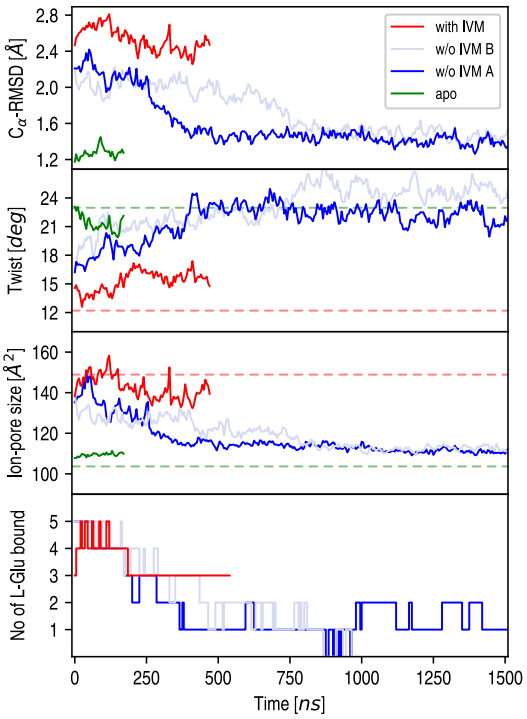


Figure 2: quatre observables d'intérêt sont présentées au cours du temps pour les quatre systèmes étudiés.

2.2 *Proposition d'un nouveau mécanisme d'ouverture / fermeture*

Une fois la transition complète d'un état ouvert à un état de repos (canal ionique fermé) observée, nous avons pu nous intéresser au mécanisme précis de cette fermeture. Nos anal-

yses ont tout d'abord confirmées que la fermeture du pore se faisait bien en position 9′ et qu'elle était associée à un mouvement vers la lumen du pore des hélices M2 et donc des prolines 268 très conservées au sein des pLGICs[5,9,11], comme déjà souligné par plusieurs études[3,4,12,13]. Dans un second temps nous avons montré que la fermeture du pore implique un changement des interactions entre certains résidus situés à l'interface entre les domaines transmembranaires et extracellulaires. Plus précisément, la fermeture du pore est permise par le déplacement vers le haut de la boucle β1-β2 qui libère le passage pour que la boucle M2-M3 puisse se déplacer vers le centre du canal (cf. Fig 3). Grâce aux mêmes simulations nous avons aussi montré que le déplacement vers le haut de la boucle β1-β2 était corrélé avec le twist de l'ECD. Enfin nous avons découvert que l'isomérisation de twisting était contrôlée à la fois par le L-Glu et par l'IVM[10]. En effet quand GluCl est lié à la fois à l'IVM et le L-Glu, les valeurs du twist environnent 14 deg alors si le L-Glu uniquement est lié ces valeurs augmentent à 18 deg et continuent d'augmenter jusqu'à plus de 22 deg dans le cas où aucun ligand n'est lié. Cette observation tendrait à expliquer l'action de modulation allostérique positive de l'IVM par la stabilisation d'un état "super-untwisté" pour lequel la barrière à franchir pour fermer le récepteur serait plus haute.

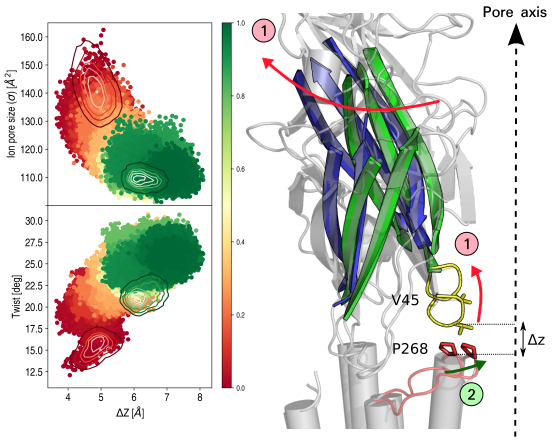


Figure 3: illustration du mécanisme proposé. La corrélation entre la fermeture du pore et le twist ainsi que celle entre le déplacement vertical (ΔZ) et le twist sont aussi illustrés. Le gradient de couleur représente le temps. Les lignes de densité représentent les simulations à l'équilibre de GluCl actif (rouge) et apo (vert).

2.3 Modulation allostérique des pLGICs

Une fois l'observable d'intérêt isolée grâce à des simulations libres de notre système nous nous sommes attelés à la compréhension de la régulation allostérique de cette famille de récepteurs et plus particulièrement au site allostérique de l'IVM. Les calculs d'énergie libre réalisés ont pour but de nous aider à mieux comprendre les effets des ligands allostériques

sur les barrières énergétiques présentes entre état actif et de repos. Dans un premier temps un potentiel de force moyenne (PMF) de la structure active en l'absence de ligand ainsi qu'un PMF de la même structure en complexe avec ses deux ligands ont été réalisés. Dans un second temps le même calcul a été réalisé en complexe avec le L-Glu seulement pour étudier l'effet de l'IVM sur le twist. Nous avons alors pu montrer qu'en présence des deux ligands (cf. Figure 4 +IVM+GLU) le twist était bloqué à une valeur d'environ 14 degrés et que lorsque l'IVM était retirée cette valeur augmentait pour se stabiliser aux environs de 17 degrés. Ces résultats en adéquation avec les observations faites lors de simulations libres nous confortent dans la justesse de notre choix quant à la coordonnée de réaction choisie, i.e., le twist, ainsi que sur la cohérence de notre protocole. D'autre part lorsque le récepteur n'est lié à aucun ligand (cf. Figure 4 -IVM-GLU) le système présente un minimum d'énergie correspondant à un twist de 18-20 degrés contrairement aux 24 degrés suggérés par nos simulations libres. Ces deux constats suggèrent que la transition d'un état activé par l'IVM et le L-Glu à un état activé par le L-Glu uniquement se fait par simple twisting alors que la transition de ce dernier état à un état de repos nécessiterait plus que le twist. L'omission de possibles degrés de liberté orthogonaux d'importance dans notre calcul d'énergie libre (tels que la fermeture du pore ou le blooming) pourrait expliquer le décalage observé entre simulations libres et PMF. Cependant ceci ne nous empêche pas de discriminer entre un état actif et de repos par l'étude du PMF le long du twist. Pour mieux comprendre comment l'IVM bloque l'isomérisation de twist nous avons réalisé un calcul de PMF avec un fragment d'IVM. Ce dernier a été choisi pour conserver la majorité des interactions observées entre l'IVM et le récepteur lors de simulations libres et dans la structure cristallographique (cf. IVM-FRAG, Figure 4). Nous avons pu montrer que le fragment choisi restait non seulement lié au récepteur lors de simulations de MD courtes mais permettait aussi de bloquer le twist de façon similaire à l'IVM. Enfin nous avons effectué les mêmes calculs avec différents ligands suggérés par la revue du Dr. Brannigan[14], i.e., l'allopregnanolone (ALLO), l'hormone triiodothyronine (T_3) et le cholestérol (CHOL). Après avoir été placés dans le site actif de l'IVM par docking et que la stabilité des complexes ait été confirmée par une courte MD, les calculs de PMF obtenus sont montrés sur la Figure 4. Ces trois ligands sont capables de bloquer le twist en stabilisant un état untwisté entre 13 et 14 degrés. Les légères variations de degré de twist qu'ils stabilisent ne semblent pas directement liées à leur taille. En effet on pourrait s'attendre à ce que plus le ligand est gros plus il bloque un degré de twist proche de la structure cristallographique, i.e., 12 degrés.

3. CONCLUSIONS

Nous avons étudié le mécanisme d'ouverture/fermeture d'un récepteur d'intérêt pharmacologique et proposé un possible mécanisme qui a fait l'objet d'une publication[10]. Le mécanisme proposé permet de faire le lien entre la liaison du neurotransmetteur au niveau du site extracellulaire et l'ouverture du pore dans le domaine transmembranaire. Nous avons aussi montré que les deux sites de liaison, i.e., orthostérique et allostérique, contrôlaient le twist. En nous appuyant sur une approche basée sur des calculs d'énergie libre nous avons

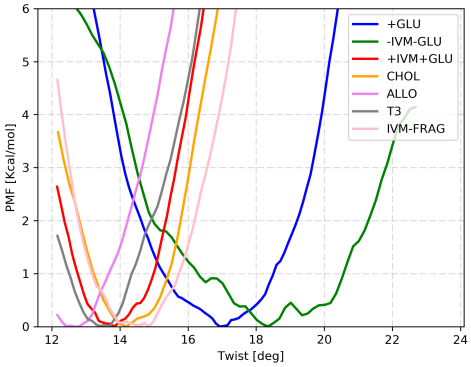


Figure 4: PMF en fonction de l'angle de twist pour sept systèmes différents.

développé une méthode qui permet d'étudier l'activité de petites molécules sur le twist et pas seulement leur capacité à se lier à une cible comme le ferait une étude de docking. Nous avons confirmé que le cholestérol était capable de bloquer le twist et donc possiblement de stabiliser les pLGICs en position ouverte, confirmant ainsi le lien entre la composition des membranes cellulaires et l'activité des pLGICs. Nous avons aussi montré que plusieurs molécules, i.e., un fragment d'IVM, ALLO et T3 étaient aussi capables de stabiliser un état untwisté du récepteur et donc potentiellement d'être des potentiateurs allostériques positifs. Nous souhaitons appliquer ce protocole à d'autres membres de la famille des pLGICs, structurellement plus proches des cibles pharmacologiques humaines, comme GlyR ou nAChR. En association avec une méthode d'échantillonnage accélérée comme l'aMD pour isoler de possibles états intermédiaires ou produire des états encore impossibles à cristalliser, notre approche pourrait aider à concevoir de nouvelles molécules dont l'activité pourrait alors être testée *in vitro*.

Les avancées dans la compréhension du fonctionnement et de la régulation des pLGICs présentées dans cette étude pourraient conduire à la conception de nouveaux médicaments pour des pathologies qui, à l'heure actuelle, restent incurables.

BIBLIOGRAPHY

[1] G. Brannigan. "Direct Interactions of Cholesterol With Pentameric Ligand-Gated Ion Channels: Testable Hypotheses From Computational Predictions". 2017.

[2] N. Calimet, M. Simoes, J.-P. Changeux, M. Karplus, A. Taly, and M. Cecchini. "A gating mechanism of pentameric ligand-gated ion channels." *Proceedings of the National Academy of Sciences of the United States of America*, 110(42), pp. E3987–96, 2013.

[3] M. H. Cheng and R. D. Coalson. "Energetics and ion permeation characteristics in a glutamate-gated chloride (GluCl) receptor channel". *Journal of Physical Chemistry B*, 116(46), pp. 13637–13643, 2012.

[4] J. Hénin, R. Salari, S. Murlidaran, and G. Brannigan. "A predicted binding site for cholesterol on the GABAA receptor". *Biophysical Journal*, 106(9), pp. 1938–1949, 2014.

[5] S. A. Heusser, Ö. Yoluk, G. Klement, E. A. Riederer, E. Lindahl, and R. J. Howard. "Functional characterization of neurotransmitter activation and modulation in a nematode model ligand-gated ion channel". *Journal of Neurochemistry*, pp. 243–253, 2016.

[6] A. Taly, M. Delarue, T. Grutter, M. Nilges, N. Le Novère, P.-J. Corringer, and J.-P. Changeux. "Normal mode analysis suggests a quaternary twist model for the nicotinic receptor gating mechanism." *Biophysical journal*, 88(6), pp. 3954–65, 2005.

[7] R. E. Hibbs and E. Gouaux. "Principles of activation and permeation in an anion-selective Cys-loop receptor". *Nature*, 474(7349), pp. 54–60, 2011.

[8] M. Jaiteh, A. Taly, and J. Hénin. "Evolution of Pentameric Ligand-Gated Ion Channels: Pro-Loop Receptors." *PloS one*, 11(3), p. e0151934, 2016.

[9] C. L. Morales-Perez, C. M. Noviello, and R. E. Hibbs. "X-ray structure of the human $\alpha 4 \beta 2$ nicotinic receptor". *Nature*, 538(7625), pp. 411–415, 2016.

[10] H. Nury, F. Poitevin, C. Van Renterghem, J.-P. Changeux, P.-J. Corringer, M. Delarue, and M. Baaden. "One-microsecond molecular dynamics simulation of channel gating in a nicotinic receptor homologue." *Proceedings of the National Academy of Sciences of the United States of America*, 107(14), pp. 6275–80, 2010.

[11] Ö. Yoluk, T. Brömstrup, E. J. Bertaccini, J. R. Trudell, and E. Lindahl. "Stabilization of the GluCl ligand-gated ion channel in the presence and absence of ivermectin". *Biophysical Journal*, 105(3), pp. 640–647, 2013.

[12] O. Yoluk, E. Lindahl, and M. Andersson. "Conformational gating dynamics in the GluCl anion-selective chloride channel". *ACS Chemical Neuroscience*, p. 150520172255005, 2015.

[13] F. Zhu and G. Hummer. "Pore opening and closing of a pentameric ligand-gated ion channel". *Proceedings of the National Academy of Sciences*, 107(46), pp. 19814–19819, 2010.

ACKNOWLEDGMENTS

Durant ces trois années (et demi) j'ai été accompagné, conseillé, soutenu et guidé par de nombreuses personnes. Un doctorat demande à la foi un soutient moral, scientifique et financier, c'est pourquoi je remercie vivement toutes les personnes qui ont pu être impliquées dans cette incroyable aventure.

Tout d'abord, je souhaite remercier le Dr. Marco Cecchini qui a eu confiance en moi et m'a donné la chance de pouvoir réaliser mon rêve de devenir docteur dans un environnement idéal. J'ai appris de toi l'exigence que requiert l'excellence.

Ensuite, je souhaite remercier les membres de mon jury. Le Dr. Annick Dejaegere mais aussi le Dr. Mounir Tarek qui ont su me conseiller et m'orienter avec justesse et bienveillance, tout d'abord lors de mon entretien de mi-parcours, puis lors de ma soutenance de thèse. Enfin, je souhaite remercier le Dr. Marc Baaden pour ses retours constructifs sur mon travail de recherche et mon manuscrit.

Autant d'un point de vue scientifique que personnel, je tiens aussi à remercier mes collègues et amis. Nico et Jerem merci de m'avoir appris quand je savais si peu. Simon, well thanks for everything! You were a model and I admire your devotion to science like no other. Adri et Flo merci pour vos conseils, vos explications, votre soutient et pour toutes les discussions scientifiques passionantes, au labo ou ailleurs. Joelito, incluso si no soy David Wozniack, it was a great pleasure to have you as a friend and colleague during all these years. Diego, thanks you for making me feel like the Python god and for all your advices.

Je voudrais aussi remercier tout ceux qui n'ont pas été directement impliqué dans mon travail de recherche mais qui ont rendu ces trois années plus douces. Amaury, merci pour tes leçons particulières de maths qui m'ont permises d'être un peu moins perdu face aux terribles quaternions. Alex, merci de croire que je connais toujours le Vidal par coeur. Merci aussi à mes potes de toujours ou presque, Dridri et Coco, Gaëlle et Gabin, Jiji et Nini, tous les pharmaciens.

Enfin, je souhaiterais remercier ma famille. Ma mère pour son soutien sans faille, son incroyable force et sa détermination qui sont pour moi des examples. Mon père pour son soutient aussi et pour avoir toujours été convaincu que je pouvais accomplir l'impossible. Ma famille pour avoir été là quand j'en avais besoin. Enfin, je voudrais remercier Marion, qui partage ma vie depuis 4 ans 1 mois et 5 jours. Merci de croire en moi et de me soutenir tous les jours. Promis, c'est la dernière thèse!

CONTENTS

χ	Dihedral angle
ϕ	Spinangle
σ	Pore $C_\alpha$cross-section
τ	Twist angle
θ_a	Azimuthal tilting angle
θ_p	Polar tilting angle
ξ, RC	Reaction coordinate
AB	Antibody
ABF	Adaptative biasing force
ACh	acetylcholine
AChBP	Acetylcholine binding protein
AD	Alzheimer's disease
ALLO	Allopregnalonone
aMD	Accelerated molecular dynamics
AP	Action potential
ATSM	Allosteric two states model
AVG	Average
BBB	Blood brain barrier
BEUS	Bias Exchange umbrella sampling
CG	Coarse grain
CHOL	Cholesterol
CNS	Central nervous system
COMT	Catechol-O-methyltransferase
CryoEM	Cryo-electron microscopy
CV	Collective variable

EC Extra-cellular

ELIC *Erwinia chrysanthemi* Ligand-gated Ion channel

FF Force field

GA General anaesthetic

GABA γ-aminobutyric acid

$GABA_AR$ γ-amino-butyric acid receptor type A

GLIC *Gloeobacter violaceus* ligand-gated ion channel

GluCl Glutamate gated chloride channel

GlyR Glycine geceptor

GPCR G-protein coupled receptor

gWHAM Generalized weighted histogram analysis method

IBS Irritable bowel syndrome

ICD International commmon denomination

IVM Ivermectin

L-Glu L-Glutamate

MBAR Multistate Bennett acceptance ratio

MD Molecular dynamics

metaD Metadynamics

nAChR Nicotinic acetylcholine receptor

NVE Microcanonical ensemble

NVT Canonical ensemble

PB Poisson-Boltzmann

PD Parkinson's disease

pLGICs Pentameric ligand gated ions channels

PME Particule mesh Ewald

PMF Potential of mean force

POPC 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine

PUFA Polyunsaturated fatty acid

RE Replica exchange

RMSD Root mean square deviation

RMSF Root mean square fluctuation

SM Sphingomyelin

SMD Steered molecular dynamics

T_3 Triiodothyronine

TM Trans-membrane

US Umbrella sampling

VMD Visual molecular dynamics

WHAM Weighted histogram analysis method

SUMMARY

Throughout millions of years of evolution the living organisms evolved in size and complexity creating the need for a nervous system. The latter progressed from simple dispersed nervous cells to organized structures such as nerves. The nervous system is composed of various cellular types among which the most known is the neuron. In the human nervous system, both central and peripheral, neurons are as numerous as 21 000 000 000 and the number of synapses, i.e., connections between the neurons themselves or with other cell types, are of the order of 10^{14}.

The human brain, through its neurons and synapses, where are found neuroreceptors, controls many vital and basic functions. Neuroreceptors can be of many kinds and among them one may cite the ligand gated ion channels. The former group is composed of the tetrameric glutamate receptors, the trimeric ATP-gated receptors also called P2X and finally the pentameric ligand gated ion channels (pLGICs). The latter group, on which we will be focusing in this thesis, comprises receptors of major importance such as the nicotinic acetylcholine receptors (nAChR) or the γ-aminobutyric acid receptors type A ($GABA_A$).

Due to the fact that they control vital mechanisms, when dysfunctionning, pLGICs may be involved in major disorders such as Alzheimer's (AD) or Parkinson's disease (PD) and for this reason they represent major pharmacological targets.

Thanks to the development of X-ray crystallographic techniques jointly with electrophysiology, but also of numerous computational studies, our global understanding of the function and structure of pLGICs has greatly increased. Nevertheless, despite the significant research effort made during the past decades, the relation between the structure and the function of pLGICs is still not well established.

To address this question we used an approach based on advanced computational methods, making use of the latest high resolution structures of the pLGICs. We performed a total of 6 μs of full-atom simulations of GluCl, the first eukaryotic pLGIC elucidated in high resolution by X-ray crystallography.

In a second time we focused on the modulation of the activation of the pLGICs by allosteric ligands aiming at providing useful insights for the design of putative new modulators.

This thesis will be articulated in four main chapters. In the first, I will contextualize the object of this study, i.e., the pLGICs, discussing their function but also their structure and mechanism of action. In the second chapter the theoretical framework will be detailed. I will in particular present some of the major computational methods which were either directly used to obtain the presented results or envisaged during the various phases of my thesis. The third and fourth chapters will be dedicated to exposing and then discussing the original work accomplished during my three years of doctoral studies.

1 INTRODUCTION

1.1 PENTAMERIC LIGAND GATED ION CHANNELS: FROM THE DISCOVERY TO THE ATOMIC STRUCTURE

1.1.1 *Brain and neurons*

Long before the apparition of modern sciences, Greek philosophers debated on if the soul originated from the brain or the heart. Hippocrates (460-370 BC) was the first to settle on the brain as the host of the soul.
Until the 18th century where Luigi Aloisio Galvani started working on nerve conduction in frogs, very little advances in the field of neurosciences were made. The discovery of Galvani allowed to better understand the role of electricity in nerve conduction and thus how the information was transported from the brain to the peripheral organs and muscles.

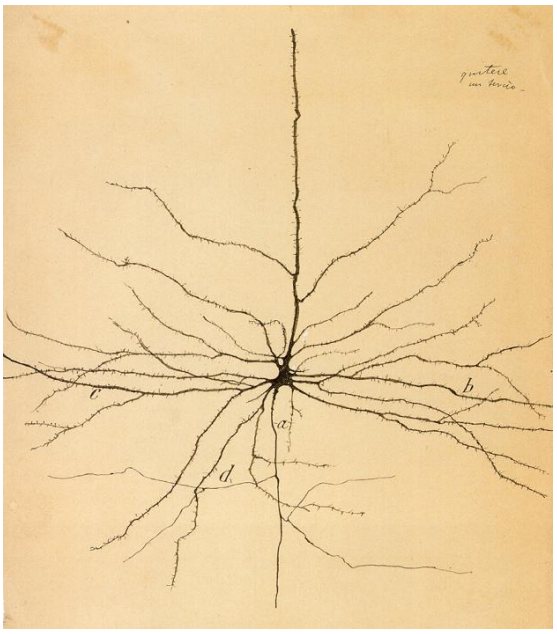


Figure 1.1: One of the first drawings of a neuron by Santiago Ramón y Cajal (1852-1934). Property of the Museo Cajal (Madrid).

Later in the 20th century, the development of advanced methods of visualization such as microscopes or cell staining, discovered by Camillo Golgi (see a review of his contribution to neuroscience in ref[15]), also allowed for a significant advance of neurosciences. It is indeed in that period of time that started to emerge the first images of neurons. Scientists were

then capable of studying the anatomy of neurons and thus described the different parts of a neuron such as the *soma*, the dendrites, the axon and the synapse, as illustrated by figure 1.2. Neurons are nervous cells specifically dedicated to the transmission of electrical signals from the brain to distant organs or muscles. They are composed of three main domains. First of all the *soma* or body, which contains the nucleus of the cell and carries the axon and dendrites. The *soma* is the place in which most of the cell's basal life processes occur such as for instance protein synthesis.

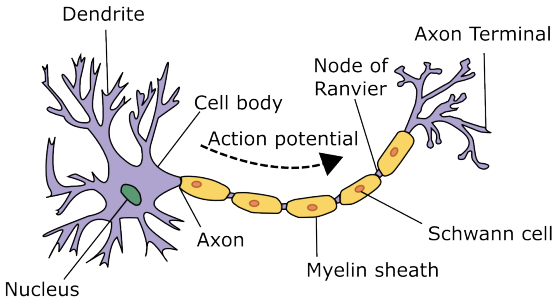


Figure 1.2: Schematic representation of a neuron featuring the myelin sheath and the Schwann cells. (from ref[16])

The dendrites are a tree-shaped construction which is mainly involved in receiving the input signals from other neurons.
The axon can be pictured as the message carrier as it transports an electrical influx generated in the *soma* toward the axon's termination where, is located the synapse. There exist several types of neurons classified upon their function. Among the most commons, the motoneurons present an axon around which is wrapped the so called myelin sheath which allows for a significant increase of the conduction speed of the electrical signal or action potential (AP) through the axon of the neuron. The myelin sheath is composed of Schwann cells, named after their discoverer, Theodor Schwann.

The synapse is described as the contact zone between two neurons or a neuron and another cell. It consists of a pre synaptic neuron, a post synaptic neuron and the synaptic cleft. In order to propagate from one neuron to another for instance, the AP must cross the synaptic cleft. To do so the pre-synaptic neuron liberates small molecules (referred to as neurotransmitters) in the synaptic cleft which are then captured by the post synaptic neuron through the intervention of neuronal receptors such as for instance pLGICs.

Because it allows the brain to control basic functions such as breathing motion, skeletal muscles activation or blood pH regulation the mechanisms of neurotransmission are vital and must be understood at all levels. In fact, a wide variety of diseases can affect the neurotransmission by either damaging the myelin sheath, such as multiple sclerosis, or by disturbing the secretion of neurotransmitters as for instance Parkinson's disease.

1.1.2 *Discovery of the pLGICs*

In 1905, as he was working on the response to various stimuli of extracts from kitten adrenal glands[17], John Newport Langley first proposed the concept of reactive substance and from there was deduced the concept of receptors.
The first identified member of the pLGICs, the muscle type nAChR was in fact named the substance which allowed for its discovery. Only later will be made the discovery that the nAChR is natively activated by acetylcholine. As a consequence of the first member of the pLGICs was discovered they were first named nicotinic receptors.
As science and in particular molecular biology and biophysics evolved the classification and thus the naming of receptors changed from empiric observations on either the activating molecules or the physiological effect to their tri-dimensional organization or even genetic origins. Nicotinic receptors' name was then changed to cys-loop receptors. The latter because of the presence of a very highly conserved loop carrying a cysteine bridge in the extracellular domain. They are now referred to as pentameric ligand gated ion channels in relation to both their function and structure.
Lately, the phylogenetic of the pLGICs suggested to rename them Pro-loop receptors as a proline was found to be conserved in 100% of the structures studied when the cys-loop was significantly less conserved[11].
In 1970, the nicotinic acetylcholine receptor, was isolated from fish electric organ[18]. In fact, the discovery of the first pLGICs was a major breakthrough in the field of neurophysiology, allowing for the exploration of its function and response to diverse stimuli. However, its detailed tri-dimensional structures remained unknown. The first atomistic structures of pLGICs were elucidated half a century later from prokaryotic homologues; GLIC from *gloebacter violaceus*[19] and ELIC from *Erwinia chrysantemi*[20]. It is only very recently that high resolution 3D structures of eukaryotic homologues[2,21] were published which represented a breakthrough in the field.
After the years 2010s the significant ameliorations of the cryoEM techniques, both in term of hardware and software, allowed for the publication of a swarm of new structures at high resolutions of membrane proteins, including pLGICs.
In fact, this method allows to keep the protein in an environment which is relatively closer to the physiological one than the one required by X-ray crystallography experiments. Indeed, in theory, membrane protein could be resolved in cryoEM keeping the native lipid bilayer in which they are embedded. Moreover, by freezing the samples before exposing them to the electron beam, this methods allow to lower the damages that the high speed electrons can cause to the structure of the protein and therefore increase the accuracy of the produced images.

1.1.3 *Phylogenetic of the pLGICs*

Until the discovery of Tasneem et al. in 2005 of a prokaryotic homologue of the pLGICs family it was thought that these membrane proteins were only present in complex eukaryotic

organisms. In fact, they are present in most of the living organisms, i.e., from bacteria to human brain[22].

Thanks to modern sequencing techniques the amount of genomic data increased and the pLGICs alone now represent more than 50 known sequences. These various sequences have been analyzed to draw several important conclusions on the proximity between species but also on the evolutionary pressure put on some key residues. In fact, phylogenetic analyzes allow to understand which residues play a key role in the various functions of the receptor such as the selectivity for anions or cations.

In an article published in 2016 Jaiteh et al.[11] demonstrated that pLGICs can be split in groups and subgroups, as shown on Figure 1.3, namely: i) AChBP which are very similar to the rest of the members in the extracellular (EC) domain but the transmembrane (TM) domain and the intracellular domain (IC) are completely absent; ii) nAChR and 5-HT3; iii) $GABA_{A,C}R$ and GlyR and iv) bacterial homologues.

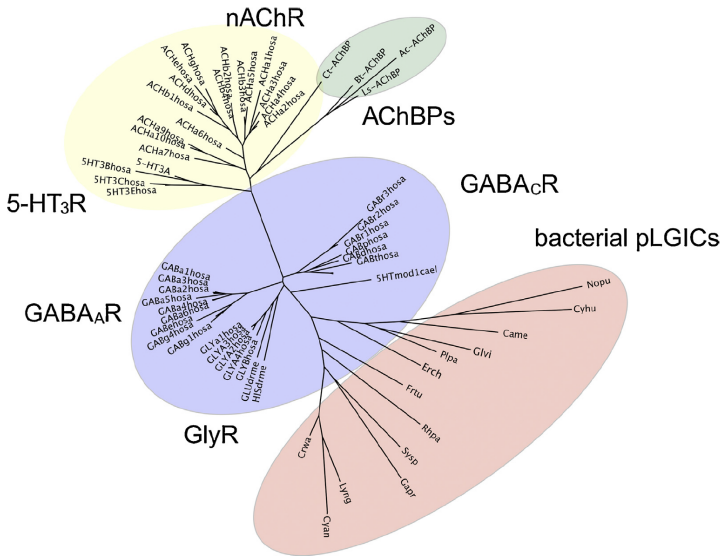


Figure 1.3: Phylogenetic tree of the pLGICs family as proposed in ref[23], including prokaryotic and eukaryotic receptors. Are shown in yellow the cation-selective eukaryotic; in green the acetylcholine binding proteins; in blue the anion selective eukaryotic; and in pink the prokaryotic homologues.

1.1.4 *Human members of the pLGICs*

Human representatives of the pLGICs are involved in major functions to maintain homeostasis. Interestingly, despite their resemblance in term of amino acid sequence and tridimensional organization they can be either selective for anions or for cations.

- $GABA_AR$ are anionic receptors selective for chloride ions. They are activated by the γ-aminobutyric acid which is the main inhibitory neurotransmitter in the central nervous system (CNS). They should not be mistaken with the $GABA_BR$ which are also activated

by GABA but are G-protein coupled receptors (GPCR) instead of ion channels. A recent structure of a human receptor was elucidated by X-ray crystallography in complex with benzamidine (a new antagonist)[24].

- 5-HT_3R are cationic receptors selective for sodium and potassium ions. They are activated by serotonin as suggested by their name. A recent structure of a mouse receptor was elucidated in X-ray crystallography. This structure was crystallized in the presence of nanobodies which are used in crystallography to reduce the entropy of the protein in order to facilitates the formation of crystals.

- GlyR are anionic receptors selective for chloride ions and activated by glycine but also by β-alanine and taurine. A recent set of structures obtained by X-ray crystallography and cryo-electron-microscopy (cryoEM) was published[25–27].

- nAChR are non selective cationic receptors as they allow for the passage of potassium and sodium cations and even sometimes calcium ions. Until very recently, no atomic-resolved structure of this receptor was elucidated. In 2016, this challenge was adressed by Morales-Perez et al. who published the first high-resolution structure of nAChR ever elucidated by X-ray crystallography[1]. Prior to 2016, only theoretical models based on prokaryotic homologues of this receptor were available to study nAChR. However the acetylcholine-binding-protein (AChBP), a homologue of the EC domain of the nAChR, was extensively used to explore the binding of small molecules to the nAChR by experimental techniques. Although, it cannot describe neither the effect of binding on the activity of the receptor nor the activation or deactivation mechanisms as it misses the TM domain. They should not be confused with the muscarinic AChR which are GPCR.

1.1.5 *Non-human members of the pLGICs*

Non human members of the pLGICs gather the prokaryotic and the non human eukaryotic homologues. They are more important in number and were more investigated than the human one so far. Indeed human structures only started appearing lately whilst the first structure of prokaryotic homologues where published in 2008. They present several minor differences in term of structure which will be discussed further in upcoming sections but overall the sequence and global organization remain the same than the one of the human pLGICs. Even thought they are less important pharmaceutical targets they still represent good models for understanding the function of pLGICs.

- ELIC: (Erwinia Ligand Gated Ion Channel) isolated from the bacteria *Dickeya dadantii* (previously known as *Erwinia chrysanthemi*) is responsible for the soft rot disease in foliage plants. It was the first structure ever elucidated of a pLGIC at high resolution (3.3 Å)[28].

- GluCl: (Glutamate Gated Chloride Channel) the two structures published of this protein were expressed in *Caenorhabditis elegans* and were the first eukaryotic structure of a pLGIC. Although the open state is believed to be an active state due to its ligation with

IVM and glutamate, respectively a positive allosteric modulator (PAM) and an endogenous neurotransmitter it was recently stated that it could be a desensitized state[27]. The resting structure was crystallized[21] in the absence of ligand and is thus believed to be representative of a closed state of the channel.

- GLIC: (Gloeobacter Ligand gated Ion Channel) isolated from the bacteria *Gloeobacter violaceus*. It presents the peculiarity to be activated by protons, thus it is sensitive to changes in pH. Therefore, the active site is far less trivial to define as there can be many different located anywhere on the receptor.

1.1.6 *General organization*

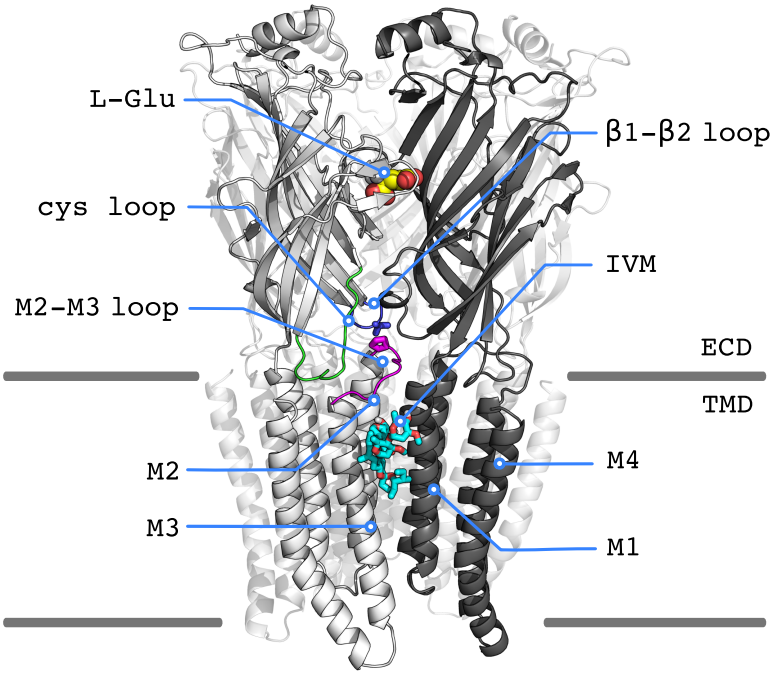


Figure 1.4: General presentation of a pLGICs (GluCl). Two of the five subunits are highlighted. Both ligands L-Glu (yellow) and IVM (cyan) bind at the interface between subunits. Some of the key segments of the protein discussed in this work are highlighted.

pLGICs are trans-membrane proteins with a molecular weight of about 290 kDa, comprising five identical or homologous subunits symmetrically arranged around a central ionic channel with a five-fold axis perpendicular to the membrane plane.
They are found in bacteria as well as across the full human body, from myocytes to neurons with a surprising similarity in their structure and sequence. The overall structure of the pLGICs is well known thanks to crystallographic structure. Each subunit consists in 3 well-defined domains.
First a large hydrophilic extracellular domain (EC), mainly composed of β-sandwiches sec-

ondary structures. The former is attached via linkers the lipophilic transmembrane domain (TM) composed in vast majority of four α-helices crossing the entire membrane slab. Finally the TM domain in some eukaryotic pLGICs extends in the cytoplasm of the cell in an intracellular domain of which the structure is much less known. In fact, its flexibility makes the resolution by X-ray crystallography more challenging. For this reason it is often removed from the crystallographic construct when present in the native protein.
As it is composed of five subunits it comprises five interfaces which host the orthosteric sites, located in the EC domain. Moreover several other binding sites, referred to as allosteric have been isolated at the interface between adjacent subunits, both in the TM and EC domains.
Moreover, pLGICs can either be homopentamers or heteropentamers in function of their composition in the different subunits types. Indeed, each subunit of the pentamer can be slightly different (subtype) which has an effect on the function of the receptor.
In the upcoming sections, we will discuss in details the structure and known function(s) of the different key segments of the pLGICs, highlighting when necessary the differences among the pLGICs.

1.1.7 *Extracellular domain*

It is the domain facing the outside of the cell. It is involved into neurotransmitter recognition and is key to understand the effect of neurotransmitters on pLGICs as it contains several important components such as the orthosteric binding site, the loop C and F and the cys-loop.

1.1.7.1 *orthosteric binding site*

No X-ray structure of the main pharmaceutical target, i.e. nAChR, was elucidated until the work of Morales-Perez and coworkers in 2016[1] who crystallized an heteromeric structure of the nAChR. Up to the first structures of human pLGICs, all the knowledge on the orthosteric binding site structure was extrapolated from the crystal structures of prokaryotic homologues such as GLIC, ELIC, eukaryotic such as GluCl or from electronic microscopy (EM) or even using the AChBP.
As the orthosteric site is capable of selectively binding a given neurotransmitter, its topology is of crucial importance. In fact small variations in the amino acids forming the binding pocket may change the selectivity from, for instance, glutamate to glycine.
The orthosteric binding site is lodged in the upper part of the extracellular domain, at the interface between two subunits. The subunit holding the A, B, and C loops is referred to as (+) or principal subunit, whereas the E, F, and G loops are carried by the (-) or complementary subunit (see figure 1.5).
The binding pocket is barely accessible to solvent and formed by the loops of the (+) subunit on one side and by the β-strands of the (-) subunit on the other.
In GluCl the binding of L-Glutamate to the orthosteric binding site forms a bridge between the principal and complementary subunits. L-Glu is tightly sandwiched between 2 tyrosines of the (+) subunit and two arginines on the (-) subunit. Despite the diverse sources of struc-

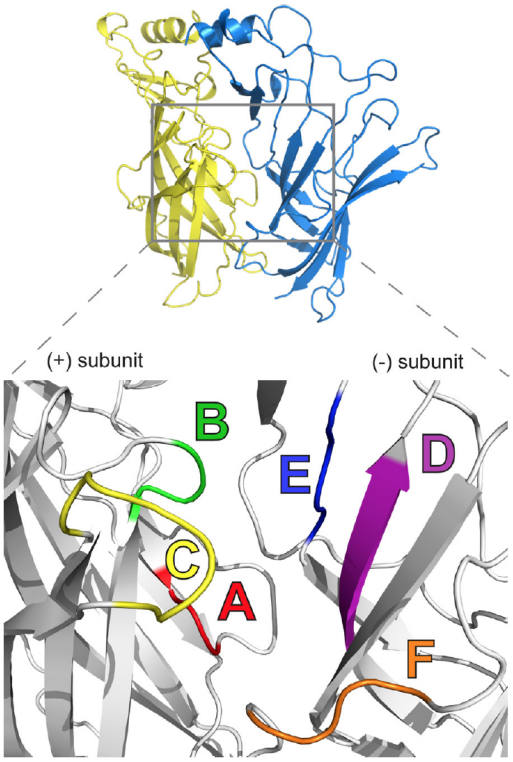


Figure 1.5: Description of the six loops forming the orthosteric binding site. Here is shown the Acetylcholine binding protein (AChBP) which is composed of only the EC domain. The two different subunits are shown in yellow and blue on the upper panel.(from ref[23])

tures, the orthosteric neurotransmitter-binding site of pLGICs seems to be overall remarkably conserved from bacteria to humans.

LOOP D AND E These two loops located at the back of the active site have been shown to play a key role in the desensitization of the receptor by varenicline. Indeed, by mutating only two residues located on these loops, it was shown that the effect of the varenicline (indicated in aid for smoking cessation) on the nAChR could be abolished [29].

LOOP F Located in the lower part of the orthosteric site, the flexible loop F is thought to undergo significant structural rearrangements upon ligand binding.

LOOP C The binding pocket is capped by a flexible segment of amino acids called loop C, which is stabilized in a closed configuration by interactions with the ligand placed in the active site. When no ligand is bound, the loop C is much more flexible and can sample a wide range of conformations going from completely open to completely closed. While in an open configuration, the active site is exposed to the solvent and can thus facilitate the spontaneous binding of a ligand.

CYS-LOOP The loop between the β6 and β7 strands is located near the interface between the TM and EC domains and is often referred to as cys-loop (see fig 1.4). It owes his name to the two cysteines it contains and which forms a disulfide bridge.
As shown by Jaiteh et al. [11], the two cysteines at the base of the cys-loop are missing in the bacterial homologues of the pLGICs family and are often replaced by a polar residue for the first cysteine and by a hydrophobic residue for the second. Surprisingly, several eukaryotic species also lack the two cysteines.
In fact, it is thought that the cysteine bridge present in most eukaryotic pLGICs was acquired later in evolution and then kept, suggesting the relative importance of this disulfide bridge.
Moreover, regardless of the presence or absence of the disulfide bridge on the cys-loop, the latter carries a very conserved proline (present in 100% of the sequences analyzed by Jaiteh et al. [11]) believed to be a key residue in the opening or closing mechanism of the pLGICs.
These results evidenced that this proline residue underwent a strong evolutionary pressure leading to a high conservation across the members of the family, both in bacteria and in human. Interestingly, the cysteines residues located on the cys-loop, after which pLGICs were initially named are less conserved. The latter would suggest renaming the Cys-loop in Pro-loop receptors for a more consistent nomenclature.

1.1.8 *Transmembrane domain*

1.1.8.1 *Allosteric binding sites*

Allosteric sites are defined as topographically distinct sites from the orthosteric or endogenous ones. They can potentially be anywhere on the molecule and can lead to various effects such as the positive or negative modulation but also activation or inhibition of the response

of the target protein.
Several allosteric sites have been identified from pLGICs' crystallographic structures. The first allosteric potentiator discovered for pLGICs was Ca^{2+} for $\alpha 7$ and $\alpha 4\beta 2$ nAChRs[30,31].

MD simulations has shown that an intrasubunit binding site located in the TM domain existed and was capable of binding cholesterol in the nAChR. In addition, a wide variety of molecules such as ethanol and other alcohols, general anaesthetics, lipids, cholesterol, neurosteroids, ivermectin and other synthetic compounds[22,32–37] can bind to these sites.
These binders typically have low intrinsic activity but can drastically increase or decrease the physiological answer expected from a given receptor. For this reason they are referred to as allosteric modulators, either positive (PAM) or negative (NAM).
The allosteric binding sites are located about 40 Å away from the orthosteric site, in the upper

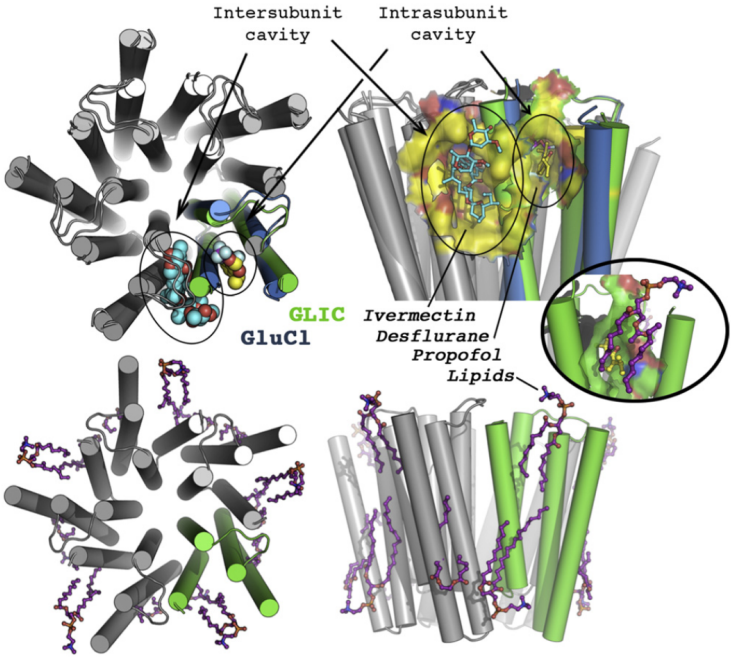


Figure 1.6: Representation of the different allosteric sites of the pLGICs.(from ref[22])

part of the TM domain. In fact, two major allosteric binding sites were identified, although they partially overlap. The first one is a cleft and is formed by the helix M1 of the (-) subunit and M3 of the (+) subunit, and hosts IVM and is thought to be the location where lipids from the membrane but also neurosteroid[38,39] or cholesterol[5] bind (reviewed in ref[40]).
Unlike the aforementioned site, the second one is lodged inside a subunit, is much smaller and was identified when bound to general anesthetics (GA). The effect of GA was investigated[41,42] long before the discovery of their binding site in 2011 thank to the elucidation of a structure of GLIC bound to propofol and desflurane[43].
Interestingly, a third potential allosteric site which binds a PAM named AM-3607 was discovered in GlyRα_3 in 2017[44]. Unlike the previously described allosteric sites it is located at the very top of the EC domain next to the binding pocket of glycine. Moreover and like the IVM

pocket it is place at the interface between adjacent subunits.
Finally, the resolution of a structures of a chimera GLIC/GABAAR[45] or GABAA[46] suggested the existence of a fourth allosteric binding site at the interface between subunits and located at the very bottom of the TM domain. This site was shown to be able to bind various neurosteroids which act as agonist on this receptor.

In conclusion, there exist a wide variety of binding sites and it is likely that many other are still to discover. Both computational and experimental methods are valid approaches to explore the allosteric pockets.

1.1.8.2 *The ion channel*

The physiological answer to ligand binding of pLGICs consists in the opening of the channel to let ions permeate through the cell membrane, which may lead to different endings such as excitation, inhibition, or metabolic signalling.
The ion channel is formed by five alpha helices, referred to as M2, oriented perpendicularly to the membrane plane and arranged symmetrically around the axis of the pore.
In order to compare the structure of pLGICs from different organisms and therefore with differences in the residues lining the ion pore, the former residues were numbered using a prime notation from -2′ (inner part of the ion pore, facing the cytoplasm of the cell) to 20′ (outer part of the ion pore, near the junction with the EC domain).
The ion channel of pLGICs consists in a hydrophobic domain in the middle of the ion channel, which is well conserved among the pLGICs and in two charged residues located at its edges. The ion channel's narrowest portion is referred to as the constriction point. In most pLGICs, i.e., GluCl, ELIC, GlyR, GABA and $5HT_3$ [47–49] a ring of five leucine residues pointing toward the lumen of the pore forms the gate of the resting or closed states in position 9′.
The inward or outward displacements of these residues are responsible for the closing or the opening of the channel to the ion flux, respectively. Indeed, previous experiments showed that if these key leucine residues were mutated to hydrophilic residues, the opening time of the channel was increased[50–52]. This could be explained by the fact that, the barrier that represents hydrophilic residues for ion permeation compared to that of hydrophobic residues is significantly lower and not sufficient anymore to close the ion pore. It is also possible that the replacement of the leucine residues by significantly smaller residues, i.e. a threonines, increases dramatically the diameter of the pore and therefore increases the opening time.
Even though the pLGICs family is homogeneous in terms of global architecture and in the overall composition of the ion channel, the selectivity for ions diverges. $GABA_AR$, GlyR and GluCl are anionic channels, when 5-HT_3R, nAChR and GLIC are cationic channels. The mechanism of the ion selectivity is not fully elucidated. It is however clear that unlike highly selective channels such as KcsA, the selectivity of this type of receptors lies in charge interactions and not in the ion radius. The residues thought to be responsible for the ion selectivity are called selectivity filter residues.

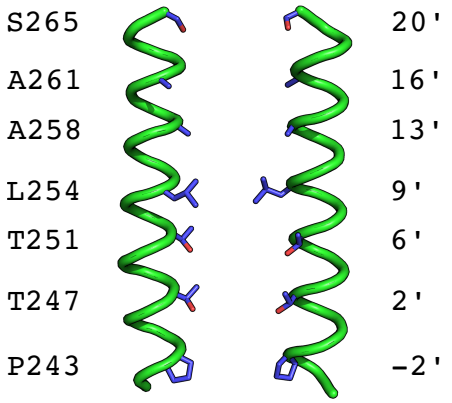


Figure 1.7: Representation of the ion channel. The prime notation used in pLGICs for the M2 helices and the corresponding residues shown for GluCl.

1.1.8.3 *The ion channel binding site*

In addition to the allosteric site aforementioned, pLGICs can bind small molecules in their ion channel. In fact, in order to control the activity of pLGICs, different strategies can be adopted. Instead of trying to impact the open and closing transitions triggered by the binding of the endogenous neurotransmitter, one can design channel blockers, which occlude the pore by binding directly inside the lumen.
The binding sites of channel blockers are widely spread over the full transmembrane part of the channel. A structure of GluCl bound to picrotoxin[2] has shown that this channel blocker binds between position -2′ and 2′ (see Fig 1.7). In addition, it was also shown that tetraethylammonium and tetrabutylantimony bind in the middle of the channel at position 6′ and that lidocaine binds near positions 9′ and 6′[53].
Interestingly, due to their mechanism of action, channel blocker can stabilize open structures, as well exemplified by the picrotoxin. Indeed, it was shown to stabilize an open but non conductive state of GluCl[2]. For this reason it is used in X-ray crystallography in order to stabilize open structures which could otherwise desensitize due to a prolongated exposure to the endogenous neurotransmitter.
In addition to their intrasubunit binding site facing the lipids in the upper part of the TM domain (see Fig 1.6) and due to their small size, GA such as bromoform[54] are also able to bind inside the ion pore, more precisely in the hydrophobic portion. It was also suggested that the GA were more likely to block the ion channel of cationic receptors as their pore is more hydrophobic than the anionic ones[55].

1.1.9 *Intracellular domain*

The inner domain of these transmembrane proteins is absent from the prokaryotic members of the pLGICs family and thus is much less known than the EC and TM domains. It was hypothesized to act as a selectivity filter allowing only for the permeation of specific ions[56].

Due to the flexibility of the intracellular (IC) domain, very few data exist on its structure. In

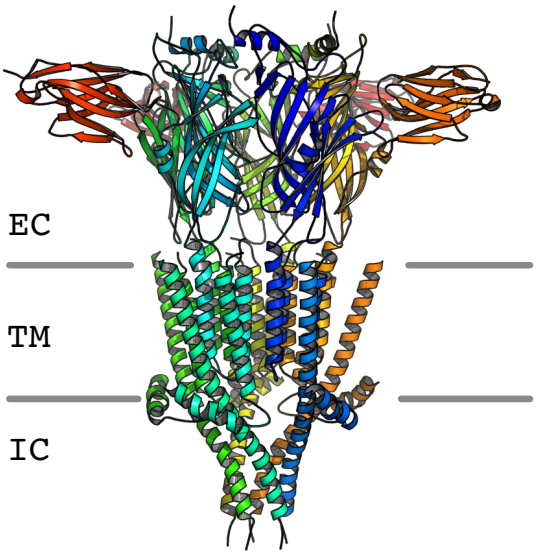


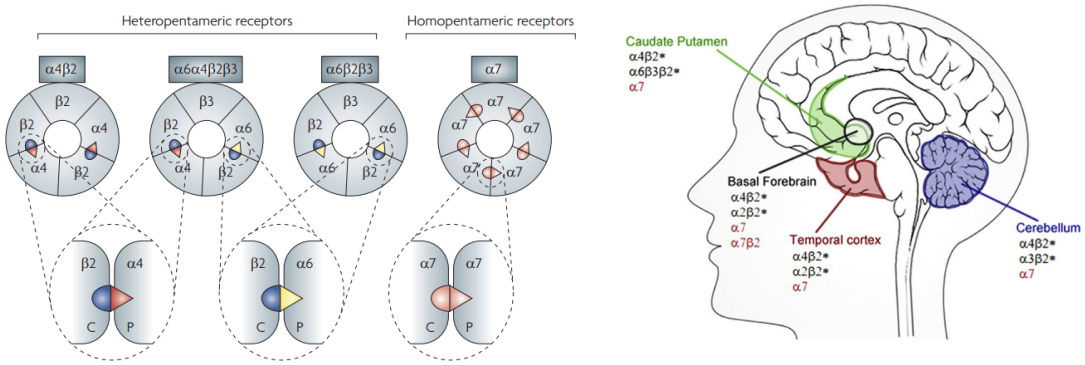
Figure 1.8: X-ray crystallographic structure of the 5-HT_3 receptor or serotonin receptor displaying a partial intracellular domain. Nanobodies (VHH15) binding at the level of the EC domain are also represented as they were used from the crystallization process. One out of five VHH15 was hidden to ease the visualization of the EC domain of the receptor. The IC shown is not complete as it could not be fully resolved by X-ray crystallography.

fact, in most of the few vertebrate structures present in the protein data bank (PDB) the IC domain has been cut out prior to the crystallization experiments.
In 2014, Hassaine et al. crystallized a 5-HT_3R with a partial IC domain, which showed that the IC domain helices formed a narrower constriction point that the one present at position 2′. Despite the interest of the IC domain and the recent data published on its structure, its function is still unknown and its absence from bacterial homologues makes the elucidation of its precise role very challenging. As a remark, a small part of the intracellular domain, the MX helix, was resolved by Morales-Perez et al. in the α4-β2 nAChR matching the corresponding helix in the strcture of the $5HT_3$R.

1.1.10 *Homomeric or heteromeric pentamers*

Homomeric describes the fact that the five subunits of a given receptor originate from the same subtype. Contrarily, heteromeric indicates that at least two different subtypes of subunits are found in a given pLGIC.
In mammals, nAChRs can be composed of different types of subunits, namely nine β-subunits and three α-subunits. They assemble into various compositions, for instance $(\alpha4)_2$ $(\beta2)_3$ or $(\alpha7)_5$, which impacts the pharmacological and physiological properties[57].
Indeed, upon their location in the human body or even in the different parts of the brain (see

Fig 1.9), nAChRs were shown to have different subunit compositions[58], which suggests that the subtype plays an important role on the function of the receptor.
Moreover, upon its composition in subunits, a given subtype of nAChR may bind from two to five neurotransmitters. Indeed only the interfaces between two α type subunits or between a β and an α subunit form a viable binding site for acetylcholine (see Fig 1.9a).
In addition, some pathologies are exclusively related to given subtypes of pLGICs, either because they affect areas of the brain where a certain subtypes of pLGICs are mainly located or because of the difference in the pharmacology of the subtypes.



(a) (b)

Figure 1.9: (a) Possible combinations of subunits for the nAChR illustrating the presence and absence of orthosteric binding site(from ref[59]). (b) Distribution of different subtypes of nAChR in the human brain.(from ref[60]).

1.2 PHARMACOLOGY OF THE PLGICS

In this section we will start by discussing the different pharmacological classes of ligands in general. Then we will be introducing the basic concepts of allostery and what is known to this date on the opening and closing or so called gating mechanism of the pLGICs. Finally we will focus on presenting some of the major diseases involving pLGICs and their treatments.

1.2.1 *Pharmacological classification of the ligands*

An agonist is defined as a molecule capable of binding to a receptor and which triggers a biological response, such as the cleavage of an intracellular messenger, or an ion flux through the cell's membrane. Several kind of agonists can be described:

- Partial agonist: even when occupying 100% of the binding sites the biological response triggered by the partial agonist is, in intensity, lower than the one obtained by the binding of the natural agonist. This class of agonist represents a major interest as it

allows for accurately modulating the response of a given protein. Unlike an antagonist it allows to decrease the intensity of the response without completely blocking it.

- Inverse agonist: when binding to the target receptor it triggers the opposite biological response to the agonist and should not be mistaken with the antagonist.

An antagonist is, on the other hand, defined as a molecule capable of blocking the action of the agonist. There are several types of antagonists:

- Reversible antagonist: ligands of that subtype can bind to their target (blood protein, nuclear receptor etc...) and unbind freely. The typical interactions involved in such cases are hydrogen bonds and Van der Waals interactions.

- Irreversible antagonist: here the ligands bind covalently to their target. The target activity will be restored as soon as the cell produces new ones. One may site as an example the action of aspirin on COX-1 and COX-2.

In order to quantify the potency, describe in a more rigorous manner the different classes of ligands and define some new subtypes, one can use the measure of the ligand efficacy which is defined as follows:

$$\alpha = \frac{[AR^*][R]}{[AR][R^*]} \tag{1.1}$$

where $[R]$ the concentration of the receptor, $[AR]$ the concentration of the complex ligand receptor. The * marks the fact that the receptor is in an active configuration.
Let us define K_A the association constant of the ligand to the inactive state of the receptor and K_A^* the association constant of the ligand to the active state of the receptor:

$$K_A = \frac{[AR]}{[A][R]} \qquad K_A^* = \frac{[AR^*]}{[A][R^*]} \tag{1.2}$$

from which one can write the following equations :

$$[R] = \frac{[AR]}{K_A[A]} \qquad [R^*] = \frac{[AR^*]}{K_A^*[A]} \tag{1.3}$$

when replaced into the former equation of α one obtains :

$$\alpha = \frac{[AR^*][AR][A]K_A^*}{[AR][AR^*][A]K_A} \tag{1.4}$$

$$\alpha = \frac{K_A^*}{K_A} \tag{1.5}$$

The latter expression of α clearly illustrates why when considering a natural antagonist (see Fig 1.14) which is defined to bind equally to the active and resting state of the protein $K_A^* = K_A$, $\alpha = 1$.
On the other hand, if the ligand is an agonist, which binds preferentially to the active state, $K_A^* >> K_A$ and therefore $\alpha >> 1$. As a remark, the partial agonism described here corresponds to a case in which the partial character of the agonist lies only in its binding constant being lower than the one of a full agonist. It should also be noted that partial agonism can

be reached through a different mechanism in which the binding constant is identical to the one of the full agonist but the molecular mechanism by which it activates the receptor leads to a decreased response of the receptor. In fact, this approach cannot describe such a case as it does not consider the physiological response to the binding explicitly.
Finally if the ligand binds preferentially to the resting state (antagonist) $K_A^* << K_A$ and therefore $\alpha \approx 0$

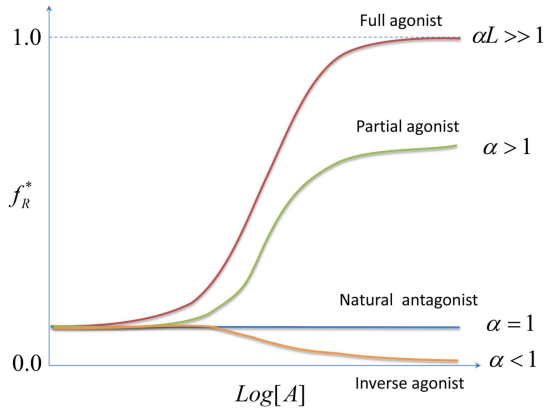


Figure 1.10: Different types of ligands and their associated alpha values. f_R represents the fraction of the receptor in an active state and $[A]$ the concentration of the ligand. (from ref[61])

1.2.2 *Allostery*

Fifty years ago from now, with the mention of the adjective "allosteric" by Jacques Monod and Francois Jacob[62], a new fields of interest was born. Allosteric comes from Greek *allos* meaning "other, different" and *stereos* standing for "solid". This concept was first brought to explain the experiments of Jean-Pierre Changeux, on end product of the enzyme L-threonine deaminase[63].
In this section we will first present the two models of allostery, KNF and MWC and then we will discuss the example of the allosteric two states model (ATSM) and its implications in term of free energy. Finally we will present a way to classify the allosteric ligands similarly to what was describe in the former section.
As it allows for an easier formalization of the allostery we will be discussing free energy profiles in this chapter. See section 2.2 for more details.

1.2.2.1 *Different models of allostery*

In 1965, the publication of a landmark paper by Monod, Wyman and Changeux[64] revolutionized the field of allostery by proposing the MWC model. A year later, the premises of the theory initiated by Pauling in 1935[65] were taken over by Koshland et al. to propose the KNF (Koshland Nemethy Filmer) model[66].

- MWC model (also referred to as conformational selection): in this description of receptor activation it is stated that both the resting and the active states pre-exist. This implies that the probability of populating a resting state in the absence of a ligand is different from zero, even though it might be very small. The binding of a ligand is, in this model, not responsible for the conformational changes but is the key that will lock or stabilize the pre-existing active conformation. It is assumed that the receptor is in dynamic equilibrium between the two conformations and that the ligand only selects one or the other, leading to opposite physiological outcomes. In this model the capability of being active or resting lies in the protein sequence and 3D conformation and the ligand only shifts the equilibrium between the two states. The discovery of rare but existing ionic currents even in the absence of ligand in GluCl strengthened the claims of this model.

- KNF model (also referred to as induced fit or population shift): here is postulated that the activation process is not a property of the protein but is due to the binding of the ligand. In fact, the latter triggers a conformational wave that goes from the orthosteric or allosteric binding site to the action site and changes the conformation of the protein. In this model, a resting state of the receptor cannot exist in the absence of a ligand.

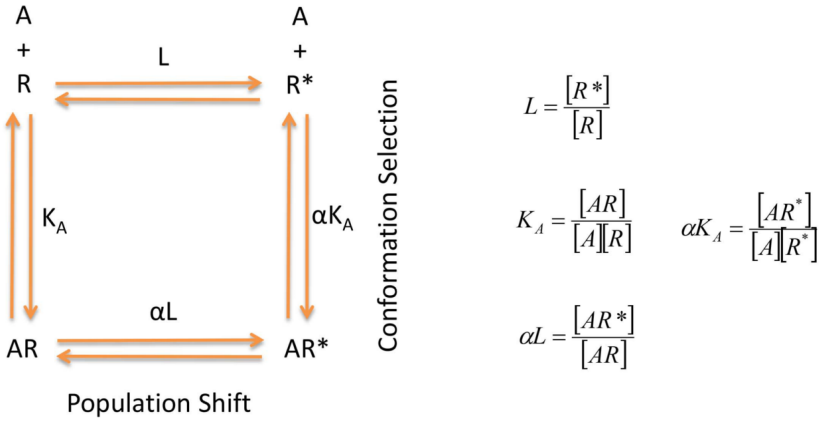


Figure 1.11: Illustration of the allosteric two states model and its measurable quantities in the context of the models aforementioned, i.e., KNF and MWC. (from ref[61])

In the context of the allosteric two states model (ATSM) where it is assumed that the receptor can only be active or inactive one can describe several quantities.
First of all αL is the equilibrium constant describing the population shift or KNF model. Indeed the receptor binds the allosteric activator as it is in a resting position and then activates. On the other hand αK_A describes the equilibrium between the active conformation not bound to a ligand on one side and the active conformation stabilized by the binding of the ligand. In this model, i.e., MWC, the ligand can only bind to the active structure and allow to stabilize this pre existing state.

1.2.2.2 *Classification of allosteric ligands*

Upon the shape or size or even the binding site of the allosteric ligands and like described in a former section for orthosteric ligands, allosteric ligands can be agonist or antagonist. Moreover, the allosteric ligands can also modulate the response to the binding of the orthosteric ligand. In other words, the allosteric modulators do not have any effect on their own as they only increase or decrease the response triggered by the binding of the orthosteric ligands. When they increase the response of the receptor to the orthosteric ligand they are referred to as positive allosteric modulators (PAM) and oppositely when decreasing it they are referred to as negative allosteric modulators (NAM).
It is crucial to stress the difference between a PAM and an allosteric agonist as it will be discussed in details later on in this thesis.

1.2.3 *Mechanism of action*

pLGICs are ion channels which open upon activation and allow for the passage of ions through the cell's membrane. They can populate three different functional states, i.e., active, resting or desensitized. As they are ion channels, an active state is defined as bound to the activator and with an open ion channel when a resting state is free from any activator and shows a closed ion channel.
Interestingly, the desensitization process, which concerns most of the neurological receptors, is characterized by a structure which is bound to the activator but displays a closed ion channel. The latter state is observed when the receptor is exposed to the activator for a sufficiently long time.
Pentameric ligand gated ion channels can be activated by two different mechanisms. The first one is the binding of the endogenous neurotransmitter to its site. It can trigger alone the activation or deactivation of the receptor by respectively binding or unbinding to the orthosteric site. Each receptor can bind its endogenous neurotransmitter plus some other small molecules at this site, e.g., L-glutamate for GluCl, acetylcholine and nicotine for nAChR, Glycine or taurine for GlyR, serotonin for the $5HT_3R$ and γ-amino-butyric-acid for the $GABA_AR$.
The second mechanism is referred to as allosteric potentiation and can be triggered by PAMs and NAMs. This category of ligand binds to different sites, for instance localized in the TM domain of the receptor.
There exist two possible ways to increase the current of ions going through the channel and triggered by the orthosteric agonist; i) the allosteric modulator can stabilize the active state which as for consequence to elongate the time for which it stays open, therefore delaying the desensitization of the receptor associated with the closing of the ion pore ii) enlarging the diameter of the transmembrane pore. The former gas for effect to increase the quantity of ions going through the cell's membrane per unit of time.
According to the MWC model described earlier and which seems to be the most suitable to describe the mechanism of action of the pLGICs, one can describe the following; the cycle of activation starts with a receptor in a resting configuration in which the orthosteric binding site is not capable of binding the ligand. The receptor then spontaneously go through a

transition populating the active state which is capable of binding the orthosteric ligand even though it is not yet bound. The binding of the orthosteric ligand then blocks the receptor in the active configuration leading to the creation of an ion flux across the cell's membrane and the physiological answer related. Finally, the receptor desensitizes and the orthosteric ligand unbinds.
Interestingly, the intervention of allosteric modulators in this cycle may delay the desensitization and therefore change the entire thermodynamic equilibrium between the functionally open and closed states.

1.2.4 *Gating mechanism*

The molecular rearrangement allowing for the transition of the receptor from an open to closed configuration is referred to as gating. Strictly speaking, the gating mechanism describes the activation of the receptor and the ungating its deactivation. Although, it is possible that both mechanisms are identical. The term gating will be used throughout this thesis to refer to both activation and deactivation mechanisms.
In the understanding of the gating mechanism lies the knowledge of how to modulate at will the response of these receptors. In fact a better understanding could lead to the possible discovery of new mechanisms to play with by various means ranging from mutations to ligand binding.

1.2.4.1 *Role of the loop C*

Based on the analysis of crystal structures of the AChBP[67,68], it was initially suggested that the loop C played a key role in the gating mechanism of the pLGICs. Later, it was hypothesized that the position of the loop C changed upon the pharmacological type of ligand, i.e., agonist or antagonist[69]. With the increasing number of available structures of pLGICs Brams et al. performed an extensive analysis of the loop C dynamics[70] from which they concluded that antagonist bound structures would depict a more open position of the loop C than the one bound to an agonist. These results were then confirmed by the work of Yu et al. on the glycine receptor on 2014[71], also suggesting that ligand bigger in size are likely to be antagonists and contrarily smaller being most likely agonists. However it seems likely that the position of the loop C is a consequence of the interactions between the ligand and the loop C and is not what triggers the opening or closing of the channel 60 Å away from the orthosteric binding site.
Finally, it was suggested by Yoluk et al.[6] that when bound to IVM the loop C of GluCl experienced more flexibility. The former suggests that not only the orthosteric ligand can have an impact on the flexibility or the position of the loop C. This peculiar, nonetheless interesting result gives a hint on a the possible coupling between allosteric and orthosteric sites.

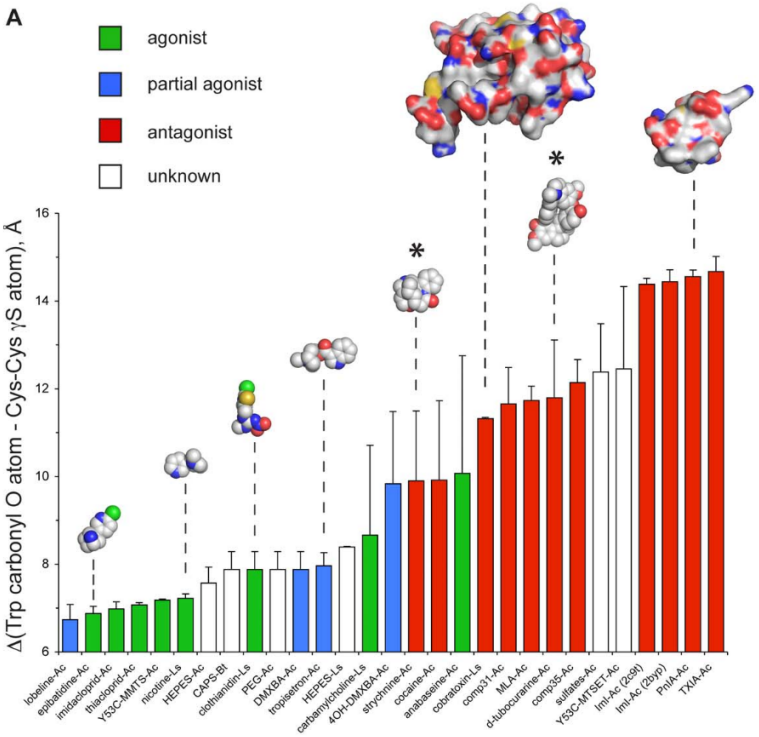


Figure 1.12: Correlation between the opening of the loop C and the pharmacological class of the various ligands (from ref[70]). One can clearly see that bigger ligands are most likely to be antagonist and smaller agonists.

1.2.4.2 *Role of the loop F*

Several experimental studies[72–78] have investigated the role of this loop in both ligand recognition and channel activation. Even though its role on the gating was subjected to debate, it was agreed that it had most likely no effect on gating and was rather involved in ligand binding and recognition. However, it was not shown capable of discriminating between antagonist and agonist binding to the orthosteric site.

1.2.4.3 *Twisting*

At first, the understanding of the gating mechanism was extrapolated from experimental studies such as electrophysiology and low resolution structures. It is only in a second time, when experimental techniques improved sufficiently to reach an atomic resolution and that computational studies came into play that the gating mechanism started to be better understood.

The existence of a twisting isomerization, which was initially described as a rigid body rotation of the EC and TM domain in opposite directions, during gating was first evidenced by normal mode analysis in 2005[9]. It was later then confirmed by the elucidation, thank to X-ray crystallography, of a structure of GLIC in 2009[19]. With the elucidation of several other structures of pLGICs such as GLIC pH4[43] or ELIC[28] (open channel), GLIC pH7 (closed channel)[79], GluCl open[2] and closed[21] the twist was confirmed to be of major importance in the activation process.

Moreover, at the time it was hypothesized that due to its nature, the twist was capable of changing the interfaces between subunits and therefore was probably involved in propagating the binding signal to the ion channel 60 Å away.

Thanks to the availability of high resolution structures, computational studies[3,6,7,12] brought a deeper understanding of the gating mechanism as they made possible to record and analyze the gating at the atomic scale.

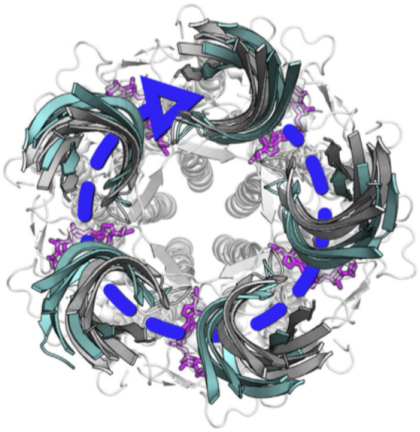


Figure 1.13: Top view of GluCl illustrating the twisting motion. (from ref[80])

1.2.4.4 *Role of the M2 helices*

The computational explorations of gating[3,12] allowed a better understanding not only of the twisting but also of the role of the M2 helices. Indeed an indirect mechanism of gating inferred on observations of partial channel closing started to emerge. The former stated that the twisting isomerization promoted the inward displacement of the M2 helices which were then able to close the ion pore. This very interesting hypothesis could not however be extensively verified as full closing could not be captured by simulation in any of the cited work.

1.2.4.5 *Blooming*

In 2013, Calimet et al.[3] inferred based on simulations of GluCl that a radial expansion of the EC domain or blooming was involved in the transition from open to close. Briefly after that Sauguet et al.[79] based on the observation of multiple structures of GLIC confirmed the observations made a year before, mentioning for the first time the blooming as an important part of the gating mechanism. In a recent review by Cecchini and Changeux[80], the blooming isomerization and its role in the gating mechanism were discussed in details. This motion is not mandatorily coordinated which means that all subunits can bloom with different kinetics. The presented model proposed that the radial contraction of the EC domain promoted the opening of the ion pore at the transmembrane level and was followed by the un-twisting of the EC domain over the TM domain locking the ion pore in an open configuration.

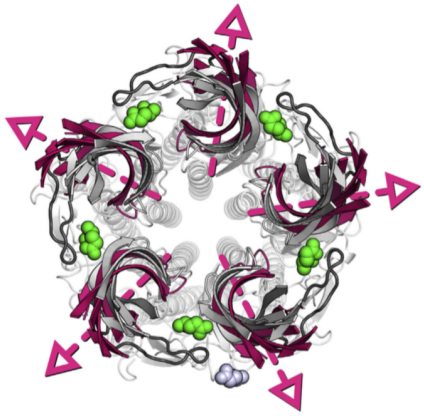


Figure 1.14: Top view of GluCl illustrating the blooming motion. (from ref[80])

1.2.4.6 *What is left to be understood*

When the work presented in this thesis started in 2013, only a few computational studies focused on the gating mechanism of pLGICs and even though X-ray crystallography is key in elucidating tri-dimensional structures it is rather limited to study the dynamics of proteins. Despite that, rather advanced hypothesis on the possible gating mechanisms were made. The twisting was already isolated and its function started to be understood. The blooming also was discovered and started rising interest regarding its possible implication in ligand binding and recognition or gating.

However, there still exist no clear description of the allosteric coupling between the ligand binding in the EC domain and the pore opening in the TM. The sequence of events leading to the opening of the pore even though it started to be sketched is not fully elucidated. A full closing of the ion pore was never captured making the exploration of gating very challenging if not impossible. The exact role of the blooming is also left to be understood. The possible hydrophobic gating mechanism described for some pLGICs must also be studied in details. In conclusion, endless questions remain to be answered regarding the gating mechanism of the pLGICs and more generally on how to control it to propose possible treatments to life threatening diseases.

1.2.5 *Modulation by membrane composition*

The composition of the surrounding lipids influences the pharmacological answer of pLGICs[57]. Indeed, in nAChR, decreasing the cholesterol concentration (below 15 to 20%) in bilayer membranes can lead to a sharp drop of conductance, which can be restored by addition of cholesterol analogs[81–83]. This mechanism, by which lipids and cholesterol activate or potentiate this family of receptors, remains unknown even though two main tracks are emerging; i) Ligands could directly bind deeply at specific allosteric sites lodged in the transmembrane domain, ii) their effect could be indirect by modifying the physical properties of the membrane. Nevertheless, other studies[84–87] have shown that in some cases, the complete removal of cholesterol can increase the response of nAChR to external stimuli. At the light of these contradictory results, one can conclude that the current knowledge on the modulation of pLGICs by lipids and by-products can greatly be improved and is of major interest.

1.2.6 *Associated pathologies*

In humans pLGICs can be located in cell types as different as muscle cells and neurons. Therefore they are found in the central and peripheral nervous system. As the members of the pLGICs assume key roles is maintaining homeostasis their healthy functioning is crucial. Pathologies involving ions channels are often refered to as channelopathies, name which also encompasses diseases which are not related to pLGICs, but to other types of ion channels.
When dysfunctionning, pLGICs are involved in diseases ranging from psychological to physical disorders.
In practice, these pathologies can either be due to; i) mutations of the amino acids sequence of the receptors concerned, thus changing the response to normal stimulation; ii)the lack or excess of these proteins at the surface of the cells' membrane.
In the upcoming section we will discuss the origins and the possible strategies of treatment for some of the most health-threatening and widely spread diseases involving pLGICs.

1.2.6.1 *Alzheimer's disease*

AD accounts for about 70% of the senile dementia in the world, it is thought to be characterized by accumulation of misfolded amyloid-β peptide in the shape of extracellular plaques,

	Torpedo[a]	Postsynaptic [b]	*Xenopus*[c]	Mammalian [d]
Phosphocholine (PC)	42	41	28	38
Phosphatidylethanolamine (PE)	34	35	17	22
PS + PI + PA[e]	16	18	10	12
Sphingomyelin (SM)	5	5	20	20
Other phospholipid	3	1	8	8
Saturated	42	49	44	53
Monounsaturated	26	15	39	20
Polyunsaturated	32	35	17	27
Cholesterol mole fraction	32	39	21	29

Figure 1.15: Composition of lipids for plasma membranes originated from various organisms from ref[14]. [a]*Torpedo California* electric organ[88], [b]Rat brain synaptosomes[89], [c]*Xenopus* oocytes [90], [d]"Idealized" (average) mammalian plasma membrane [91], [e]Sum of total phosphatidylserine, phosphatidylinositol, and phosphatidic acid. First group of rows indicates mole fraction of phospholipids among glycerophospholipid headgroups or sphingomyelin (SM); second group is distribution of fatty acyl chains. *Torpedo* and postsynaptic membranes have a higher abundance of PUFAs and also have PUFAs with longer chains and more double bonds. Final row has overall mole fraction of cholesterol, relative to the total number of phospholipids.

intracellular polymerization of tau protein, a loss of cholinergic tone and low levels of acetylcholine in the brain. The latter was hypothesized to be responsible for the cognitive decline observed in AD[92,93].

It was first postulated that the memory loss shown in patients with AD was due to a deficit in muscarinic receptors[94,95] mostly. Although, autoradiographic and histochemical of autopsy brain tissues [96–98] and brain imaging studies of AD patients lead toward a more specific loss of nicotinic receptors rather than muscarinic acetylcholine receptors. These data also show that muscarinic receptors including M2 receptors, are much less, if at all, reduced in AD.

To this date the only treatments commercialized for AD (galantamine, donepezil, rivastigmine) aim at increasing the level of acetylcholine in the synaptic cleft by inhibiting its degradation at the exception of the galantamine which in addition is a PAM of the nAChR. These treatments even though they help to slow down the progression of the disease are completely inefficient in juvenile or advanced type of AD for which there exist no treatment.

1.2.6.2 *Parkinson's disease*

PD is the second most common neurodegenerative disorder[99,100] after AD and plays a role on the central nervous system mainly affecting motor neurons in the brain. In patients with PD the level of dopamine in the brain is decreased. This is due to the abnormal death of the neurons generating dopamine in the part of the brain called *substantia nigra*.

For this reason one of the treatment of PD is the administration of dopamine which compensate the loss associated with PD. Unfortunately, after the patient experienced a short "honey moon" period where the symptoms are gone the treatment becomes less effective if not ineffective.
As an alternative to the administration of dopamine were developed inhibitors of the catechol-O-methyltransferase (COMT), enzyme responsible for the degradation of dopamine. As the inhibitors of AChE described for AD they allow for an increase of the dopamine in the synaptic cleft.
One can also mention the dopamine agonist which do not act on the pLGICs but rather on the dopamine receptors and are mainly used in early stages of the PD[101].
Finally, as the dopaminergic neurons inhibit the secretion of acetylcholine (ACh), their diminution in number leads to an hypercholinergic tone. To correct this problem in the regulation balance between dopaminergic and acetylcholinergic neurons anticholinergic drugs such as biperiden or tropatepine can be used.
Overall, even if some treatments are very efficient, they are only symptomatic and last for a given period of time. Therefore, it exists undoubtedly room for improvement in the management of PD.

1.2.6.3 *Schizophrenia*

Schizophrenia is a disorder principally characterized by cognitive deficits and disordered thoughts. It is manifested by hallucinations and paranoia. It can have either genetic or environmental etiologies. Schizophrenic brain tends to show a robust decrease in number of $\alpha 7$ nAChRs in the hippocampal region. Thus enhancing the response of nAChR to the binding of ACh using allosteric positive modulators could compensate the lack of $\alpha 7$ nAChRs.

1.2.6.4 *Addiction*

Tobacco use is one of the most important health issues in the world nowadays. Its use is increasing in the less developed countries and thus is responsible for an increase of the mortality. It was estimated in 1997 by the World Healthcare Organization (WHO)[102] that half of the adolescents who continued smoking throughout their whole life will die from smoking related diseases.
Tobacco, due to its content in nicotine, is very addictive. Indeed, 80% of the attempts to quit smoking fail within a year. The effect of nicotine is believed to be due specifically to the $\alpha 4\beta 2$ subtype of nAChR. The smoking cessation drug varenicline, acting on the nAChR as a partial agonist has been shown efficient[103]. In fact it binds to the receptor and creates a lower response than the one observed with nicotine. This has for effect to inhibit, without completely annihilating, the subsequent secretion of dopamine. Indeed, the lack of dopamine is associated with the withdrawal syndrome[104] and keeping a low level allows to prevent the former whilst it protects the patient from the pharmacological reward bound to the event of smoking. Designing novel allosteric modulators could lead to the development of alternative treatments possibly more receptor-specific and thus with less side effects.

Table 1.1: Recapitulation of the drugs commercially available or withdrawn from market targeting human pLGICs.

Target	ICD	Indication
nAChR	Fluoxetine, Sertraline, Paroxetine, Citalopram	Depression
	Targacept	Alzheimer's disease
	Varenicline	Smoking cessation
	Carbachol	Glaucoma
$GABA_AR$	Phenobarbital	Anticonvulsant
	Diazepam	Anxiety
	Zolpidem, Zopiclone	Insomnia
	Etomidate, Propofol	General anesthesia
$5\text{-}TH_3$	Tropisetron, Ondansetron, Granisetron, Dolasetron	Antiemetic
	Mitrazapine	Depression
	Alosteron	Irritable bowel syndrome

1.2.7 *Drugs acting on human pLGICs*

The potency and the specificity are two crucial properties to consider when designing drug binding to neuronal receptors. Ideally, one wants a drug to bind specifically to a given subtype of receptor and regulate accurately its physiological response. In fact, the design of allosteric ligands is a promising approach as they are mainly potentiators, i.e., they cannot activate the receptor on their own. Because the pLGICs endorse major basal functions in the maintain of homeostasis but also because they are involved in upregulation processes, their action should be carefully modulated.

The non-specific binding of the allosteric drugs used in human is responsible for most of the side effects. On the other hand, specific binding is also responsible for side effects as the potency of the drugs can be so high that it not only treat the physiological dysfunction but also perturb the healthy functioning of the receptor. In fact one wants to be able to control the potency of a drug accurately enough so that, when possible, only the main effect of the drug can be kept and the adverse effects minored.

The allosteric ligands can potentially bind anywhere on pLGICs but in an orthosteric site. In fact, due to the inferior evolutionary pressure on the rest of the protein compared to the orthosteric binding site, the variability in sequence is increased. The former provides a very good, but challenging, approach to design ligands that are specifically binding to given members of the pLGICs.

2

METHODS

2.1 MOLECULAR DYNAMICS SIMULATIONS

2.1.1 *Quantum mechanics*

Light has been described throughout history as both a particle and a wave. Among others, De Broglie has rationalized this dual description and has extended this concept to any particle. This has been ideally formulated with the quantum mechanics formalism, as particularly well exemplified by the time independent Schrödinger equation:

$$\hat{H}\psi = E\psi \tag{2.1}$$

Where $\hat{H}$, the Hamiltonian operator, contains all information related to the energetics of the particle(s) and is applied on ψ, the wave-function associated to the particle(s) and E the energy of the studied particle(s).

In quantum mechanics, the movement of electrons is explicitly considered, which allows to study chemical reactions (e.g., formation or breaking of covalent bonds). Schrödinger's equation can be solved analytically only for molecular systems containing one electron. Therefore, to tackle problems of major biological relevance such as protein functions or ligand binding energies, several approximations must be made.

2.1.1.1 *Born-Oppenheimer approximation*

Considering that the mass of electrons is negligible with respect to that of nuclei, the Born-Oppenheimer approximation[105] suggests to decouple the motion of electrons and nuclei. Among other consequences of this approximation one may consider that the energy can be split into two independent contributions:

$$E_{tot} = E_{electrons} + E_{nuclei} \tag{2.2}$$

2.1.2 *Molecular mechanics*

Due to the current computer's power limitations, the very accurate methods based on the quantum theory cannot be used for large molecular systems such as proteins. Molecular mechanics (MM), by definition neglects electronic motion, considers atoms as constant volume spheres, bonds as springs and obeys to Newtonian physics. These assumptions make MM much faster than solving Schrödinger's equation. Within the MM formalism, the potential energy of a system is written as the sum of several contributions: (i) bond stretching; (ii)

angle opening and closing; (iii) dihedral angle rotation; (iv) van der Waals interactions; (v) electrostatic interactions. The MM formalism can be used in molecular dynamics and to perform energy minimization of a given system.

2.1.3 *Molecular dynamics*

MD is a widely used method for studying problems ranging from material to biological sciences. It was first used by Alder and Wainwright in 1957[106] to study the phase transition of hard spheres and is now used to study systems of millions of atoms with a much more complex energy function. Unlike QM, particles are here treated as classical objects limiting the amount of information that one can obtain from them but decreasing drastically the computational cost. When molecular mechanics methods only allow for the calculation of static properties, molecular dynamics allows to study the evolution of a given set of particles and their properties along time by integrating Newton's second law of motion in discrete steps:

$$-\frac{\partial U}{\partial x_i} = m_i \frac{d^2 x_i}{dt^2} \tag{2.3}$$

where U is the potential energy, x_i and m_i are respectively the coordinates and the mass of the particle i.

To obtain a trajectory, i.e., the variation of x along time, we first compute the potential energy of the system and the resulting forces for all the particles. Next, the displacement is computed after a given Δt and new coordinates are assigned. This cycle is repeated as many times as required to accumulate data and create a longer trajectory.

The choice of an adequate Δt is crucial. Indeed, the applied forces on the system are assumed to be constant within an integration step. This has for consequence that one should set Δt so it accounts for the fastest degree of freedom of a system (1 fs), i.e., vibration of covalent bonds involving hydrogen atoms.

2.1.3.1 *Integrators*

The number of time the former cycle needs to be repeated is about 10^{11} if one wants to reach biologically relevant time scales. For this reason one must design integrators that do not accumulate errors, which makes the use of the simple Euler algorithm impossible. Unlike the former, the integrators commonly used by MD packages, i.e., velocity Verlet or Verlet[107] are reversible and allow for the conservation of energy along time. They also are computationally efficient and allow for a longer integration time step.

2.1.3.2 *Increasing the time step*

Computing the energy, forces and new coordinates every 1 fs is time consuming and fortunately methods have been developed to increase the time step. The SHAKE[108], LINCS[109] or RATTLE[110] algorithms allow to restrain the covalent bonds involving hydrogen atoms to allow for increasing the time step. Several other algorithms such as for example multiple

time scales r-RESPA[111], which allows to compute the electrostatics (computationally very costly) less frequently than the other interactions, or the addition of cutoffs to compute the electrostatics only below a given threshold can be used to maximize the efficiency without scarifying the accuracy.

2.1.3.3 *Thermostat and barostat*

Because of its formulations, MD is performed in the microcanonical ensemble (NVE), unlike experiments that are done in the canonical ensemble (NVT). Therefore, to reproduce *in vitro* conditions in molecular dynamics simulations one should introduce a thermostat to control the temperature all along the simulation. Among the most used, one may cite the ones of Langevin[111] or Nosé-Hoover[112,113]. In addition to the thermostat one should use a barostat to maintain the pressure at a target value during the simulation as for instance the one proposed by Berendsen[114].

2.1.4 *Force fields*

Molecular dynamics simulations are highly parameter dependent calculations, and therefore rely on the use of force fields (FF). A force field is built as the combination of a set of parameters (reference values for most common bond length, angle, dihedral angles etc ...) and an energy function which allows to compute the energy of a given configuration based on these parameters such as described bellow in Eq.2.5. Most force fields available use fixed point representation of charges but there exist more and more FF capable of handling polarizability by the surroundings such as for instance AMOEBA[115] (see Ref[116] for a detailed review). Even thought they may vary in some aspects, most fixed point-charges FF share common features. The Hamiltonian used to compute energy and forces is for instance in CHARMM[117–119] divided in two terms, i.e., the potential (V) and the kinetic energies (K):

$$H = V + K \tag{2.4}$$

where V is function of the atomic coordinates and K is function of the momenta of each particles. The potential energy term returns the interaction energy that is itself divided in bonded and non bonded terms as follows for the CHARMM force field :

$$\begin{aligned} V(r) &= V_{bonded}(r) + V_{nonbonded}(r) \\ V(r) &= \sum_{bonds} K_b(b-b_0)^2 + \sum_{angles} K_\theta(\theta-\theta_0)^2 + \\ &\sum_{torsions} K_\gamma(1+\cos(n\gamma-\delta)) + \sum_{impropers} K_\zeta(\zeta-\zeta_0)^2 + \\ &\sum_{Urey-Bradley} K_{UB}(c_{1-3}-c_{1-3}^{min}) + \sum_{\gamma,\psi} V_{CMAP} + \\ &\sum_{non-bondedpairs} \left\{ \epsilon_{ij}^{min} \left[\left(\frac{r_{ij}^{min}}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{ij}^{min}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\epsilon_0 \epsilon R_{ij}} \right\} \end{aligned} \tag{2.5}$$

where, r_{ij} is the observed distance between atoms i and j and r_{ij}^{min} is the distance corresponding to the minimum energy as computed by more accurate QM methods. q_i and q_j are the partial charges of atoms i and j, respectively, ϵ_0 is the permittivity of the vacuum, and ϵ is the relative dielectric permittivity of the system.
Unlike some others, CHARMM FF includes an additional term called improper dihedral. It is mainly used to control out of plan motions as it is constructed around a central atom referred to as branching atom which is bound to 3 others, and is therefore different from a conventional dihedral angle.
In addition to the improper dihedral, another unconventional term is used to restrain the out of plain motion. The Urey-Bradley term controls by an harmonic potential the distance between the atom 1 and 3, c_{1-3} of a given angle compared to a reference c_{1-3}^{min}. Unlike the improper dihedral term it is applied case by case based on observations made in QM to specific atomic organizations.
Finally the CMAP correction aims at improving the conformational properties of protein backbones[120]. It is based on *ab initio* QM calculations and corrects some of the systematic error of the protein backbone introduce by CHARMM.
Most of the force fields available use computationally expensive quantum calculations for the determination of bonded parameters such as bond length and angles. In fact, the major varying point among the different force fields is the way non bonded parameters are computed. Some will fit the parameters so they reproduce electrostatic potential obtained from quantum-chemical calculations (AMBER[121], CHARMM) when others will fit these parameters so they reproduce given thermodynamic properties (OPLS[122,123] and GROMOS[124,125]).

The choice of an adequate force field lies in finding the optimum balance between sampling efficiency and computational cost for a particular system. Indeed FF are specialized in different types of molecules therefore one should find the most appropriate for a given system.

Because they are computationally costly, all atom FF are nowadays limited to systems of a millions of atoms[126]. Indeed, if one wanted to simulate a full cell at the atomistic scale it would not be possible regarding the computer resources and technologies available. To overcome that limitation, coarse-grained (CG) models and the corresponding force fields were developed. They decrease the number of degrees of freedom of a system by grouping several atoms in a pseudo atom, also referred to as coarse-grained bead. The latter are parametrized so the properties of the bead reproduce at best the one of the group of atoms it contains. Among all the models available for CG simulations, one of the first and most used is MARTINI[127].

Despite their numerous advantages in term of sampling efficiency and system size, CG approaches do not allow to study specific interactions such as for example ligand binding. To benefit from the accuracy of all-atom models and the efficiency of CG models, multiscale techniques have been developed. There exist two different approaches to multiscale CG models :

- Use of various resolutions in different part of the system[128,129]. This method was successfully applied for instance to simulate the ion channel Gramicidin A at the atomic scale as the lipids and solvent molecules were treated by a CG model[130].
- Modification of the resolution in function of time. An example of the use of this technique is the study of the dynamics of the villin headpiece where the starting point for all-atom simulations were taken from a previous CG simulation of the same system[131]. Another approach, also known as resolution exchange[132–134] consists in changing the level of resolution, i.e., moving back and forth between CG and all-atom over time, to allow for a better barrier crossing rate, in the spirit of what is done in simulated annealing[135–137].

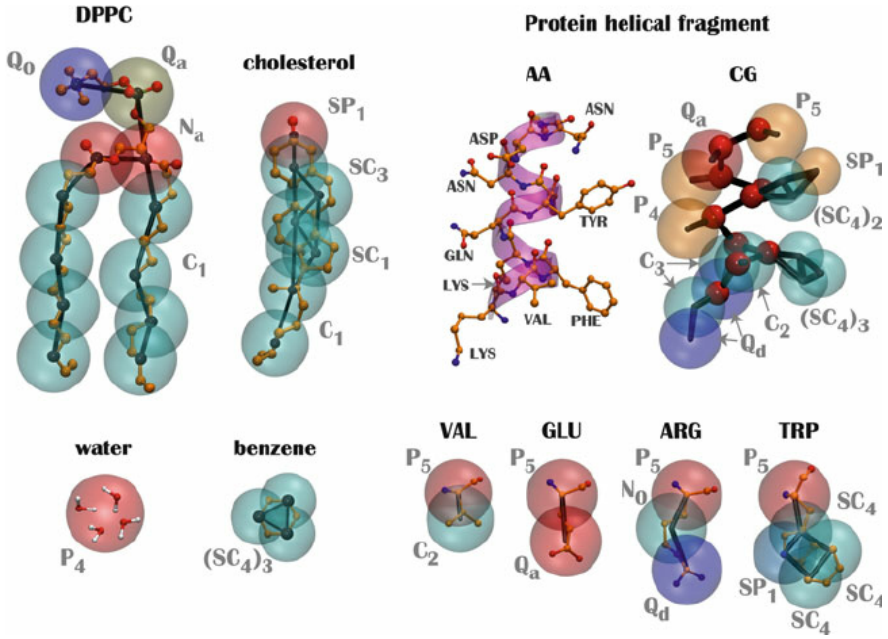


Figure 2.1: Representation of the mapping between all atomic structures of some example molecules and the corresponding CG beads (in transparent) for the MARTINI FF. (From Ref[138])

2.2 FREE ENERGY CALCULATIONS

2.2.1 *Definition and applications of the free energy*

Defined as the energy available to perform thermodynamic work and therefore different from the potential energy of a system, the free energy is a quantity of major interest in chemistry, biology and physics. Even though the absolute free energy is inaccessible for most systems of interest in biology, it can be calculated analytically for very simple systems such

as H_2 or butane.
The free energy $A(\xi)$ can be defined by the following equation :

$$A(\xi) = V(\xi) - TS \tag{2.6}$$

where $V(\xi)$ is the potential energy, T the temperature, S the entropy and ξ a collective variable describing the atomic coordinates.
Computing the free energy profile of a given system allows to assess major quantities such as the difference in free energy between two stable states (1 and 3) as represented on Figure 2.2 by ΔG_2. In fact, thanks to the Boltzmann distribution we know that the lower is the free energy, the higher is the probability of existence a given state, as illustrated by the following equation :

$$P(\xi) \propto \exp\left(-\frac{A(\xi)}{k_B T}\right) \tag{2.7}$$

From this we can conclude that the state corresponding to the lower basin and thus with the lower energy (state 1 on Figure 2.2) is the most stable as it is the one with the highest probability.
Moreover, the barrier between two states can also be measured from free energy calculations. This will give crucial indications on the transition rate between the two basins. For instance, in the case of a protein, such calculation allows to compute the time for the activation or deactivation of the protein. The transition rate $k_{2,1}$ describing the transition from state 2 to state 1 can be estimated thanks the Arrhenius equation as follows :

$$k_{2,1} = a \exp\left(-\frac{\Delta G_1}{k_B T}\right) \tag{2.8}$$

In addition, the transition from basin 1 to 2 displays a barrier which is significantly higher that the one from basin 2 to 1. In the context of a protein such observation would lead to the conclusion that activation and deactivation paths are not equivalent.
Finally, the shape of a given basin will give indications on the population that it describes. At equivalent free energy levels a wider basin is stabilized by entropic effect and a narrower by enthalpic effects.
Despite their very challenging complexity, free energy calculations have been intensively used to understand biological systems thanks to the development of various protocols. In general and regardless of the protocol used, free energy calculations can be divided in three main steps[139] :

- The choice of an adequate Hamiltonian
- The use of a sampling protocol, which generates an ensemble of configurations of a given system
- The choice of a mean to estimate the free energy differences

These three points will be discussed in more details in the upcoming sections.

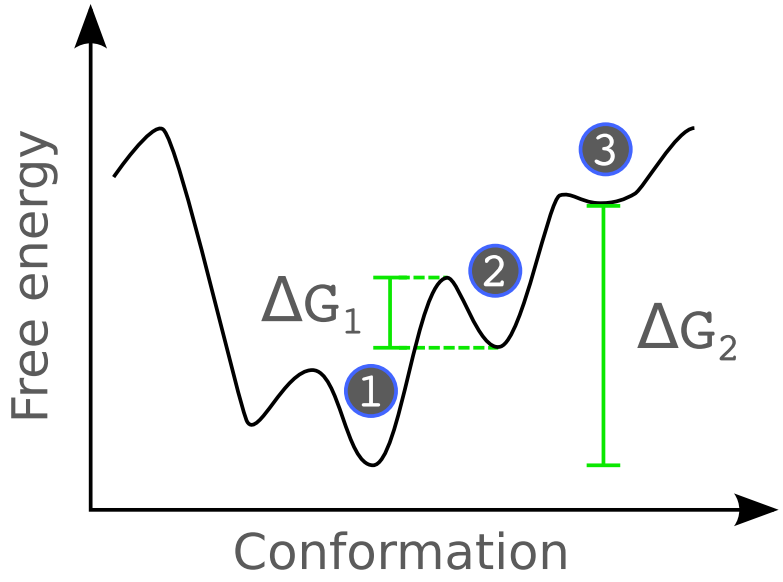


Figure 2.2: Bi-dimensional representation of the free energy as a function of the atomic coordinates. ΔG_1 represents the height of a barrier and ΔG_2 the difference between two states

2.2.2 *Enhanced sampling methods*

Sampling is the most time consuming step of a free energy calculation as it can require months of calculation for complex systems when using traditional methods such as molecular dynamics (see section 2.1). Indeed, when sampling in the canonical ensemble, the probability of sampling a given state decreases exponentially as its energy increases (see Eq 2.7). In order to compute absolute probabilities for a given configuration one needs to explore the full ensemble of possible configurations of a given system, which in the case of a system of, for instance, more than 100 000 atoms is challenging if not impossible. Although the sampling of the most important configurations, i.e., the ones with the highest contributions, might be enough, in practice it is often made impossible by high energetic barriers between states. Thus, one must think of different ways to sample the full conformational space.
Enhanced sampling methods propose a solution to the lack of sampling by various means. They can either try to modify/flatten the potential energy surface or force the systems to sample higher energy configurations either by using harmonic restrains or by temperature variations[140–142].

2.2.2.1 *Umbrella sampling*

Most of the biological processes occur within time scales ranging from microseconds to minutes. With the computer resources available nowadays it is not possible to reach such sampling time for system of thousands of atoms. Umbrella Sampling (US) is based on the concept that if one cannot observe a phenomenon in a computationally reasonable amount of time, one can force the system, by applying a harmonic restraint, to undergo the expected

transition and observe the variation of its free energy. The applied restraining potential is of the following form:

$$W(\xi) = \frac{k}{2}(\xi - \xi_0)^2 \tag{2.9}$$

where ξ is the current value of the variable restrained, ξ_0 the target value set by the user and k the force constant. This method consists in dividing the space of an observable of interest in so called windows. Each window will be assigned a different target value of the observable and the harmonic bias will restrain the system around this value. $W(\xi)$ basically quantifies how strongly the system wants to move away from the reference values. When running US calculations one should thus set several key parameters such as the number and the width of each window as long as their respective centers (or target values). Unfortunately these choices often rely on trials and errors, making the setup of US calculations somewhat complex. One of the biggest limitations of the US method, and more generally of all the collective variable (CV) based enhanced sampling methods, is that it requires *a priori* knowledge of an observable to bias. Indeed, in practice, other methods to explore the conformational space such as for instance accelerated MD or targeted MD are used to isolate one or several observable of interest on which US calculations will then be run.

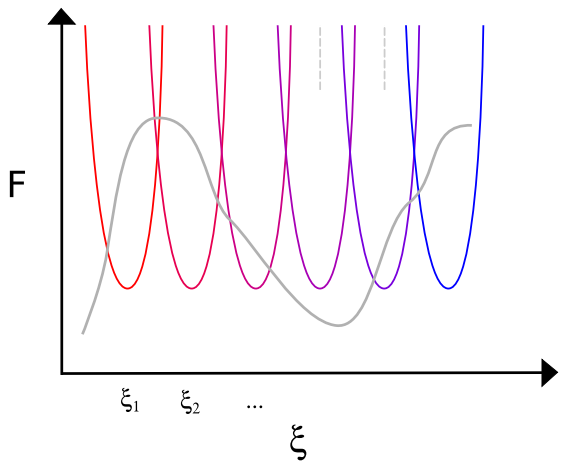


Figure 2.3: Schematic representation of the harmonic potentials applied along the reaction coordinate ξ. The gray curve represents the evolution of the free energy along ξ.

2.2.2.2 *Replica exchanged based methods*

The replica-exchange (RE) based approaches are widely used to increase sampling. The exchanged quantity may vary as we will illustrate with the two following examples. On the one hand, a major limitation of these methods is that all simulations must be ran in parallel to allow for the exchange between replicas, which increases dramatically the computational effort required. On the other hand, this technique allows to spread events of barrier crossing over the different replicas making for a faster sampling of the configurational space.

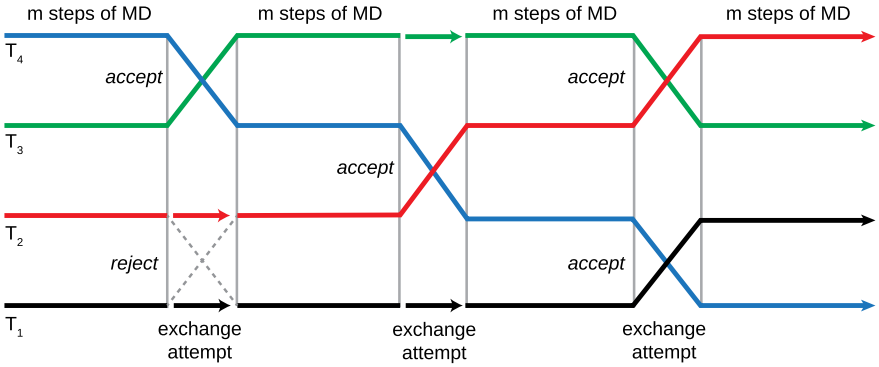


Figure 2.4: Schematic representation of the replica exchange algorithms (from ref[143]).

TEMPERATURE REPLICA EXCHANGE This method was initially developed by Sugita et al.[144] to tackle the problem of protein folding. It consists in running in parallel n simulations at different temperatures, which can exchange their temperature along time. The exchange between two replicas is accepted or rejected based on the Metropolis criterion. A major limitation of this method is that FF are not designed to work outside a physiological range of temperatures. Thus, if the chosen range is too large the system might experience non physical behaviors.

BIAS REPLICA EXCHANGE The quantity exchanged can also be the center of the bias between two umbrella sampling windows. Like temperature RE, the exchange is accepted based on a Metropolis criterion. This method has been shown useful to increase the sampling of the restrained degree(s) of freedom and of orthogonal degrees of freedom when running US[145–147].

2.2.2.3 *Metadynamics*

Metadynamics (MetaD) belongs to the class of non Boltzmann sampling methods. As most methods it aims at reducing the dimensionality of the conformational space. To do so the user defines one or several collective variables, which ideally describe the transition of interest. In order to sample high energy conformations, a bias that takes the shape of a sum of Gaussians (see Equation 2.10) is added on-the-fly to the collective variable(s). This has for effect to fill the potential energy wells with so called "computational sand" that prevents the system from visiting again the same state. Unlike US (see section 2.2.2.1) metaD uses a history-dependent bias. This technique allows for visiting more efficiently the full conformational space and especially possible metastable states since it facilitates barriers crossing. The time dependent bias applied to the potential energy in metaD is of the following form :

$$V_b(s,t) = w \sum_{t=\tau,2\tau,\ldots} \exp\left(-\frac{|s - s(t)|^2}{2\sigma^2}\right) \tag{2.10}$$

where σ is the width of the Gaussian and w its height, τ is the Gaussian deposition stride and s is the vector containing the values taken by the collective variable(s).

A major pitfall of this method is that it is highly parameter-dependent and even though there exist rules of thumb to give an initial guess to these parameters[148,149], one should proceed by trials and errors to find the best trade-off between accuracy and computational cost.

One of the tunable parameters is the height of the Gaussian. If chosen too big, the resulting free energy profile will be rugged and with high error values as the energy basin will be filled too rapidly. If chosen too small, the free energy profile obtained will be accurate but the time required to fill the basins will drastically increase making the calculation more costly. Moreover, the time between two additions of Gaussians should be wisely chosen as one must consider that after introducing a bias the system needs some time to relax before another can be added. On the other hand, if Gaussians are added too slowly the gain of speed that this method provides is decreased. Because of its formulation, a MetaD simulation never converges as Gaussian repulsive terms added every τ move the system away from equilibrium, making the monitoring of the convergence of such calculations challenging.

2.2.2.4 *Well-tempered metadynamics*

To solve the convergence issue rising in metaD, well-tempered metaD[150] was developed. It is different from the former method as the biasing potential added along the simulation is scaled down along time by a specific factor to prevent the system from visiting non physical high free energy regions. The height of the Gaussian repulsive terms is rescaled according the following formula:

$$w(t) = w_0 \exp\left(-\frac{V_b(s,t)}{k_B \Delta T}\right) \tag{2.11}$$

where $w(t)$ is the current height of the Gaussian, w_0 is the initial value and ΔT is an input parameter with the dimension of a temperature. When $\Delta T \to +\infty$ it corresponds to running normal metaD, thus the advantage of well tempered is canceled. On the opposite, when $\Delta T \to 0$ the system undergoes normal MD. For these reasons well-tempered metaD is more often used than metaD.

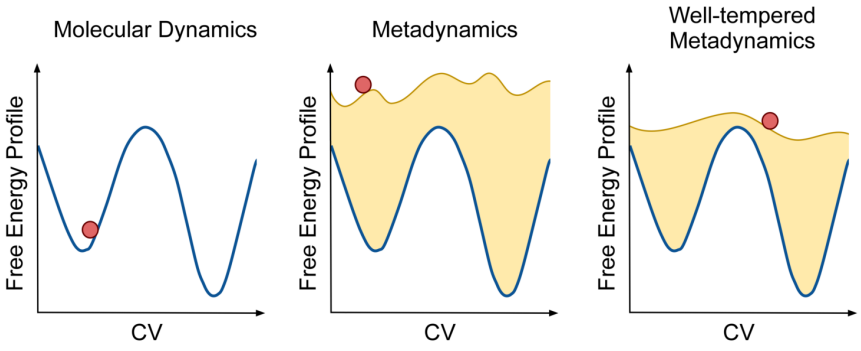


Figure 2.5: Schematic representation of the three different algorithms described in this section, i.e., MD, metaD and well-tempered metaD (from Ref[151]).

2.2.2.5 *Accelerated molecular dynamics*

Also known as aMD[152], Accelerated Molecular Dynamics belongs to the class of CV independent biasing methods since it acts directly on the potential energy, unlike the methods described above. It has been shown to sample the full configurational space more efficiently than conventional MD and to converge to the true canonical ensemble. Unlike CV based methods, aMD does not require any prior knowledge of a given transition and allows in theory to sample the entire configurational space. For this reason it is often used prior to docking studies to generate ensemble of conformations[153]. This technique allows to sample according to a modified potential of the following form:

$$V^*(r) = V(r) + \Delta V(r) \tag{2.12}$$

where $V(r)$ is the true potential and $\Delta V(r)$ is the boosting potential described as follows:

$$\Delta V(r) = \frac{(E - V(r))^2}{\alpha + (E - V(r))} \tag{2.13}$$

In aMD, the potential energy surface is flattened according to a α parameter when value is bellow a user defined threshold E. The choice of a value for these two parameters is arbitrary and based on observations made on, for instance, prior free MD simulations. E should be greater than the minimum energy observed in the true potential energy surface otherwise one is performing classical MD. In practice the biasing potential can be added to the dihedral term of the potential energy, the total potential energy or to both. Biasing only the dihedral term of the potential energy allows for faster calculations and is correct as we assume that state transitions in proteins mainly involve dihedral angle rotations.

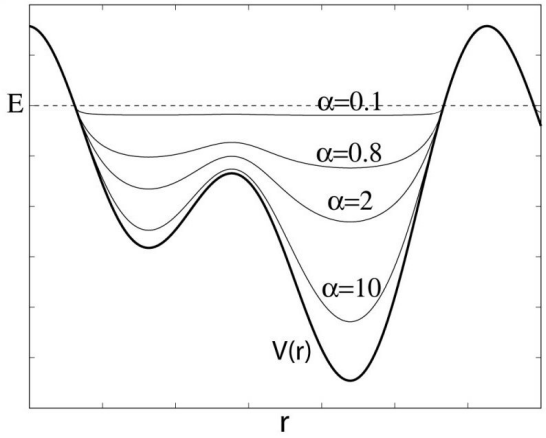


Figure 2.6: Schematic representation of the algorithm of accelerated MD where α is the acceleration factor, $V(r)$ the true potential and E the threshold energy (from ref[154]).

2.2.2.6 *Reweighting methods for biased probability distributions*

Running biased simulations using either US or other biasing methods such as metadynamics[155] adaptive biasing force (ABF)[156], or Accelerated Molecular Dynamics[152] allows to compute biased probabilities of a given quantity.

WEIGHTED HISTOGRAM ANALYSIS In order to unbiase the biased probabilities obtained in US and recover the unbiased free energy along a given RC (PMF), one may use an automated and reproducible procedure based on the algorithm WHAM developed by Kumar et al.[157]. Using umbrella sampling to explore the energetics of a system along a RC, one runs several restrained simulations called windows i and each of them experiences a different biasing potential $U'(\xi)$. The unbiased free energy $A(\xi)$ of a given window can be expressed as follows:

$$A(\xi) = -k_B T \ln P'(\xi) - U'(\xi) + F \tag{2.14}$$

where $P'(\xi)$ is the biased probability to observe the value ξ, k_B the Boltzmann constant and F an unknown constant referred to as offset. The latter cannot be easily measured unlike $U'(\xi)$ and $P'(\xi)$, which can be computed directly for every window.
By iteratively solving the following system of two equations until self-consistency is reached, WHAM allows to find the best F and $A(\xi)$ values.
Let us define the biased probability of finding the system at a value $\xi = \xi_j$ in a window i:

$$\rho_i^{(b)}(\xi_j) = \frac{n_i(\xi_j)}{n_i} \tag{2.15}$$

where j is the index of the bin and ξ_j the value corresponding to the center of the bin, $n_i(\xi_j)$ is the total number of occurrences of the value ξ_j in window i, and n_i is the total number of sampled frames in windows i.

$$\begin{cases} \rho(\xi_j) = \dfrac{\sum_{i=1}^{N_{wind}} n_i \rho_i^{(b)}(\xi_j)}{\sum_{i=1}^{N_{wind}} n_i e^{\beta\left(F_i - U_i'(\xi_j)\right)}} \\ F_i = \sum_{j=1}^{N_{bin}} \rho(\xi_j) e^{-\beta U_i'(\xi_j)} \end{cases} \tag{2.16}$$

where N_{wind} is the total number of windows and i the windows index, N_{bin} is the number of bins, j the bin index and $\beta = \dfrac{1}{k_B T}$.
The self consistent solvation of the former equation lies on several steps. First of all, when the number of iterations j is equal to zero, the user defines a guess value for the so called offset, i.e., F_i. This value is then placed into the first equation of the system and $\rho(\xi)$ calculated. The latter is then used to calculate F_i using the second equation of the system. The guessed calculated values of F_i are then compared by calculating their difference ΔF_i. Then two situations arise; i) if $\Delta F_i \leqslant$ user defined threshold, the iterations stop as the algorithm has reached convergency, ii) if $\Delta F_i >$ user defined threshold a new cycle starts with the guessed F_i replaced by the value calculated at iteration $j-1$.
The RC chosen is divided in n bins within which the PMF is assumed to be constant. This implies that the more the bins the better the resolution of the final PMF, assuming a sufficient number of events per bin. This method also implies that the probability distribution of the

RC corresponding to a given window should overlap with its neighbors so WHAM is able to accurately reconstruct the full PMF.

GENERALIZED WHAM This approach is comparable to the one of MBAR (Multistate Bennett Acceptance Ratio)[158] and was proposed initially by Souaille and Roux[159] and further detailed by Bartels[160]. It is generalized in the sense that the bin size tends to zero. Like WHAM it will consist in solving iteratively the two following equations until self-consistency is reached.

$$\begin{cases} \frac{1}{p_i^t} = \sum_j N_j f_j \exp(-\beta U_j(\xi_i^t)) \\ \frac{1}{f_j} = \sum_i \sum_{t=1}^{N_i} p_i^t \exp(-\beta U_j(\xi_i^t)) \end{cases} \tag{2.17}$$

where p_i^t is the weight of a given configuration, N_i is the total number of configurations in window i, f_j an unknown constant, X_i^t a given configuration of window i, and U_j the energy of the bias in window j. Unlike WHAM this approach allows to obtain the statistical weight for each configuration of a set of trajectories. In theory the weights obtained through this method can be used to compute the PMF along any RC, assuming that the sampling is sufficient in the chosen RC space.

2.2.3 *Potential of Mean Force*

A potential of mean force is described as the variation of the free energy along a given RC, in other words it represents the projection of the free energy profile on a given reaction coordinate. It can be expressed as follows:

$$F(\xi) = -k_B T \ln p(\xi) + cste \tag{2.18}$$

where ξ is a reaction coordinate. The constant is usually taken as the opposite of the minimum value of the PMF. A PMF allows one to isolate possible states and to measure barriers between them. Having access to the latter allow for instance to compute transition rate between states.

2.2.4 *Convergence of PMF calculations*

The convergence of these calculations was assessed from different approaches.

OBSERVABLE DISTRIBUTION The homogeneous coverage of the full range of values of interest of the studied observable was visually checked. It was also verified that adjacent windows displayed a sufficient overlap to allow for an accurate reconstruction of the PMF by WHAM.

DISTANCE TO REFERENCE PMF The sampled data sampled within a given window was split into N cumulative bins. The PMF was then computed for each bin and the average distance to the reference PMF (last) was computed. It was assumed that the PMF calculations were converged once the $\langle \Delta PMF \rangle$ went bellow a given threshold.

$$\langle \Delta PMF \rangle = \frac{\sum_{k=1}^{K} (PMF(n,k) - PMF(ref,k))}{K} \tag{2.19}$$

where n is the index of the cumulative bin, K the number of bins used to compute the PMF and k the index of each PMF bin.

2.3 OBSERVABLE OF INTEREST

A number of observables were designed to best monitor the transition from open to close of the pLGICs or to assess the convergence of the performed calculations. The one which are not trivial to compute or that requires a significant computational effort were implemented in the trajectory analysis software WORDOM[161].

2.3.1 *Twist*

The global twisting (τ) was evaluated per subunit and defined as the angle spanned by the projections of the geometrical centers of the EC and the TM portions of each subunit on the pseudo-symmetry axis of the receptor. Geometrically, this angle measures the torsion of the EC domain relative the TM domain of the receptor around the pore axis. For the analysis, receptor twisting per snapshot was evaluated by averaging over the twist angle of its five subunits. τ is defined as positive in the clockwise direction.

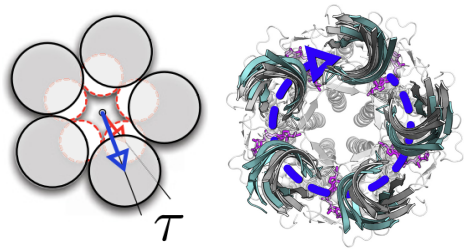


Figure 2.7: Schematic representation of the twist angle of a given subunit on the left hand side and superimposition of two structures with a different twisting angle is shown on the right hand side (from Ref[3]).

2.3.2 *Polar and azimuthal tilting*

The tilting of both the EC domain and TM domains can be measured per subunit and decomposed into polar (θ_p) and azimuthal (θ_a) components. These two angles were measured in the reference frame of each EC subunit with the Z-axis perpendicular to the plane of the membrane and the X axis pointing outwards along the radial direction. Denoting by $\vec{v}$ the vector defining the principal axis of the EC subunit, the polar (radial) tilt was measured as the angle between the Z-axis and the projection of $\vec{v}$ on the XZ plane, whereas the azimuthal (tangential) tilt as the angle between the Z-axis and the projection of $\vec{v}$ on the XZ plane.

Similar to the twisting angle, the polar and azimuthal tilt used for the analysis correspond to averages over the five subunits per snapshot.

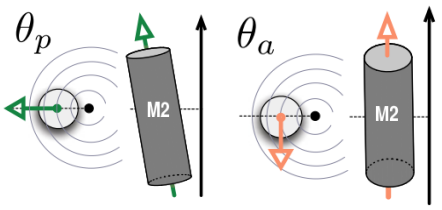


Figure 2.8: Schematic representation of the polar and azimuthal tilt angle for the an alpha helix (from Ref[3]).

2.3.3 *Pore opening*

Pore opening is a crucial feature of ions channels since it controls their function. Therefore, one should carefully design accurate observables to quantify the opening. The latter was probed by measuring the pore radius at the constriction point (residue 9′ , LEU 254) by HOLE[162] or the C_α-cross section at position 9′ (σ) using a simple geometric definition. The latter is referred to as the ion-pore size throughout the text. Unlike HOLE, σ is only computed for given positions of the ion pore, i.e., -2′, 2′, 9′, 13′. It consists in measuring the five distances between the C_α of two adjacent subunits. The averaged distance is then computed to measure the surface of a regular pentagon. In some specific cases, i.e, when the pore is not symmetric anymore, this approach can be inaccurate. Indeed, in the case in which the pore has an elongated instead of a round shape, σ will overestimate the size of the pore. By visually checking the simulations we estimated that the overall shape of the ion pore at the constriction point remained spherical all along the transition, therefore making the use of the previously described approximation valid. Compared to HOLE this observable allows to study the variation in size of a given section of the ion pore on very long trajectories with a reasonable computational effort. In fact because the HOLE algorithm includes a Monte Carlo exploration at every slice of the pore it is much slower than computing σ and often impractical to use on μs long trajectories.

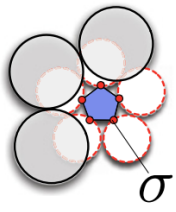


Figure 2.9: Schematic representation of the C_α-cross section (from Ref[3]).

2.3.4 *Pore hydration*

The presence of water molecules and ions were probed by counting their number within a cutoff distance of 2 and 5 Å, respectively, from the constriction point (9′). The latter provides an orthogonal and not structure-based measure of pore opening. This analysis was done using the toolbox of VMD 1.9.

The pore hydration was also monitored by measuring the dehydrated stretch at the constriction point. The latter was defined by the length of the void between to limit values, which were arbitrary defined as a hundredth of the maximum water density (bulk).

The water distribution was also studied by projecting the coordinates of all the water molecules of the system on the Z-axis at each frame has shown by Trick et al[163].

2.3.5 *Flux*

We developed an algorithm to efficiently monitor the flux of particles such as water molecules or ions through the pore. No distinction is made between the inward and outward flux, the net flux outputted by the program being the sum of the two. The same algorithm can be applied to any selection, e.g. water, ions, etc... The algorithm counts the number of particles that goes from an upper compartment to a lower compartment through the channel or the opposite. This algorithm was designed to prevent from counting the particles going from one compartment to another through the periodic box. An option was also added to compute the flux only in a given cylinder in space in which the ionic pore is contained. Indeed, rare events of permeation across the lipid bilayer and not inside the ion pore may be observed and thus bias the results.

2.3.6 *Proline position*

In order to monitor the position of the P268 conserved residue located on the M2-M3 loop (see Figure 1.4) the coordinate of the C_α of each of the five P268 residues were projected on the membrane plane. To study the inward motion along time of the proline, the distance between the center of geometry of a given proline to the center of the pore was computed for all five prolines.

2.3.7 *ΔZ dist*

It is the distance between the C_α of residue V45 located on the $\beta1 - \beta2$ loop and P268 on the M2M3 loop projected on the Z-axis. This distance was chosen because it measures the vertical displacement of two important segments of the protein ($\beta1\beta2$ loop and M2M3 loop) located at the interface between the TM and EC domains. We chose to project this distance on the pore axis because we were mainly interested in the vertical displacement and its possible correlation with the twist.

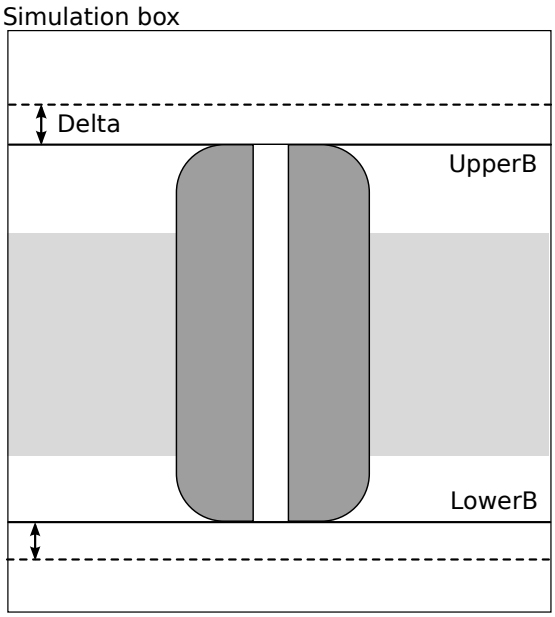


Figure 2.10: Schematic representation of the flux parameters in the case of an ion channel. The gray slab represents the membrane and the central dark shape the channel. The upper and lower borders are defined by the user as is the delta parameter. The latter allows to filter transition events going through the pore and not through the periodic box.

2.3.8 *Number of L-Glu bound*

The number of L-Glutamate bound was computed with two different approaches.

NATIVE POCKET BINDING The interface of two adjacent subunits forming the orthosteric binding site were superimposed to their initial positions to remove the possible rotation or translation of the protein. Then the RMSD, from its initial position, of the corresponding L-Glu was computed. The RMSD threshold to determine if a ligands was bound or unbound was set to 10 Å.

SYMMETRY CORRECTED REBINDING GluCl has five homologous binding site for L-Glu and looking at the simulations we observed spontaneous rebinding events to orthosteric pockets different from the crystallographic ones. For instance, the glutamate binding at the level of the orthosteric site formed by the subunits A and B could rebind to the homologous site formed by subunits C and D. The previously described algorithm to compute the number of L-Glu bound is not designed to detect possible rebinding to a different orthosteric site, i.e., not the interface in which it was bound in the crystal structure. To account for that, we also computed the RMSD of each ligand to all the other pockets, after alignment and set that the ligand was bound whenever the symmetry corrected RMSD was bellow 10 Å.

2.3.9 *Autocorrelation function*

The autocorrelation function of the RMSD was computed (separately for EC and TM domains) using the correlate function of the Numpy[164] python package to measure the convergence of free MD simulations assuming that decorrelated systems are converged.

2.4 DOCKING AND RESCORING

Docking consists in predicting the bests binding modes of a ligands into a binding pocket. Docking methods have been intensively developed in the past decades and gradually became the method of choice to start *in silico* drug design studies or to screen libraries of compounds. There exists several methods for generating poses and a wide variety of scoring functions to rank the produced poses. Due to the fact that scoring functions are meant to be fast to allow for screening databases of several million of compounds, it is generally highly recommended to rescore[165] the poses produced by docking, with a more accurate and computationally costly method, such as MMPBSA or QM.

In this work, the docking and rescoring protocol we have used is based on PLANTS[166,167] for the docking and on CHARMM for the rescoring of the poses. We first performed a few steps of MD in vacuum to explore the stability of the complexes generated by PLANTS and then performed a single point calculation to obtain the interaction energy between ligands and protein with the CHARMM36 FF in order to rescore the poses and keep only the best.

2.5 SYSTEMS PREPARATION

2.5.1 *Building*

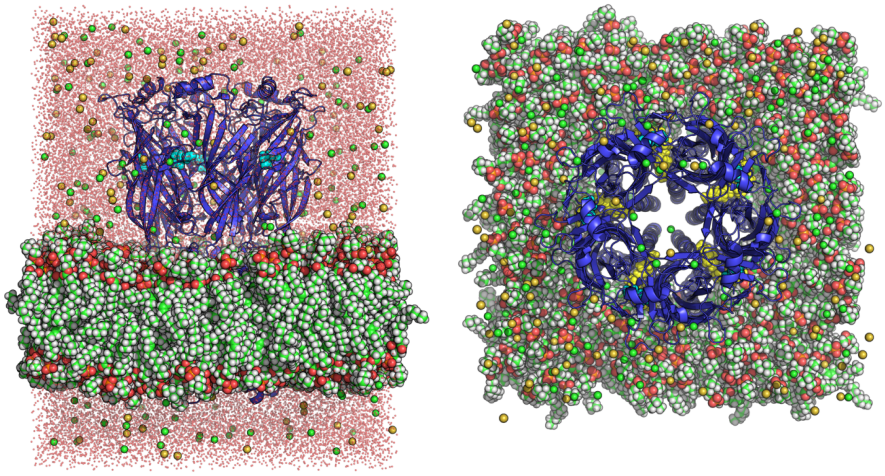
The structures produced by X-ray crystallography are not suitable for running molecular dynamics simulations. Indeed, they are chemically imperfect in the sense that they miss atoms or even residues that must be added prior to MD. Moreover the crystallization conditions are far from the physiological environment of such proteins. One may for example cite the use of detergents, which may perturb the electrostatic interactions between ligand and protein or between the residues of the protein crystallized. The crystal packing is also a concern when studying biological systems since it can induce a significant strain energy in the system which must also be taken care of before MD. Finally in some cases, to ease the crystallization process by diminishing the entropy of the protein, one may use antibodies interacting with the structure and which can also induce strain energy and should be removed before MD production.

To prepare the crystal structures of GluCl apo (4TNV) we first added the missing residues with MODELLER[168]. The missing residues in GluCl apo, i.e., three residues (103-105) forming a loop in the lumen of the pore in the EC domain, were reconstructed based on the sequence provided and using GluCl active (3RIF) as a template. The poly-histidines at the

end of the five chains were removed in both structures. The active structure did not have missing residues within range of interest.

The protonation states of the titrable residues were predicted using Poisson-Boltzmann (PB) calculations[169] and the multi-site titration approach[170]. The same protonation states were predicted and introduced in both resting and active structures. The intra-subunit di-sulfide bridges specified in the corresponding publications for apo and active structures, i.e. between cysteines 130 and 140 and between 191 and 202, were introduced in the 5 subunits.

Once the model built, it was sent to the webserver MolProbity[171] and all the side chains flips suggested were introduced in the structures. The latter models were inserted in a bilayer membrane of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC). The initial coordinates for the lipid bilayer were obtained from a membrane model pre-equilibrated with the CHARMM36 lipid parameter available from J. Klauda's resources webpage[172]. Explicit water molecules (TIP3P[173]) were added to fill an orthorhombic box of 115x115x150 Å centered around the protein. The systems were neutralized by addition of sodium or chloride ions. Then an excess of the same ions was added to reach a physiological concentration of 150 mM/L. To reproduce the bulk conditions while sparing computational resources, periodic boundary conditions were added at the limits of the box. The systems built consisted in about 200 000 atoms. The topology and parameters for the ligands were obtained from the CGenFF software[174,175].



(a) Side view (b) Top view

Figure 2.11: Built model of GluCl (blue cartoon) bound to L-Glu (cyan) and IVM (yellow) embedded in a POPC bilayer. Water (red dots) and ions (green and yellow spheres) are also displayed to ease the visualization of the periodic box.

Table 2.1: Sum up of the three main systems studied in this work.

Receptor	PDBID	Ligand(s)	Atoms	Water	Na^+, Cl^-	Lipids	Total simulated time (ns)
GluCl	3RIF	L-Glu	199925	41957	118, 138	347	5183.38
GluCl	3RIF	L-Glu, IVM	195134	41608	117, 137	314	475.22
GluCl	4TNV	-	197056	42034	119, 134	324	178.44

2.5.2 *Minimization, heating, equilibration*

To prepare for MD simulations all systems used were subjected to energy minimization using a similar protocol. A first 5000 steps of minimization was run to release the possible strain energy induced by the crystallographic conditions and the building of the system. Indeed if this step is skipped, the first calculation of forces during the heating procedure could lead to non physically high velocities for the clashing atoms, which would destroy the system at the next iteration. This first step is run with harmonic restrain on the backbone atoms to allow for the relaxation of the side chains only.

The coordinates produced at the end of the minimization were used to start the heating procedure. The systems were heated to a physiological temperature of 300 K in 600 picoseconds and under backbone atoms restraints.

The equilibration was carried out in the NPT ensemble starting from the coordinates and velocities generated at the end of the heating simulation. The restrains on the backbone atoms were gradually turned off following a cubic decay. The equilibration phase was run for 2 ns allowing the system to relax before production, e.g. to allow the water molecules to evenly distribute inside the pore and around the protein.

The production phase was carried at 300 K and 1 bar. Depending on the simulation the FF CHARMM27 or CHARMM36 were used. The pressure was maintained constant by the Berendsen barostat[114] and the temperature was controlled by the Langevin thermostat[111]. A cutoff of 12 Å was used for the electrostatic interactions with a switch[176] at 10 Å to allow for a smooth decrease of the interactions. The multiple time step r-RESPA[111] was used to compute the long range electrostatic interactions every two steps (4 fs) and the short range electrostatics plus bonded interactions every step (2 fs). Long-range electrostatic interactions were taken into account by the Particle Mesh Ewald (PME) method set up with a grid spacing of 1 Å and a sixth-order B-spline charge interpolation scheme. All covalent bonds involving hydrogen atoms were constrained with the SHAKE algorithm[108].

All four phases, minimization, heating, equilibration and production were run using the highly scalable molecular dynamics code NAMD[177].

2.5.3 *Restrained simulation*

A simulation of GluCl bound to L-Glu only was run with restrains on the interactions between L-Glu and the protein. This simulation was started from the X-ray structure of GluCl active (PDBID 3RIF) and prepared as discussed in section 2.5.

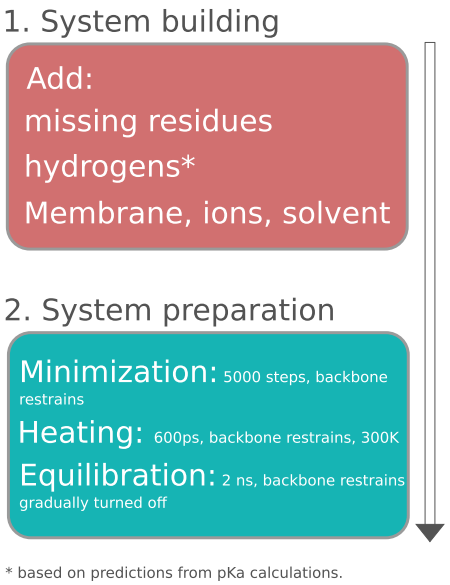


Figure 2.12: Flowchart presenting the typical steps to prepare a structure for MD.

The restrains were set to the distances measured in the crystal and were taken as reference for the harmonic potential. Four distances involving three residues and a glutamate were set as shown in Figure 2.13

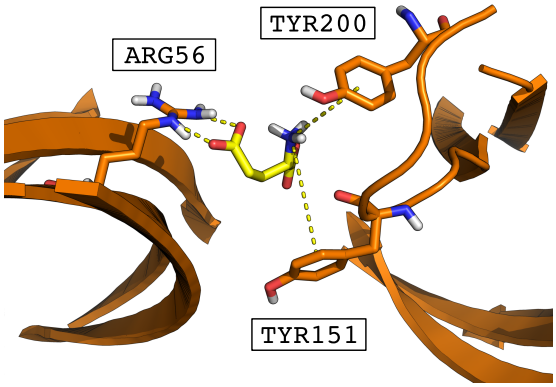


Figure 2.13: Representation of the distances restrained in the simulation LGLU*

2.5.4 *PMF calculations setup*

All PMF calculations were carried using the active structure of GluCl (PDBID 3RIF), which is bound to IVM and L-Glu. Depending on the simulations these ligands were kept or removed and others were added to the binding site of IVM when the orthosteric binding site was only bound to L-Glu. All calculations were performed using US with or without a RE scheme (BEUS).

The range of interest for the twist (τ) was defined based on the crystal structures, respectively 12.23 and 22.9 deg for active and apo and from free simulations. It was set from 12 to 24 degree. In order to restrain τ, which does not exist as we defined it in section 2.3.1, we had to adapt the Spinangle (ϕ) collective variable. The latter is not strictly defined as the twist angle but is strongly correlated with it. ϕ represents a global and relative value of τ. Indeed when τ is measured per subunit ϕ is a property of the 5 subunit. It is defined as the angle of rotation around a given axis of a selection. In order to compute an angle of rotation between the EC domain and the TM domain, as is τ, we had to first align the structure on the TM domain and then compute the angle of rotation over the EC domain (a detailed algorithm is presented in Fig 4.12). The alignments were made on the core atoms of both TM and EC domains as defined by Calimet et al.[3] The force constant used to restrain the twist angle was set to 10 Kcal/mol/A^2. The collective variable space was divided in 24 windows of 0.5 degree. This setup was shown to be adequate to reconstruct accurate PMF using WHAM. The sampling within each window was variable upon the system studied and the window but never was below 1 ns.

3

ELUCIDATING THE GATING MECHANISM OF PLGICS

The mechanism of action of dynamic proteins such as pLGICs can hardly be understood through methods such as X-ray crystallography and designing meaningful homology models of such complex proteins is very challenging. Using high definition structures from X-ray crystallography to compute dynamic properties thanks to MD seems to be a very promising approach, although it may be challenging as well. When studying the conformational dynamics of proteins it is much simpler to sample deactivation than activation as going uphill in energy (against the free energy gradient) is significantly more demanding in computational resources than going downhill. Nevertheless, it is not always possible to follow the energy gradient because of the lack of active structures. For these reasons and because only the active structure of GluCl was available when this study started in 2013, we used the active structure of GluCl in which we removed the bound ligands aiming at capturing a spontaneous transition to a resting state.

By studying the deactivation to understand the activation mechanism one makes the approximation that both mechanisms are identical[178], which might not be true in some cases. Ideally, one would want to study both activation and deactivation separately and thus compares the two paths taken by the system.

As only the structure of the active state of GluCl was elucidated producing, by MD, a structure with a closed channel was very challenging. Fortunately, in 2014 was published the resting structure of GluCl, providing us a convenient way to validate the resting model produced by MD.

The transition from active to rest (un-gating) of a pLGICs (GluCl) was triggered by the removal of the five IVM molecules bound at the level of the transmembrane domain. As a consequence of the removal of IVM, the five L-Glu molecules binding in the EC domain site, which stayed bound in the first part of our simulations, experienced spontaneous unbinding.

A total of 7 µs of simulation at the atomic scale of GluCl in a native lipid-membrane environment was produced, including the two simulations of the stable states of GluCl active and resting. The two former systems were used as reference states to characterize the starting and end points of the transition from open to close.

In a first time, we will describe the observed transition and characterize the resting state produced by MD. Then, we will propose a gating mechanism based on the analysis of our data and will discuss the influence of the orthosteric and allosteric ligands on this mechanism.

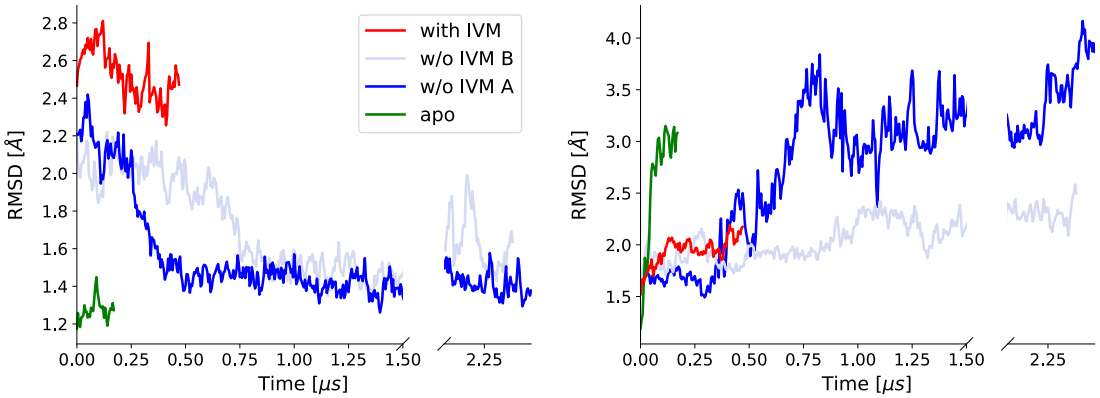
3.1 CONVERGENCE OF GLUCL ACTIVE TO GLUCL APO UPON REMOVAL OF IVM

3.1.1 *Results*

Two simulations of GluCl in which IVM was removed were run for about 2.5 µs each. They will be referred to as GluCl w/o IVM run A and run B throughout this thesis. A first step in understanding the gating mechanism of pLGICs is showing that the transition we captured is a true deactivation, i.e., it leads to the closing of the channel and to a resting state compatible with the one produced by X-ray crystallography. To do so, a wide variety of observables were monitored along the relaxation of GluCl both local such as the distance between two key residues or more global such as the RMSD.

3.1.1.1 *Global structural changes upon IVM removal*

RMSD TO APO To quantify the convergence of GluCl w/o IVM to the X-ray structure of GluCl apo we computed the RMSD of both TM and EC domain to the apo structure. The results are presented in Fig. 3.1. As the behavior of the C_α-RMSD is significantly different for the EC and TM domain it was computed separately.



(a) TM (b) EC

Figure 3.1: C_α-RMSD to the apo X-ray structure computed separately for the EC and TM domains. The x axis was cut as both RMSD showed similar trends after 1.5 µs.

Interestingly, in the first 400 ns for run A and 1 µs for run B the C_α-RMSD of the TM domain decreases from 2.2 Å to 1.4 Å, indicating that the structure of the TM converges to the apo structure. Indeed, 1.4 Å of C_α-RMSD corresponds to the equilibrium fluctuations sampled for the TM domain of GluCl apo (se the green lin in Fig. 3.1a). The C_α-RMSD of both run A and run B is relatively stable before the twisting transition is complete, respectively 400 ns and 1 µs for run A and run B, and then sharply decreases. This indicates that the twisting most probably unlocks an (or several) orthogonal transition which then allows the relaxation of the TM domain from the active to the resting structure.

The EC domain of the apo state undergoes a large transformation in the first nanoseconds of the simulations. This can be easily explained by the absence of ligand at the inter subunit binding sites but also by the removal during the building of the system of the antibodies (AB) binding at the level of the EC domain.

Moreover, the C_α-RMSD of the EC domain of GluCl active is significantly lower than the one of run A or of the apo simulation. This could be explained by the stronger coupling between adjacent subunits in the simulation of GluCl active. Indeed, the presence of both IVM and L-Glu increases the contacts between adjacent subunits. Run A and apo simulations show a significantly higher value of the C_α-RMSD for the EC domain which can be explained by the decoupling of the subunits or by the higher blooming angle (see Figure 3.3).

Finally, the C_α-RMSD of the EC domain of run A, unlike run B, increases a lot after the twisting is complete. This increase likely corresponds to the blooming isomerization, which is not observed in run B. Moreover the absence of ligand in run A and the blooming of the EC domain favor the decoupling of the subunits, the interfaces in contact in the contracted form being then loosely interacting.

This analysis shows that upon removal of IVM the TM domain of GluCl spontaneously converges to the crystallographic structure of the apo state.

TWIST ISOMERIZATION This observable was often described as one of the major component of the gating mechanism and it is significantly different in active and resting crystal structures. We have monitored the twist for the four simulations previously discussed and the results are shown in Fig. 3.2.

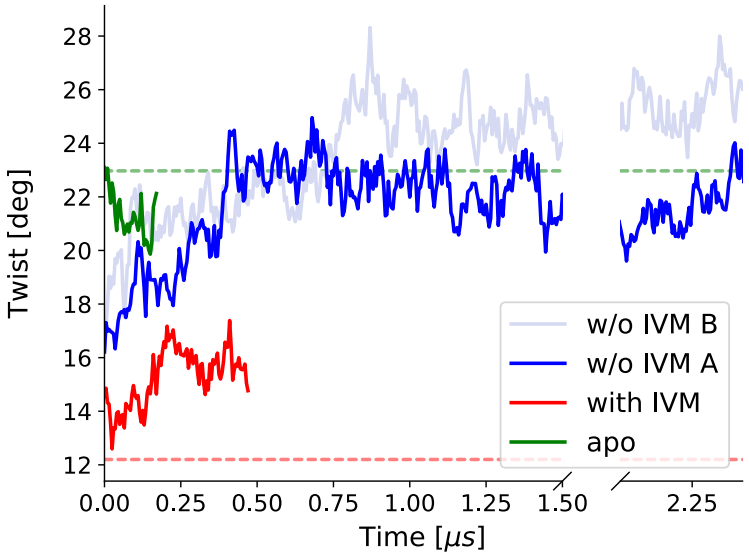


Figure 3.2: Evolution of the twist angle averaged over the 5 subunits for four different simulations. The dashed lines represents the X-ray values for the active (red) and apo (green) structures.

It is clear from the simulations of the stables state, i.e., GluCl bound to IVM and L-Glu (referred to here as "with IVM") and GluCl apo that the value of twist sampled are differ-

ent in both systems. Indeed the apo state shows a twist angle of about 22 deg when the active state is significantly less twisted with values around 15 deg. Moreover, both active and apo simulations deviates slightly from the X-ray values, which is expected as the conditions of crystallization and the simulated medium are different. Nevertheless, even though they slightly deviate from X-ray value they seem stable all along the simulated time unlike run A and B in which IVM was removed.

Indeed, in both run A and B a sharp increase is observed when IVM is removed. In fact, the twist angle increases from 12 (X-ray value) to almost 15 during the equilibration only, which underlies a large gradient of free energy toward a more twisted state. Run A undergoes a full twisting transition, i.e., from initial values compatible with an active state to final values compatible with a resting state, in about 400 ns. Interestingly, run B takes about twice as much time to reach a full twisted state, i.e., 750 ns.

BLOOMING ANGLE The blooming angle of the EC domain was also monitored for the four simulations aforementioned. The results presented in Figure 3.3 show that, run A and run B do not behave the same way. Indeed, when run A blooming values are stable until the twisting isomerization is finished (400 ns) and then sharply increase to reach values compatible with a resting state, run B does not show such blooming of the EC domain.

Interestingly, the timing of the blooming transition corresponds to both the end of the twisting (see Figure 3.2) and the spontaneous unbinding of the 4th L-Glu (see Figure 3.9). It is possible that the full twist isomerization unlocks the radial expansion of the EC domain. It is also likely that the unbinding of most of the L-Glu from the interfaces binding sites in the EC domain allows for the decoupling of the subunits and therefore the radial expansion of the EC or blooming. Nevertheless, even though all L-Glu unbind from GluCl in run B (see Figure 3.9), no radial expansion of the EC domain can be observed, suggesting that the presence of L-Glu is not the only barrier preventing from blooming the EC domain.

Even though our two repeats of the transition from open to close differ in the EC domain, they both produce after a minimum of 1 µs of simulation a twisted state. To our knowledge, it is the first time the full twisting isomerization of a pLGIC is ever sampled thanks to computational techniques. These three general observables suggest that our simulations have captured the full isomerization of the active state of GluCl to a resting state of GluCl upon removal of IVM. Interestingly, this conclusion holds even when looking at more detailed observables as will be explained in the upcoming sections.

3.1.1.2 *Pore closing*

This major feature of ion channels was monitored using different observables. σ measured at the level on the constriction point, i.e., 9′, the HOLE profile, the water and ion count around the constriction point and the water flux. Since no ionic flux was observed in any of our simulations, no striking data can be shown. This set of observable represents an accurate way of monitoring pore closing.

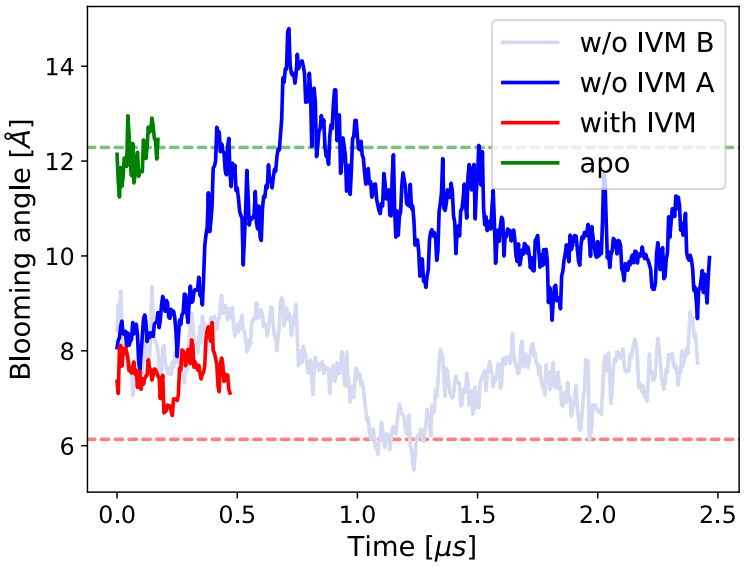


Figure 3.3: Evolution of the blooming angle (θ_p) averaged over the 5 subunits for four different simulations. The dashed lines represents the X-ray values for the active (red) and apo (green) structures.

HOLE PROFILE The HOLE profile was computed every 10 ps and averaged over specific time windows. Results shown in Figure 3.4 illustrate the gradual shrinking of the pore radius along time as IVM is removed. The profile of the two crystal structures are significantly different and corroborate their classification as open and closed structures. Our simulations of IVM removed show a gradual closing bridging the gap between the two crystal structure and showing that 50% of the closing occurs in less than 10 ns. Interestingly, even though the first half of closing is rather quick the second half takes about 400 ns to be complete, i.e., reach values compatible with a resting state.

It is clear from this analysis that the constriction point when the channel is closed is located at 9′ in both the MD produced apo state and the X-ray crystallographic structure. Moreover, in the active structure the constriction point location is different as it is located at -2′ position, which might also be the desensitization gate.

Both apo structures (MD and X-ray) are non conductive for chloride ions as the radius of their constriction point is inferior to the radius of a chloride ion. Moreover, when trying to discriminate between open and closed states of a channel, one should also consider the radius of a hydrated chloride ion (3.4 Å[179]) which is significantly higher than the one of a dehydrated chloride ion (1.8 Å[179]), as the dehydration of the ion could represents a significant barrier on permeation.

Despite the fact that the radius of the pore in position 9′ of our simulation of GluCl +IVM +L-Glu, here referred to as MD active, significantly decreases compared to the one of the crystal structure, the average radius is style compatible with dehydrated ion permeation. This could be explained by possible imprecisions in the FF parameters used in our simulations for

IVM or for the protein. It is also possible that the physiological conductive state of GluCl is only permeable to partially dehydrated ions.

The conclusion made in this section for run A are fully applicable to run B.

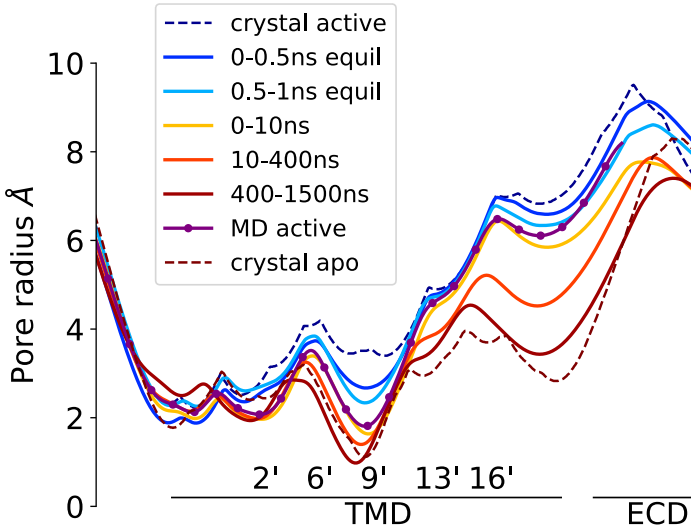


Figure 3.4: Evolution of the HOLE profile of the pore along a simulation in which IVM was removed (run A).

C_α CROSS SECTION: σ First of all the position of the constriction points were indicated by HOLE profiles and allowed to analyze more in detail the evolution of the radius at the constriction point. It is also clearly shown by Figure 3.6 that the apo state of GluCl produced by MD and the one elucidated by X-ray crystallography are equivalent. Indeed, the position of the side chain of the residues at the constriction point (LEU254), the radius and the orientation of the M2 helices are very similar.

Moreover, the active and resting states are also significantly different in the variations of σ. Indeed, the active, even though it is bound to IVM seems to be more flexible in term of ion pore size. In fact, this is most probably due to the fact that the packing of the TM helices is more important in the resting state compared to the active state preventing from fluctuations of the pore radius. This can be explained by the absence of IVM in the apo structure, which acts as a wedge between the subunits, in the apo structure.

Finally, the analysis of the trajectory of σ confirms that the channel fully closes to a state similar to the apo crystallographic structure when IVM is removed.

WATER AND ION FLUX Both flux and presence of water around the constriction point were monitored along time for the four simulations aforementioned. The total flux was measured as the sum of the inward and outward flux (see section 2.3.5 for a detailed description of the algorithm).

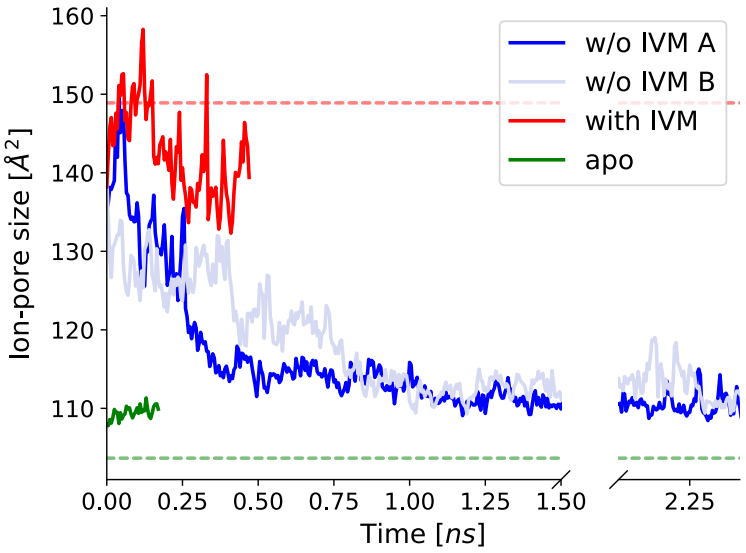


Figure 3.5: Evolution of σ describing the pore size at position 9′. The dashed lines represents the X-ray values for the active (red) and apo (green) structures.

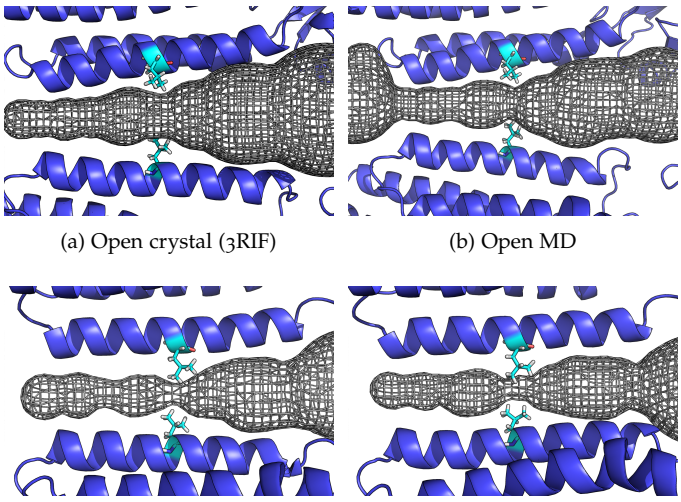
As it was not possible to observe ion permeation events we looked at the water flux which can be seen as a proxy for the permeation of ions[163]. Indeed, the number of ions being several order of magnitude lower than the one of water molecules, the probability to sample such event is much lower. Moreover, the absence of concentration gradient or electric field around the bilayer membrane could also be an explanation of this observation.

The water flux decreases quickly to reach a fully non water-permeable state in 300 ns or 750 ns for run A and run B respectively. The water flux for the active state is stable around 3 wat/ns. The decrease of the number of water molecules around the constriction point (water count in Figure 3.8) correlates with the drop of the flux or the decrease of the constriction point radius. Indeed, as the pore radius decreases, the water molecules are pushed away and therefore the water count drops, creating a vacuum around the constriction point. The ion count follows the same direction as the constriction point shrinks.

Interestingly, despite the fact that the channel is fully closed in the apo simulation, two events of permeation are recorded at 40 and 100 ns. Similar measurements made on a simulation of ELIC, believed to be in a desensitized state, were done and no water permeation event have been recorded. Even though these are very preliminary results, this could suggest that water permeation of pLGICs could potentially be used to discriminated between resting and desensitized states.

As pLGICs are thought to close a mechanism called hydrophobic gating, we also decided to study the density of water molecules around the constriction point and all along the ion pore.

The results shown in Figure 3.7 make clear that the shrinking of the ion pore leads to the dewetting of the constriction point. Indeed, GluCl bound to IVM and L-Glu shows an homogeneous distribution of the waters inside the pore all along the simulation, unlike GluCl



(a) Open crystal (3RIF) (b) Open MD

(c) Closed MD (d) Closed crystal (4TNV)

Figure 3.6: Comparison of the constriction point between MD and crystal structures, both open and closed. The mesh represents the volume of the ion pore as detected by MOLE2.0[180]

apo simulation where the dewetted stretch spreads over approximately 4 Å around the constriction point. Both simulations of GluCl in which IVM was removed show a first portion of completely hydrated pore. Interestingly, the time corresponding to the dehydration of the constriction point for run A and run B corresponds to what is observed for the twist and HOLE profiles, i.e., 400 ns for run A and 800 ns for run B. So far we could not explain why the dehydrated stretch was slightly smaller in the simulations without IVM compared to the simulation apo.

Overall, the pore closing occurs gradually and is correlated with a dewetting of the pore. The closing of the pore also seems to be time correlated with the twist. Finally, even though it happens with slightly different kinetics for run A and run B both of the reached structures are very similar to the apo structure.

3.1.2 *Discussion*

At the global scale, when IVM is removed, a twist of the EC domain over the TM domain is observed. The values of twist sampled by both run A and B correspond to an active state at the very beginning of the simulation, i.e., when the system still has memory of the presence of IVM, and to a resting or apo state at the end of the simulations. One out of two simulations of GluCl without IVM does not show a blooming of the EC domain even though the channel fully closes. This suggests that blooming the EC domain represents a significant barrier to cross and that it is not strictly required for closing.

The convergence to the apo state is locally evidenced by the closing of the pore in 9′, which occurs in the same time as the twist in both run A and B. The final diameter of the pore

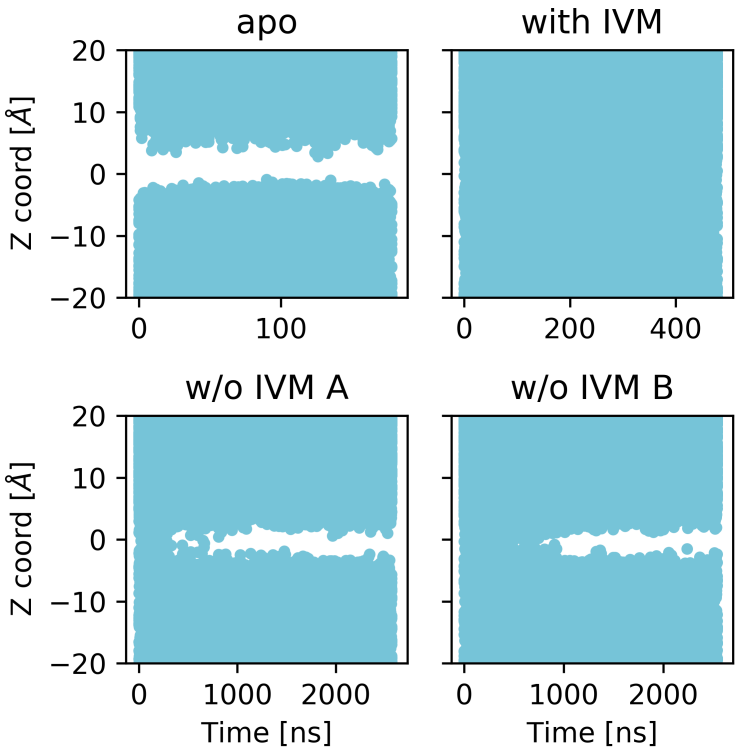


Figure 3.7: Density of water molecules projected on the Z axis along time, results shown for four simulations of GluCl. This graph lets one visualize the evolution of the water density inside the ion channel. The zero of the Z axis was aligned with the LEU254 constituting the constriction point.

is smaller than a chloride ion. Moreover, the shriking of the ion channel creates a vacuum stretch which increases the barrier for ion permeation.

In conclusion, we have shown in this section that upon IVM removal, the conformationally strained state of GluCl converges toward the resting structure. These experiments also illustrate that, MD can be used as a predictive tool to model unknown states, such as resting states of proteins. Indeed, without prior knowledge of the structure of GluCl apo our simulations have produced, repeatedly, a state compatible with the apo state both locally and globally.

3.2 A NEW GATING MECHANISM

3.2.1 *In depth study of the transition from open to close*

In the previous section we have shown that the transition sampled upon removal of IVM was meaningful as it bridged the gap between the active and the resting states as illustrated by the comparisons made with the X-ray structures and MD simulations. In this section we will focus on analyzing, in details, the captured transition to better understand the gating mechanism and propose a new one, coherent with our data.

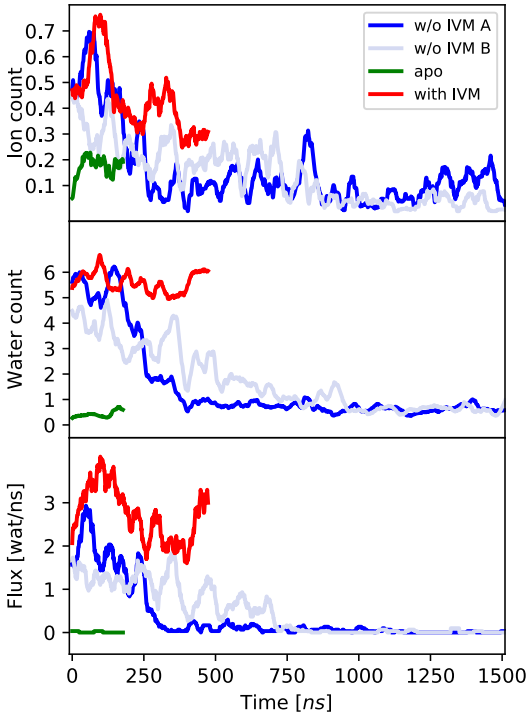


Figure 3.8: Ion and water permeability in the four simulations of GluCl. On top and middle panels, the number of water molecules and chloride ions sitting in the pore region is monitored over time. On bottom, the time series of the water flux through the pore in shown.

We will first describe the evolution of the most important observables among all the one we analyzed, discussing first the EC domain then the interface between EC and TM domain and to finish, the TM domain.

In a second time we will discuss some of the main correlations between the chosen set of observables and start to draw a new gating mechanism.

3.2.1.1 *The orthosteric site*

NUMBER OF L-GLU BOUND The binding of a ligand to the receptor can be simply defined by an RMSD based approach. The RMSD from the starting position of a given ligand is monitored along the simulation and a bound state is defined bellow a given threshold, i.e., 6 Å.

The results presented in Figure 3.9 show that the number of L-Glu bound to their initial active site drastically decreases in the first microsecond of simulations. Interestingly, the unbinding of the fourth L-Glu seems to be correlated with the end of the twisting in both run A and B. Indeed, the full twisting takes 400 ns or 750 ns for run A and B (see Figure 3.2) and the unbinding of the fourth L-Glu happens at about 400 ns and 800 ns for run A and B respectively.

Spontaneous events of rebinding to the initial binding site are captured mainly in run A. Nevertheless, the spontaneous unbinding or rebinding occurring after 1 µs of L-Glu does

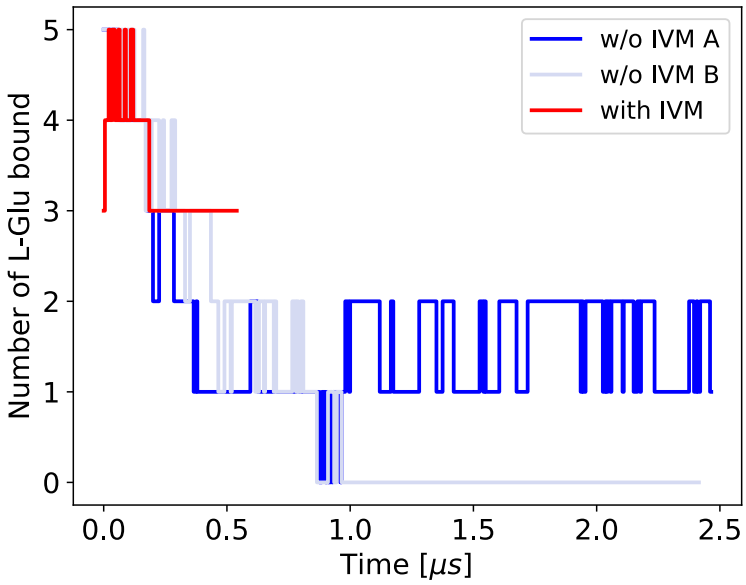


Figure 3.9: Number of L-Glu bound to GluCl as defined by a 6 Å cutoff on the RMSD.

not seems to be correlated with a reactivation of the receptor, which might be due to an insufficient residence time of L-Glu in its binding site or to a non-specific rebinding. In fact when binding to a bloomed state, L-Glu may interact with only one of the two faces of the interface, which does not stabilize an active state of the receptor.

GluCl is a homopentamer, which means that it is composed of five homologous subunits and therefore of five identical binding sites for L-Glu. As discussed in chapter 2, the algorithm used to compute these results do not allow to monitor events of rebinding to a different binding site from the one of the X-ray structure (see section 2.3.8 for the description of the symmetry corrected version of the algorithm).

Expectedly, the symmetry corrected rebinding algorithm allows to record more events of rebinding. However, the conclusions drawn in the former paragraph are in agreement with the new data. Indeed, the spontaneous unbinding of all the L-Glu is observed in a similar time frame, i.e., 400 ns for run A and 700 ns for run B. Interestingly, the first algorithm could not detect any rebinding events for run B after 1 µs when it could for run A. This suggests, that only in run A, L-Glu were capable to rebind to an orthosteric site. This observation could be explained by the absence of blooming (see Figure 3.3) in run B, which leads to significantly more packed interfaces and could prevent rebinding of L-Glu.

Perhaps surprisingly, the unbinding of nearly all L-Glu is spontaneous and is also observed, to a lesser extent, in the simulation of GluCl active, in which IVM remains bound (as shown by both Figure 3.9 and Figure 3.11). In fact two out of five L-Glu unbind from their initial binding site in the receptor in less that 400 ns. The spontaneous unbinding of L-Glu could be explained by imprecisions in the parametrization of the interactions between L-Glu and the receptor or by the parametrization of L-Glu itself. In addition to that, one could also envision

that the removal of IVM triggers a conformational transition of the protein which expels the L-Glu molecules out of their binding sites.

As our main goal was to investigate the deactivation of the receptor this issue was not investigated further. When the need arose to keep L-Glu bound to GluCl we added harmonic restrains to specific interactions between ligands and protein (see section 2.5.3)).

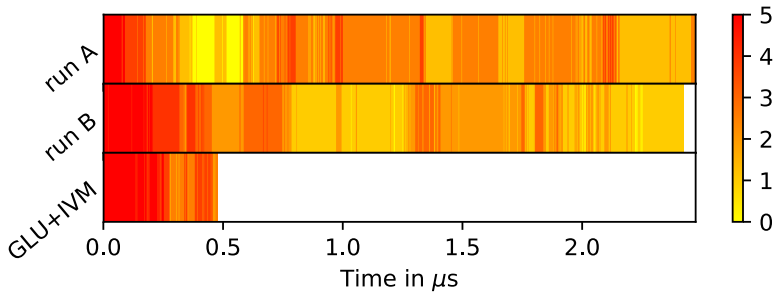


Figure 3.10: Number of L-Glu bound to GluCl taking into account the possible rebinding events to a different pocket than the one occupied in the X-ray structure.

RMSD OF THE BINDING SITE The topology of the orthosteric binding site is prone to change during both twisting and blooming transitions as it is located at the interfaces between subunits. To better understand the variation of its topology, its RMSD was computed in the four simulations available. The binding pocket was defined by the residues 91, 150, 151 and 200 on the (+) subunit and 37, 56 and 121 on the (-) subunit. These residues were chosen because of their known interactions with the ligands, either by hydrogen bonding or electrostatics. We chose not to include residues carried by the loop C as its flexibility could have made the interpretation of the results more complex. The C_α of the former residues were used to first align each frame and compute the RMSD of the pocket for the four simulations of interest.

First of all, we can see that the binding pocket is significantly more stable when L-Glu is bound as the active simulation shows a RMSD of less than 1 Å and the resting climbs to more than 1.5 Å. Interestingly, run A RMSD increases drastically after 400 ns to remain stable around 2 Å. Unlike run A, run B RMSD slowly increases during the entire length of the simulation.

CHARACTERISTIC DISTANCES Characteristic distances were monitored to describe its changes in shape and size along time, providing a complement of information to the RMSD analysis. The first distance is measured between the C_α of residues TYR151 on the (+) subunit and SER121 on the (-) subunit. It describes a distance between two adjacent subunits in the active site of L-Glu and at the back of the pocket. The second distance is very similar to the first one as it represents the same spreading but across the entire pocket. It is defined between C_α of residues SER150 on the (+) subunit and ARG56 on the (-) subunit. The third distance accounts for the position of the loop C relative to the active site. It is measured between the

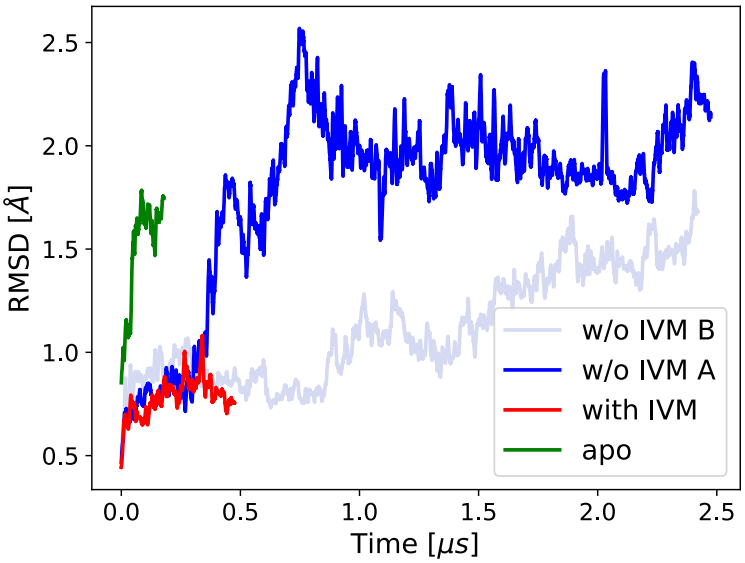


Figure 3.11: RMSD of the residues defining the binding pocket of L-Glu along time for four simulations.

C_α of residues ARG56 on the (-) subunit and THR197 on the (+) subunit.

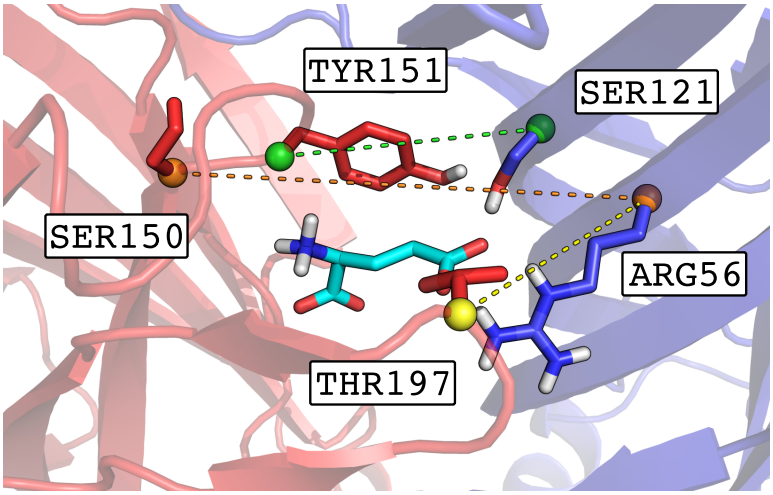


Figure 3.12: Representation of the thee distances of interest. L-Glu is represented in cyan sticks. Distance one is materialized by green dashes, distance two by orange ones and distance three by yellow ones.

These data show that distance one and two are very similar if not equivalent, suggesting that the back and front of the pocket move in a similar fashion. The active state shows a shorter distance on average over the five subunits than the one of run A (see Table 3.1), which underlies the contracted state of the orthosteric site in the active structures. Indeed, these two distances are significantly different between the active simulation and run A, observation that is not true for run B. In fact, a possible explanation for the former is that the blooming of the EC domain observed in run A and the apo simulation, which is responsible for a significant

stretching of the interfaces between subunits is not observed in run B where the interfaces stay close together.

Moreover, as shown on Figure 3.13 both resting and run A populate contracted form of the orthosteric site. This very interesting finding suggest the existence of a pre-active state. The former would present an EC domain and an orthosteric binding site similar to an active structure, i.e. contracted and untwisted, but presenting a closed ion pore. Therefore, the binding of L-Glu would stabilize this structure and eventually promote the opening of the ion pore.

The third distance undoubtedly shows the most variations as observed values range from 10 to 25 Å. Since it involves the loop C the captured variations are wider. Also, more discrete states, i.e., 4, seem to be isolated. Moreover, the average distance observed in run A seems to be slightly larger than the one in run B, which could be explained by the more contracted EC domain in run B.

Surprisingly, the apo simulation with an average distance of 13 Å shows the more contracted loop. It might seem counter intuitive as the presence of the ligand is thought to stabilize the loop C in a closed position. When looking at the simulation of GluCl apo, we could see the formation of stabilizing interactions between the tip of the loop C and the β-sandwiches of the (-) subunit. The conformation then taken by the loop C seemed to be even more closed than in the presence of ligand, suggesting that the absence of ligand stabilize an hyper-contracted form of the loop C instead of an open one. Even though these results are very preliminary and must be further investigated they offer an alternative interpretation of the role of the loop C.

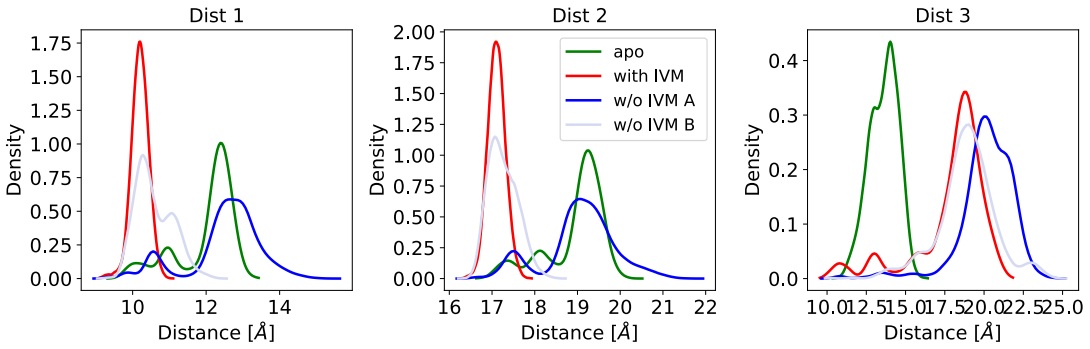


Figure 3.13: The probability density for the three characteristic distances is plotted for four different simulations. The density is normalized so the area under the curve is equal to one.

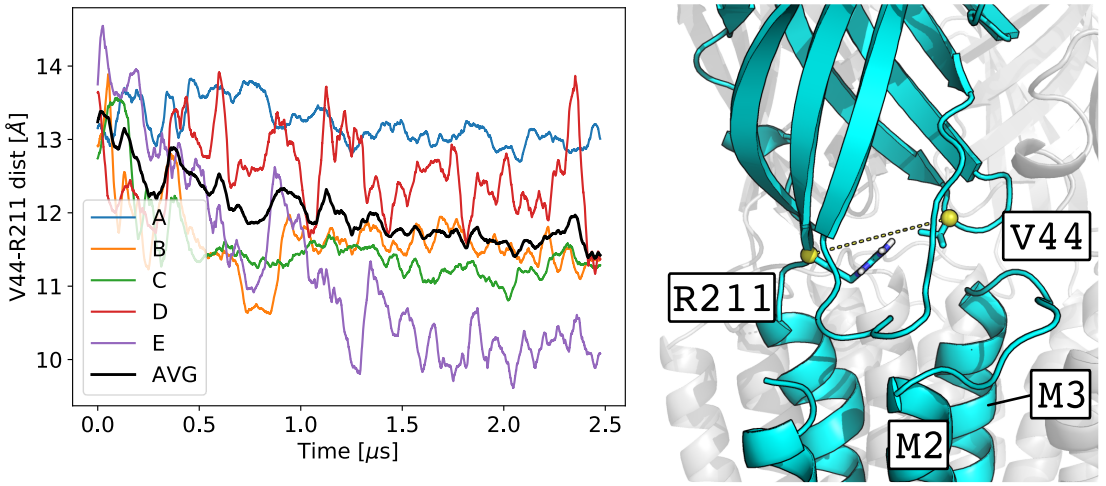
3.2.1.2 *EC and TM domains interface*

β-SHEET EXPANSION The bottom of the β-sandwiches is located near the upper part of the TM. This interfacial domain of the protein was suggested to be crucial in the control of the gating as it might be involved in the transmission of the L-Glu binding signal to the TM domain.

Table 3.1: Average value and corresponding standard deviation for the three characteristic distances of interest computed for four different simulations.

Distance	w/o IVM A		w/o IVM B		Active		Apo	
	mean	std	mean	std	mean	std	mean	std
1	12.51	1.01	10.58	0.52	10.19	0.24	12.00	0.81
2	19.06	0.84	17.26	0.35	17.11	0.20	18.95	0.69
3	20.18	1.76	18.88	1.87	17.92	2.29	13.61	0.92 1

It was shown by Lev et al. in 2017[181] that the closing of the ion pore in GLIC involved a stretching of the contact between the lower part of the β-sandwiches, which could allow for the inward passage of the M2-M3 loop and thus close the pore. Even though the salt bridge present in GLIC in this area of the protein is absent in GluCl, we monitored the distance between equivalent residues, i.e., V44 and R211. Our analysis of run A, shown on Figure 3.14a,



(a) (b)

Figure 3.14: (a) Distance between the C_α of the residues R211 and V44 for run A of the simulation of GluCl w/o IVM, highlighting the intersubunit variability. (b) Illustration of the distance and its location in the protein. The C_α are represented by yellow spheres.

shows that the average distance decreases upon closing. Indeed when IVM is removed from the active site this distance decreases on average of 1 Å in about 1 µs. Nevertheless the variability among the subunits is rather large, i.e., 5 Å. Therefore, it seems that unlike GLIC where an expansion is captured upon closing, we observed the exact opposite. Similar results are obtained for run B, see Figure 3.15.

In addition, the analysis of the active and resting states did not show a significant difference in the expansion of the β-sheets. Surprisingly, this distance decreases also in the simulation of GluCl active, suggesting that this observable is not suitable for describing the gating in GluCl. Moreover, the convergence of the RMSD and of the twist suggested that both simulations of

GluCl without IVM were converged after 1 µs, conclusion which seems to be supported by the convergence of the distance here analyzed as it also takes about 1 µs to converge. Nevertheless, the difference in timing observed on the twist for instance is here not observed as both run A and B converge to the same values at the same speed.

Overall this analysis show significant differences in the proposed gating mechanism of GLIC and GluCl regarding the β-sheet expansion. Differences which are most likely due to the absence of a critical salt bridge in GluCl compared to GLIC.

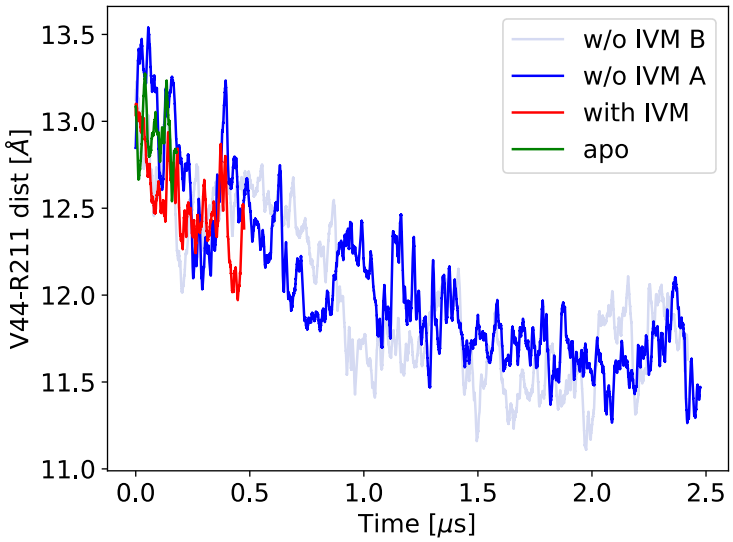
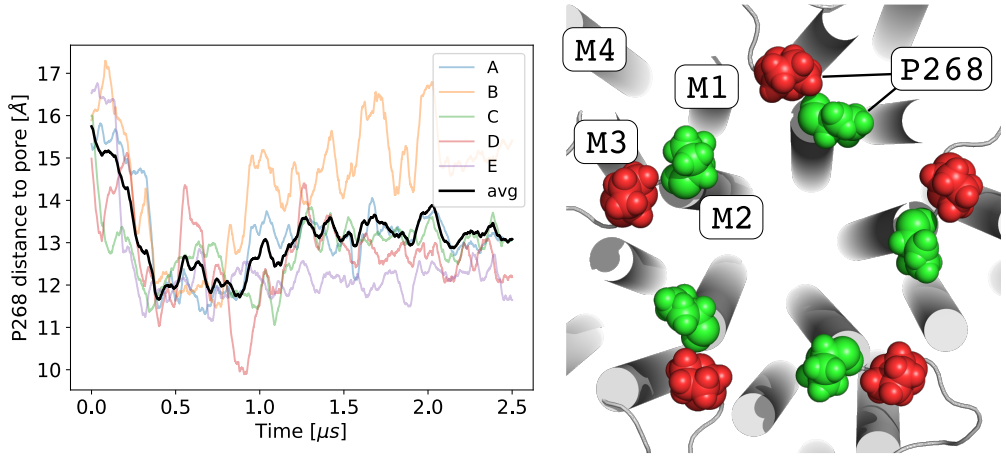


Figure 3.15: Measure of the expansion of the bottom part of the β-sandwiches for the four simulations of interest. The distance between the C_α of residues V44 and R211 is averaged over the five subunits.

INWARD DISPLACEMENT OF THE PROLINE 268 The proline 268 located on the M2-M3 loop is believed to be of crucial importance, partly because of its conservation among the member of the pLGICs. We have monitored both its position projected on the membrane plane and its distance to the center of the ion pore along the four simulations of interest to understand its role in the transition from open to close. One can see on Figure 3.16 that the average distance to the center of the pore rapidly decreases, i.e., in less than 500 ns, from 16 Å to 12 Å. Interestingly, it seems that during the twist, all prolines undergo a coordinated inward motion, note the low standard deviation of the distances between 300 and 400 ns in Figure 3.16a.

Moreover, the behavior of the P268 residues carried by different subunits varies a lot in the second part of the simulation of GluCl w/o IVM A. If the proline in subunit B seems to be far form the center of the pore (16 Å) all the other are significantly closer to the lumen (between 12 and 14 Å). In fact the significant increase of the average distance after 1 µs is mostly related to the behavior of the proline of subunit B. This illustrate that despite the fact that GluCl is a symmetric homomeric system, each subunit undergoes its own transition.



(a) Distance to the pore residue P268

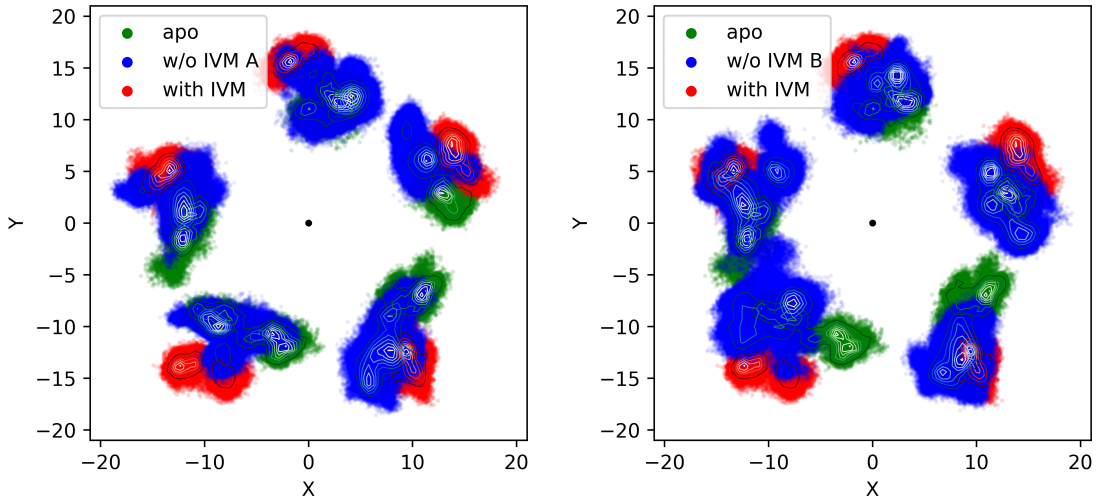
(b) Illustration of inward and outward positions

Figure 3.16: On the left panel is shown the distance to the center of the pore for the five prolines and for the average for the simulation of GluCl w/o IVM A. A similar trend is observed for run B. On the right is shown an illustration of the inward (green) or outward (red) position of the proline 268.

The illustration presented on Figure 3.16b shows the two possible positions, i.e., "in" and "out", that the P268 residues can take. "Out" corresponding to a large distance to the center of the pore and "in" to a smaller one. One can clearly see how the coordinated inward displacement of these residues can lead to the reduction of the pore diameter lower in the ion channel.

Moreover, even though the exact role that the P268 residues play in the gating mechanism is not fully elucidated yet, this analysis show that they can be considered as a good observable for deactivation of the channel. The results shown on Figure 3.17 illustrate the dynamic behavior of the P268 residues. At the beginning of the simulation they adopt an outward position which is compatible with an active state of the receptor (see the red cloud of points) and then move toward the center of the pore to occupy a position compatible with a resting state. These results also illustrate nicely the relative asymmetry of the system, which will be discussed in greater details later on in this thesis, because the distance between P268 of adjacent subunits is significantly different. In addition to this radial displacement, azimuthal motions are also captured both in the "out" and "in" positions of the prolines, although their function is not clear yet.

As a remark, the two analyses presented in this section emphasize two different kinds of information. When the time series, which gives a clear indication of the time evolution of the distance, do not allow to visualize the displacement of the prolines in two dimension, the scatter plot allows to do so in a clearer manner. Moreover, unless one uses a complex color scale, scatter plots do not allow to represent the time evolution. For this reason, both plots are required to understand in detail the motion of the prolines upon deactivation.



(a) GluCl w/o IVM A

(b) GluCl w/o IVM B

Figure 3.17: Position of the 268 residues projected on the XY or membrane plane for both run A and B for 4 simulations.

Z DISTANCE BETWEEN THE $\beta 1 - \beta 2$ LOOP AND THE M2-M3 LOOP To better understand the interaction between the TM and EC domains several analysis were run. The distance between two residues, one located on the tip of the $\beta 1 - \beta 2$ loop (V45) part of the EC domain and the P268 aforementioned, has risen our interest.

Indeed the interface between the two domains plays a key role in the propagation of the signal from the orthosteric site to the TM domain where the channel closes.

Interestingly, results presented on Figure 3.18 show that the distance between the two residues is significantly different between the active and the resting states, i.e., 5 Å and 6 Å respectively. When IVM is removed from the active structure, and consistently in both run A and B, this distance increases significantly. In fact, it takes 400 ns for run A to reach 6.5 Å and about 1 µs for run B to reach the same value of the distance. This seems to be correlated with the timing of the twist isomerization.

Even though this analysis cannot determinate if both residues are moving away from each other thus increasing the computed distance or if only the motion of one of the two is responsible of the recorded increase, it confirms the importance of the interfacial domain between the EC and TM segments of the receptor.

3.2.1.3 *Transmembrane domain*

M2 TILT The closing of the ion pore involves an inward motion of the proline carried by the M2-M3 loop, which suggests that the full M2 helices could move inward to close the ion pore. Indeed, Figure 3.6 clearly shows the un-tilting of the lining pore helices which are straight when the channel is closed and outward tilted when it is open. It was suggested that in GLIC[79,182] the closing of the ion pore happened through the un-tilting of the M2

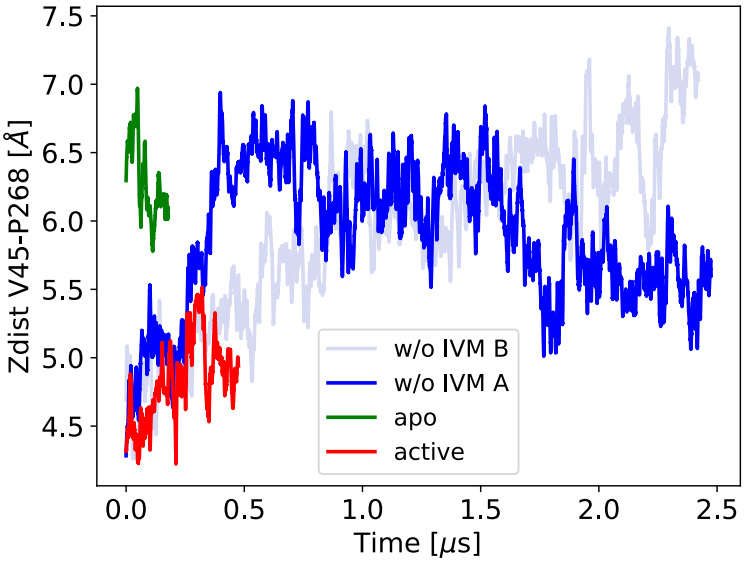


Figure 3.18: Averaged distance between the C_α of the residues V45 and P268

Table 3.2: Ranges of residues used to compute the M2-M3 distance and their correspondence between three different pLGICs.

Protein	GLIC	ELIC	GluCl
Residue	238-244	245-250	258-265
	253-259	259-264	274-280

helices and the breaking of the contacts between the former helices and the M3 helices. To study the decoupling of the M2 and M3 helices we monitored the distance between the center of geometry of two specific selections, the first including the upper part of the M2 and the second the upper part of the M3 helices (see Table 3.2). By measuring the distance between the center of geometry of two selections we aimed at monitoring more general changes averaging out the possible fluctuations of a side chains or local movements of given residues.

The results presented on Figure 3.19 show that there is no significant difference between active and resting state. Interestingly, upon closing a slight increase, i.e., from 12 to 13 Å of the M2-M3 distance is monitored before the 500 ns mark. Despite this increase, the separation between the M2 and M3 helices never reaches the values measured for the locally closed state of GLIC (14.2 Å using 4NPP) or for GLIC resting (14.04 Å using 4NPQ).

We can conclude from these results that pore shutting in GluCl is the result of a coordinated movement of the M2 and M3 helices toward the lumen of the pore with no apparent uncoupling. Moreover, these results along with the X-ray structure of GluCl apo could question the significance of the locally-closed channel structures of GLIC or at least suggest that GLIC has a unique pore-closing mechanism compared to the other member of the pLGICs.

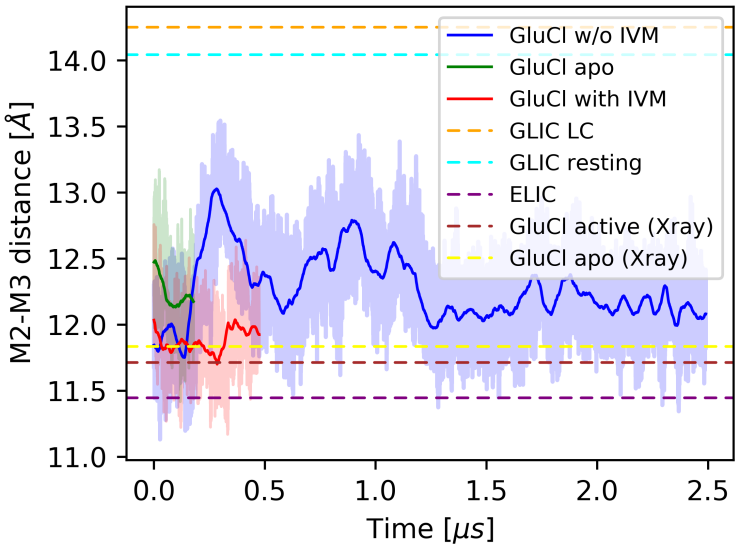
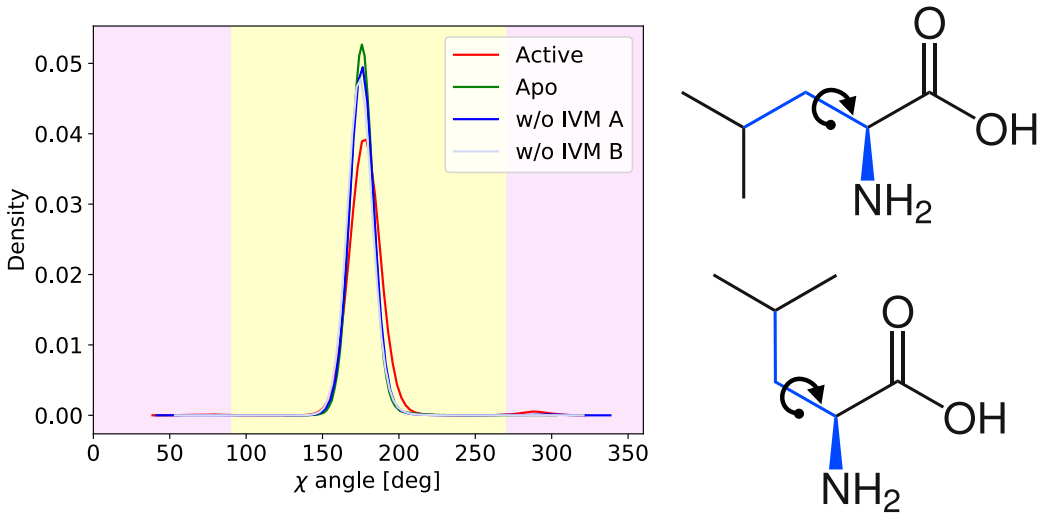


Figure 3.19: The averaged distance between transmembrane helices M2 and M3 along three simulations of GluCl is shown as a function of time. The simulation of the active state is represented in red, w/o IVM A in blue and resting in green. Run B leads to similar conclusions. The colored shadows correspond to the standard deviation of each series.The dotted lines correspond to the M2-M3 distance measured in the crystal structures of the locally closed (LC) state of GLIC (PDB 4NPP), the resting state of GLIC (PDB 4NPQ) and the closed channel state of ELIC (PDB 2VLo)

LEU 254 ROTATION The residues LEU 254, located at the position 9′, form the constriction point of the resting structure. It was suggested by Yuan et al. in 2016[183] that the rotation of the χ angle (see Figure 3.20b) of these residues could control the closing of the ion pore. Indeed, in function of the conformation of the χ angle, the side chain of these residues can point toward the lumen of the pore or toward an adjacent subunit, freeing space for the passage of ions by increasing the pore radius.
Our analysis shows that all resting state simulations, i.e., end of run A and B and the apo state, sampled only the *trans* conformation of the χ angle, which is associated with the side chain facing toward the lumen of the pore. Interestingly, the active simulation of GluCl, shows a marginal but existing population of *cis* conformations of the χ angle.

In conclusion, our results show that when the pore is closed, the side chain of the LEU 254 residues can only be facing toward the lumen of the pore and this is easily explained by the packing of the helices of the TM domain in the resting state. Indeed, in the resting state the pore lining residues are so close together that they can only arrange toward the center of the ion channel reducing the pore diameter. Interestingly, this analysis also shows that in the active state and unlike what happens for the closed state, the increased space between adjacent subunits allows for the rotation of the side chain of the LEU 254 residues during a small fraction of the time.



(a) (b)

Figure 3.20: (a) The density of the value taken by the χ angle along the four simulations of interest. The *cis* configuration is highlighted by a pink background and the *trans* by a yellow one. (b) The two configurations of the the leucine residue are drawn to show the χ angle (highlighted is blue).

HYDROPHOBIC GATING The definition of hydrophobic gating was established based on the analysis of simulations of nanopores. It describes the fact that a channel can be closed to the passage of ions or even water molecules but not physically close, i.e., with a diameter above the one of an ion. Indeed, the void created by the repulsion of the water molecules by hydrophobic residues represents a non-negligible barrier which prevents permeation.

This rather elegant mechanism of gating suggests that a simple study of the ion pore size is not always sufficient to conclude on the functional annotation of a receptor. Hence, one should also look at the distribution of water molecules inside the ion pore seeking for a vacuum stretch in a physically open portion of the channel.
Our analysis shows that in the case of GluCl bound to IVM and L-Glu and as discussed before the pore is significantly more open than in the three other simulations, i.e., GluCl apo and the last portion of GluCl w/o IVM A and B.

As the water density shown in Figure 3.21 is normalized, values cannot be compared across simulations, hence one should look at the area between the red and blue curve at the level of the constriction point. Indeed, if there is no difference between the water density and the ion pore, one can safely conclude that the space not occupied by the protein is occupied by water molecules and therefore that there is no hydrophobic gating.

Interestingly, in the three closed systems the dehydration stretch matches closely the HOLE profile, which indicates that the dehydration at the constriction point is due to steric hindrance and not to hydrophobic effects. Moreover, the diameter of the pore is smaller than the one of a chloride ion demonstrating that the closing mechanism described by our simulations is not compatible with hydrophobic gating.

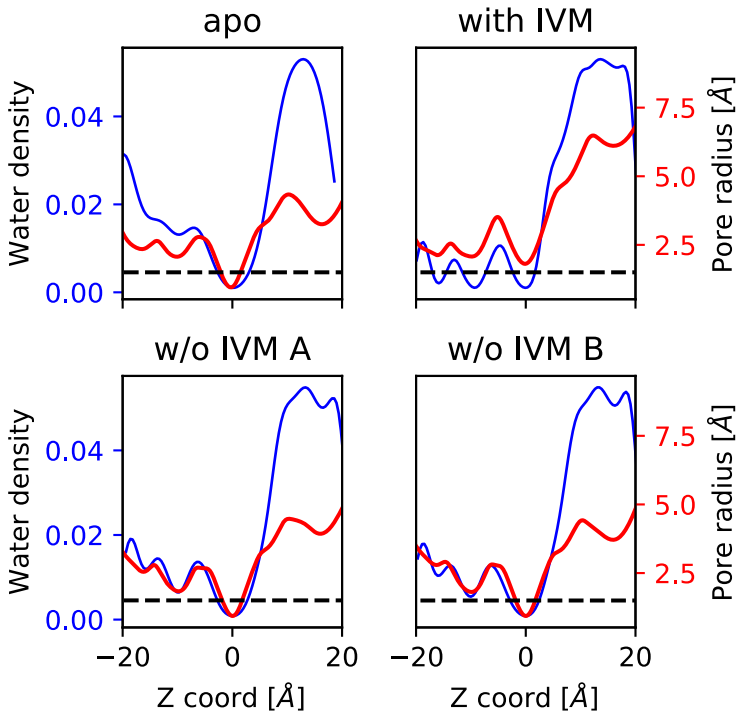


Figure 3.21: Comparison between HOLE profiles (in red) and water densities (in blue) along the pore for four different systems. The zero of the x axis was aligned on the LEU 254 residues in all systems to allow for a fair comparison. For run A and B only the last 100 ns of the simulations where used to compute both observables. For active and apo the whole simulations were considered.

IVM BINDING SITE IVM binds at the interface between two adjacent subunits in a cleft situated in the TM domain. The size of the IVM pocket is measured by the distance between residues GLY 281 on the (+) subunit and residue LEU 218 on the (-) subunit. Our analysis shows that the IVM pocket is significantly wider in the active state (10.5 Å) than in the resting state (8.5 Å) as it is occupied by IVM. Interestingly, the production phase of the simulation A and B in which IVM was removed starts with an IVM binding site already significantly smaller (9 Å) than the one of the active state. Indeed, during the equilibration the TM helices and side chains rearrange quickly to fill the void created by the removal of IVM.
Moreover, distances measured for run A and B decrease from 9 Å to 8 Å in 400 and 700 ns respectively, timing which is correlated with the twist of the EC domain.
In addition, after reaching a minimum at the end of the twist, this distance increases again to 9 Å (more striking on run A) but never reaches the initial values of the active structure. It seems that in order to close the ion pore, the binding site of IVM has to be over compressed and that once the closing is complete it can relax to a wider configuration. This compression could be explained by the coordinated motion of the M2-M3 helices toward the center of the pore.

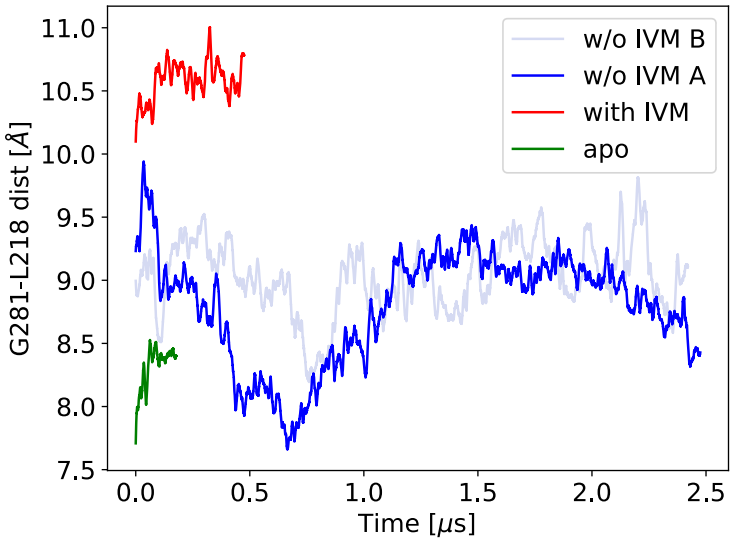


Figure 3.22: Evolution of the IVM pocket. The distance between the C_α of residues GLY 281 on the (+) subunit and residue LEU 218 on the (-) subunit is shown for four simulations.

3.2.1.4 *Correlation between the different observables*

In the former section we have looked at time series of various observables, approach which is not suited to investigate the correlation between events. In this section we will analyze the possible correlations between the various observables we described and try to draw the picture of a new gating mechanism for the pLGICs, based on the analysis of our two repeats capturing the full transition from open to close.

NUMBER OF BOUND L-GLU AND TWISTING L-Glu activates the receptor by a mechanism which is still not yet fully elucidated. To understand the role of the endogenous neurotransmitter on the twist isomerization we have computed the average twist angle in function of the number of ligands bound. Our analysis show that both run A and run B experience a significant increase of the twist angle upon unbinding of L-Glu. It is also evident from the data shown in Figure 3.23 that run B samples higher values of the twist. Interestingly, the lower values of the twist are only sampled for highly ligated states (more than 4 L-Glu), which is coherent with the fact that L-Glu blocks the twisting of the receptor. In addition, one can conclude from our analysis that this receptor samples a continuous range of twist angles. In fact, each event of unbinding seems to allow a higher increase of the twist angle. Consequently, these multiple twisting states may or not be related to different values of pore diameter.

NUMBER OF BOUND L-GLU AND BLOOMING In a similar fashion as what we have done for the correlation between twisting and number of L-Glu bound we analyzed the correlation between the latter and the blooming.
Interestingly, our data show that there exists no obvious correlation between the number of L-Glu bound and the blooming. Indeed, when run A's blooming angle increases as L-Glu

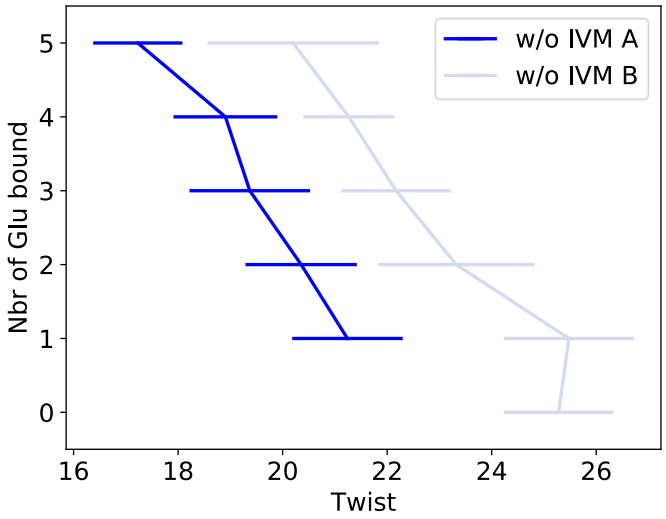


Figure 3.23: Average value of the twist angle in function of the number of L-Glu bound to the receptor for two simulation of GluCl in which IVM was removed. The binding is defined as described in section 2.3.8. The standard deviation is represented by horizontal lines.

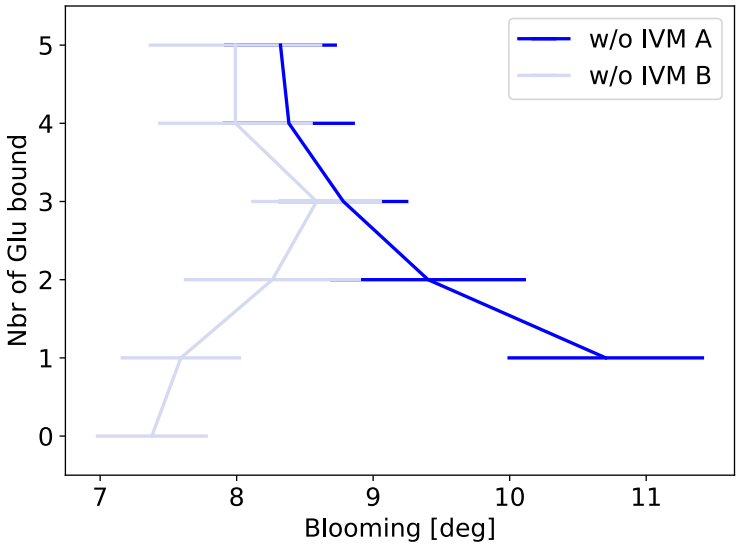


Figure 3.24: Average value of the blooming angle in function of the number of L-Glu bound to the receptor for two simulations of GluCl in which IVM was removed. The binding is defined as described in section 2.3.8. The standard deviation is represented by horizontal lines.

unbinds, the one of run B shows no significant variation upon L-Glu unbinding. As a full blooming isomerization is recorded only for one of the two repeats, it suggests that there exists a significant barrier to bloom the EC domain of the receptor. In addition, these results also suggest that the aforementioned barrier is not due to the presence of the ligand but is

rather encoded in the protein. Indeed, the breaking of contacts between adjacent subunits as the receptor blooms has a significant energetic cost which could be compensated by entropic stabilizing effects.

NUMBER OF BOUND L-GLU AND ION PORE SIZE As the binding of the neurotransmitter controls the function of the receptor, the correlation between the former and the size of the pore must also be investigated. This analysis evidences the correlation between the function

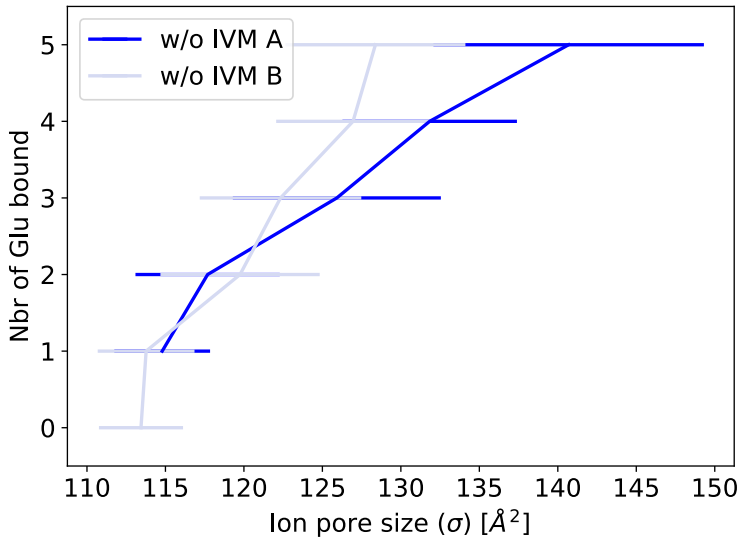


Figure 3.25: Average value of the ion pore size as defined by σ in function of the number of L-Glu bound to the receptor for two simulation of GluCl in which IVM was removed. The binding is defined as described in section 2.3.8. The standard deviation is represented by horizontal lines.

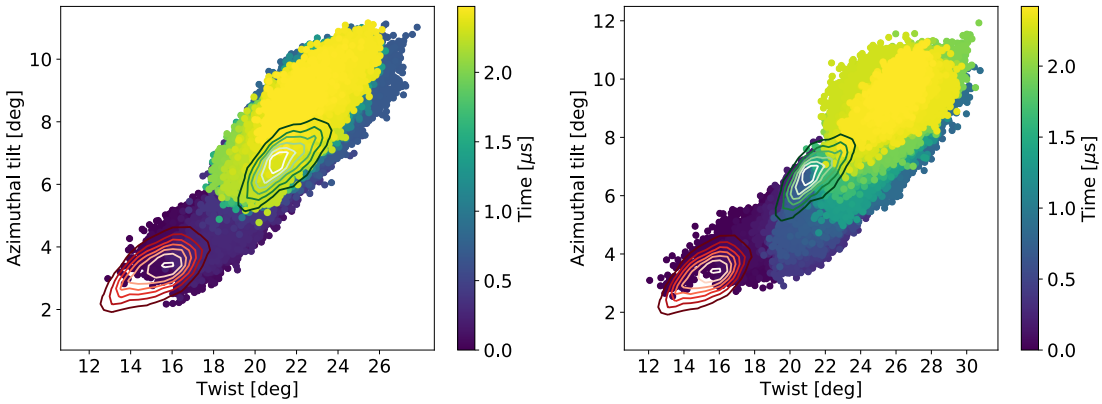
of the receptor and the binding of ligands. Our simulations show that when no ligand is bound, an open state of the channel is never sampled. This may seem peculiar as it is known that pLGICs can populate an open state even in the absence of ligands. Nevertheless, such events are rare and the time scale accessible through all-atom molecular dynamics hardly allows for the sampling of these spontaneous activation events.

Moreover our data show that when 5 ligands are bound to the receptor a fully closed state cannot be sampled. On the other hand, when the number of ligands bound is bellow three, our analysis suggests that the opening of the channel is unlikely, setting the minimum number of ligands required for activation to three.

Finally this analysis also suggests that there exist an ensemble of open states, depending on the number of L-Glu bound.

TWISTING AND AZIMUTHAL TILTING The twisting has often be described as the key isomerization for the gating of pLGICs. As defined in this thesis, the twist is a rigid body rotation of the entire EC over the TM domain and thus can not by directly responsible for the

modification of the interfaces between adjacent subunits. Nevertheless, because EC and TM domains are linked in each subunits, the twisting of the EC over the TM domain may lead to the tilting of the EC domain or even of the full subunit.



(a) GluCl w/o IVM A

(b) GluCl w/o IVM B

Figure 3.26: Correlation between the twisting of the receptor and the azimuthal component of the tilting angle.. The values sampled by the two simulations of GluCl active and apo are represented by red and green isocontour lines, respectively.

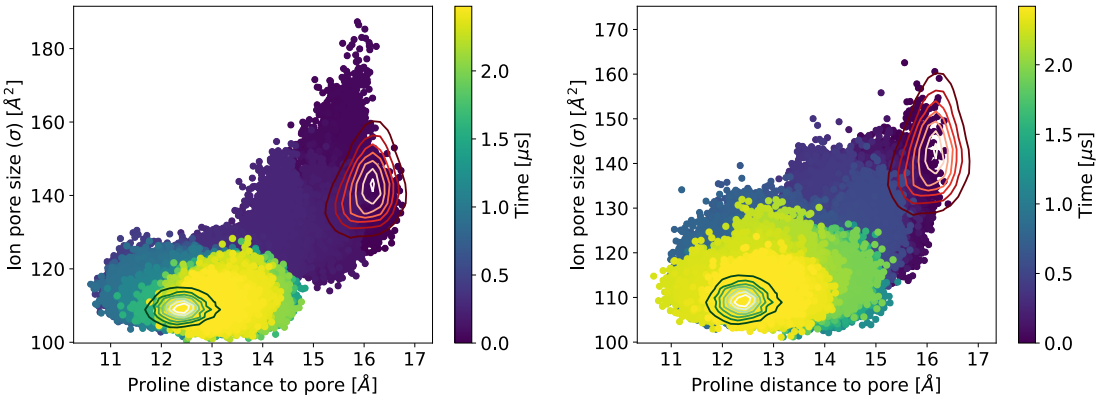
Our analysis shows that twisting of the EC domain over the TM domain and the azimuthal component of the tilting of the EC domains are strongly correlated. Indeed, states with high degree of azimuthal tilt (10-12 deg) and low twist (12-16 deg) are not sampled. Oppositely, our analysis shows that the receptor cannot be fully twisted and keep straight β-sandwiches.

In fact one could see the twist as a proxy for the true observable, whilst the azimuthal tilting of the subunits describes it best. While twisting, the upper part of the EC domain of a given subunit will travel more and the lower part less. The former is nicely represented by the azimuthal or tangential component of the tilting of the β-sandwiches.

Finally, even though this analysis provides few insights on the sequence of events occurring during the gating, it enlightens the relevance of the azimuthal tilting compared to the twisting. Therefore, it suggests that when describing the gating mechanism of pLGICs one could rather use azimuthal tilting instead of twisting.

PROLINE 268 POSITION AND PORE CLOSING We have isolated from time series of both pore closing and the average distance of the residues P268 to the center of the pore a time correlation as both events happens within the same time frame, i.e., 400 ns and 750 ns for run A and run B respectively.

The correlation plot between these two observables confirms that such a correlation exists. Indeed, in both runs, the ions pore size in 9′ decreases with the distance to the pore. Moreover, our data show that it is unlikely to reach full closing if the prolines are not in an *in* position. On the other hand, it also seems unlikely to observe a configuration in which all the proline residues would be outward and the channel closed. This observation confirms that the closing



(a) GluCl w/o IVM A

(b) GluCl w/o IVM B

Figure 3.27: Correlation between the the ΔZ distance and the twisting of the receptor. The values sampled by the two simulations of GluCl active and apo are represented by red and green isocontour lines, respectively.

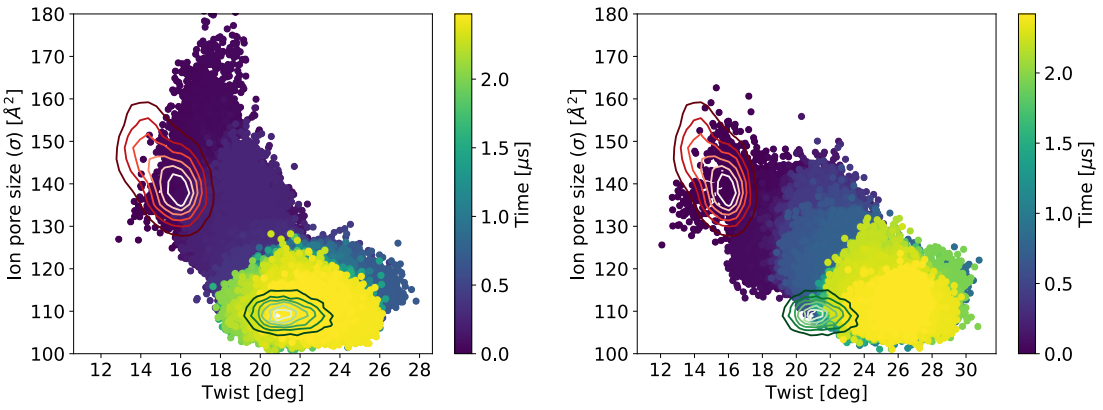
of the pore occurs thanks to an inward displacement (toward the lumen of the pore) of the P268 residues.

TWISTING AND PORE CLOSING The correlation between twist and pore closing is evidenced by the data shown on Figure 3.28. Our data show that the receptor cannot be fully twisted and in an open state and that it can neither be fully untwisted and in a closed state. Expectedly, both repeats of the transition confirm the correlation suggested by the analysis of times series.

Moreover, this analysis compares with the former correlation made between the closing of the ion pore and the position of the P268 residues, indicating that if twist and pore closing are correlated, then the twist and the inward displacement of the proline residues are as well.

TWISTING AND ΔZ DISTANCE The overall twisting of the receptor and its correlation with the distance between the C_α of residues V45 located on the $\beta 1 - \beta 2$ loop and P268 on the M2-M3 loop were studied for the four simulations of interest. Surprisingly, this analysis shows a good correlation between the two variables. As shown by time series, the values for the active and resting simulations are significantly different and the two repeats of the simulation of GluCl without IVM bridge the gap between the active and resting structures. Despite a slightly different behavior of the two runs, the conclusion can be made that as the twist increases the ΔZ distance increases as well.

In fact, it is likely that the twisting of the EC domain of a given subunit pulls upward the $\beta 1 - \beta 2$ loop increasing the ΔZ distance, thus explaining the observed correlation. Indeed, and as discussed in an earlier paragraph, the twisting corresponds to an azimuthal rotation of the EC domain of a given subunits. If one thinks of a rotation around the center of geometry



(a) GluCl w/o IVM A

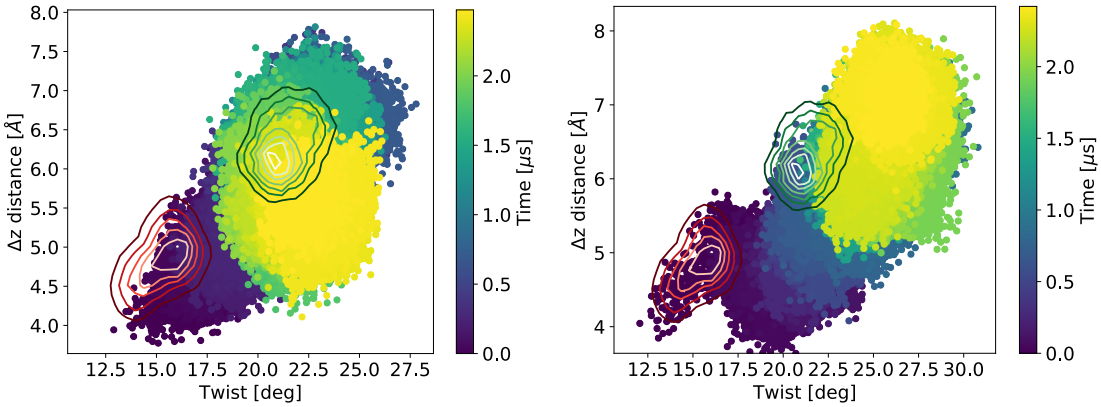
(b) GluCl w/o IVM B

Figure 3.28: Correlation between the the closing of the pore (σ) and the twisting of the receptor. The values sampled by the two simulations of GluCl active and apo are represented by red and green isocontour lines, respectively.

of the EC domain (and not around the lower part of this domain, which would correspond to a hinge rotation), the lower part of the EC domain would be pulled upward, therefore increasing the ΔZ distance.

TWISTING AND BLOOMING As these two observables are key to understand the gating mechanism of pLGICs their correlation was carefully studied for the four simulations of interest. First of all, our results show that both active and resting simulations are distinct and stable. When the active simulation populates conformations which are globally untwisted and unbloomed, the resting simulation populates bloomed and twisted configurations of the system. This first observation confirms the relevance of the twisting and blooming angles to discriminate between active and resting like structures.
Our analysis shows that both repeats of the simulation in which IVM was removed from GluCl behave differently. From run A only we can make the conclusion that both blooming and twisting are correlated, whilst considering run B leads to a different conclusion. These results indicate that the receptor can be fully twisted (24 to 28 deg of twist angle) without experiencing a significant blooming (8 deg of blooming angle).
On the other hand, it shows that when fully untwisted (12 to 14 deg of twist angle) it cannot be bloomed. This is most likely due to the fact that ligand binding controls both the twisting and the blooming. Nevertheless, ligand unbinding, if it triggers the twisting isomerization, does not necessarily triggers the blooming one.

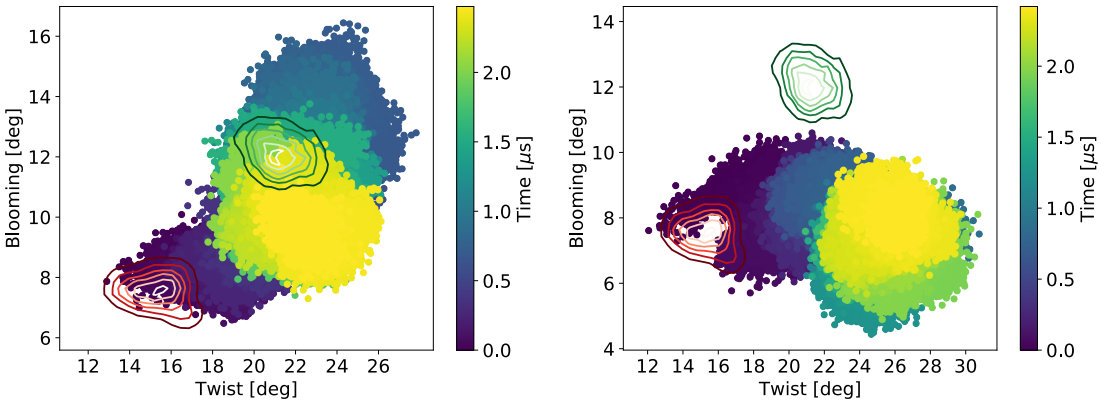
POCKET RMSD AND BLOOMING Since active and resting structures are different in the EC domain, one may hypothesize that one of the main transformation experienced by this domain changes significantly the orthosteric site. Moreover, the changes in the binding site of a given ligand is responsible for the selectivity of the structure to which it will bind. Hence,



(a) GluCl w/o IVM A

(b) GluCl w/o IVM B

Figure 3.29: Correlation between the the ΔZ distance and the twisting of the receptor. The values sampled by the two simulations of GluCl active and apo are represented by red and green isocontour lines, respectively.

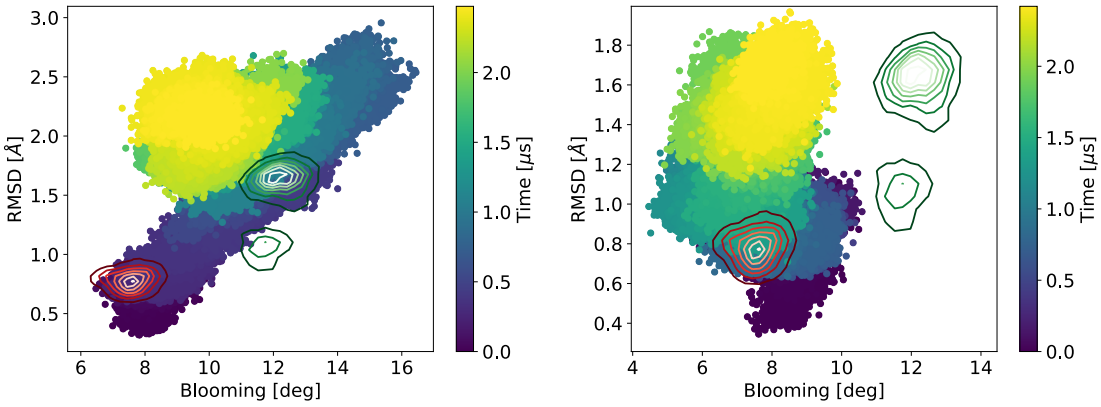


(a) GluCl w/o IVM A

(b) GluCl w/o IVM B

Figure 3.30: Correlation between the the blooming angle as described by the polar component of the tilt (θ_p) and the twist of the receptor. The values sampled by the two simulations of GluCl active and apo are represented by red and green isocontour lines, respectively.

an agonist will stabilize an active configuration of the protein and thus of the binding site and reversely for an antagonist. As we have shown that the blooming angle is significantly different in the active compared to the resting state, one may think that the blooming transition changes the shape of the orthosteric pocket. Moreover, as the blooming angles of run A and B do not undergo the same transition, different results regarding the correlation studied are expected. Indeed, run A shows a strong correlation during the first microsecond of the simulation between the two observables suggesting that the RMSD changes up to 3 Å of



(a) GluCl w/o IVM A

(b) GluCl w/o IVM B

Figure 3.31: Correlation between the the blooming angle as described by the polar component of the tilt (θ_p) and RMSD of the resides constituting the orthosteric pocket. The values sampled by the two simulations of GluCl active and apo are represented by red and green isocontour lines, respectively.

the orthosteric binding pocket are due to the blooming of the receptor. On the other hand, run B, which shows a lower increase of the RMSD over time, i.e, 1.8 Å, shows no apparent correlation between the blooming and the RMSD of the orthosteric binding pocket.

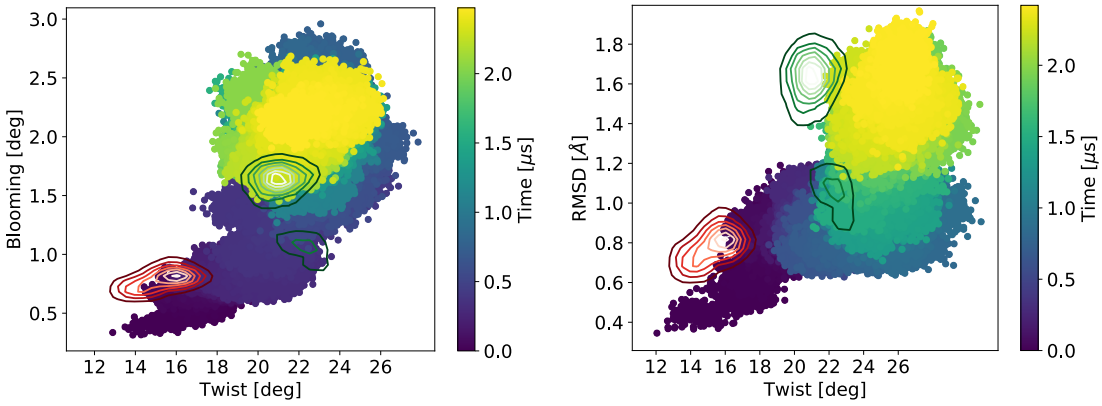
POCKET RMSD AND TWISTING For similar reasons as the one previously described for the blooming we computed the correlation between the twisting angle and the RMSD of the orthosteric pocket. Our results show that during the first part of run A (0 to 500 ns), which corresponds the to full twisting isomerization, the RMSD of the pocket increases of about 1Å. In the second part of the simulation the RMSD increases from 1 to 3 Å with not apparent increase of the twist. Run B leads to a similar observation for the first part of the simulation as very little changes in the RMSD are observed. Nevertheless, the increase of RMSD observed during the second part of the simulation (after 1µs) cannot be attributed to the blooming or the twisting of the receptor, suggesting that another reaction coordinate could impact on the shape of the orthosteric site.

The former statements are in agreement with the conclusion made in the previous section and lead us to conclude that twisting the receptor changes mildly the topology of the active site whilst blooming is responsible for its major modifications.

3.2.2 *Mechanistic description of the gating*

Based on the analysis made in the two former sections we proposed a detailed gating mechanism for GluCl that we believe applicable to most pLGICs.

Interestingly, the proposed mechanism shares common features with the previously described ones[3,6,181,183,184] and brings important new insights.



(a) GluCl w/o IVM A

(b) GluCl w/o IVM B

Figure 3.32: Correlation between the the blooming angle as described by the polar component of the tilting (θ_p) and the twisting of the receptor. The values sampled by the two simulations of GluCl active and apo are represented by red and green isocontour lines, respectively.

Although the latest proposed models generally agree on that gating involves a global isomerization composed of both twisting and blooming, they provide little details on the mechanistic role of these movements, or of their modulation by ligand binding, both orthosteric and allosteric.

Our analysis illuminates how the twisting controls the gating of GluCl and how blooming plays an important role in ligand recognition. In fact, we show that blooming is not required for channel closing and that instead, twisting is the only molecular requirement for closing.

The proposed mechanism offers a unified view of gating, making the link between quaternary changes such as the twisting and local displacements as best exemplified by the one observed at the interface between EC and TM domains. Indeed, as the receptor twists, the $\beta1-\beta2$ loop moves upward freeing the way for the prolines in an *out* position to move inward. Therefore, the position of the $\beta1-\beta2$ loops at the interface between the TM and EC domains controls the closing of the ion pore 30 Å away by interacting with the conserved P268 residues on the M2-M3 loops. Remarkably and in agreement with our interpretations, the important mechanistic role of the EC/TM interfacial loops in the gating of pLGICs has recently been discussed in a study presenting a string method optimization of the gating pathway in GLIC[181]. Nevertheless, we could not identify a direct causality between the expansion of the lower part of the β-sandwiches and the closing of the pore in our simulations capturing the un-gating of GluCl.

Moreover our captured transition from open to close of GluCl confirmed that closing of the ion pore involved the inward motion of the M2 helices and the M2-M3 loops without the need for decoupling, unlike described for GLIC locally-closed.

In addition, our analysis allowed for the detailed investigation of the mechanism of closing. Our simulations confirmed that the channel closed at the $9'$ position as observed in the *apo* structure of GluCl elucidated by X-ray crystallography.

The gating mechanism of pLGICs has also been described as hydrophobic[185], which means that the pore is not physically closed but a vacuum stretch is created at the constriction point, making the passage of ion impossible as it significantly increases the barrier for permeation. Our analysis has shown that the pore was physically shut and therefore that a hydrophobic gating mechanism was incompatible with our data.

The former conclusions naturally leads to the following gating mechanism as represented on Figure 3.33 ; i) unbinding of the ligands ii) twisting of the receptor pulling upward the $\beta 1 - \beta 2$ loop and freeing the way for the M2-M3 loop to move inward iii) inward motion of the M2-M3 loop and of the M2 helices leading to the shrinking of the ion pore which ultimately closes in $9'$.

In conclusion, our proposed mechanism bridges the gap between ligand unbinding and pore closing 60 Å. We proposed a series of events and their coupling that shows in the clearest fashion how, mechanically, the signal of unbinding is propagated to the constriction point.

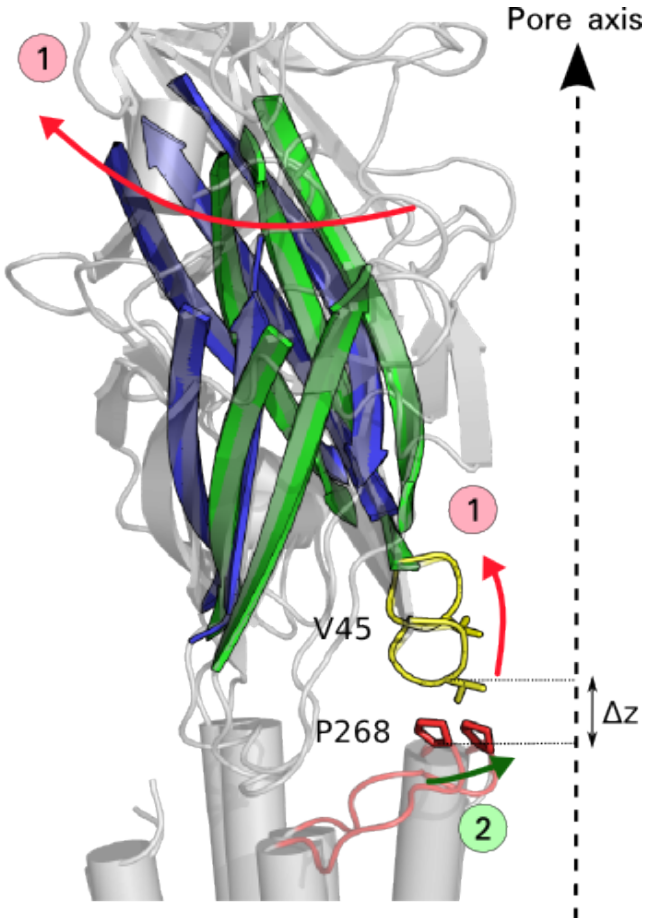


Figure 3.33: Proposed mechanism based on the analysis of four simulations of GluCl and in particular two repeats of the full closing transition. The active state is shown in green and the resting in blue.

3.3 MODULATION BY LIGAND BINDING

In the former section we have presented a new gating mechanism which clarifies the role of the blooming and twisting in the closing of the ion pore. In this section we will discuss the modulation of the receptor's response in function of its binding to L-Glu alone or L-Glu and IVM.

First of all, our analysis of the transition from open to close of GluCl suggests the existence of a metastable state on closing, which we think to be the active state stabilized by L-Glu and that will be described in details below.

Nevertheless, because of the spontaneous unbinding of L-Glu in the two simulations in which IVM was removed we could not study in details this state. Hence, in order to better understand the role of each ligands, i.e., L-Glu and IVM, a simulation of GluCl bound to L-Glu only was setup as described in section 2.5.3 and analyzed in a similar fashion as done for the four other simulations.

3.3.1 *Metastable state on closing*

The states sampled during the first 300 ns of the simulation of GluCl w/o IVM A show numerous characteristics of an open state whilst IVM has already been removed. Moreover, the kinetic stability of this metastable state is not negligible as it is sampled for about 300 ns and is time-correlated with the presence of L-Glu.
The isolated ensemble of states is characterized by an ion pore size compatible with an open state as shown by the measure of σ in Figure 3.5 for the first 300 ns of run A. In addition, it is also compatible with an open state while looking at the HOLE pore profile (see Figure 3.34D) and the water flux (see Figure 3.8). Altogether, these results confirm that like when GluCl is bound to IVM and L-Glu, the isolated metastable state, which is bound to L-Glu only, can be annotated as an active state. Interestingly, the metastable state's twisting angle is halfway (18 deg) between the State bound to L-Glu and IVM (14 deg) and resting or *apo* (22 deg) states, when its blooming angle (8 deg) is more compatible with an active configuration (8 deg). This state is thus partially twisted and unbloomed as illustrated on Figure 3.34A.

In addition, the position of the prolines, i.e., their distance to the center of the pore is also compatible with an open state (see Figure 3.16).

Moreover, the size of the binding pocket of IVM as measured by the distance from the C_α of residues LEU 218 on the (-) subunits and GLY 281 on the (+) subunit, in the isolated metastable state, is compatible with a resting state of the receptor (illustrated on Figure 3.34C). This result suggests that the receptor, at least locally, as no memory of the presence of IVM as it presents a closed binding site.

Taken altogether these results suggest that the metastable state isolated is open and activated by L-Glu only.

Interestingly, even though the IVM molecules have been removed from their active sites, the diameter of the pore remains constant (see Figure 3.34D). The former suggests that the IVM does not potentiate the channel by increasing its ion pore size but by stabilizing further

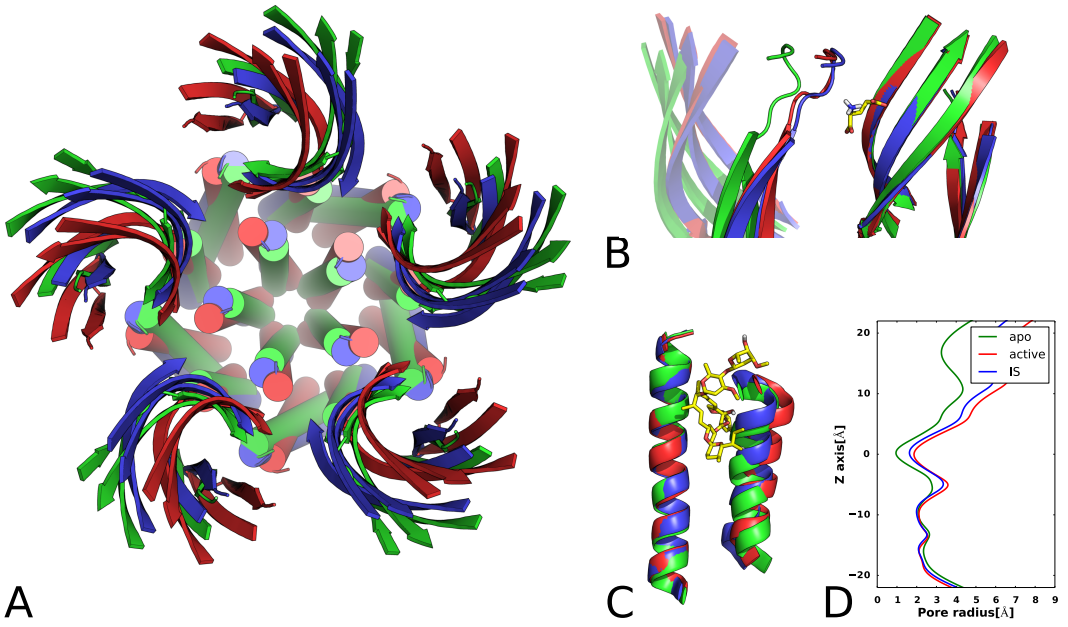


Figure 3.34: Characterization of the metastable state on closing. The color code used in all four panels is red for the active structure, blue for the metastable state investigated and green for the resting state. Panel A is a top view of the three structures, aligned on their TM domains, evidencing the twist of the metastable state and its contracted EC domain. Panel B shown the configuration of the orthosteric site with L-Glu represented in yellow and panel C the one of the allosteric site of IVM (in yellow). Panel D is the average pore profile for the three states zoomed on the TM domain and centered on the 9′ residues.

in twisting the active state and therefore delaying its desensitization. In other words, the presence of IVM increasing the twist angle could lock the receptor in a super-untwisted state (compared to the untwisted state of GluCl bound to L-Glu only) which would represent a higher barrier for desensitization and thus increase the overall flux of ion through the cell's membrane.

As one could argue that the stability of the isolated state is marginal and that it rather is a transition state upon closing, we simulated and analyzed a simulation of GluCl bound to L-Glu only.

3.3.2 *GluCl stabilized by L-Glu*

The simulation of GluCl bound to L-Glu only was set up with restrains on the interactions between L-Glu and the protein, as spontaneous unbinding of the endogenous ligand was observed (see section 2.5.3).

Even though our simulation is significantly shorter than the two simulation of GluCl w/o IVM, its length of 300 ns seems to be sufficient for concluding that the simulation is converged. Indeed the RMSD of both EC and TM domains (see Figure 3.35) remains stable below 1.5 Å from the crystallographic structure indicating that no major transition occurs within

the simulation time. Moreover, the fact that a transition, partial or total, is observed upon removal of IVM in run A and B in less than 300 ns and not in the simulation of GluCl bound to L-Glu is another argument in favor of the stability of this shorter simulation.

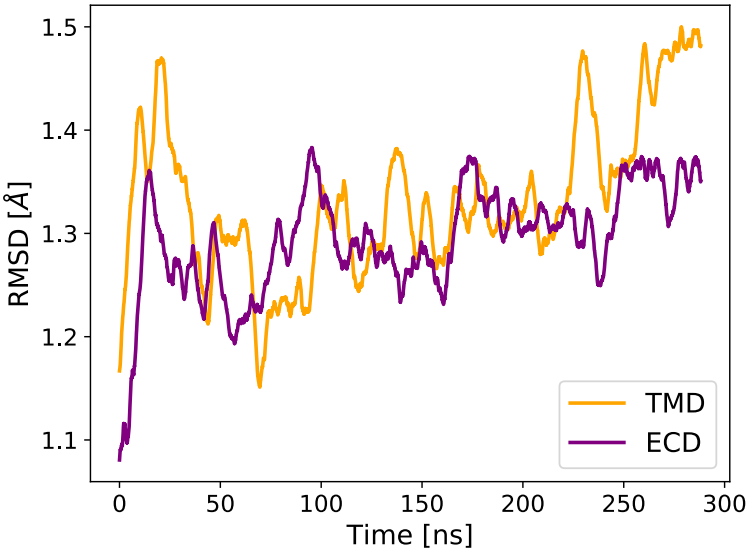


Figure 3.35: C_α RMSD of the core domain for the EC and TM domains for the simulation of GluCl bound to L-Glu only.

Moreover the twist angle is stable over the length of the simulation at an average value of 16.7 deg, which is consistent with the values of the twist angle measured for the metastable state. In addition, the value observed for the twist is significantly different from the one measured for GluCl bound to IVM and L-Glu and for GluCl apo.
The value measured for the blooming angle (7.8 deg) is consistent with the one of an active state (8 deg) and thus with a contracted form of the EC domain. Expectedly, the number of ligands bound to the receptor is always of 5 as restrains are applied on the interaction between the ligands and the receptor. However, despite the use of restrains, we monitored a non negligible increase of the RMSD of the ligands, indicating that our restrains, while keeping the ligand bound to the receptor allow to keep the intrinsic flexibility of the EC domain (see EC RMSD in Figure 3.35). Finally, the analysis of this simulation confirmed the relevance of the metastable state isolated in the simulation of GluCl in which IVM was removed.

We have shown that upon removal of IVM the receptor evolves to a state which is partially twisted (18 deg) but which keeps a contracted EC domain. Our analysis has also shown that this metastable state despite the closing of the IVM binding site kept an open pore configuration.

Our simulation of GluCl bound to only L-Glu showed a lot of similarities with the metastable state discussed in the former section, such as an open ion pore, a contracted and partially twisted EC domain and a closed IVM binding site.

Altogether these results strongly suggests that the discussed configurations are most representative of an active state of GluCl activated by L-Glu.

3.3.3 *Conclusion*

The modulation of the function or activation of GluCl is made possible by the binding and unbinding of L-Glu, the endogenous neurotransmitter, on one side, and allosteric ligands such as IVM on the other side.

Our results have shown that L-Glu plus IVM bound to the receptor stabilized twist angle values of 12 to 14 deg, L-Glu only of 18 deg and that the absence of ligand stabilized a twist angle above 22 deg. Hence, we concluded that both IVM and L-Glu impacted on the same mechanism, i.e., the twisting isomerization, even though they bind at topographically distinct sites.

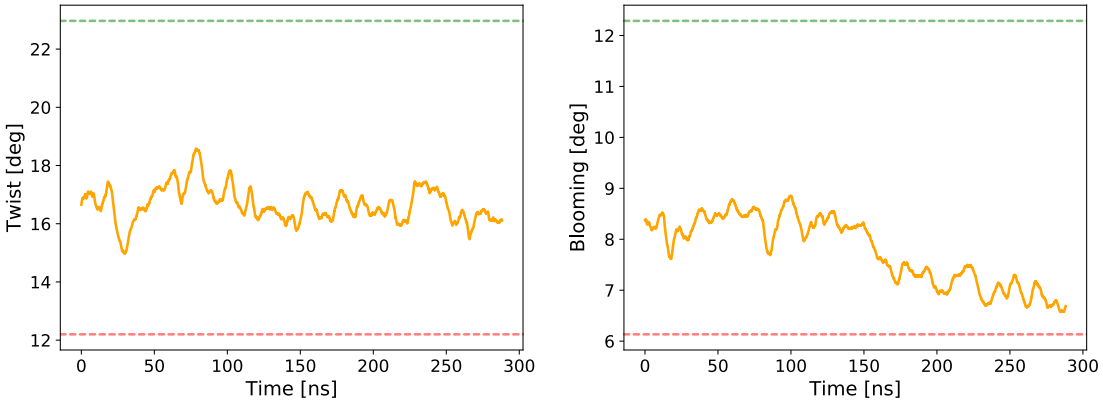
Moreover, we have shown that the blooming of the receptor was controlled by the binding of the endogenous ligand at the orthosteric site whilst the unblooming was encoded in the protein itself. Indeed, when L-Glu was bound we did not observe a blooming transition in any of our simulations. The conclusion was also made that even though L-Glu can lock the receptor into a contracted state, the blooming of the EC domain in the absence of the endogenous ligand represents a significant cost in energy and therefore is not always sampled, as exemplified by run B of our simulation of GluCl without IVM.

Our mechanism therefore suggests that twisting is key in the activation and deactivation of the receptor as it controls, indirectly, the opening and closing of the gate, but that blooming is also crucial in the activation mechanism. Indeed, it seems reasonable to think that a pre-active state of the receptor, i.e. an unbloomed or contracted EC domain and a closed TM domain exists and is capable of binding the endogenous neurotransmitter.

The binding of both IVM and L-Glu is associated with a *super*-untwisted state of the receptor compared to the one of GluCl bound to L-Glu. This observation suggests that the binding of IVM stabilizes further the active state, increasing the barrier for desensitization of the receptor. The former indicates that the capability of IVM to positively modulate the response of GluCl and thus increase the flux of ion through the cell membrane is not due to an increase of the ion pore size but instead to an increase of the time before desensitization.

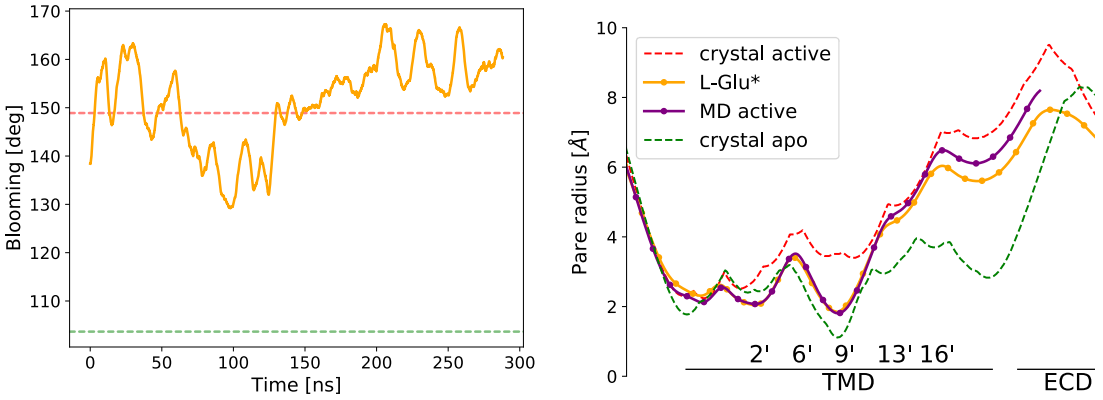
In the presented picture, the blooming isomerization can be seen as a ligand recognition mechanism. Indeed, our data suggests that a bloomed configuration of the EC domain is not compatible with an open state of the channel. Therefore, a ligand capable of locking the receptor into a bloomed configuration could prevent the channel from opening and then lead to a diminution of the receptor's response to the endogenous neurotransmitter. Either by blocking the access of the active site to the endogenous neurotransmitter or by hindering the mechanism by which the receptor is activated, i.e., the twisting, a ligand could modulate the receptor's response.

Our findings open the door to the design on new drug candidates, which could block the twist isomerization, or contrarily stabilize, as IVM does, a *super*-untwisted configuration to potentiate the response of the receptor.



(a) (b)

Figure 3.36: Evolution of the twist angle (a) and of the blooming angle (b) or polar tilt for a simulation of GluCl bound to L-Glu. The green and red dashed line represent the values for the apo and active X-ray structures, respectively.



(a) (b)

Figure 3.37: Evolution of four examples of observable illustrating the functional state of the simulation of GluCl bound to L-Glu. The green and red dashed line represent the values for the apo and active X-ray structures, respectively.(d) shows the HOLE profile averaged over the full length of the simulation1 and zoomed on the TM domain. The zero is placed on the position of the constriction point (LEU 254).

4

EXPLORING THE ALLOSTERIC MODULATION OF PLGICS BY SMALL MOLECULES

We have proposed in the previous chapter a new gating mechanism for the pLGICs in which both orthosteric and allosteric ligands controlled the twisting isomerization. pLGICs being major pharmaceutical targets it is crucial to better understand how allosteric ligands can modulate their action. Indeed, allosteric binding sites present the advantage of being numerous as they can be located anywhere in the protein. Moreover, and unlike orthosteric binding sites which undergo a strong evolutionary pressure to remain specific of the same binder, allosteric binding sites are much more diverse in sequence and thus allow for a better selectivity between for instance different subunit types. Allosteric binders offer the possibility to have a more subtle effect than orthosteric ones which often have a binary effect, i.e., agonist or antagonist. For these reasons designing allosteric modulators seems promising (more details can be found in Ref[186,187]).

Our approach to understand the allosteric modulation of pLGICs is based on free energy calculation in the presence and absence of various ligands. We aim at understanding if the twist can be controlled by molecules which are not known binders. As a first step we will compute the PMF of three different states; i) GluCl +IVM+L-Glu, ii) GluCl +L-Glu and iii) GluCl active in which both ligands were removed. As we want to understand the effect of allosteric modulators we will compare mainly the PMF of systems i) and ii). The transition between system ii) and iii) being considered as the effect of the orthosteric ligand (L-Glu) in which we are not interested in this study.

In a second time we will try to identify a fragment of IVM which is significantly smaller than the full IVM and which remains able to bind to the receptor and to block its twist. This fragment could then be used as a base for rational drugs design, or to screen databases of putative binders sharing the same scaffold.

Finally, we will use the reference PMF computed before and will use them to discriminate between putative potentiators.

4.1 PROTOCOL DEVELOPMENT

4.1.1 *Twisting*

The twist was initially described as the average of 5 characteristic dihedral angles (see Figure 4.2). However, because of its implementation in the NAMD simulation package, it is not possible to use straight forward dihedral angles to restrain accurately the twist.
Therefore, we decided to use the spinangle collective variable, which seemed adequate, with some minor adaptations, to describe the twist. The choice of the spinangle CV had to be validated as a first step. To do so, we chose the butane molecule as a model because it is

small enough to allow for the efficient sampling of the spontaneous rotation of its main dihedral angle. In addition to that, the butane molecule is suited to compute both dihedral and spinangle collective variables. The PMF of the full rotation (360 deg) of the dihedral angle formed by the four carbons of the butane molecule (referred to as χ) was computed using either the spinangle or the dihedral angle as restrained collective variables[188].
The results presented in Figure 4.1 clearly show that restraining either the spinangle (ϕ) or the twist angle (τ) leads to a similiar PMF and may therefore be considered as equivalent. Moreover, the barrier between *cis* and *trans* basins ($\Delta F_{cis/trans}$) computed using either of the two collective variables is very similar, i.e., 0.1794 and 0.1802 Kcal/mol/deg^2 for ϕ and τ respectively.

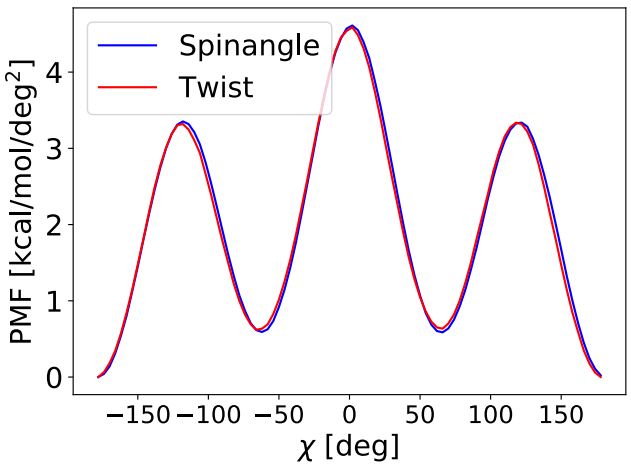


Figure 4.1: PMF along the main dihedral angle of the butane molecule computed with spinangle or dihedral as a collective variable for the US calculations.

Once these promising results obtained for the butane molecule, we analyzed a free simulation of GluCl in which IVM was removed and a twisting transition was observed to assess for the correlation of ϕ and τ in a bigger and more complex system. The results presented in Figure 4.3 show a very strong correlation between these two values and allow to define the mathematical relation between spinangle and twist angle as follows :

$$\tau = \phi + 12.3 \tag{4.1}$$

where 12.3 corresponds to the value of τ for the reference structure used to compute ϕ (see section 2.5.4).

4.1.2 *Blooming*

The blooming angle also referred to as polar tilting angle (θ_p) throughout this work describes the radial contraction or expansion of the EC domain. As all five subunits may undergo a different and complex blooming transition one must design an observable which

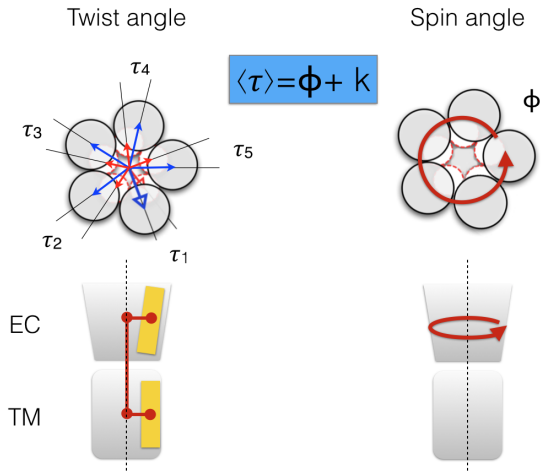


Figure 4.2: Scheme representing the conceptual differences between ϕ the spinangle and τ the twist angle. $\langle\tau\rangle$ is the average over the 5 subunits and k is the average twist angle of the reference structure, i.e., crystallographic structure.

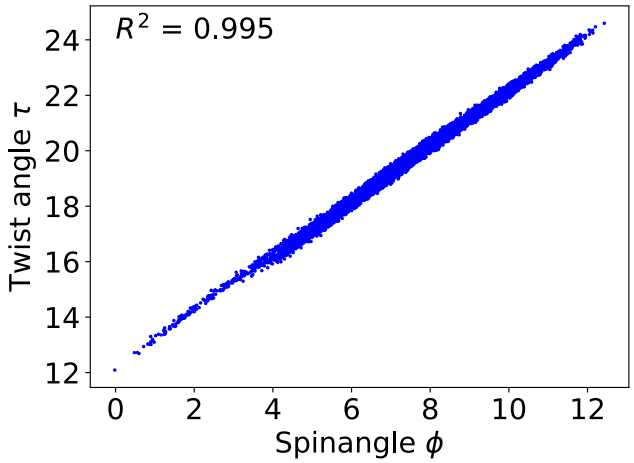


Figure 4.3: Correlation between the twist and spinangle calculated in a 2.5 µs long free simulation of GluCl in which IVM was removed.

will independently restrain each subunit. As a first approach we thought to restrain the inter-subunit distances, either between adjacent or opposite subunits. As previously done for the spinangle, we computed the correlation between such distances and the blooming angle in order to validate the choice of this rather simple observable. The correlation being low we assumed that the approximation of restraining a distance to explore all θ_p values was incorrect. It is also not suitable to use the spinangle, defined as a self-rotation around a given axis that goes through the center of mass of a given selection, because the blooming is a hinge motion of the EC domain of a given subunit.
In fact in order to make the best use of enhance sampling methods based on RC such as US, one should design an observable that best describes the transition of interest as each crucial

orthogonal degree of freedom will have to be spontaneously sampled, making the computational effort significantly higher. In other word, the approximations made in the choice of a collective variable to describe a complex transition such as the blooming must be paid in computer time.
In conclusion, we had to use a more complex but very flexible approach based on quaternions, which allows for accurately restraining the blooming isomerization.

4.1.2.1 *Introduction to quaternions*

Quaternions are mathematical objects first introduced by William Rowan Hamilton[189] and often used to describe rotations. For instance NAMD uses this formalism to compute the RMSD or to monitor the overall orientation of a protein in the Cartesian referential. By analogy with the complex number, one can describe the quaternion q as the combination of a real and a complex or imaginary part. It can be written as follows $q = a + bi + cj + dk$ where a is the real part and i, j, k are the fundamental quaternion's units (complex part). Let us define $q1$, $q2$, $q3$, $q4$ the normalized components of the quaternion q, respectively corresponding to a, bi, cj, dk and Q the normalized quaternion. Q can then also be expressed as one Euler angle ($2 \times \arccos(q1)$) and one Euler axis of rotation $q2, q3, q4$. The former expression is more convenient to visualize quaternions in molecular visualization programs.

4.1.2.2 *Protocol*

To best reproduce the motion which transforms the EC of the active structure into the one of the resting structure, which involves a blooming, we computed the quaternion q^{ref} which defines the best rotation fit between two specifically chosen structures. Indeed, if the former was done on crystal structures, and the transition between active and resting states involving both twisting and blooming, the computed quaternion would describe both reaction coordinates, which we want to avoid. In order to compute a quaternion which only describes the blooming we selected two structures from molecular dynamic simulations so they would have similar twist angles but very different blooming angles. Since the starting structure for the PMF calculations along the blooming is symmetrical, computing $qref$ for one out of five subunits is sufficient. One can then simply rotate q^{ref} by one fifth of a turn around the axis of the pore to produce the four other quaternions.

In order to assess the correlation between the restrained quaternions and the polar tilt (θ_p) we computed the latter during a simulation in which the former was restrained (results shown in Figure 4.4). The correlation coefficient between the quaternion angle and θ_p is 0.97 (see Figure 4.4 AVG). It is interesting to mention that the fluctuation within subunits are compensated in the average, underlying the fact that the subunits are still coupled.

Let us define $s = 1 \ldots 5$ the index of a given subunit of our system and $n = 1 \ldots N$ the index of the US windows. q_s^{ref} is then used to generate all the required q_s^n which will later be used as the target centers of each individual US window. The generated quaternions only vary in their real part, the angle, the complex part defining the axis of rotation remains unchanged, so only the amplitude of the initially computed rotation changes.

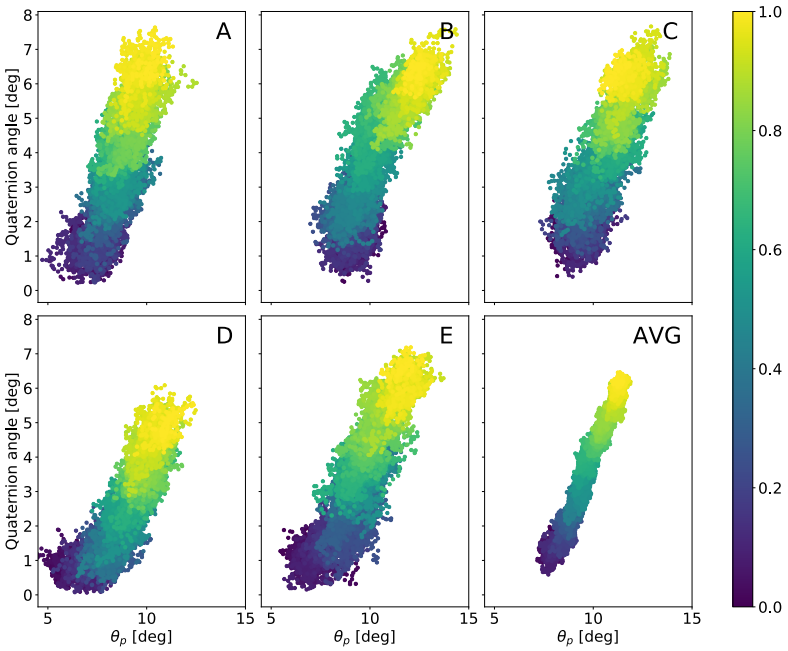


Figure 4.4: Correlation between the angle of the restrained quaternions and the polar tilting for each subunits and averaged over the five subunits of the EC domain. The color scale represents the time.

In order to prevent the system form being pulled too quickly out of its equilibrium position, which may induce non-physical behavior as the bias used increases quadratically, N calculations were run in series, i.e. the starting coordinates of window n were taken from the last frame of window $n-1$, to create a starting structure for each window. The coordinates and velocities of the n starting points produced thanks to this "stepped SMD" (Steered Molecular Dynamic) were then used to run BEUS calculations along the blooming angle.

As discussed by Moradi et al[147], the overall biasing potential per window can be seen as the sum of the five biasing potential applied on each subunits, which makes the unbiasing step performed using WHAM or gWHAM trivial. The force constant was set to 10,000 Kcal/mol/rad^2.

In conclusion, the approach we have developed to compute the PMF along the blooming angle using quaternions as the collective variable restrained seems accurate and promising. Unfortunately, only very preliminary results were obtained regarding this reaction coordinate partly due to time limitations but also because free MD have shown that closing of the ion pore was possible without the blooming of the receptor. The former changed the focus of our work toward the twisting as this development was left aside.

4.1.3 *Umbrella sampling on GluCl*

Umbrella sampling calculations can be challenging to set up since several parameters have to be optimized by the user. To find the best set of parameters we proceeded by trial and errors

aiming at obtaining a full coverage of the range of interest of the twist and a sufficient overlap between adjacent windows, to reconstruct the PMF as shown on an example PMF calculation in Figure 4.5. Indeed, if the chosen force constant is too high the sampling around the target value of a given window might be too restricted and the system might start experiencing harmonic motion, which will lead to increasing the number of windows to cover the full range of interest. On the other hand, if the force constant is too low, it might not be sufficient to sample high energy states and therefore lead to a PMF poorly converged around maxima.

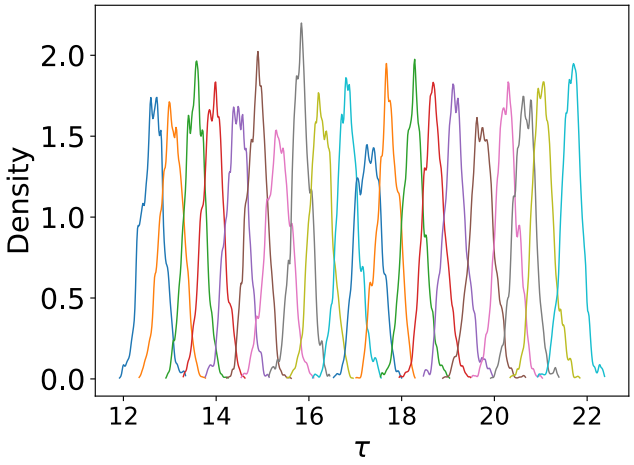


Figure 4.5: Density of probability plotted for all windows of a given PMF calculation. This plot illustrates the full sampling coverage of the range of τ of interest. Each color represents a different windows/simulation.

4.2 REFERENCE PMF FOR THE ACTIVE AND RESTING STATES

From free simulations of GluCl with and without IVM as long as from observations made on structures elucidated by X-ray crystallography, it can be stated that an active state is stabilized at values of $\tau \approx 14$ deg when a resting state is stabilized at values around 22 deg. The first step of this part of the study was to confirm that our method was able to discriminate between an active and a resting state and that ideally, it could identify energy minima corresponding to the values of τ observed in free MD.

To these ends, three different systems were studied, GluCl bound to both IVM and L-Glu, GluCl in which IVM and L-Glu were removed and GluCl bound to L-Glu only. We chose not to study the case in which IVM only was bound because we wanted to understand the allosteric modulation and not the possible activation by allosteric binders or the orthosteric activation. Indeed, even though it was claimed that GluCl is activated by IVM and potentiated by L-Glu unlike other pLGICs in which IVM is a potentiator, this very question is still under debate.

To compute the PMF of the twist we decided to go in the forward direction, i.e., 12 to 22 deg, and thus started with the active structure of GluCl. Moreover, starting from the

aforementioned structure offers the possibility to observe free energy changes upon ligand removal when it is more challenging to observe the same upon ligand binding. Indeed the binding of a ligand to a receptor is very challenging to sample in free MD as it requires to sample; i) the spontaneous activation of the receptor in the absence of ligand and ii) the stabilization of the receptor in its active configuration by the biding of the ligand. The results of the PMF are presented in Figure 4.9.

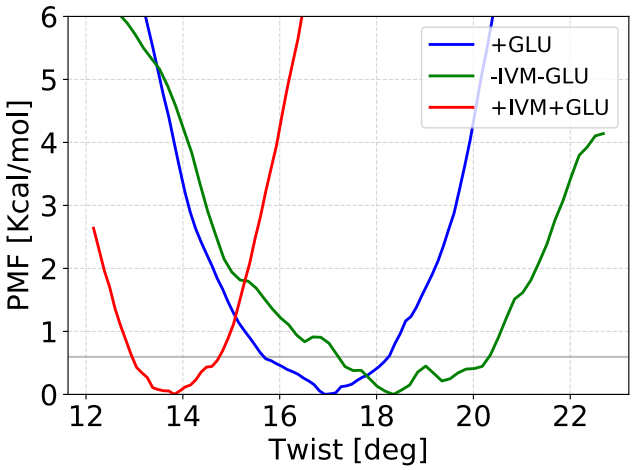


Figure 4.6: PMF for the three reference states of GluCl. The gray horizontal line represent the value of K_bT at 300 K.

Interestingly when both IVM and L-Glu remain bound to the receptor, the minimum in free energy is located around 14 deg of twist, which is in agreement with free MD simulations. Moreover, the variability of the twist when both ligands are bound is rather low (2 deg). When only L-Glu is bound to the active structure the free energy minimum is around 17 deg, which is also in agreement with free simulations. Nevertheless, this system shows a slight increase in the variability of the twist compared to the former. Finally, and maybe surprisingly, the minimum observed when both ligands are removed does not compare with the one sampled in free MD, 18-20 deg versus 22 deg respectively, and the variability of τ observed here reaches almost 4 deg.

These results allow to identify three distinct states, stabilized by either both ligands, L-Glu only or no ligand. Both IVM and L-Glu rigidify the overall structure and stabilize the twist at different values as suggested by free MD. The presence of IVM represents a significant barrier on twisting, i.e., about 7 Kcal/mol, as does the presence of L-Glu. The increased variability of the twist in the absence of ligand and of the overall structure goes in the direction of a model where the resting state is stabilized entropically, and the active state is stabilized enthalpically, i.e., by the contacts formed between the two ligands and the protein but also by the compact interfaces.

The shift observed between free MD and free energy calculations in the case of GluCl without ligand is not a direct consequence of our methodology but is most likely due to important orthogonal degrees of freedom which are not sampled in this particular system. Contrarily,

the absence of this shift in the two simulations of GluCl +IVM+L-GLU and GluCl +L-Glu indicates that the transition from a state that is activated by both L-Glu and IVM and a state activated by only L-Glu can be described accurately with the twisting alone.
Regarding the simulation of GluCl without ligand, the shift could be explained by the fact that the closing of the channel could represent a significant energetic barrier, which is not directly controlled by twisting. In fact, during our PMF calculations on twisting, we did not observed spontaneous channel closing. A simple explanation for such event is that each windows was simulated for a maximum of 10 ns, time which is not sufficient, by at least an order of magnitude, to sample the closing of the ion pore. A possible solution to this problem could be to design and restrain an observable able to force channel closing as the PMF along the twist is computed. Another possible explanation of this shift is the absence of blooming in these PMF calculations. Indeed, the blooming angle was left un-restrained during the PMF calculations. Therefore, it is possible, even thought our free simulations have shown that blooming is not required for closing, that twisting the receptor when it is not bloomed represents a significantly higher cost than when it is bloomed and that the two RC would be correlated, directly or indirectly. Despite the fact that the results obtained from the simulation of GluCl without ligand do not correlate with the ones obtained in free MD we can still conclude on the existence of 3 states and use these results as references to discriminate between active and non active allosteric ligands.

4.3 AN ACTIVE FRAGMENT OF IVERMECTIN

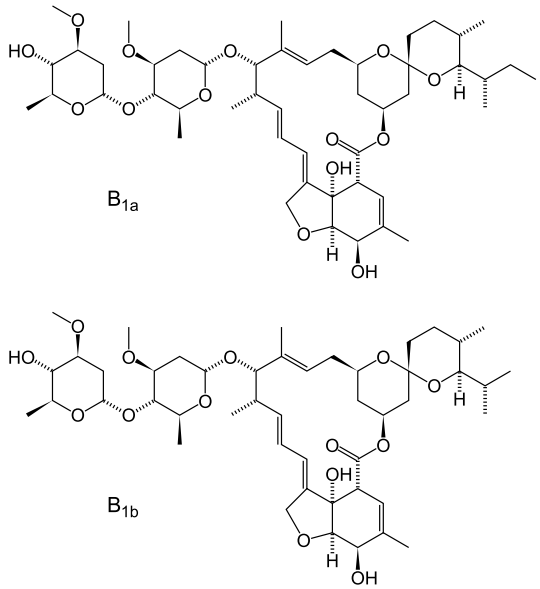


Figure 4.7: Structure of two molecules of ivermectin.

4.3.1 *Designing a meaningful fragment*

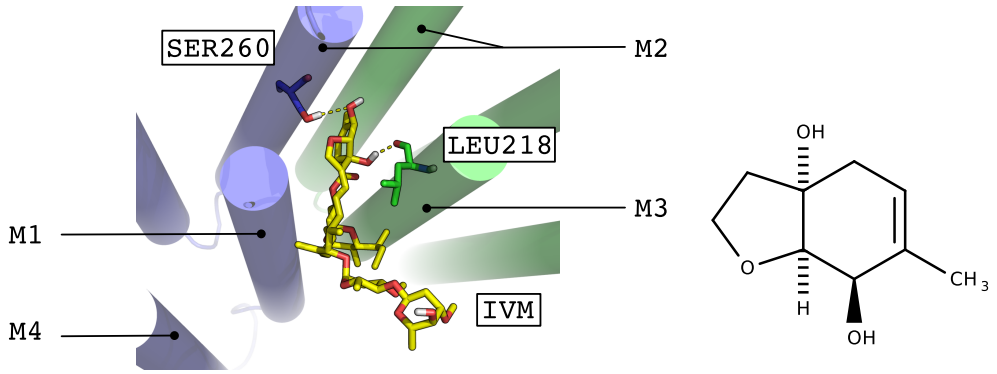
The IVM is a macrocyclic lactone and a member of the avermectin family. It is composed of 80 % of ivermectin B1a and 20 % of ivermectin B1b (see Figure 4.7). Throughout this study we will always refer to the Ivermectin B1a when mentioning IVM, the B1b will be left aside. It is a widely commercially available drug that is used as anthelmintic to treat parasites infestation such as schistosomiasis or scabies. Fortunately, it is too big in size to massively cross the blood brain barrier (BBB) and has no active transporter known so far. This prevents IVM from having the same lethal effect on human and parasites, which do not have a BBB. In order to use IVM for the design of new drugs active on the central nervous system (CNS) one could for example find the smallest active fragment of IVM and thus increase the chances of BBB crossing. This basic fragment could be then functionalized to optimize the pharmacological properties, e.g., increase intake or decrease toxicity.

The protocol of PMF calculations that we have developed allows to find out which values of the twist are stabilized by a particular ligand. By comparing with the PMF of the reference states, i.e., GluCl +L-Glu and GluCl +IVM+L-Glu, and knowing that twist controls the closing of the receptor, conclusions on the activity of the ligand studied can be made.

IVM is composed of several portions with different roles. Indeed, the two pyranose moieties, functionalized by alcohol groups and therefore hydrophilic, are most likely not involved in controlling the twist isomerization. Indeed, the X-ray structure of GluCl bound to IVM and L-Glu shows that this fragment of the allosteric ligand is lying at a position corresponding to the upper limit of the membrane, next to the hydrophilic heads of the lipids and only interact with the surface of GluCl. On the other hand, the macrocycle and particularly the moiety comprising a fused tetrahydrofuran cycle deeply inserted between the subunits, is very likely to be involved in the stabilization of the twist, possibly through a wedge mechanism.

To verify this hypothesis, the occupancy of hydrogen bonds between IVM and the receptor was computed along a free simulation of GluCl bound to IVM and L-Glu. The two most frequent hydrogen bonds regardless of the interface of binding are displayed in Figure 4.8a and the exact occupancy values are shown in Table 4.1. Interestingly, both hydrogen bonds involve the most deeply buried fragment of the protein suggesting that the hypothesis previously made was correct. The moiety comprising a fused tetrahydrofuran cycle seems strongly involved in the binding of IVM to GluCl. Nevertheless, this analysis is far from enough to conclude on the possible activity of the aforementioned fragment. To answer this question, PMF calculations along the twisting were run using the setup described in section 2.5.4.

The initial coordinates for the chosen fragment of IVM were taken from the equilibrated structure of GluCl bound to IVM and L-Glu. Several short simulations of this complex in vacuum were run to assess for its stability. No spontaneous unbinding was observed in any of the 5 ligands suggesting that the complex was stable.



(a) Main hydrogen bonds (b) Chosen fragment

Figure 4.8: Choosing a meaningful fragment of IVM. (a) shows an illustration of the two main hydrogen bonds isolated from a simulation of GluCl bound to IVM. The two adjacent subunits forming the binding site of IVM are shown in blue and green. .(b) shows the fragment of IVM which conserved the moieties involved in the main hydrogens bonds isolated from free MD

Table 4.1: The two most frequent hydrogen bonds between IVM and its receptor isolated in a simulation of GluCl bound to IVM and L-Glu

IVM	1		2		3		4		5	
Residue number	218	260	18	260	218	260	218	260	218	260
Occupancy (%)	35.8	36.51	14.63	0.88	0	0	0.37	29.8	17.27	12.64

4.3.2 *PMF of the fragment*

Interestingly, the fragment of IVM chosen stabilizes a minimum which is located around 14 to 15 deg of twist, value which is compatible with an active state. No significant shrinking of the ion channel was monitored in the corresponding PMF calculations. This fragment of IVM is, like IVM, capable of blocking the twist even though it is significantly smaller in size. This result is a proof of concept to show that it is possible to design small molecules which will bind to the receptor in the IVM site, block the twist and potentially activate the receptor. This fragment represents a base for understanding which parts of IVM are essential to block the twist and could be optimized considering factors such as the bioavailability or metabolization. In a second time, one could identify a fragment of IVM which is able to bind to the receptor but as no effect on the twist. This ligand could for instance be used to compete with IVM when binding to GluCl and therefore offer a solution to IVM intoxications. These results also rise the questions of the correlation between the size of the ligand and the putative activity on the twist. Indeed, it is possible that, up the a certain limit, the bigger the ligand inside this allosteric pocket the more restrained would be the twist. Nevertheless, this hypothesis must be more carefully investigated.

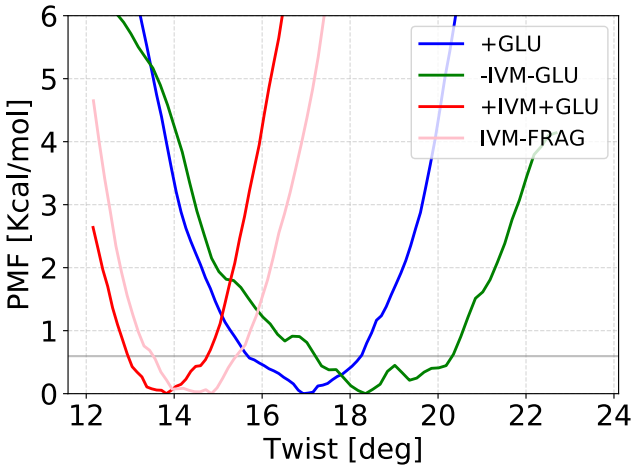


Figure 4.9: PMF for the fragment of IVM compared to the reference PMF as presented in section 4.2. The gray horizontal line represent the value of K_bT at 300 K.

4.4 SET OF PUTATIVE PAMS

A wide variety of molecules is known to bind to pLGICs allosterically, ranging from general anesthetics such as isoflurane with a molecular weight of about 150 g/mol to much bigger molecules such as IVM (875 g/mol), cholesterol (386 g/mol) or POPC (760 g/mol). Although several crystal structures of pLGICs in complex with allosteric ligands exist, computational methods can bring a deeper understanding of the mechanism of action of allosteric binders on pLGICs and propose structure of complexes unknown so far.

Triiodothyronine or T3 hormone is known to inhibit the response of $GABA_A$ receptors unlike allopregnalonone (ALLO) which can activate the same receptor or cholesterol (CHOL) which is thought to be a potentiator of the pLGICs. Moreover, it was suggested in 2015 by Westergard et al.[39] that theses ligands could bind in the same pocket as IVM.

In addition to the aforementioned ligands, lipids are thought to be able to potentiate the pLGICs as suggested by the X-ray structure of GluCl apo bound to a POPC[21]. Studies, both computational[5,40] and experimental have also suggested that cholesterol could play a role in activating the receptor (reviewed in Ref.[14]). The former makes the composition of the membrane crucial in the putative basal activation of theses receptors.

As no crystal structure of ALLO, T3 or cholesterol (CHOL) bound to GluCl exist we had to place them by docking in the active site of IVM.

The 10 best poses produced by PLANTS were kept and re-scored with a single-point energy calculation (see section 2.4 for details). All the poses where the ligand was not inside the pocket but interacting with the membrane-facing side of the protein were not considered for the next steps of our study. Fortunately, all the first ranked poses by our rescoring method were deeply inserted inside the IVM binding cleft and were therefore selected for the next steps of this study.

Table 4.2: Sum up of the docking scores rescoring for the 3 putative ligands in the IVM pocket of GluCl active. IE stands for interaction energy.

Ligand	PLANTS rank	PLANTS score	CHARMM rank	CHARMM IE	Comments
CHOL	1	-141.36	4	-43.32	
	2	-141.06	7	-40.31	
	3	-140.23	2	-45.66	
	4	-139.24	3	-44.52	
	5	**-128.53**	**1**	**-48.61**	
	6	-126.73	8	-39.90	
	7	-124.30	9	-36.58	Identical to #8
	8	-122.39	10	-31.77	Outside of the pocket.
	9	-121.62	6	-40.42	
	10	-121.51	5	-42.33	
ALLO	1	-114.48	5	-32.52	
	2	-109.09	2	-35.75	
	3	-109.07	6	-32.02	
	4	-104.68	4	-33.57	
	5	-102.04	3	-33.61	
	6	-102.03	7	-27.75	
	7	**-100.52**	**1**	**-42.00**	
	8	-94.88	10	-22.38	Outside of the pocket.
	9	-90.71	8	-26.22	Outside of the pocket.
	10	-88.92	9	-26.19	
T_3	1	-102.02	9	-35.82	Outside of the pocket.
	2	-101.02	2	-43.61	
	3	-100.16	5	-38.74	
	4	-97.62	10	-35.04	Outside of the pocket.
	5	-97.27	3	-42.80	
	6	-97.18	8	-36.50	
	7	-96.92	4	-39.73	
	8	-96.28	6	-38.66	
	9	-95.90	7	-36.90	
	10	**-95.85**	**1**	**-50.45**	

The PMF along the twist of the active structure bound to those ALLO, T3 and Cholesterol is presented in Figure 4.10.

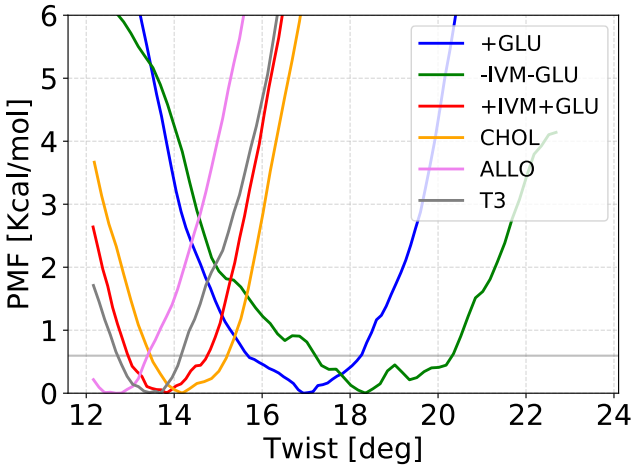


Figure 4.10: PMF along the twist isomerization from the active structure of GluCl for ALLO, T3 and CHOL compared to the three reference states as described before.

Interestingly, all three ligands stabilize the twist angle at values compatible with an active state. No significant difference, on the twist angle at least, was observed between positive (CHOL and ALLO) and negative (T3) allosteric modulators, suggesting that the mechanism by which T3 inhibits the activity of the receptor might not be related to the twisting isomerization. It must also be noted that the effect of T3 on GluCl is unknown and could be different than the one on $GABA_A$. The cholesterol and ALLO are capable of blocking the twist like IVM, suggesting that these two molecules are positive modulators. Indeed, and as already discussed before the putative role on cholesterol on pLGICs is confirmed by these results. Cholesterol molecules in the membrane could bind in the IVM allosteric pocket and have a similar effect on the channel activation. This could explain why the composition of the membrane has an effect on the physiological response of the pLGICs.

Finally the convergence of the PMF calculation for the 4 ligands aforementioned are displayed Figure 4.11. Even though the simulations are rather short they all seem to be close to convergence.

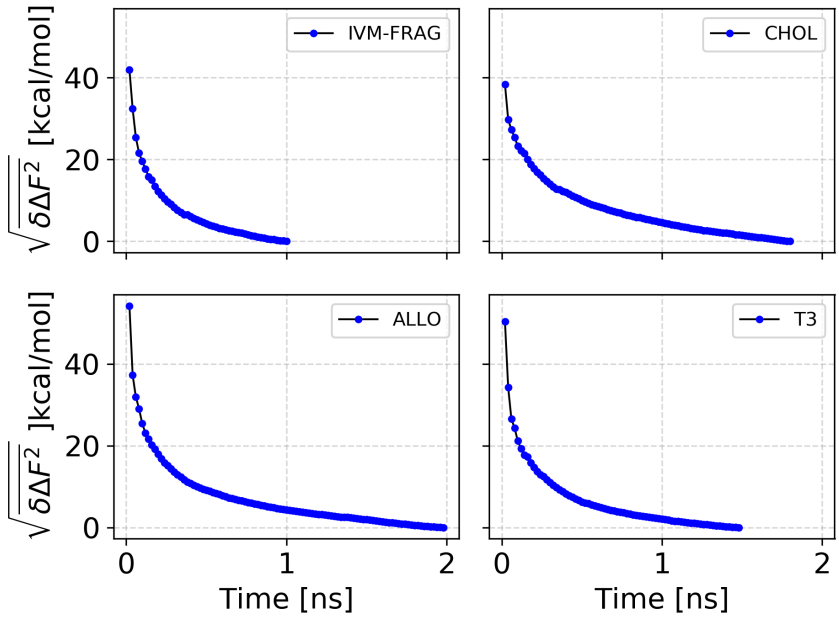


Figure 4.11: Convergence of four set of PMF calculations along the twisting isomerization of GluCl bound to different ligands. The convergence is calculated as described in section 2.2.4.

CONCLUSIONS AND PERSPECTIVES

As a first step to understand the mechanism of action of ion channels we studied the spontaneous deactivation of one of them upon the removal of an allosteric modulator. One of the main discovery that we have made was that we could produce a resting structure with no *a priori* knowledge of the former by MD simulations. It demonstrates that the use of computational techniques such as MD could allow to produce meaningful structures of complex membrane proteins. In fact, this indicates that in the numerous cases in which experimental techniques fail at obtaining high resolution structures, computational techniques and in particular MD can be an adequate alternative. In addition, the identification of a metastable state of GluCl upon closing which was not resolved in X-ray crystallography is also a good example of the predictive power of MD. One can even imagine that someday, thanks to the improvement of the protocols and softwares, the structures produced by MD could be considered of better quality than the one obtained through other experimental techniques.

Moreover, MD also presents the significant advantage over the experimental techniques to give an easy access to the atomic scale. Atoms can be easily manipulated and studied with a precision that is not yet reachable through experimental techniques. Even thought to this date we are limited to study systems of a maximum of several millions of atoms, it is not inconceivable that it will soon be possible to simulate an entire neuron at the atomic scale. This would drastically increase the amount of data to analyze but also the strength of the conclusions drawn.

Our important findings on the gating mechanism of pLGICs also indicate that it is now possible to reach time scales allowing to make physiologically relevant interpretations which will eventually help to the design of more efficient and less harmful treatments. This also demonstrates that MD can be used to study the mechanisms of action of complex proteins with an unprecedented precision.

The place taken by computational techniques in the design of new drugs has undergone a tremendous increase throughout the past decades and is now routinely used in the early stages of the drug design process. With the increasing power of computer and the development of many new methods this trend does not appear to be ready to change. Our study on the effect of several ligands on the function of an ion channel illustrates how one can use advanced computational techniques to make prediction on the potency of given ligands.

As of today, we are limited to make prediction which are validated by experiments, but it is reasonable to think that sooner or later the predictions made by computational models will be so accurate that the need for experimental validation will drastically decrease if not disappear completely. Even if major advances were made in the understanding of neurotransmission in general and in particular in the mechanism of action of ion channels, every discovery seems to rise an endless list of fascinating questions, illustrating the famous quote of Albert Einstein "The more I learn, the more I realize how much I don't know."

In the future we aim at analyzing the latest published structures to determine in what extent

they fit in our gating mechanism and to identify possible differences among the pLGICs.
Moreover, in order to better understand the role of the blooming in ligand recognition and its putative effect on the gating we wish to perform a bi-dimensional PMF of twisting against blooming.
In addition, one could use the framework of free energy calculations developed for either the twisting, the blooming or both combined to better understanding by which mechanism a agonist or an antagonist affects the function of an ion channel.
To tackle the question of the allosteric modulation one could also rely on the increased sampling provided by aMD simulations in order to isolate transient binding pockets and try to design ligands binding in these pockets. In a second time one could run free energy calculations to see if the ligand of interest is capable of blocking the twist and thus keeping the channel open or classical MD to study the stability of the complex.

STRUCTURE SYMMETRIZATION pLGICs are pentameric channel arrange in a fivefold symmetry around a pore axis. Although they are symmetric and some of them have homologous subunits all subunits may undergoes different changes, particularly for the simulations in the absence of ligands.

In fact, the variability among subunits of a given quantity makes very challenging for the human eye to have a glance of the average value of the given quantity.

When we wanted to evidence the overall twisting or blooming of the receptor visually we proceeded to a symmetrization of the average structure.

To do so we first picked the one of the 5 subunits which was the best representative of the average value of the twist for instance. The subunits was then replicated 5 times and rotated by 72 degrees (one fifth of 360 degrees) around the pore axis.

Table 4.3: Table of the average and standard deviation of the main observable discussed throughout this Text for the 5 simulations of interest.

	Twisting [deg]		Cα cross section [Å^2]		Water count	
GluCl apo	21.3 +/- 1.1		109.4 +/- 2.7		0.4 +/- 0.2	
GluCl with IVM	15.3 +/- 1.2		142.5 +/- 7.7		5.6 +/- 0.8	
L-Glu restrained	16.7 +/- 0.9		151.2 +/- 10.9		5.2 +/- 1.0	
GluCl w/o IVM A	0-400ns	last 100ns	0-400ns	last 100ns	0-400ns	last 100ns
	19.1 +/- 1.5	22.2 +/- 1.2	128.1 +/- 10.4	110.4 +/- 3.0	3.9 +/- 1.9	0.4 +/- 0.2
GluCl w/o IVM B	0-1000ns	last 100ns	0-1000ns	last 100ns	0-1000ns	last 100ns
	22.5 +/- 2.2	25.7 +/- 1.2	122.2 +/- 7.1	112.2 +/- 3.6	2.3 +/- 1.4	0.4 +/- 0.2

	Polar tilt [deg]		Azimuthal tilt [deg]		Ion count	
GluCl apo	12.1 +/- 0.6		6.7 +/- 0.7		0.2 +/- 0.2	
GluCl with IVM	7.6 +/- 0.6		3.3 +/- 0.6		0.4 +/- 0.3	
L-Glu restrained	7.8 +/- 0.7		4.2+/- 0.6		0.2 +/- 0.2	
GluCl w/o IVM A	0-400ns	last 100ns	0-400ns	last 100ns	0-400ns	last 100ns
	8.8 +/- 0.8	10.0 +/- 0.7	5.3 +/-1.1	8.3 +/- 0.9	0.3 +/- 0.3	0.1 +/- 0.1
GluCl w/o IVM B	0-1000ns	last 100ns	0-1000ns	last 100ns	0-1000ns	last 100ns
	8.1 +/- 0.6	7.8 +/- 0.6	6.1 +/- 1.2	9.3 +/- 0.6	0.2 +/- 0.2	0.0 +/- 0.1

	Water flux [wat/ns]		ΔZ [Å]		Number of L-Glu bound	
GluCl apo	0.0 +/- 0.1		6.3 +/- 0.4		0.0	
GluCl with IVM	2.7 +/- 1.7		4.8 +/- 0.4		3.4 +/- 0.6	
L-Glu restrained	3.1 +/- 1.8		4.9 +/- 0.3		5.0 +/- 0.0	
GluCl w/o IVM A	0-400ns	last 100ns	0-400ns	last 100ns	0-400ns	last 100ns
	1.2 +/- 1.4	0.0 +/- 0.1	4.2 +/- 0.2	5.8 +/- 0.2	3.2 +/- 1.2	1.6 +/- 0.5
GluCl w/o IVM B	0-1000ns	last 100ns	0-1000ns	last 100ns	0-1000ns	last 100ns
	0.7 +/- 1.0	0.0 +/- 0.1	4.9 +/- 0.3	7.1 +/- 0.2	2.4 +/- 1.5	0.0 +/- 0.0

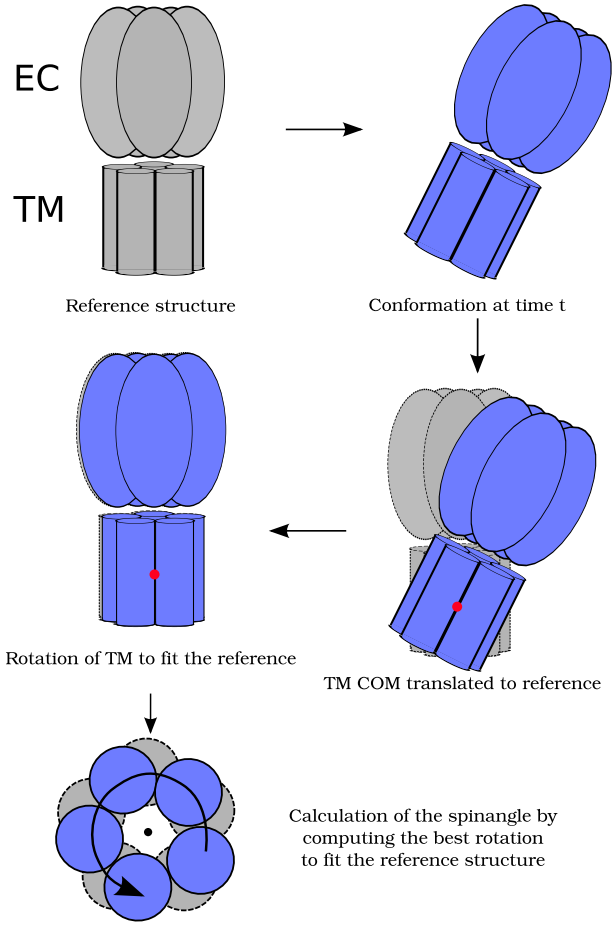


Figure 4.12: Flowchart showing the different steps to compute the spinangle ϕ.

SPINANGLE COMPUTATION ALGORITHM

EXAMPLE SCRIPT TO RESTRAIN THE TWIST ANGLE USING THE SPINANGLE COLVAR IN NAMD

spinangle.in was shown to be compatible with the version 2.9, 2.10 and 2.11 of NAMD. However, recent changes in the latest implementation of the collective variable module might break this compatibility.

```
# set the frequency to whatever value suits your needs
colvarsTrajFrequency       1

# Create a colvar
colvar {
        # general parameters of the colvar
            outputAppliedForce on       # optional, use if needed for debugging
        name spinangle
        # definition of the colvar
        spinangle {
                axis (0,0,1)    # Z axis to match the reference structure in which the
```

```
                        # the pore axis was alined with the Z axis
                # define on which group of atoms will be computed the colvar
                atoms {
                        # using a PDB file with 1.0 in the beta column value to select
                            the core
                        # atoms of the EC domain only
                        atomsFile ../input/all.pdb
                        atomsCol B
                        atomsColValue 1.0
                        # remove the possible diffusion of the protein by aligning on the
                             TMD
                        # of the reference structure (X-ray)
                        centerReference on
                        # remove the possible rotation of the protein yb aligning on the
                            TMD
                        # of the reference structure (X-ray)
                        rotateReference on
                        # set the selection on which will be made the fit at every step,
                            i.e.,
                        # the core atoms of the TMD
                        # /!\ The fit is done only for the calculation of the colvar and
                            there
                        # is no actual cahnge of the coordinate during the simulation
                        refPositionsGroup { #..reference group
                                atomsFile ../input/all.pdb
                                atomsCol B
                                atomsColValue 2.0
                        }
                        # reference positions for the fit
                        refPositionsFile ../input/all.pdb
                }
                # reference positions for the calculation of the spinangle
                refPositionsFile ../input/all.pdb
        }
}

# Apply and hormonic bias on the colvar previously created
harmonic {
        name harmonic
        colvars spinangle
        # force constant expressed in kcal\mol\deg
        forceConstant 0.006
        centers XXXX
}
```

EXAMPLE SCRIPT TO RESTRAIN THE BLOOMING ANGLE USING THE QUATERNION COLVAR IN NAMD blooming.in was shown to be compatible with the version 2.9, 2.10 and 2.11 of NAMD. however, recent changes in the latest implementation of the collective variable module might break this compatibility.

```
## the following is a example for restraining only a single chain
## and should be repeated for each subunir
## chain X
colvar {
  name chainX-q
  orientation {
    atoms {
      atomsFile constr.pdb
      atomsCol B
      atomsColValue 2.0 # ECD to measure q
      centerReference on
      rotateReference on
      refPositionsGroup { #..reference group
        atomsFile constr.pdb
        atomsCol B
        atomsColValue 1.0 # TMD cause twist is similar in both structures
      }
      refPositionsFile constr.pdb
    }
    refPositionsFile constr.pdb
  }
}
## The harmonic bias must be applied individually on each subunit as the
## target centers are different for all of them
harmonic {
  name restrainX
  colvars chainX-q
  centers (1.0 , 0.0 , 0.0 , 0.0) # the identity quaternion
  targetcenters (0.991444827373 , 0.0304491490769 , -0.126815328226 , 0.00527978429424)
  targetNumSteps 500000
  forceConstant 10000
}
```

MISCELLANEOUS In addition to the two example scripts presented here various scripts to help the preparation or the analysis of MD simulations can be found at the following address: *https://github.com/niko97320/misc-IFM*

BIBLIOGRAPHY

[1] C. L. Morales-Perez, C. M. Noviello, and R. E. Hibbs. "X-ray structure of the human $\alpha 4 \beta 2$ nicotinic receptor". *Nature*, 538(7625), pp. 411–415, 2016.

[2] R. E. Hibbs and E. Gouaux. "Principles of activation and permeation in an anion-selective Cys-loop receptor". *Nature*, 474(7349), pp. 54–60, 2011.

[3] N. Calimet, M. Simoes, J.-P. Changeux, M. Karplus, A. Taly, and M. Cecchini. "A gating mechanism of pentameric ligand-gated ion channels." *Proceedings of the National Academy of Sciences of the United States of America*, 110(42), pp. E3987–96, 2013.

[4] M. H. Cheng and R. D. Coalson. "Energetics and ion permeation characteristics in a glutamate-gated chloride (GluCl) receptor channel". *Journal of Physical Chemistry B*, 116(46), pp. 13637–13643, 2012.

[5] J. Hénin, R. Salari, S. Murlidaran, and G. Brannigan. "A predicted binding site for cholesterol on the GABAA receptor". *Biophysical Journal*, 106(9), pp. 1938–1949, 2014.

[6] Ö. Yoluk, T. Brömstrup, E. J. Bertaccini, J. R. Trudell, and E. Lindahl. "Stabilization of the GluCl ligand-gated ion channel in the presence and absence of ivermectin". *Biophysical Journal*, 105(3), pp. 640–647, 2013.

[7] O. Yoluk, E. Lindahl, and M. Andersson. "Conformational gating dynamics in the GluCl anion-selective chloride channel". *ACS Chemical Neuroscience*, p. 150520172255005, 2015.

[8] S. A. Heusser, Ö. Yoluk, G. Klement, E. A. Riederer, E. Lindahl, and R. J. Howard. "Functional characterization of neurotransmitter activation and modulation in a nematode model ligand-gated ion channel". *Journal of Neurochemistry*, pp. 243–253, 2016.

[9] A. Taly, M. Delarue, T. Grutter, M. Nilges, N. Le Novère, P.-J. Corringer, and J.-P. Changeux. "Normal mode analysis suggests a quaternary twist model for the nicotinic receptor gating mechanism." *Biophysical journal*, 88(6), pp. 3954–65, 2005.

[10] N. E. Martin, S. Malik, N. Calimet, J.-P. Changeux, and M. Cecchini. "Un-gating and allosteric modulation of a pentameric ligand-gated ion channel captured by molecular dynamics". *PLOS Computational Biology*, 13(10), p. e1005784, 2017.

[11] M. Jaiteh, A. Taly, and J. Hénin. "Evolution of Pentameric Ligand-Gated Ion Channels: Pro-Loop Receptors." *PloS one*, 11(3), p. e0151934, 2016.

[12] H. Nury, F. Poitevin, C. Van Renterghem, J.-P. Changeux, P.-J. Corringer, M. Delarue, and M. Baaden. "One-microsecond molecular dynamics simulation of channel gating in a nicotinic receptor homologue." *Proceedings of the National Academy of Sciences of the United States of America*, 107(14), pp. 6275–80, 2010.

[13] F. Zhu and G. Hummer. "Pore opening and closing of a pentameric ligand-gated ion channel". *Proceedings of the National Academy of Sciences*, 107(46), pp. 19814–19819, 2010.

[14] G. Brannigan. "Direct Interactions of Cholesterol With Pentameric Ligand-Gated Ion Channels: Testable Hypotheses From Computational Predictions". 2017.

[15] U. Muscatello. "Golgi's contribution to medicine". 55(1), pp. 3–7, 2007.

[16] Wikipedia. "Neuron — wikipedia, the free encyclopedia", 2017. [Online; accessed 23-October-2017].

[17] J. Langley. "On the reaction of cells and of nerve-endings to certain poisons, chiefly as regards the reaction of striated muscle to nicotine and to curari". *Journal of Physiology*, 33(4-5), pp. 374–413, 1905.

[18] J. P. Changeux, M. Kasai, and C. Y. Lee. "Use of a snake venom toxin to characterize the cholinergic receptor protein." *Proceedings of the National Academy of Sciences of the United States of America*, 67(3), pp. 1241–7, 1970.

[19] N. Bocquet, H. Nury, M. Baaden, C. Le Poupon, J.-P. Changeux, M. Delarue, and P.-J. Corringer. "X-ray structure of a pentameric ligand-gated ion channel in an apparently open conformation". *Nature*, 457(7225), pp. 111–114, 2009.

[20] R. J. C. Hilf and R. Dutzler. "Structure of a potentially open state of a proton-activated pentameric ligand-gated ion channel". *Nature*, 457(7225), pp. 115–118, 2009.

[21] T. Althoff, R. E. Hibbs, S. Banerjee, and E. Gouaux. "X-ray structures of GluCl in apo states reveal a gating mechanism of Cys-loop receptors". *Nature*, 512(7514), pp. 333–337, 2014.

[22] P. J. Corringer, F. Poitevin, M. S. Prevost, L. Sauguet, M. Delarue, and J. P. Changeux. "Structure and pharmacology of pentameric receptor channels: From bacteria to brain". *Structure*, 20(6), pp. 941–956, 2012.

[23] M. Nys, D. Kesters, and C. Ulens. "Structural insights into Cys-loop receptor function and ligand recognition". *Biochemical Pharmacology*, 86(8), pp. 1042–1053, 2013.

[24] P. S. Miller and A. R. Aricescu. "Crystal structure of a human GABAA receptor." *Nature*, 512(7514), pp. 270–5, 2014.

[25] X. Huang, H. Chen, K. Michelsen, S. Schneider, and P. L. Shaffer. "Crystal structure of human glycine receptor-$\alpha 3$ bound to antagonist strychnine". *Nature*, 526(7572), pp. 277–280, 2015.

[26] X. Huang, H. Chen, and P. L. Shaffer. "Crystal Structures of Human GlyR$\alpha 3$ Bound to Ivermectin". *Structure*, 25(6), pp. 945–950.e2, 2017.

[27] J. Du, W. Lü, S. Wu, Y. Cheng, and E. Gouaux. "Glycine receptor mechanism elucidated by electron cryo-microscopy." *Nature*, 526(7572), pp. 224–9, 2015.

[28] R. J. C. Hilf and R. Dutzler. "X-ray structure of a prokaryotic pentameric ligand-gated ion channel". *Nature*, 452(7185), pp. 375–379, 2008.

[29] B. Billen, R. Spurny, M. Brams, R. van Elk, S. Valera-Kummer, J. L. Yakel, T. Voets, D. Bertrand, A. B. Smit, and C. Ulens. "Molecular actions of smoking cessation drugs at $\alpha4\beta2$ nicotinic receptors defined in crystal structures of a homologous binding protein." *Proceedings of the National Academy of Sciences of the United States of America*, 109(23), pp. 9173–8, 2012.

[30] C. Mulle, C. Léna, and J. P. Changeux. "Potentiation of nicotinic receptor response by external calcium in rat central neurons." *Neuron*, 8(5), pp. 937–945, 1992.

[31] S. Vernino, M. Amador, C. W. Luetje, J. Patrick, and J. A. Dani. "Calcium modulation and high calcium permeability of neuronal nicotinic acetylcholine receptors". *Neuron*, 8(1), pp. 127–134, 1992.

[32] D. Bertrand and M. Gopalakrishnan. "Allosteric modulation of nicotinic acetylcholine receptors". *Biochemical Pharmacology*, 74(8), pp. 1155–1163, 2007.

[33] C. J. B. DaCosta, C. R. Free, J. Corradi, C. Bouzat, and S. M. Sine. "Single-channel and structural foundations of neuronal $\alpha7$ acetylcholine receptor potentiation." *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 31(39), pp. 13870–13879, 2011.

[34] I. A. Lobo and R. A. Harris. "Sites of Alcohol and Volatile Anesthetic Action on Glycine Receptors". *International Review of Neurobiology*, 65, pp. 53–87, 2005.

[35] S. J. Mihic, Q. Ye, M. J. Wick, V. V. Koltchine, M. D. Krasowski, S. E. Finn, M. P. Mascia, C. F. Valenzuela, K. K. Hanson, E. P. Greenblatt, R. A. Harris, and N. L. Harrison. "Sites of alcohol and volatile anaesthetic action on GABA(A) and glycine receptors." *Nature*, 389(6649), pp. 385–9, 1997.

[36] S. A. Forman, D. C. Chiara, and K. W. Miller. "Anesthetics target interfacial transmembrane sites in nicotinic acetylcholine receptors". *Neuropharmacology*, 96(PB), pp. 169–177, 2015.

[37] A. Taly, P.-J. Corringer, D. Guedin, P. Lestage, and J.-P. Changeux. "Nicotinic receptors: allosteric transitions and therapeutic targets in the nervous system". *Nature Reviews Drug Discovery*, 8(9), pp. 733–750, 2009.

[38] A. M. Hosie, M. E. Wilkins, H. M. A. da Silva, and T. G. Smart. "Endogenous neurosteroids regulate GABAA receptors through two discrete transmembrane sites". *Nature*, 444(7118), pp. 486–489, 2006.

[39] T. Westergard, R. Salari, J. V. Martin, and G. Brannigan. "Interactions of L-3,5,3′-triiodothyronine, allopregnanolone, and ivermectin with the GABAA receptor: Evidence for overlapping intersubunit binding modes". *PLoS ONE*, 10(9), pp. 1–18, 2015.

[40] G. Brannigan, J. Hénin, R. Law, R. Eckenhoff, and M. L. Klein. "Embedded cholesterol in the nicotinic acetylcholine receptor." *Proceedings of the National Academy of Sciences of the United States of America*, 105(38), pp. 14418–23, 2008.

[41] D. C. Chiara, L. J. Dangott, R. G. Eckenhoff, and J. B. Cohen. "Identification of Nicotinic Acetylcholine Receptor Amino Acids Photolabeled by the Volatile Anesthetic Halothane". *Biochemistry*, 42(46), pp. 13457–13467, 2003.

[42] K. Nishikawa and N. L. Harrison. "The actions of sevoflurane and desflurane on the gamma-aminobutyric acid receptor type A: effects of TM2 mutations in the alpha and beta subunits." *Anesthesiology*, 99(3), pp. 678–684, 2003.

[43] H. Nury, C. Van Renterghem, Y. Weng, A. Tran, M. Baaden, V. Dufresne, J.-P. Changeux, J. M. Sonner, M. Delarue, and P.-J. Corringer. "X-ray structures of general anaesthetics bound to a pentameric ligand-gated ion channel." *Nature*, 469(7330), pp. 428–431, 2011.

[44] X. Huang, P. L. Shaffer, S. Ayube, H. Bregman, H. Chen, S. G. Lehto, J. A. Luther, D. J. Matson, S. I. McDonough, K. Michelsen, M. H. Plant, S. Schneider, J. R. Simard, Y. Teffera, S. Yi, M. Zhang, E. F. DiMauro, and J. Gingras. "Crystal structures of human glycine receptor $\alpha3$ bound to a novel class of analgesic potentiators". *Nature Structural & Molecular Biology*, 24(2), pp. 108–113, 2016.

[45] D. Laverty, P. Thomas, M. Field, O. J. Andersen, M. G. Gold, P. C. Biggin, M. Gielen, and T. G. Smart. "Crystal structures of a GABAA-receptor chimera reveal new endogenous neurosteroid-binding sites". *Nature Structural & Molecular Biology*, (October), pp. 1–14, 2017.

[46] P. S. Miller, S. Scott, S. Masiulis, L. De Colibus, E. Pardon, J. Steyaert, and A. R. Aricescu. "Structural basis for GABAA receptor potentiation by neurosteroids". *Nature Structural & Molecular Biology*, p. nsmb.3484, 2017.

[47] A. Miyazawa, Y. Fujiyoshi, and N. Unwin. "Structure and gating mechanism of the acetylcholine receptor pore". *Nature*, 423(6943), pp. 949–955, 2003.

[48] B. H. White and J. B. Cohen. "Agonist-induced changes in the structure of the acetylcholine receptor M2 regions revealed by photoincorporation of an uncharged nicotinic noncompetitive antagonist". *Journal of Biological Chemistry*, 267(22), pp. 15770–15783, 1992.

[49] E. Arevalo, D. C. Chiara, S. A. Forman, J. B. Cohen, and K. W. Miller. "Gating-enhanced accessibility of hydrophobic sites within the transmembrane region of the nicotinic acetylcholine receptor's δ-subunit: A time-resolved photolabeling study". *Journal of Biological Chemistry*, 280(14), pp. 13631–13640, 2005.

[50] Y. Chang, R. Wang, S. Barot, and D. S. Weiss. "Stoichiometry of a recombinant GABAA receptor". *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 16(17), pp. 5415–5424, 1996.

[51] C. Labarca, M. W. Nowak, H. Zhang, L. Tang, P. Deshpande, and H. a. Lester. "Channel gating governed symmetrically by conserved leucine residues in the M2 domain of nicotinic receptors." *Nature*, 376(6540), pp. 514–516, 1995.

[52] G. N. Filatov and M. M. White. "The role of conserved leucines in the M2 domain of the acetylcholine receptor in channel gating." *Molecular pharmacology*, 48(3), pp. 379–384, 1995.

[53] R. J. C. Hilf, C. Bertozzi, I. Zimmermann, A. Reiter, D. Trauner, and R. Dutzler. "Structural basis of open channel block in a prokaryotic pentameric ligand-gated ion channel". *Nature Structural & Molecular Biology*, 17(11), pp. 1330–U184, 2010.

[54] R. Spurny, B. Billen, R. J. Howard, M. Brams, S. Debaveye, K. L. Price, D. A. Weston, S. V. Strelkov, J. Tytgat, S. Bertrand, D. Bertrand, S. C. R. Lummis, and C. Ulens. "Multisite binding of a general anesthetic to the prokaryotic pentameric erwinia chrysanthemi ligand-gated ion channel (ELIC)". *Journal of Biological Chemistry*, 288(12), pp. 8355–8364, 2013.

[55] D. N. LeBard, J. Hénin, R. G. Eckenhoff, M. L. Klein, and G. Brannigan. "General anesthetics predicted to block the GLIC pore with micromolar affinity". *PLoS Computational Biology*, 8(5), 2012.

[56] G. Hassaine, C. Deluz, L. Grasso, R. Wyss, M. B. Tol, R. Hovius, A. Graff, H. Stahlberg, T. Tomizaki, A. Desmyter, C. Moreau, X.-D. Li, F. Poitevin, H. Vogel, and H. Nury. "X-ray structure of the mouse serotonin 5-HT3 receptor". *Nature*, 512, pp. 276–281, 2014.

[57] E. X. Albuquerque, E. F. R. Pereira, M. Alkondon, and S. W. Rogers. "Mammalian Nicotinic Acetylcholine Receptors: From Structure to Function". 89(1), pp. 73–120, 2009.

[58] M. Zoli, F. Pistillo, and C. Gotti. "Diversity of native nicotinic receptor subtypes in mammalian brain." *Neuropharmacology*, 96, pp. 302–311, 2014.

[59] C. Gotti, M. Zoli, and F. Clementi. "Brain nicotinic acetylcholine receptors: native subtypes and their relevance". *Trends in Pharmacological Sciences*, 27(9), pp. 482–491, 2006.

[60] M. Zoli, F. Pistillo, and C. Gotti. "Diversity of native nicotinic receptor subtypes in mammalian brain". *Neuropharmacology*, 96(PB), pp. 302–311, 2015.

[61] C. J. Tsai and R. Nussinov. "A Unified View of "How Allostery Works"". *PLoS Computational Biology*, 10(2), 2014.

[62] F. Jacob and J. Monod. "Genetic regulatory mechanisms in the synthesis of proteins". *Journal of Molecular Biology*, 3(3), pp. 318–356, 1961.

[63] J. P. Changeux. "The feedback control mechanisms of biosynthetic L-threonine deaminase by L-isoleucine." *Cold Spring Harbor symposia on quantitative biology*, 26, pp. 313–318, 1961.

[64] J. Monod, J. Wyman, and J. P. Changeux. "On the nature of allosteric transitions: A plausible model". *Journal of Molecular Biology*, 12(1), pp. 88–118, 1965.

[65] L. Pauling. "The Oxygen Equilibrium of Hemoglobin and Its Structural Interpretation." *Proceedings of the National Academy of Sciences of the United States of America*, 21(4), pp. 186–91, 1935.

[66] D. E. Koshland, G. Némethy, and D. Filmer. "Comparison of experimental binding data and theoretical models in proteins containing subunits." *Biochemistry*, 5(1), pp. 365–385, 1966.

[67] P. H. N. Celie, S. E. Van Rossum-Fikkert, W. J. Van Dijk, K. Brejc, A. B. Smit, and T. K. Sixma. "Nicotine and carbamylcholine binding to nicotinic acetylcholine receptors as studied in AChBP crystal structures". *Neuron*, 41(6), pp. 907–914, 2004.

[68] K. Brejc, W. J. van Dijk, R. V. Klaassen, M. Schuurmans, J. van der Oost, A. B. Smit, and T. K. Sixma. "Crystal structure of an ACh-binding protein reveals the ligand-binding domain of nicotinic receptors". *Nature*, 411(6835), pp. 269–276, 2001.

[69] R. E. Hibbs, G. Sulzenbacher, J. Shi, T. T. Talley, S. Conrod, W. R. Kem, P. Taylor, P. Marchot, and Y. Bourne. "Structural determinants for interaction of partial agonists with acetylcholine binding protein and neuronal alpha7 nicotinic acetylcholine receptor." *The EMBO journal*, 28(19), pp. 3040–51, 2009.

[70] M. Brams, A. Pandya, D. Kuzmin, R. van Elk, L. Krijnen, J. L. Yakel, V. Tsetlin, A. B. Smit, and C. Ulens. "A structural and mutagenic blueprint for molecular recognition of strychnine and d-tubocurarine by different Cys-loop receptors". *PLoS Biology*, 9(3), 2011.

[71] R. Yu, E. Hurdiss, T. Greiner, R. Lape, L. Sivilotti, and P. C. Biggin. "Agonist and antagonist binding in human glycine receptors". *Biochemistry*, 53(38), pp. 6041–6051, 2014.

[72] D. A. Wagner and C. Czajkowski. "Structure and dynamics of the GABA binding pocket: A narrowing cleft that constricts during activation." *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 21(1), pp. 67–74, 2001.

[73] T. L. Kash, J. R. Trudell, and N. L. Harrison. "Structural elements involved in activation of the gamma-aminobutyric acid type A (GABAA) receptor." *Biochemical Society transactions*, 32(Pt3), pp. 540–6, 2004.

[74] A. J. Thompson, C. L. Padgett, and S. C. R. Lummis. "Mutagenesis and molecular modeling reveal the importance of the 5-HT3 receptor F-loop." *The Journal of biological chemistry*, 281(24), pp. 16576–82, 2006.

[75] C. L. Padgett and S. C. R. Lummis. "The F-loop of the GABA A receptor gamma2 subunit contributes to benzodiazepine modulation." *The Journal of biological chemistry*, 283(5), pp. 2702–8, 2008.

[76] J. Zhang, F. Xue, and Y. Chang. "Agonist- and antagonist-induced conformational changes of loop F and their contributions to the ρ1 GABA receptor function". *The Journal of Physiology*, 587(1), pp. 139–153, 2009.

[77] A. Khatri, A. Sedelnikova, and D. S. Weiss. "Structural rearrangements in loop F of the GABA receptor signal ligand binding, not channel activation." *Biophysical journal*, 96(1), pp. 45–55, 2009.

[78] S. A. Pless and J. W. Lynch. "Ligand-specific conformational changes in the alpha1 glycine receptor ligand-binding domain." *The Journal of biological chemistry*, 284(23), pp. 15847–56, 2009.

[79] L. Sauguet, a. Shahsavar, F. Poitevin, C. Huon, a. Menny, a. Nemecz, a. Haouz, J.-P. Changeux, P.-J. Corringer, and M. Delarue. "Crystal structures of a pentameric ligand-gated ion channel provide

a mechanism for activation". *Proceedings of the National Academy of Sciences*, 111(3), pp. 966–971, 2014.

[80] M. Cecchini and J.-P. Changeux. "The nicotinic acetylcholine receptor and its prokaryotic homologues: Structure, conformational transitions & allosteric modulation". *Neuropharmacology*, 96, pp. 137–149, 2014.

[81] H. E. M. Criado and F. J. Barrantes. "Effects of Lipids on Acetylcholine Receptor. Essential Need of Cholesterol for Maintenance of Agonist-Induced State Transitions in Lipid Vesicles". *3622 Biochemistry*, 1982.

[82] A. W. Dalziel, E. S. Rollins, and M. G. McNamee. "The effect of cholesterol on agonist-induced flux in reconstituted acetylcholine receptor vesicles". *FEBS Letters*, 122(2), pp. 193–196, 1980.

[83] O. Quesada, C. González -Freire, M. C. Ferrer, J. O. Colón -Sáez, E. Fernández-García, J. Mercado, A. Dávila, R. Morales, and J. A. Lasalde-Dominicci. "Uncovering the lipidic basis for the preparation of functional nicotinic acetylcholine receptor detergent complexes for structural studies". *Scientific Reports*, 6(1), p. 32766, 2016.

[84] W. S. Leibel, L. L. Firestone, D. C. Legler, L. M. Braswell, and K. W. Miller. "Two pools of cholesterol in acetylcholine receptor-rich membranes from Torpedo". *BBA - Biomembranes*, 897(2), pp. 249–260, 1987.

[85] J. R. Zabrecky and M. A. Raftery. "The role of lipids in the function of the acetylcholine receptor". *J. Recept. Res.*, 5(5-6), pp. 397–417, 1985.

[86] V. Borroni, V. Borroni, C. Baier, T. Lang, I. Bonini, M. M. White, I. Garbus, and F. J. Barrantes. "Cholesterol depletion activates rapid internalization of submicron-sized acetylcholine receptor domains at the cell membrane". *Molecular Membrane Biology*, 24(1), pp. 1–15, 2007.

[87] C. A. Báez-Pagán, Y. Martínez-Ortiz, J. D. Otero-Cruz, I. K. Salgado-Villanueva, G. Velázquez, A. Ortiz-Acevedo, O. Quesada, W. I. Silva, and J. A. Lasalde-Dominicci. "Potential role of caveolin-1-positive domains in the regulation of the acetylcholine receptor's activatable pool: Implications in the pathogenesis of a novel congenital myasthenic syndrome". *Channels*, 2(3), pp. 180–190, 2008.

[88] F. J. Barrantes. "The lipid environment of the nicotinic acetylcholine receptor in native and reconstituted membranes." *Critical reviews in biochemistry and molecular biology*, 24(5), pp. 437–78, 1989.

[89] W. C. Breckenridge, G. Gombos, and I. G. Morgan. "The lipid composition of adult rat brain synaptosomal plasma membranes". *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 266(3), pp. 695–707, 1972.

[90] W. G. Hill, N. M. Southern, B. MacIver, E. Potter, G. Apodaca, C. P. Smith, and M. L. Zeidel. "Isolation and characterization of the Xenopus oocyte plasma membrane: a new method for studying activity of water and solute transporters". *American Journal of Physiology - Renal Physiology*, 289(1), pp. F217–F224, 2005.

[91] H. I. Ingólfsson, M. N. Melo, F. J. Van Eerden, C. Arnarez, C. A. Lopez, T. A. Wassenaar, X. Periole, A. H. De Vries, D. P. Tieleman, and S. J. Marrink. "Lipid organization of the plasma membrane". *Journal of the American Chemical Society*, 136(41), pp. 14554–14559, 2014.

[92] S. Lombardo and U. Maskos. "Role of the nicotinic acetylcholine receptor in Alzheimer ' s disease pathology and treatment". *Neuropharmacology*, 96, pp. 255–262, 2015.

[93] R. Giniatullin, A. Nistri, and J. L. Yakel. "Desensitization of nicotinic ACh receptors: Shaping cholinergic signaling". *Trends in Neurosciences*, 28(7), pp. 371–378, 2005.

[94] K. L. Davis and H. I. Yamamura. "Cholinergic underactivity in human memory disorders". *Life Sciences*, 23(17-18), pp. 1729–1733, 1978.

[95] A. I. Levey. "Muscarinic acetylcholine receptor expression in memory circuits: implications for treatment of Alzheimer disease." *Proceedings of the National Academy of Sciences of the United States of America*, 93(24), pp. 13541–13546, 1996.

[96] A. Nordberg and B. Winblad. "Reduced number of [3H]nicotine and [3H]acetylcholine binding sites in the frontal cortex of Alzheimer brains". *Neuroscience Letters*, 72(1), pp. 115–120, 1986.

[97] M. A. Smith, L. M. Sayre, V. M. Monnier, and G. Perry. "Radical ageing in Alzheimer's disease". *Trends in Neurosciences*, 18(4), pp. 172–176, 1995.

[98] P. J. Whitehouse, A. M. Martino, P. G. Antuono, P. R. Lowenstein, J. T. Coyle, D. L. Price, and K. J. Kellar. "Nicotinic acetylcholine binding sites in Alzheimer's disease." *Brain research*, 371(1), pp. 146–151, 1986.

[99] S. C. Yao, A. D. Hart, and M. J. Terzella. "An evidence-based osteopathic approach to Parkinson disease". *Osteopathic Family Physician*, 5(3), pp. 96–101, 2013.

[100] O. B. Tysnes and A. Storstein. "Epidemiology of Parkinson's disease". 124(8), pp. 901–905, 2017.

[101] N. C. C. f. C. C. (UK). *Parkinson's Disease*. 2006.

[102] W. H. Organization and Others. "Tobacco or Health: a global status report. Geneva; 1997". 1997.

[103] J. W. Coe, P. R. Brooks, M. G. Vetelino, M. C. Wirtz, E. P. Arnold, J. Huang, S. B. Sands, T. I. Davis, L. A. Lebel, C. B. Fox, A. Shrikhande, J. H. Heym, E. Schaeffer, H. Rollema, Y. Lu, R. S. Mansbach, L. K. Chambers, C. C. Rovetti, D. W. Schulz, F. D. Tingley 3rd, and B. T. O'Neill. "Varenicline: an $\alpha 4\beta 2$ nicotinic receptor partial agonist for smoking cessation". *J Med Chem*, 48(10), pp. 3474–3477, 2005.

[104] J. R. HUGHES, S. T. HIGGINS, and W. K. BICKEL. "Nicotine withdrawal versus other drug withdrawal syndromes: similarities and dissimilarities". *Addiction*, 89(11), pp. 1461–1470, 1994.

[105] M. Born and W. Heisenberg. "Zur Quantentheorie der Molekeln". *Annalen der Physik*, 379(9), pp. 1–31, 1924.

[106] B. J. Alder and T. E. Wainwright. "Phase Transition for a Hard Sphere System". *J. Chem. Phys.*, 27(5), pp. 1208–1209, 1957.

[107] L. Verlet. "Comyuter "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard, -Jones Molecules". *Phys. Rev.*, 159(5), p. 98, 1967.

[108] J. P. Ryckaert, G. Ciccotti, and H. J. Berendsen. "Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes". *Journal of Computational Physics*, 23(3), pp. 327–341, 1977.

[109] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije. "LINCS: A linear constraint solver for molecular simulations". *Journal of Computational Chemistry*, 18(12), pp. 1463–1472, 1997.

[110] H. C. Andersen. "Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations". *Journal of Computational Physics*, 52(1), pp. 24–34, 1983.

[111] M. Tuckerman, B. J. Berne, and G. J. Martyna. "Reversible multiple time scale molecular dynamics". *The Journal of Chemical Physics*, 97(3), pp. 1990–2001, 1992.

[112] W. G. Hoover. "Canonical dynamics: Equilibrium phase-space distributions". *Physical Review A*, 31(3), pp. 1695–1697, 1985.

[113] S. Nosé. "A unified formulation of the constant temperature molecular dynamics methods". *The Journal of Chemical Physics*, 81(1), pp. 511–519, 1984.

[114] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. "Molecular dynamics with coupling to an external bath". *The Journal of Chemical Physics*, 81(8), pp. 3684–3690, 1984.

[115] Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder, and P. Ren. "Polarizable atomic multipole-based amoeba force field for proteins". *Journal of chemical theory and computation*, 9(9), pp. 4046–4063, 2013.

[116] C. M. Baker. "Polarizable force fields for molecular dynamics simulations of biomolecules". *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 5(2), pp. 241–254, 2015.

[117] B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus. "CHARMM: The biomolecular simulation program". *Journal of Computational Chemistry*, 30(10), pp. 1545–1614, 2009.

[118] A. D. MacKerell, B. Brooks, C. L. Brooks, L. Nilsson, B. Roux, Y. Won, and M. Karplus. "CHARMM: The Energy Function and Its Parameterization". In "Encyclopedia of Computational Chemistry", 2002.

[119] R. B. Best, X. Zhu, J. Shim, P. E. Lopes, J. Mittal, M. Feig, and A. D. MacKerell Jr. "Optimization of the additive charmm all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain $\chi 1$ and $\chi 2$ dihedral angles". *Journal of chemical theory and computation*, 8(9), pp. 3257–3273, 2012.

[120] A. D. Mackerell, M. Feig, and C. L. Brooks. "Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulation". *Journal of Computational Chemistry*, 25(11), pp. 1400–1415, 2004.

[121] P. K. Weiner and P. A. Kollman. "Amber: Assisted model building with energy refinement. a general program for modeling molecules and their interactions". *Journal of Computational Chemistry*, 2(3), pp. 287–303, 1981.

[122] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives. "Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids". *J. Am. Chem. Soc*, 118(45), pp. 11225–11236, 1996.

[123] G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, and W. L. Jorgensen. "Evaluation and reparametrization of the opls-aa force field for proteins via comparison with accurate quantum chemical calculations on peptides". *The Journal of Physical Chemistry B*, 105(28), pp. 6474–6487, 2001.

[124] W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. Scott, and I. G. Tironi. "Biomolecular simulation: the {GROMOS96} manual and user guide". 1996.

[125] W. Van Gunsteren and H. Berendsen. "Groningen molecular simulation (gromos) library manual, biomos, groningen, the netherlands, 1987". *There is no corresponding record for this reference*, pp. 1–221.

[126] Y. Miao, J. E. Johnson, and P. J. Ortoleva. "All-atom multiscale simulation of cowpea chlorotic mottle virus capsid swelling." *The journal of physical chemistry. B*, 114(34), pp. 11181–95, 2010.

[127] L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman, and S. J. Marrink. "The MARTINI coarse-grained force field: Extension to proteins". *Journal of Chemical Theory and Computation*, 4(5), pp. 819–834, 2008.

[128] M. Praprotnik, L. Delle Site, and K. Kremer. "Adaptive resolution molecular-dynamics simulation: Changing the degrees of freedom on the fly". *Journal of Chemical Physics*, 123(22), 2005.

[129] S. Izvekov and G. A. Voth. "Modeling real dynamics in the coarse-grained representation of condensed phase systems". *Journal of Chemical Physics*, 125(15), 2006.

[130] Q. Shi, S. Izvekov, and G. A. Voth. "Mixed atomistic and coarse-grained molecular dynamics: Simulation of a membrane-bound ion channel". *Journal of Physical Chemistry B*, 110(31), pp. 15045–15048, 2006.

[131] G. M. De Mori, G. Colombo, and C. Micheletti. "Study of the villin headpiece folding dynamics by combining coarse-grained Monte Carlo evolution and all-atom molecular dynamics". *Proteins: Structure, Function and Genetics*, 58(2), pp. 459–471, 2005.

[132] M. Christen and W. F. Van Gunsteren. "Multigraining: An algorithm for simultaneous fine-grained and coarse-grained simulation of molecular systems". *Journal of Chemical Physics*, 124(15), 2006.

[133] E. Lyman, F. M. Ytreberg, and D. M. Zuckerman. "Resolution exchange simulation". *PHYSICAL REVIEW LETTERS*, 96(2), p. 2, 2005.

[134] W. Kwak and U. H. E. Hansmann. "Efficient sampling of protein structures by model hopping".

Physical Review Letters, 95(13), 2005.

[135] S. V. Semenovskaya, K. A. Khachaturyan, and A. G. Khachaturyan. "Statistical mechanics approach to the structure determination of a crystal". *Acta Crystallographica Section A*, 41(3), pp. 268–273, 1985.

[136] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. "Optimization by Simulated Annealing". *Science*, 220(4598), pp. 671–680, 1983.

[137] V. Černý. "Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm". *Journal of Optimization Theory and Applications*, 45(1), pp. 41–51, 1985.

[138] X. Periole and S. J. Marrink. "The martini coarse-grained force field". *Methods in Molecular Biology*, 924, pp. 533–565, 2013.

[139] C. Steffen, K. Thomas, U. Huniar, A. Hellweg, O. Rubner, and A. Schroer. "TmoleX–a graphical user interface for TURBOMOLE." *Journal of computational chemistry*, 31(16), pp. 2967–2970, 2010.

[140] D. Bassolino-Klimas, R. Tejero, S. R. Krystek, W. J. Metzler, G. T. Montelione, and R. E. Bruccoleri. "Simulated annealing with restrained molecular dynamics using a flexible restraint potential: theory and evaluation with simulated nmr constraints". *protein Science*, 5(4), pp. 593–603, 1996.

[141] R. E. Bruccoleri and M. Karplus. "Conformational sampling using high-temperature molecular dynamics". *Biopolymers*, 29(14), pp. 1847–1862, 1990.

[142] P. Auffinger and G. Wipff. "High temperature annealed molecular dynamics simulations as a tool for conformational sampling. application to the bicyclic "222" cryptand". *Journal of computational chemistry*, 11(1), pp. 19–31, 1990.

[143] W. Commons. "File:schematic of a replica exchange molecular dynamics simulation.svg — wikimedia commons, the free media repository", 2016.

[144] Y. Sugita and Y. Okamoto. "Replica-exchange molecular dynamics method for protein folding". *Chemical Physics Letters*, 314(1-2), pp. 141–151, 1999.

[145] S. Park, T. Kim, and W. Im. "Transmembrane helix assembly by window exchange umbrella sampling". *Physical Review Letters*, 108(10), pp. 1–4, 2012.

[146] S. Park and W. Im. "Two dimensional window exchange umbrella sampling for transmembrane helix assembly". *Journal of Chemical Theory and Computation*, 9(1), pp. 13–17, 2013.

[147] M. Moradi and E. Tajkhorshid. "Mechanistic picture for conformational transition of a membrane transporter at atomic resolution". *Proceedings of the National Academy of Sciences*, 110(47), pp. 18916–18921, 2013.

[148] A. Laio and F. L. Gervasio. "Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science". *Reports on Progress in Physics*, 71(12), p. 126601, 2008.

[149] A. Barducci, M. Bonomi, and M. Parrinello. "Metadynamics". *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 1(5), pp. 826–843, 2011.

[150] A. Barducci, G. Bussi, and M. Parrinello. "Well-tempered metadynamics: A smoothly converging and tunable free-energy method". *Physical Review Letters*, 100(2), 2008.

[151] F. Comitani. "Computational modelling of activation mechanisms in ligand-gated ion channels."

[152] D. Hamelberg, J. Mongan, and J. A. McCammon. "Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules". *Journal of Chemical Physics*, 120(24), pp. 11919–11929, 2004.

[153] Y. Miao, D. A. Goldfeld, E. V. Moo, P. M. Sexton, A. Christopoulos, J. A. McCammon, and C. Valant. "Accelerated structure-based design of chemically diverse allosteric modulators of a muscarinic G protein-coupled receptor". *Proceedings of the National Academy of Sciences*, 113(38), pp. E5675–E5684, 2016.

[154] "Accelerated Molecular Dynamics - NAMD UG".

[155] A. Laio and M. Parrinello. "Escaping free-energy minima." *Proceedings of the National Academy of Sciences of the United States of America*, 99(20), pp. 12562–12566, 2002.

[156] E. Darve, D. Rodriguez-Gomez, and A. Pohorille. "Adaptive biasing force method for scalar and vector free energy calculations". *Journal of Chemical Physics*, 128(14), 2008.

[157] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman. "THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method". *Journal of Computational Chemistry*, 13(8), pp. 1011–1021, 1992.

[158] M. R. Shirts and J. D. Chodera. "Statistically optimal analysis of samples from multiple equilibrium states". *Journal of Chemical Physics*, 129(12), 2008.

[159] M. Souaille and B. Roux. "Extension to the weighted histogram analysis method: Combining umbrella sampling with free energy calculations". *Computer Physics Communications*, 135(1), pp. 40–57, 2001.

[160] C. Bartels. "Analyzing biased Monte Carlo and molecular dynamics simulations". *Chemical Physics Letters*, 331(5-6), pp. 446–454, 2000.

[161] M. Seeber, M. Cecchini, F. Rao, G. Settanni, and A. Caflisch. "Wordom: A program for efficient analysis of molecular dynamics simulations". *Bioinformatics*, 23(19), pp. 2625–2627, 2007.

[162] O. S. Smart, J. G. Neduvelil, X. Wang, B. A. Wallace, and M. S. P. Sansom. "HOLE: a program for the analysis of the pore dimensions of ion channel structural models". *Journal of molecular graphics*, 14(6), pp. 354–360, 1996.

[163] J. L. Trick, S. Chelvaniththilan, G. Klesse, P. Aryal, E. J. Wallace, S. J. Tucker, and M. S. P. Sansom. "Functional Annotation of Ion Channel Structures by Molecular Simulation". *Structure*, 24(12), pp. 2207–2216, 2016.

[164] S. Van Der Walt, S. C. Colbert, and G. Varoquaux. "The NumPy array: A structure for efficient numerical computation". *Computing in Science and Engineering*, 13(2), pp. 22–30, 2011.

[165] D. C. Thompson, C. Humblet, and D. Joseph-Mccarthy. "Investigation of MM-PBSA Rescoring of Docking Poses".

[166] O. Korb, T. Stützle, and T. E. Exner. "PLANTS: Application of Ant Colony Optimization to Structure-Based Drug Design". *Lecture notes in computer science vol. 4150: Ant colony optimization and swarm intelligence - ANTS2006 proceedings*, 4150, pp. 247–258, 2006.

[167] S. Zhong, Y. Zhang, and Z. Xiu. "Rescoring ligand docking poses." *Current opinion in drug discovery & development*, 13(3), pp. 326–334, 2010.

[168] A. Fiser and A. Sali. "Modeller: generation and refinement of homology-based protein structure models". *Methods in enzymology*, 374, pp. 461–491, 2003.

[169] D. Bashford and M. Karplus. "{pK_as} of ionizable groups in proteins: atomic detail from a continuum electrostatic model". *Biochemistry*, 29(44), pp. 10219–10225, 1990.

[170] D. Bashford and M. Karplus. "Multiple-site titration curves of proteins: an analysis of exact and approximate methods for their calculation". *The Journal of Physical Chemistry*, 95(23), pp. 9556–9561, 1991.

[171] V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, and D. C. Richardson. "MolProbity: All-atom structure validation for macromolecular crystallography". *Acta Crystallographica Section D: Biological Crystallography*, 66(1), pp. 12–21, 2010.

[172] Dr. Jeff Klauda. "Membrane model".

[173] S. R. Durell, B. R. Brooks, and A. Ben-Naim. "Solvent-Induced Forces between Two Hydrophilic Groups". *The Journal of Physical Chemistry*, 98(8), pp. 2198–2202, 1994.

[174] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A. D. Mackerell. "CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields". *Journal of Computational Chemistry*, 31(4), pp. 671–690, 2010.

[175] W. Yu, X. He, K. Vanommeslaeghe, and A. D. MacKerell. "Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations". *Journal of Computational Chemistry*, 33(31), pp. 2451–2468, 2012.

[176] P. J. Steinbach and B. R. Brooks. "New spherical cutoff methods for long range forces in macromolecular simulation". *Journal of Computational Chemistry*, 15(7), pp. 667–683, 1994.

[177] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten. "Scalable molecular dynamics with NAMD". *Journal of Computational Chemistry*, 26(16), pp. 1781–1802, 2005.

[178] P. Purohit, A. Mitra, and A. Auerbach. "A stepwise mechanism for acetylcholine receptor channel gating." *Nature*, 446(7138), pp. 930–933, 2007.

[179] E. R. Nightingale. "Phenomenological Theory of Ion Solvation. Effective Radii of Hydrated Ions". *The Journal of Physical Chemistry*, 63(9), pp. 1381–1387, 1959.

[180] D. Sehnal, R. S. Vařeková, K. Berka, L. Pravda, V. Navrátilová, P. Banáš, C. M. Ionescu, M. Otyepka, and J. Koča. "MOLE 2.0: Advanced approach for analysis of biomacromolecular channels". *Journal of Cheminformatics*, 5(8), 2013.

[181] B. Lev, S. Murail, F. Poitevin, B. A. Cromer, M. Baaden, M. Delarue, and T. W. Allen. "String method solution of the gating pathways for a pentameric ligand-gated ion channel." *Proceedings of the National Academy of Sciences of the United States of America*, 114(21), pp. E4158–E4167, 2017.

[182] M. S. Prevost, L. Sauguet, H. Nury, C. Van Renterghem, C. Huon, F. Poitevin, M. Baaden, M. Delarue, and P.-J. Corringer. "A locally closed conformation of a bacterial pentameric proton-gated ion channel". *Nature Structural & Molecular Biology*, 19(6), pp. 642–649, 2012.

[183] S. Yuan, S. Filipek, and H. Vogel. "A Gating Mechanism of the Serotonin 5-HT3 Receptor". *Structure*, 24(5), pp. 816–825, 2016.

[184] S. Gupta, S. Chakraborty, R. Vij, and A. Auerbach. "A mechanism for acetylcholine receptor gating based on structure, coupling, phi, and flip". *The Journal of General Physiology*, 149(1), pp. 85–103, 2017.

[185] P. Aryal, M. S. P. Sansom, and S. J. Tucker. "Hydrophobic gating in ion channels". *Journal of molecular biology*, 427(1), pp. 121–130, 2015.

[186] C. J. Wenthur, P. R. Gentry, T. P. Mathews, and C. W. Lindsley. "Drugs for Allosteric Sites on Receptors". *Annual Review of Pharmacology and Toxicology*, 54(1), pp. 165–184, 2014.

[187] A. F. Abdel-Magid. "Allosteric modulators: an emerging concept in drug discovery." *ACS medicinal chemistry letters*, 6(2), pp. 104–7, 2015.

[188] G. Fiorin, M. L. Klein, and J. Hénin. "Using collective variables to drive molecular dynamics simulations". *Molecular Physics*, 111(22-23), pp. 3345–3362, 2013.

[189] W. R. Hamilton. "II. On quaternions; or on a new system of imaginaries in algebra". *Philosophical Magazine Series 3*, 25(163), pp. 10–13, 1844.

Nicolas MARTIN

ALLOSTERIC MODULATION OF PENTAMERIC LIGAND GATED ION CHANNELS

Résumé

Les *Pentameric ligand gated ion channels* (pLGICs) sont des récepteurs neuronaux impliqués dans la neurotransmission rapide au niveau des synapses et comprennent les récepteurs à la nicotine, au GABA, à la glycine et à la sérotonine notamment. Lorsqu'ils ne fonctionnent pas correctement, ils peuvent être mis en cause dans des maladies telles qu'Alzheimer, Parkinson ou encore certraines formes d'épilepsie. Ils sont les cibles de traitements d'ores et déjà commercialisés mais leur méchanisme de fonctionnement reste méconnu.

Dans cette thèse je presente les résultats issus de plus de 6 µs de simulations de dynamique moléculaire tout atomes du premier récepteur eucaryote de la famille des pLGICs (GluCl) elucidé en crystallographie X-ray. Ces analyses ont conduit à la proposition d'un nouveau méchanisme de *gating* que nous pensons applicable à l'ensemble des pLGICs.

Dans un second temps, je présente une méthode basée sur des calculs d'énergie libre qui permet de mieux comprendre le rôle des ligands allostériques sur le méchanisme de *gating* des pLGICs. Enfin, je discute les possibilités d'utilisation de cette méthode pour la conception de nouveau potentiateurs et illustre celle-ci par quelques exemples.

Les résultats présentés dans cette thèse illustre notamment la capacité des simulations numériques à être prédictives et à représenter une alternative viable aux experiences.

Résumé en anglais

Pentameric ligand gated ion channels (pLGICs) are brain receptors involved in fast neurotransmission at synapses and include the nicotinic acetylcholine, the $GABA_A$, the glycine or the serotonin receptors. When dysfunctioning, they are involved in a number of diseases including Alzheimer's, Parkinson's or epilepsy. They are allosteric proteins and the targets of commercially available drugs. Nonetheless, their mechanism of action has remained elusive.

In this thesis I present the results of more than 6 µs of full atomistic molecular dynamics simulations of the first ever elucidated eukaryotic homologue (GluCl), which lead to the proposition of a new gating mechanism for the pLGICs.

A second capter of my thesis presents the developpment of a method involving free energy calculations to better understand the role of allosteric ligands on the gating mechanism of pLGICs. Finally, I discuss the possibilities to use this approach in drug design and illustrate the former by several examples.

The results discussed in this manuscripts highlight, in particular, how simulation techniques can be predictive and can reasonably be percieved as an alternative to experiments.