

Université de Montréal

**Étude numérique de la formation du complexe
protéique formé du canal potassique humain Kv4.2 et
de sa sous-unité bêta DPP6.2.**

par

Michaël Morin

Département de physique
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de
Maître ès sciences (M.Sc.)
en physique, option biophysique et physiologie moléculaire

26 octobre 2018

SOMMAIRE

L'homéostasie de tout organisme vivant résulte de l'activité concertée d'un nombre inimaginable d'acteurs biochimiques. Parmi ceux-ci, les protéines jouent un rôle fondamental de par leurs fonctions aussi diverses que cruciales, autant à un niveau structural, enzymatique que d'un point de vue de transportation et de signalisation. Par exemple, les canaux potassiques voltage-dépendants (Kv) forment une famille de protéines transmembranaires responsables du transport des ions potassium à travers la membrane cellulaire des cellules excitables. Les courants électriques résultants de leur activité constituent une composante essentielle d'une variété de processus physiologiques de grande importance. Le potentiel d'action en est un exemple omniprésent chez tous les animaux pour le relais d'informations entre le système nerveux et autres systèmes comme les muscles, les glandes, autres organes, ou encore lui-même. Cette étude s'intéresse à l'origine moléculaire de deux autres exemples de courants potassiques, référés par $I_{to,f}$ et I_{SA} , qui joue un rôle régulateur similaire dans les cellules cardiaques et neuronales, respectivement. Or, bien qu'on ait déjà identifié l'isoforme Kv4 comme étant à l'origine du pore conducteur de ces courants, les cinétiques de Kv4 sont trop lentes pour rendre compte de leur forme précise de façon isolée. En effet, ces courants sont la manifestation de complexes protéine-protéine impliquant deux sous-unités auxiliaires modulant les propriétés biophysiques de Kv4, soit KChIP (*Kv Channel Interacting Proteins*) et DPP (*dipeptidyl-aminopeptidase like proteins*). Plusieurs propriétés de Kv4 sont affectées par ces interactions, comme par exemples l'accélération des cinétiques d'activation et d'inactivation de Kv4 par DPP – un effet crucial pour la régulation que I_{SA} et $I_{to,f}$ exercent sur les autres courants des neurones et myocytes qui les expriment. Certains détails moléculaires importants à la base de cette interaction ont été identifiés, principalement pour les parties cytosoliques, mais très peu reste connu sur les interactions principales, au niveau des parties transmembranaires. Cette étude vise à combler ce vide en précisant le mode d'action de DPP6 sur Kv4.2 avec l'aide de la dynamique moléculaire.

Dans un premier temps, étant donné que seules quelques sous-structures ont été cristallisées pour les deux protéines, on devra construire des modèles pour la structure des protéines complètes. Ensuite, on utilisera une combinaison d'algorithmes d'amarrage

moléculaire, complétés de dynamique moléculaire, pour préciser leur mode d'interaction. Ce raffinement de notre compréhension passera par différentes méthodes d'analyses pour identifier le modèle d'interaction le plus probable. D'abord, on comparera la stabilité structurelle et la force d'interaction pour identifier la pose préférentielle. Cette étape permet autant de comparer notre modèle à la littérature et de formuler des hypothèses supplémentaires pour orienter les recherches en laboratoire future. Ensuite, on regardera l'effet de la liaison de la sous-unité auxiliaire sur les sous-structures de Kv4 dans l'espoir d'y trouver un indice de comment DPP6.2 perturbe le mécanisme d'ouverture et de fermeture du canal « normal ». On a ainsi pu corroborer une hypothèse, formulée par Jerng et Pfaffinger (2014), selon laquelle le pore serait contraint par l'interaction. On a également prédit que le domaine sensible au voltage était moins stable, et on suggère que cela pourrait être la base de l'activation plus rapide de Kv4.2 en présence de DPP6.2. De plus, le site d'interaction identifié par les algorithmes d'amarrage était consistant avec des données préliminaires dans la littérature (Ren et al., 2003). Quant à l'extrémité N-terminale, la seule région d'interaction (réaliste) prédite était hautement consistant avec les études fonctionnelles en laboratoire, puis on a prédit quelques autres interactions. On a également montré que les premiers acides N-terminaux sont flexibles et peuvent interagir avec le pore et son mécanisme, tel que suggéré par le modèle prévalent de type « balle et chaîne » qui s'applique aussi pour Kv4.2.

mots clés : canaux potassiques dépendants du voltage, sous-unités auxiliaires, interaction protéine-protéine, dynamique moléculaire, relation structure-fonction

SUMMARY

The homeostasis of all living organism results from the concerted action of a unfathomable number of biochemical actors. Among them, proteins play a fundamental role by range of functions as diverse as crucial that they fulfill, whether it is on a structural or enzymatic level, or on a transportation and signaling one. For instance, voltage-dependent potassium channels (Kv) form a family of transmembrane proteins responsible for the transportation of potassium ions across the membrane of excitable cells. The electric currants resulting from their activity constitute an essential component of a wide variety of vital physiological processes. The action potential is such an example, omnipresent in all vertebrates to relay information between the nervous and other systems, such as the muscles, glands, organs, or the nervous system itself. The present study looks at the molecular origin of two other potassium currents, referred by $I_{to,f}$ et I_{SA} , that play a similar regulatory role for cardiac and neuronal cells, respectively. While it has been clearly established that the isoforme Kv4 is the molecular basis for the conducting pore of these currents, the kinetics of this Kv channel are too slow to reproduce the currents on its own. Indeed, these currents are the manifestation of a larger protein-protein complex that imply two auxiliary subunits that modulate Kv4's biophysical propreties, namely KChIP (*Kv Channel Interacting Proteins*) and DPP (*dipepdyl-aminopeptidase like proteins*). Many of Kv4's properties are affected by these interactions, like the acceleration Kv4's gating and inactivation kinetics by DPP – a crucial effect for the regulation I_{SA} et $I_{to,f}$ exert on other neuronal or cardiac currents. A handful of molecular details about this interaction are known, especially for the cytosolic part, but very little is known about the main interaction, at the level of the transmembrane domains. This study aims at filling this void by precisising the mode of action of DPP6.2 on Kv4.2 with the help of molecular dynamics.

Initially, since only partial crystal structure data is available for both proteins, predictive models for the structure of the models will have to be constructed. Then, a combination of docking algorithms, complemented with molecular dynamics, will help refine the mode of interaction. The most probable model will be identified by comparing the structural stability and the strenght of the affinity of the different complexes proposed by docking algorithms.

The preferential pose thereby predicted will allow to compare with the literature and to formulate further hypotheses to test the models. Then, we will look at the effect of the binding of auxiliary subunits of Kv4's substructures to gain insight into how DPP6.2 disturb the normal gating process. We could corroborate an hypothesis, formulated by Jerng et Pfaffinger (2014), according to which the pore would be constrained by the interaction. We also predicted that the voltage sensing domain was less stable, and we suggest that this could be the basis of the faster activation of Kv4.2 in presence of DPP6.2. Moreover, the region of interaction identified by docking algorithms are consistent with preliminary data (Ren et al., 2003). As for the N-terminal part of the interaction, the only (realist) region of interaction predicted was highly consistent with functional studies available, and we predicted a few other interactions. We also showed that the first N-terminal residues form a flexible chain, readily available to interact with the pore and its mechanism, as suggested by the prevalent model of the ball and chain that applies for Kv4.2.

keywords : voltage-gated potassium channels, auxiliary subunits, protein-protein interaction, molecular dynamics, structure-function relationship.

TABLE DES MATIÈRES

Sommaire	iii
Summary	v
Liste des tableaux	ix
Liste des figures	xi
Liste des sigles et des abréviations	xiii
Dédicaces	xvii
Remerciements	xix
Introduction	1
Chapitre 1. Les canaux potassiques	5
1.1. Introduction aux canaux potassiques.....	6
1.2. Classification topologique et structure des canaux potassiques.....	9
1.2.1. La région du pore.....	9
1.2.2. Le senseur de voltage des canaux K_V	12
1.2.3. Biophysique thermodynamique des canaux K_V	17
1.2.4. Inactivation de K_V4	21
1.3. Couplage électromécanique.....	23
1.4. Les sous-unités bêta de K_V4 et les courants I_{SA} et $I_{to,f}$	28
1.5. Les sous-unités bêta de K_V4 et les courants I_{SA} et $I_{to,f}$	30
Chapitre 2. Méthodologie	37
2.1. La dynamique moléculaire classique.....	38
2.1.1. Principes généraux.....	38
2.1.2. Dynamique de Langevin et intégrateur BAOAB.....	40

2.2.	Champ de force.....	41
2.3.	Modèles atomiques in silico.....	42
2.3.1.	Modèle d'homologie de Kv4.2.....	43
2.4.	Amarrage moléculaire.....	44
2.5.	Préparation du système, pH et autres options de simulation.....	46
2.6.	Logiciels et <i>hardware</i>	48
2.7.	Méthodes d'analyse.....	48
2.8.	La validité, limite et utilité des résultats numériques.....	51
Chapitre 3.	Résultats.....	55
3.1.	Prédiction de la topologie de DPP6.2.....	56
3.2.	Amarrage moléculaire de DPP6.2 et Kv4.2.....	62
3.2.1.	Interactions entre le domaine T1 de Kv4.2 et le N-terminal de DPP6.2	62
3.2.1.1.	Prédictions de ClusPro2.0.....	62
3.2.1.2.	Prédictions de GRAMM-X.....	64
3.2.1.3.	Détails moléculaires de l'interaction N-terminale.....	67
3.2.2.	Amarrage des parties transmembranaires.....	72
3.3.	Stabilité relative des configurations MemDock.....	75
3.3.1.	Équilibre thermodynamique.....	75
3.3.2.	Orientations finales ϕ_f	77
3.3.3.	Énergies d'interaction.....	81
3.3.4.	Analyse de la stabilité par le RMSD.....	84
3.3.5.	Changements conformationnels lors de la formation du complexe et mécanisme d'action.....	88
3.3.6.	Discussion.....	91
Chapitre 4.	Conclusion.....	95
Bibliographie	103

LISTE DES TABLEAUX

3. 1	Pointage ClusPro et population des six familles identifiées par ClusPro.	64
3. 2	Pointages de MemDock en ordre décroissant et facteur de raffinement FiberDock.	74
3. 3	Rotations de corps rigide du squelette des TM de DPP6 de θ autour de \mathbf{u} par rapport à leurs orientations initiales respectives et décomposition de cette rotation en x, y et z	80
3. 4	Énergie d'interaction non-liée entre les segments transmembranaires de DPP6.2 et Kv4.2	81
3. 5	RMSD moyen et déviation standard pour les régions de contact définies par une distance d fixe du TM pour chaque configuration initiale ϕ_i	86
3. 6	Changements conformationnels lors de la liaison du complexe.	90

LISTE DES FIGURES

1.1	La fonction catalysatrice des canaux ioniques.	7
1.2	Arbre phylogénique des canaux ioniques principaux chez l'humain.	8
1.3	Structure cristalline et topologie de KcsA.	10
1.4	Le filtre de sélectivité.	11
1.5	Topologie typique des canaux Kv.	13
1.6	Sous-structures typiques des canaux Kv.	14
1.7	Polarité des sous-structures du domaine sensible au voltage.	15
1.8	Interactions électrostatiques au sein du domaine sensible au voltage.	16
1.9	Crevasses du domaine sensible au voltage accessible au milieu aqueux.	18
1.10	Courants unitaires typiques des canaux ioniques.	19
1.11	Voltages et courants pour $G(V)$ et pour l'inactivation à l'équilibre.	21
1.12	Comparaison des cinétiques d'inactivation de Kv4.2 et Kv2.1.	23
1.13	Comportement local du champ électrique près des VSD.	25
1.15	Modulation de Kv4.2 par DPP6.	33
3.1	Analyse séquentielle des cent premiers acides aminés de DPP6.2.	56
3.2	Alignement structurel des modèles QUARK pour les acides aminés de 32 à 57 de DPP6.2.	57
3.3	Alignement structurel des modèles QUARK pour le domaine N-terminal (aa. 1-32).	58
3.4	Dynamique de l'extrémité N-terminale des modèles QUARK pour DPP6.2 en solution.	60
3.5	Alignement BLAST des isoformes de DPP6.	61
3.6	Interaction de l'hélice centrale du N-terminal de DPP6.2 avec le domaine T1 de Kv4.2.	63

3.7	Catégories GRAMM-X de modèles d'interaction entre T1 et l'hélice N-terminale.	66
3.8	Détails moléculaires de l'interface putative d'interaction de T1 avec l'hélice N-terminale de DPP6.	68
3.9	Interactions électrostatiques inter-sous-unités au coeur du domaine de tétramérisation.	69
3.10	Représentation isosurfacique de la région de T1 médiant l'interaction putative avec l'hélice N-terminale.	70
3.11	Interfaces d'interactions putatives entre l'hélice N-terminale et T1.	72
3.12	Amarrage moléculaire des segments transmembranaires de K_V et DPP6 par deux algorithmes « classiques ».	73
3.13	Simulations d'amarrage moléculaire MemDock.	74
3.14	Évolution des RMSD.	76
3.15	Évolutions des angles de rotations solides.	79
3.16	Comparaison de deux paires de TM d'une même classe d'équivalence rotationnelle.	83
3.17	Région de contact entre Kv4.2 et le TM de DPP6.2.	85
3.18	Pose du segment transmembranaire de DPP6.2 dans la pochette créée par S1 et S5.	87
3.19	Relation entre l'énergie d'interaction non-liée et l'écart-type des RMSD des régions d'interaction.	88
4.1	Structure cristalline du domaine T1 de K_V 4.3 avec KChIP1.	98
4.2	Effets des différentes sous-unités β de Kv4.2 sur ces cinétiques.	99

LISTE DES SIGLES ET DES ABRÉVIATIONS

aa	acide aminé
TM	transmembranaire
Ca ²⁺	Ion calcium
(cellule) CHO	cellule ovarienne du hamster chinois, de l'anglais <i>chinese hamster ovary (cell)</i>
Cl ⁻	Ion chlorure
CM	Centre de masse
CSI	inactivation à l'état fermé, de l'anglais <i>closed state inactivation</i>
DM	Dynamique moléculaire
DPP6/10	de l'anglais <i>dipeptidyl-aminopeptidase-like protein 6/10</i>
I_{SA}	Courant transitoire sous-seuil de type A
K ⁺	Ion potassium
K _{2P}	Canal potassique à deux pores
K _{Ca}	Canal potassique activé par le calcium
K _{ir}	Canal potassique à courant rectifiant-retrant
K _V X	Canal potassique dépendant du voltage, isoforme X (où X peut aller de 1.1 à 4.3, dans ce texte)
KChIP	Protéine interactrice avec le canaux potassiques, de l'anglais <i>K+ channel interacting protein</i>
Na ⁺	Ion sodique
ns	nanoseconde

RMSD	Racine des déviations moyennes au carré, de l'anglais <i>root-mean-square deviation</i>
T1	domaine de tétramérisation T1
TEA	tetraethylammonium
VSD	Domaine sensible au voltage, de l'anglais <i>voltage-sensing domain</i>

DÉDICACES

À mon père Bruno.

REMERCIEMENTS

Je tiens d'abord à exprimer toute ma gratitude au Dr. Rikard Blunck pour son rôle de mentor lors de mon passage au sein de son laboratoire. Sa disponibilité, son ouverture d'esprit, sa flexibilité, son constant support et ses précieux conseils ont tous été des éléments cruciaux de ma réussite. Il s'est montré compréhensif durant certains moments difficiles et n'était en rien obligé de m'offrir la possibilité de prouver que je méritais une seconde chance. J'en suis très reconnaissant.

J'aimerais ensuite souligner la contribution de mes collègues de travail qui, par leur désir de partager leurs connaissances, par leur aide des plus concrètes et leurs judicieux conseils, ont tous significativement contribué à mon évolution dans ce domaine de recherche passionnant. Je pense particulièrement à Tanja Kalstrup, Oliver Fürst et Juan Zhao pour m'avoir tout appris des techniques expérimentales en biologie moléculaire et en électrophysiologie durant mon premier projet (non publié).

Mes remerciements s'adressent également aux professeurs Michèle Brochu, Nazzareno D'Avanzo et Jean-Yves Lapointe et aux techniciens Mireille Marsolais, Michel Brunette et Hélène Klein pour leur support, leurs enseignements et l'aide technique qu'ils m'ont apportés.

Finalement, ces remerciements seraient incomplets sans reconnaître l'inestimable apport de mes proches. D'abord, de ma mère et ma soeur, Nathalie McKenzie et Karine Morin, dont l'amour inconditionnel ont sculpté la fabrique même de ma réussite. Ensuite, de mon père Bruno Morin qui, je sais, serait comblé de fierté, et qui, encore aujourd'hui, continue de m'apprendre énormément. Enfin, je m'estime choyé d'être l'ami proche d'Alexandre Galarneau et de Jean-François Lavallée depuis près de vingt ans. Partager cette grande aventure cosmique avec vous tous donne un sens indéniable au mystère perpétuel qu'est le moment présent.

INTRODUCTION

C'est au 17^e siècle que les êtres humains ont pu voir pour la première fois, grâce à l'amélioration des microscopes, le plus petit système physique que l'on puisse qualifier de « vivant ». Le physicien Robert Hooke sera le premier à observer ce qu'on appelle aujourd'hui une *cellule*, nom qu'il a lui-même donné à ces unités de base du tissu végétal qu'il étudiait car la structure lui rappelait celle des cellules d'une ruche d'abeille. Cependant, ne voyant pas les multiples organelles et le noyau cellulaires, Hooke ne se doutait pas que les cellules étaient le siège de presque toutes les réactions chimiques propres à la vie et que de tous les organismes vivants s'organisaient de cette façon. Il faudra attendre près de deux cents ans pour que les physiologistes et botanistes Schwann et Schleiden comprennent le rôle fondamental qu'elles jouent. De la simple observation qu'elles étaient composées de matériaux recyclés provenant de leur cellule mère, ils formuleront leur fameuse *théorie cellulaire*. Ils ont en effet eu la perspicacité de comprendre toute la portée de leur découverte, à partir de laquelle ils formulent l'hypothèse, maintenant universellement acceptée, selon laquelle *tous* les organismes vivants sont formés de ces unités structurales et fonctionnelles de base, qu'une cellule donnée contient l'ensemble de l'information (génétique) nécessaire pour accomplir et réguler ses fonctions et que chaque cellule provient d'une cellule mère.

Cependant, la nature précise des mécanismes moléculaires à l'oeuvre derrière l'encodage, la transmission et l'utilisation fonctionnelle de l'information génétique restait un mystère pour ces biologistes précurseurs. D'autant plus que pour la majeure partie de l'histoire, l'approche matérialiste et empirique de la science moderne était loin d'être le paradigme par défaut pour comprendre le monde du vivant. Par exemple, René Descartes, contemporain de Hooke et figure de proue dans la révolution scientifique ayant marqué leur époque, croyait encore que la glande pinéale de l'humain était le siège d'une âme immatérielle qui contrôlait le véhicule matériel qu'elle habitait. Toutefois, même si Descartes accordait encore un statut divin à la conscience humaine, il reste l'un des premiers à avancer que le comportement des (autres) animaux pourrait s'expliquer selon un modèle purement mécanistique et déterministe.

Aujourd'hui, il est clair que la vie n'échappe pas aux lois de la physique ; qu'au contraire, elle n'en est qu'une propriété émergente. Plusieurs découvertes marquantes ont contribué à réduire cet apparent fossé entre l'animé et l'inanimé, entre la biologie et la physique. Luigi Galvani l'observe de façon assez spectaculaire, en 1780, lorsqu'il a réussi à faire bouger les pattes amputée d'une grenouille morte en y faisant passer un courant électrique. Le vitalisme jusqu'alors largement incontesté a pris un sérieux coup avec cette démonstration irréfutable que les mouvements des animaux ont une origine électrique. Depuis, peu de problème biologique résiste encore au réductionnisme dont se vantent avec raison les physiciens. Parmi les plus grandes applications des techniques issues de la physique à des

problèmes biologiques, pensons entre autre à Francis Crick et James Watson qui découvrent la structure tridimensionnelle exacte de l'ADN grâce à la diffraction aux rayons X. Pour leurs travaux décisifs dans le fondement de la biologie moléculaire, ils recevront le prix Nobel de médecine et physiologie en 1962. Pas plus tard que l'année suivante, le prestigieux prix sera encore remis à des biophysiciens, Alan Hodgkin et Andrew Huxley, pour l'élaboration d'un modèle mathématique expliquant la dynamique du courant électrique circulant dans les nerfs – ce même courant, nommé *potentiel d'action*, que Galvani induisant dans ses pattes de grenouilles.

Les expériences d'électrophysiologie en voltage contrôlé que Hodgkin et Huxley réalisent sur l'axone géant du calmar leur indiquent que deux conductances bien distinctes jouent un rôle clé dans la propagation du potentiel d'action : la conductance aux ions *potassiques* et aux ions *sodiques*. Ils montrent comment ces conductances varient en fonction du potentiel membranaire selon un profil respectif bien précis qu'ils parviennent à mesurer grâce à cette technique récemment inventée. L'avantage majeure de celle-ci repose dans l'ajout d'une seconde électrode injectant du courant selon un *feedback loop* dans le but de garder le potentiel membranaire constant. Ainsi, en traitant la membrane comme une capacitance semi-perméable dont on connaît la perméabilité, ils parviennent à écrire un ensemble d'équations différentielles non-linéaires dont la solution reproduit la forme particulière du courant électrique traversant les neurones. Cependant, Hodgkin et Huxley n'ont aucune idée de l'origine moléculaire de ces perméabilités variables de la membrane plasmique. Ils intuent qu'un changement dans la structure physique de la membrane doit avoir lieu, mais les mécanismes exacts restaient alors insondables. De façon étonnante, par contre, même si l'existence de petites « machines moléculaires » à la surface des cellules excitables restaient encore hautement hypothétique, les mathématiques du modèle s'avèrent refléter quelques unes des caractéristiques biophysiques, comme par exemple leur agencement homotétramérique.

Ces travaux comptent parmi les plus grands accomplissements de la biophysique. Ils établissent un précédent remarquable en appliquant les techniques de la physique au monde du vivant et en ouvrant la porte à plusieurs domaines de recherche toujours très actifs autant en biochimie, en biologie moléculaire qu'en biophysique. Depuis lors, nous avons compris que ces « machines » contrôlant le passage de certains ions sont des protéines transmembranaires. Nous avons identifié plusieurs centaines de protéines de transport similaires et la quête de la compréhension des multiples mécanismes à l'origine de leurs fonctions physiologiques vitales se poursuit toujours intensément, partout dans le monde.

Ce mémoire s'inscrit dans cet élan effréné vers une compréhension complète du « miracle » de la vie à son niveau le plus fondamental. Les recherches qu'il présente se concentrent sur un membre de la superfamille étudiés par Hudgkin et Huxley, soit les canaux potassiques dépendants du voltage. On y tente de comprendre comment le canal potassique Kv4.2 interagit avec une de ses sous-unités bêta, la *dipeptidyl-aminopeptidase-like protein 6*, ou (DPP6), pour accomplir son rôle. En effet, in vivo, les canaux ioniques forment souvent des complexes protéiques avec des sous-unités auxiliaires régulant leurs propriétés biophysiques et leur niveau d'expression. Par contre, très peu de détails est connu sur la nature de cette interaction et comment celle-ci change les propriétés du canal pour donner les courants mesurés in vivo. Un modèle moléculaire détaillé sera construit in silico pour les deux protéines et le processus de *docking* sera simulé numériquement pour mieux comprendre la dynamique de formation du complexe et comment les interactions protéines-protéines causent les changements des propriétés biophysiques du canal K_V mesurés in vivo. Dans un premier temps, cependant, l'état actuel des connaissances sur les canaux potassiques et leur sous-unités bêta sera résumé.

Chapitre 1

LES CANAUX POTASSIQUES

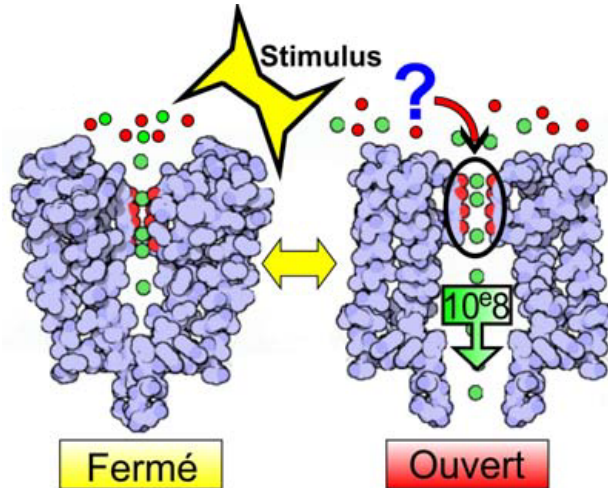
1.1. INTRODUCTION AUX CANAUX POTASSIQUES

Bien que la membrane cellulaire soit composée d'une bicouche de phospholipides hautement imperméable, il est crucial que les cellules puissent constamment échanger avec leur environnement. Cette tâche est principalement accomplie par des protéines transmembranaires qui catalysent le transport d'ions, d'eau et d'autres composés polaires à travers ce milieu hydrophobe. Parmi ces protéines régulatrices figure notamment la pompe sodium-potassium qui, grâce à l'énergie emmagasinée dans l'ATP, parvient à maintenir une concentration d'ions de potassium (K^+) et de sodium (Na^+) allant à l'encontre de leur gradient électrochimique respectif. Ainsi, de façon général, le milieu extracellulaire est riche en Na^+ et faible en K^+ , alors que l'inverse est vrai pour le milieu intracellulaire. Ces gradients de concentration sont maintenus de façon telle à garder un potentiel électrique membranaire négatif à l'intérieur de toute cellule à l'équilibre.

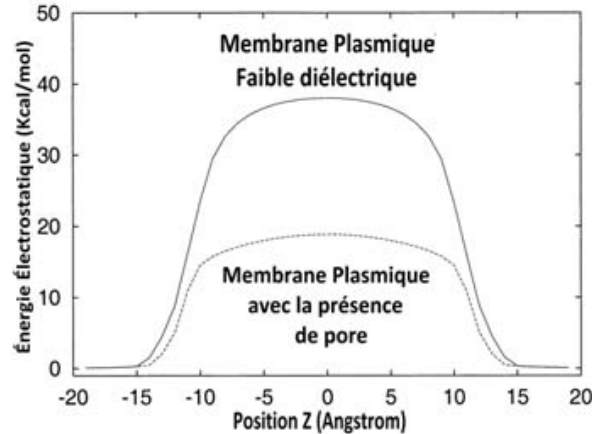
Cette énergie potentielle électrique est mise à profit par maints processus physiologiques vitaux. Par exemple, les cellules dites « excitables », comme les neurones et les cellules musculaires, ont la capacité de se dépolariser et de communiquer ce changement de potentiel aux cellules avec lesquelles elles communiquent. Cette dépolarisation (et la repolarisation qui s'en suit) constitue un événement physiologique significatif, le potentiel d'action, qui joue rôle fondamental dans le transfert d'information d'une cellule excitable à l'autre. C'est grâce à ce signal électrique que l'information neuronale est traitée et transmise et que les muscles sont contractés.

Comme l'ont montré Hodgkin et Huxley, la transduction du signal le long de l'axone des neurones est permise par l'existence de deux conductances membranaires bien distinctes, qui varient en fonction du potentiel et du temps selon un profil très différent selon que les ions transportés soient des ions K^+ ou Na^+ . Nous savons aujourd'hui que cette particularité est due à l'existence de deux grandes superfamilles de protéines transmembranaires possédant un pore hautement sélectif au passage de seulement certains ions : les canaux sodiques et les canaux potassiques. Les membres des deux familles ont la propriété de « s'ouvrir » et de se « fermer » selon les conditions physico-chimiques de l'environnement, abaissant considérablement la barrière énergétique nécessaire au passage des ions dans la membrane (fig. 1.1).

Pour cette raison, ces protéines transmembranaires sont appelées des canaux ioniques. L'arbre phylogénétique de la figure 1.2 recense les 147 canaux ioniques humains selon leur similitude structurelle (Yu et al., 2005). La présente étude se concentre sur un membre de la famille des canaux potassiques, lesquelles forment la famille de canaux ioniques la plus diversifiée et, pourrait-on argumenter, la plus importante dans le règne animal (Mazzuca et



(a) États fonctionnels d'un canal ionique.



(b) Profil énergétique des états fonctionnels.

FIGURE 1.1. La fonction catalysatrice des canaux ioniques. En (a), le changement de conformation subi par un canal ionique suite à un stimulus de l'environnement est schématisé. L'ouverture du canal provoque le passage sélectif des ions (seuls les verts) à un taux près de la diffusion simple. En (b), on voit comment l'ouverture du chemin de conduction abaisse la barrière énergétique (en y) des ions à travers la membrane plasmique (en x , avec $x = 0$ au centre de la membrane). Adapté de (Roux et Bernèche, 2000).

Lesage, 2007).

Chez l'humain seulement, il existe 77 types de canaux potassiques que l'on retrouve exprimés à des niveaux variables partout dans le corps. À l'intérieur de cette superfamille, on distingue encore quatre sous-familles : les canaux activé par le calcium (K_{Ca}), les canaux à deux pores (K_{2P}), les canaux à courant rectifiant-rentant (K_{ir}) et, finalement, les canaux potassiques dépendant du voltage (K_V). Cette étude se concentre plus particulièrement sur

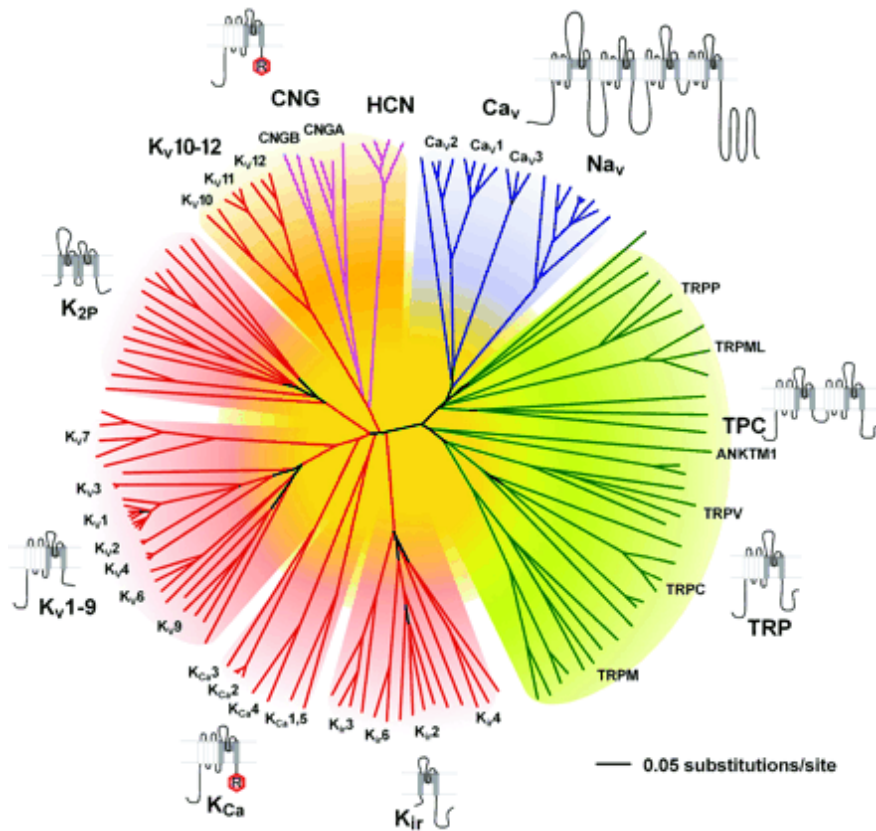


FIGURE 1.2. Arbre phylogénique des canaux ioniques principaux chez l'humain. On y voit les canaux potassiques (rouge), les canaux sodiques (bleu), les canaux HCN et CNG (mauve) et les canaux mécano-sensitif (TRP, en vert). Chaque famille est accompagnée d'un schéma représentant sa structure secondaire. Adapté de (Yu et al., 2005).

un membre de la dernière sous-famille, le $K_V4.2$. Cette grande diversité des canaux K^+ reflète une large gamme de fonctions physiologiques. Pour n'en nommer que quelques unes, citons seulement leur rôle vital dans la propagation des potentiels d'action, l'émission de neurotransmetteurs, la sécrétion d'insuline, la contraction des muscles et le maintien du volume cellulaire. Vu leur omniprésence, il n'est pas surprenant que le dysfonctionnement des canaux K_V soit relié à plus d'une centaine de pathologies génétiques graves : l'arythmie, l'épilepsie, la neuromyotonie, la myopathie, l'autisme, le syndrome du long QT, etc.

Avant de poursuivre avec la classification des canaux K^+ selon leur topologie, il convient de dire un mot ou deux sur la raison pour laquelle cette classification est pertinente. En effet, il existe un lien direct entre la structure et la fonction de toute protéine. Comprendre cette relation est l'objectif principal de la biophysique. Le lien entre la structure et la fonction découle du fait que la structure tertiaire d'une protéine – i.e. sa conformation tridimensionnelle – résulte directement du repliement de la séquence d'acides aminés sur

elle-même à mesure que le gène est transcrit par les ribosomes, uniquement via le jeu des interactions électromagnétiques. Or, pour remplir leurs fonctions, plusieurs sous-structures spécialisées s'agencent et coordonnent leur activité. Il est donc possible que des acides aminés éloignés les uns des autres dans la séquence primaire fassent partie d'une même structure fonctionnellement importante, une fois la protéine repliée. Pour prendre l'exemple des canaux K_V , il est facile d'imaginer que le mécanisme permettant l'ouverture du canal suite à une dépolarisation et le mécanisme assurant la sélectivité ne soit pas assuré par la même sous-structure. Ainsi, pour identifier les sous-structures fonctionnellement importantes, il est utile de débuter la classification à son niveau le moins raffiné, en comparant la topologie macroscopique grossière des canaux, puisque les structures conservées entre les différents isoformes rapportent souvent des fonctions conservées.

1.2. CLASSIFICATION TOPOLOGIQUE ET STRUCTURE DES CANAUX POTASSIQUES

1.2.1. La région du pore

Ce qui distingue les canaux K^+ de tous les autres est leur grande sélectivité au potassium. Les membres de cette superfamille sont tous de 50 à 10 000 fois plus perméables au K^+ qu'au Na^+ , son principal compétiteur physiologique (Doyle et al., 1998; Heginbotham et al., 1994; Heginbotham et al., 1998). De plus, cette conduction ionique préférentielle s'effectue à un taux s'approchant de près de la limite de la diffusion simple. La combinaison de ces deux propriétés est a priori très contre-intuitive étant donné que le rayon atomique de K^+ est 1.22 fois plus grand que celui de Na^+ . La question de comprendre comment un ion ayant la même charge et un rayon plus gros peut passer tellement plus facilement, et ce, à une telle vitesse, est longtemps restée un profond mystère. En effet, le pore ne peut pas être un simple trou permettant un passage passif, sans quoi les ions Na^+ passeraient au moins aussi facilement que les ions K^+ , mais, en même temps, la cinétique observée contraint significativement la nature des interactions chimiques derrière la sélectivité. Des interactions covalentes, comme celles ayant lieu avec la pompe Na^+/K^+ , ont lieu sur des échelles temporelles beaucoup plus longues. C'est en 1998 que l'équipe de Roderick MacKinnon résout ce paradoxe en obtenant la première structure cristalline d'un canal potassique procaryote, KcsA, exprimé chez la bactérie *Streptomyces lividans* (fig. 1.3). Non seulement s'agissait-il d'un accomplissement considérable d'un point de vue technique, vu la difficulté d'obtenir un cristal de macromolécules en bonne partie hydrophobes, mais, surtout, les informations structurales obtenues se sont prouvées très riches en information sur l'origine moléculaire d'un bon nombre de fonctions physiologiques. Il s'agit d'un exemple parfait illustrant l'intime relation entre la structure et la fonction d'une protéine, laquelle

relation est le leitmotiv moteur en biophysique et en biologie structurale.

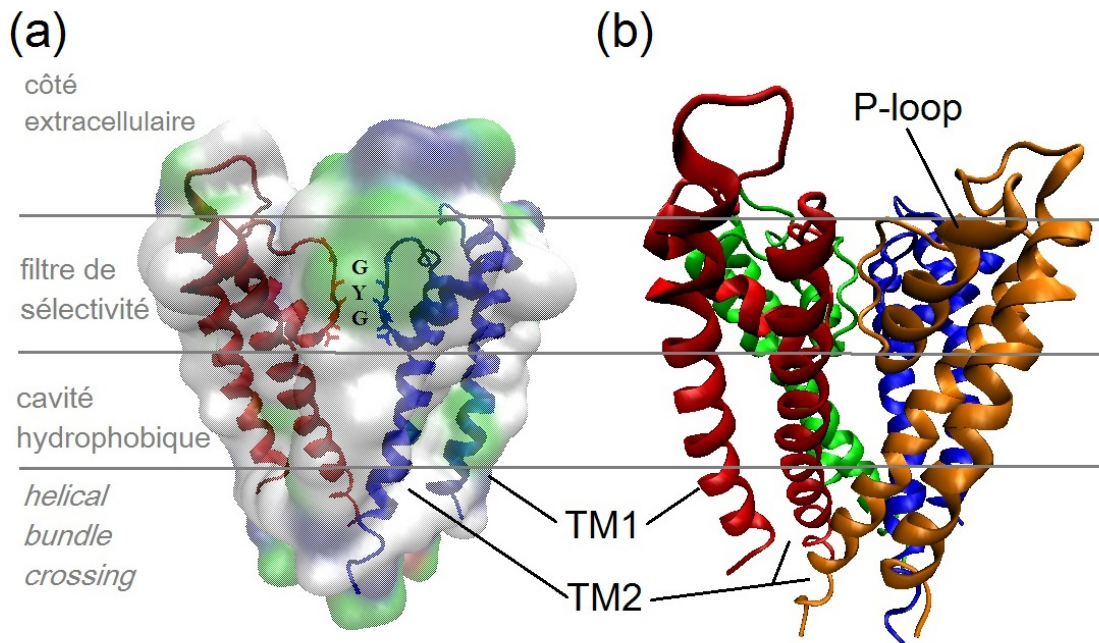


FIGURE 1.3. Structure cristalline et topologie de KcsA. En (a), sont illustrées seulement deux sous-unités sur quatre (rouge et bleue) dans le but de mettre en évidence le filtre de sélectivité (GYG). Surimposée est l'isosurface *QuickSurf* (VMD) de KcsA, colorée selon sa polarité (blanc indique une surface hydrophobique, le vert, polaire mais non-chargé; le bleu est chargé positivement et le rouge, chargé négativement). En (b), la structure tétramérique est mise en évidence avec les quatre sous-unités (rouge, orange, verte et bleue). Les sous-structures (filtre de sélectivité, cavité hydrophobique, helical bundle crossing) et sa topologie (TM1&2, P-loop) sont illustrée sur les deux représentations. Produit à partir de 1BL8.pdb.

Ces travaux révèlent une structure tétramérique où quatre paires d'hélices transmembranaires (TM) identiques s'organisent autour d'un pore central. Une des deux hélices (TM1) fait face à la membrane alors que l'autre (TM2) forme la partie interne du pore. Dans sa configuration « fermée », les hélices interne et externe sont inclinées de 25° et 10° , respectivement, de façon à former un cône inversé (Doyle et al, 1998). Sauf pour les canaux à deux pores, il s'avère que cette structure homo- ou hétérotétramérique est répliquée chez tous les canaux potassiques découverts à ce jour. La sélectivité, donc, opère au niveau de la région supérieure du pore. Chaque paire TM est reliée par une région nommée « *P-loop region* » formée principalement d'une courte hélice α ré-entrant dans la membrane, suivie d'une séquence d'acides aminés universellement exprimée chez cette famille : T/SxxTxGxG (fig. 1.3). Cette *séquence de sélectivité* est partiellement enfouie dans la partie supérieure

de la membrane et, lorsque les quatre sous-unités s'agencent, cette section forme le filtre de sélectivité du canal.

Les glycines de la séquence de sélectivité joue un rôle particulièrement important en projetant leur groupement carbonyle vers le centre du pore. Cette configuration en forme d'anneaux de huit oxygènes est cruciale pour le passage préférentiel des ions K^+ . Les données de cristallisation révèlent que le diamètre du pore (3.0 \AA) est juste assez large pour accueillir des ions K^+ (2.7 \AA) (Doyle et al., 1998). La géométrie résultante de l'agencement des anneaux d'oxygène se trouve donc à imiter de très près l'organisation des oxygènes de la sphère de solvation des K^+ en solution (fig.1.4 (b)). Cette relative « continuité » avec le milieu aqueux permet aux ions K^+ de se déshydrater sans devoir surmonter la barrière énergétique significative qu'aurait demandé le passage dans cet environnement autrement très hydrophobe (Bernèche et Roux, 2001). Il est intéressant de noter que c'est cette similitude chimique avec la sphère de solvation de K^+ qui permet une conductance s'approchant de la limite de la diffusion simple. Or, ce même environnement n'est pas aussi énergétiquement favorable pour le sodium. En effet, à cause surtout de l'interaction avec trois résidus stabilisateurs situés derrière le filtre (E71, D80 et W67), celui-ci est trop rigide pour s'adapter au diamètre plus petit de Na^+ (1.9 \AA) (Cordero-Morales, Cuello et al. 2006b) (fig.1.4(c)). Conséquemment, bien que possible, la conduction des ions sodiques par les canaux K^+ est très improbable.

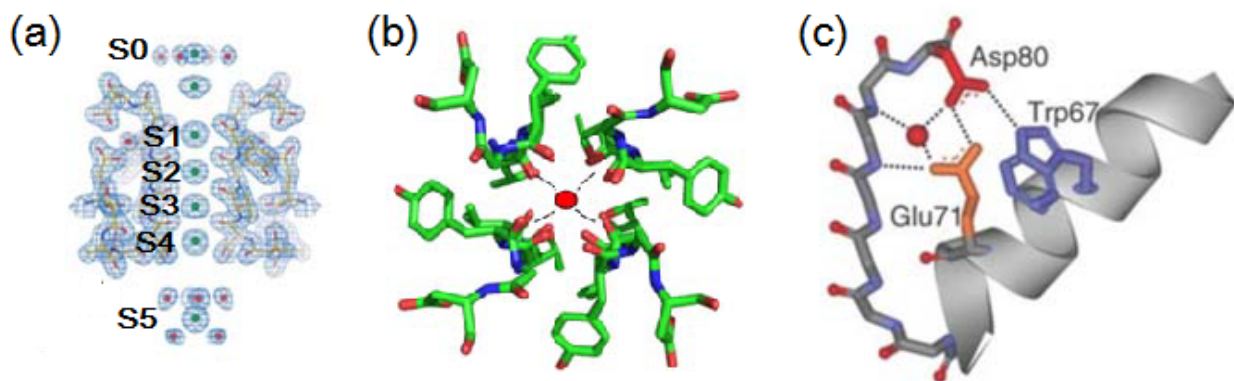


FIGURE 1.4. Le filtre de sélectivité. En (a), la densité électronique du filtre montre l'existence de quatre sites internes, S1-S4 (dont seuls deux peuvent être occupés par K^+ , séparé par un site occupé par de l'eau) et de deux sites externes, S0 et S5. En (b), les groupements carbonyles sont projetés vers le centre du pore, juste assez large pour couvrir le rayon d'hydratation du K^+ . En (c), la stabilité du filtre est assurée par les interactions entre E71 et W67 de la P-loop et D80 du haut du filtre. Adapté de (Zhou et al., 2001) et (Cordero-Morales et al., 2006b).

Le filtre de sélectivité possède quatre sites de liaisons numérotés de S1 à S4. Cependant, les données de dynamique moléculaire et de cristallisation ont révélé que seule la moitié des sites peuvent être occupés par un ion à un instant donné. En effet, pour éviter une répulsion électrostatique trop importante, seuls les sites S1 et S3 *ou* S2 et S4 peuvent accueillir un ion simultanément, alors que les deux autres sont occupés par une molécule d'eau (Berneche et Roux, 2001 ; Morais-Cabral, Zhou et al. 2001). Il existe également deux sites externes, un à la sortie (S0, côté extracellulaire) et l'autre à l'entrée (S5), où les ions n'interagissent que partiellement avec les résidus du filtre. L'arrivée d'un ion à l'un des deux sites externes repousse les occupants du filtre vers le prochain site disponible, et ce, sans devoir traverser de barrière énergétique.

Après le filtre de sélectivité, la région du pore s'élargit en une grosse cavité hydrophobique à l'intérieur de laquelle les quelques ions présents sont hydratés par les molécules d'eau pouvant s'y loger (fig. 1.3). Le bas de cette cavité, formé par la rencontre des C-termini des hélices internes du pore, constitue la première région importante pour le contrôle du passage des ions. Cet « *helical bundle crossing* » bloque physiquement la conduction des ions lorsque le canal est dans sa conformation fermée (Armstrong, 1971 ; Holmgren et al., 1997 ; Doyle et al., 1998). C'est cette partie qui s'ouvrira si les conditions physiologiques le permettent. Le processus de « *gating* » comporte plusieurs séquences et les changements de conformations associés peuvent varier d'un type de canal à l'autre, selon les stimuli externes provoquant l'ouverture. Il suffit pour l'instant de dire que la conduction est enfin complète lorsque l'hélice interne est redressée verticalement, permettant ainsi l'ouverture du diamètre de l'*helical bundle crossing* et la libre diffusion des ions cytoplasmiques vers le filtre. Les détails concernant l'origine de ces changements de conformation pour les canaux K_V seront explicités dans une sous-section à venir sur le couplage électromécanique. Il convient d'abord de décrire la topologie des structures supplémentaires impliquées dans la réponse aux stimuli externes.

1.2.2. Le senseur de voltage des canaux K_V

À la section précédente, la structure du pore et son mécanisme de sélectivité ont été décrits. Or, bien que certains canaux potassiques ne possèdent que les deux hélices α transmembranaires composant le pore — KcsA et K_{ir} étant les plus étudiés —, la vaste majorité en possède six. On compte parmi cette sous-famille à six TM les canaux KCNQ, les canaux K^+ activés par le calcium (BK, IK et SK), ainsi que les canaux dépendants du voltage (K_V), sur lesquels la présente étude se concentre. Leur pore central partagent toute la même topologie que leurs cousins à deux TM, mais ceux-ci (renommés S5 et S6 pour K_V) sont reliés à quatre autres hélices TM (S1 à S4) qui forment le domaine sensible au voltage (VSD) (fig. 1.5) (Aggarwal et MacKinnon, 1996 ; Seoh, Sigg et al., 1996 ; Jiang, Lee et al.,

2003; Chanda, Asamoah et al., 2005). Les hélices du VSD s'agencent antiparallèlement les unes à côté des autres en une structure dont la forme rappelle celle d'un sablier. Ainsi, contrairement au pore, les VSD ne sont pas un agencement quaternaire des différentes sous-unités. On les retrouve plutôt en retrait les unes des autres, en périphérie du pore, où leur interaction est quasi-inexistante.

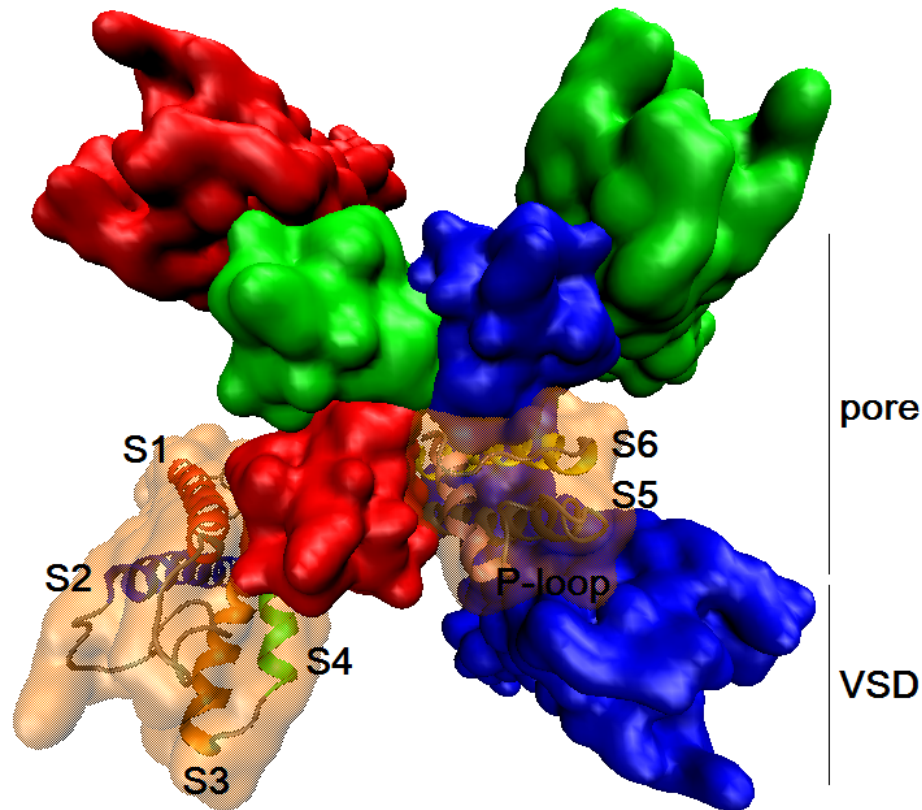


FIGURE 1.5. Topologie typique des canaux K_V . Vue de haut de $K_V4.2$, dont la topologie tétramérique consiste en un pore central (S5, S6 et P-loop) et quatre senseurs de voltage (ou VSD, S1 à S4). L'isosurface de chaque sous-unité est colorée différemment (rouge, bleue, verte et orange). L'une d'entre elle est laissée transparente pour montrer les six TM dont elle est composée. Produit à partir du modèle d'homologie de $K_V4.2$ de (Heler, Bell et Boland, 2013).

Le lien covalent entre le pore et le VSD est assuré par une structure appelée le S4-S5L (de l'anglais *S4-S5 linker*). Il s'agit d'une courte hélice α orientée quasiment à l'horizontal dans la membrane qui, comme son nom l'indique, fait le lien entre le segment S4 du VSD et l'hélice externe S5 du pore (TM1 chez KcsA) (voir fig. 1.6). Ce lien jouera un rôle de premier plan dans le transfert d'énergie mécanique du VSD au pore (voir section 1.3).

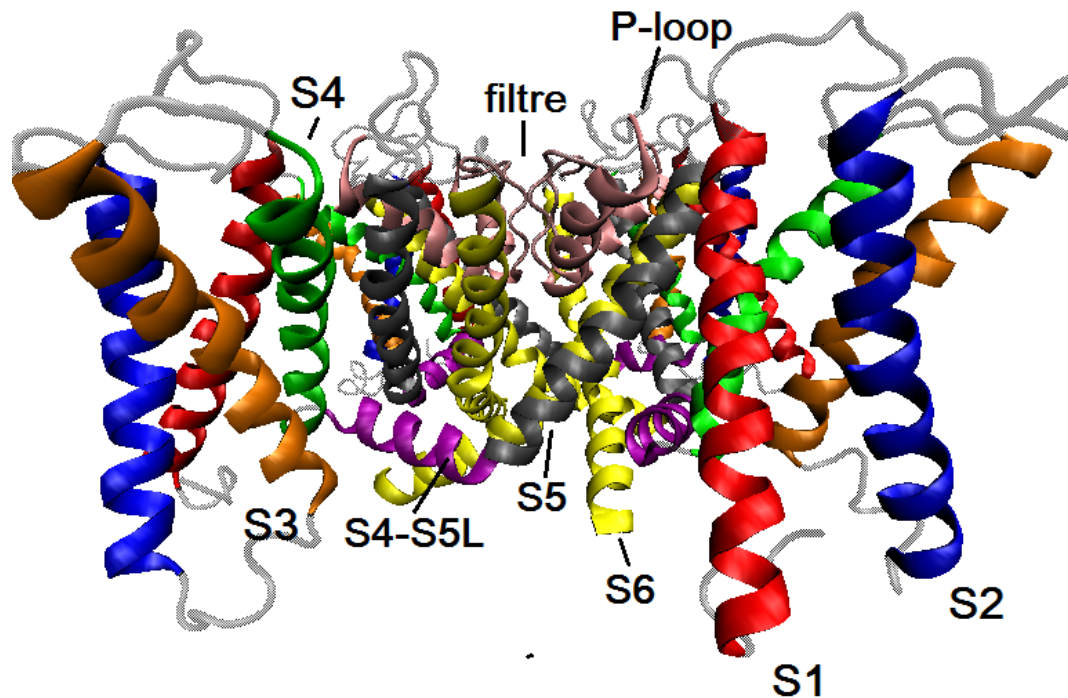


FIGURE 1.6. Sous-structures typiques des canaux K_V . Vue de profil de $K_V4.2$, dont chaque segment transmembranaire a été coloré différemment : S1, rouge ; S2, bleu ; S3, orange ; S4, vert ; S4-S5L, mauve ; S5, gris ; S6, jaune ; P-loop + filtre de sélectivité, rose. Produit à partir du modèle d'homologie de $K_V4.2$ de (Heler, Bell et Boland, 2013).

Par contre, avant de pouvoir parler des changements de conformation qui mènent à l'ouverture du pore, on doit d'abord comprendre les structures microscopiques qui les permettent. Il s'avère que c'est surtout au niveau de l'hélice S4 qu'on verra les plus grands changements conformationnels. Selon l'isoforme, chaque S4 comprend une séquence de quatre à six acides aminés positivement chargés, tous séparés par deux acides aminés hydrophobes (Noda et al., 1984). Dans le cas de $K_V4.2$, il s'agit des résidus R293, R296, R299, K302 et R305. Or, comme chaque tour d'hélice α comprend 3.6 acides aminés, cette répartition crée une surface positivement chargée au sein de la membrane lipidique (fig. 1.7).

Cette surface chargée de S4 est orientée vers le centre du VSD, enfouie dans une pochette formée par S1 à S3, où elle est stabilisée par un mélange de ponts salins (aa. chargé + avec -), de ponts H (aa. polaire non-chargé avec chargé) et d'interactions hydrophobiques. Pour les interactions électrostatiques, la figure 1.8 montre comment interagissent deux ensembles de résidus chargés. Dans sa partie supérieure, les arginines R293, R296 et R299 de S4 projettent

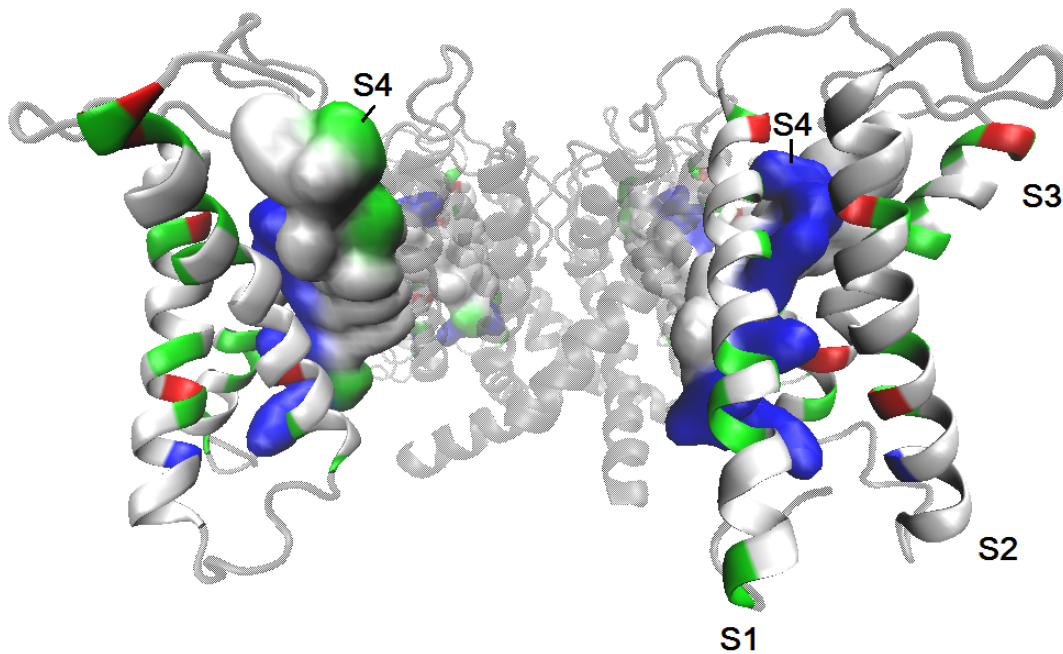


FIGURE 1.7. Polarité des sous-structures du domaine sensible au voltage (VSD). Vue de profil de $K_V4.2$, dont le pore central a été rendu transparent pour accentuer les VSD, colorés selon la polarité des acides aminés les composant. Les acides aminés hydrophobes sont blancs, les polaires mais non-chargés, verts, les chargés négativement, rouges, alors que les positifs sont bleus. Les segments S1-S3 (représentation *NewCartoon*) ont quelques résidus stabilisateurs chargés négativement (rouge), mais S4 (représentation isosurfacique) se démarque clairement par sa grande surface chargée positivement (bleue). Produit à partir du modèle d'homologie de $K_V4.2$ de (Heler, Bell et Boland, 2013).

leurs chaînes vers l'environnement négativement chargé créé par E205 (S1) et D233 (S2) (Tiwari-Woodruff et al., 2000; Khalili-Araghi et al., 2010). Dans sa partie inférieure, les résidus K302 et R305 de S4 interagissent avec D266 (S3) et E243 (S2), mais une autre charge positive, R247 de S2, fait en sorte que la charge reste non-nulle (Papazian et al., 2002; Papazian et al., 1995; Tiwari-Woodruff et al., 2000; Börjesson et Elinder, 2008).

L'observation des détails moléculaires indique qu'il existe un surplus de charge au coeur du VSD. En plus de n'être que partiellement neutralisées, le haut degré de localisation des charges positives de S4 sur une même hélice rend celle-ci hautement réactive aux changements de potentiel. Ce sera en effet surtout la réorientation de ces charges, suivant le champ électrique variable lors d'une dépolarisation, qui enclanchera les changements de conformation nécessaires à l'ouverture du canal (Sigworth, 1994; Aggarwal et MacKinnon,

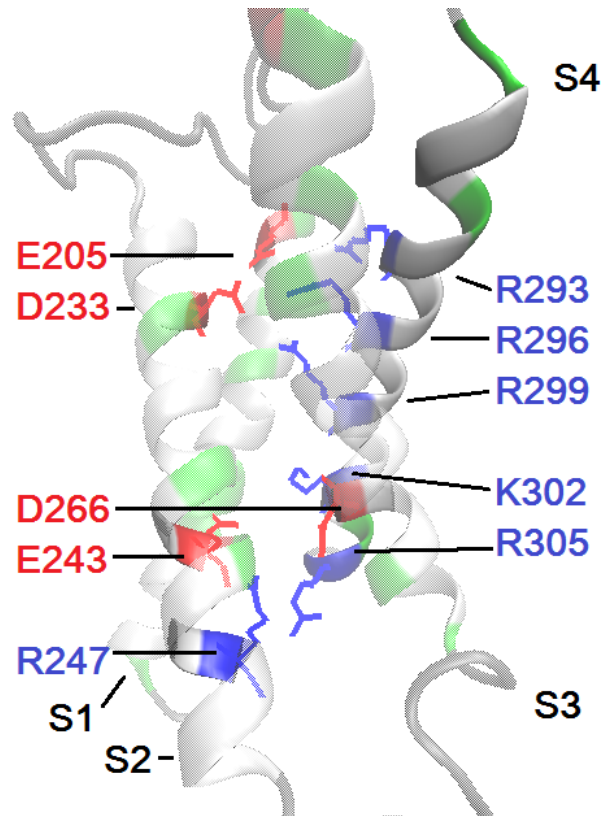


FIGURE 1.8. Interactions électrostatiques au sein du domaine sensible au voltage. Vue de profil du VSD de $K_V4.2$, colorés selon la polarité des acides aminés les composant (hydrophobe = blanc, polaire non-chargé = vert, positif = bleu, négatif = rouge). Les résidus E205 (S1) et D233 (S2) stabilise partiellement les trois résidus positifs de la partie supérieure de S4 (R293, R296 et R299) alors que les deux du bas (K302 et R305) et R247 (S2) sont partiellement stabilisés par D266 (S3) et E243 (S2). Produit à partir du modèle d'homologie de $K_V4.2$ (Heler, Bell et Boland, 2013).

1996 ; Seoh et al. 1996). Les études d'accessibilité aux agents sulfhydryles et de mutagenèse dirigée ont montré que les trois premières arginines, mieux accessibles du côté extracellulaire et donc moins enfouis dans la membrane, sont les plus actifs lors des dépolarisations (Aggarwal and MacKinnon, 1996 ; Seoh et al., 1996).

Or, bien qu'instrumental, S4 n'assure pas la sensibilité au voltage de façon isolée. Par exemple, le glutamate E233 de S2 s'avère également très actif lors des dépolarisations. En tout, on mesure que l'équivalent de 12 à 13 charge élémentaires e_0 par canal, auxquelles participent les charges négatives sus-mentionnées, traverse l'entièreté de la membrane lors de l'activation (Schoppa et al., 1992). Cependant, il est à noter que, tel que Seoh et al. (1996) l'ont fait montré, que certaines mutations neutralisantes diminuent de plus de $4e_0$ la charge déplacée, révélant donc des effets d'interaction non-linéaire. Il est possible, par

exemple, que l'absence d'une charge change légèrement la topologie locale du canal ou l'interaction avec les têtes polaires des lipides. Ces deux perturbations peuvent avoir des effets sur le champ électrique local ressenti par les autres résidus (Papazian et al. 1991).

Bref, il est important de rappeler qu'une discussion limitée aux interactions électrostatiques est inévitablement simpliste; que c'est plutôt l'action concertée des quatre segments TM du VSD qui est responsable de la sensibilité au voltage. Il est difficile de bien articuler l'ensemble des interactions, surtout quand on s'éloigne de la simplicité des ponts salins ici décrits et qu'on considère les effets non-négligeables des résidus polarisables et hydrophobes. De plus, on sait que l'environnement joue un rôle dynamique important. Les lipides environnants participent à la modulation du VSD autant par leur chaîne hydrophobe que par l'interaction de leur tête polaire avec les deux premières arginines extracellulaires de S4 (Gandhi, 2002; Cuello, 2004; Schmidt et al., 2006; Jogini et Roux, 2007; Long et al., 2007, Borjesson et al., 2008, 2010; Xu et al., 2008).

En résumé, plusieurs facteurs contribuent à façonner la réponse des VSDs au champ électrique local ressenti au sein de leur environnement hydrophobe. À cet égard, un aspect de nature davantage topologique s'ajoute aux déterminants moléculaires déjà décrits. En effet, la forme de sablier du VSD crée en son sein des régions accessibles relativement profondément dans la membrane, réduisant son épaisseur effective à une région hydrophobique de 5-10 Å, scellée par une douzaine de résidus non-polaires au centre du VSD (Goldstein, 1996; Larsson et al., 1996; George et al., 1996; Yusaf et al., 1996; Baker et al., 1998; Starace et Bezanilla, 2001) (fig. 1.9). La section suivante détaille l'effet de cette topologie particulière sur la sensibilité au voltage.

1.2.3. Biophysique thermodynamique des canaux K_V

Aux sections précédentes, la structure des canaux K_V a été décrite. On a exploré autant la topologie globale que quelques exemples des interactions microscopiques sous-jacentes à leurs fonctions principales (sélectivité au K^+ , sensibilité aux ΔV). La biophysique thermodynamique des canaux ioniques s'intéresse quant à elle aux propriétés macroscopiques et dynamiques d'ordre davantage fonctionnelles que structurelles. Cette section a l'objectif d'appliquer les principes de physique statistique à l'étude des transitions entre les états fonctionnels des canaux ioniques. D'ailleurs, ce sont les cinétiques de ces transitions que vient moduler DPP6.

En effet, on l'a vu, les canaux peuvent être dans plusieurs états fonctionnels différents. Au potentiel de repos, les canaux K_V sont non-conducteurs. À des potentiels dépolarisants,

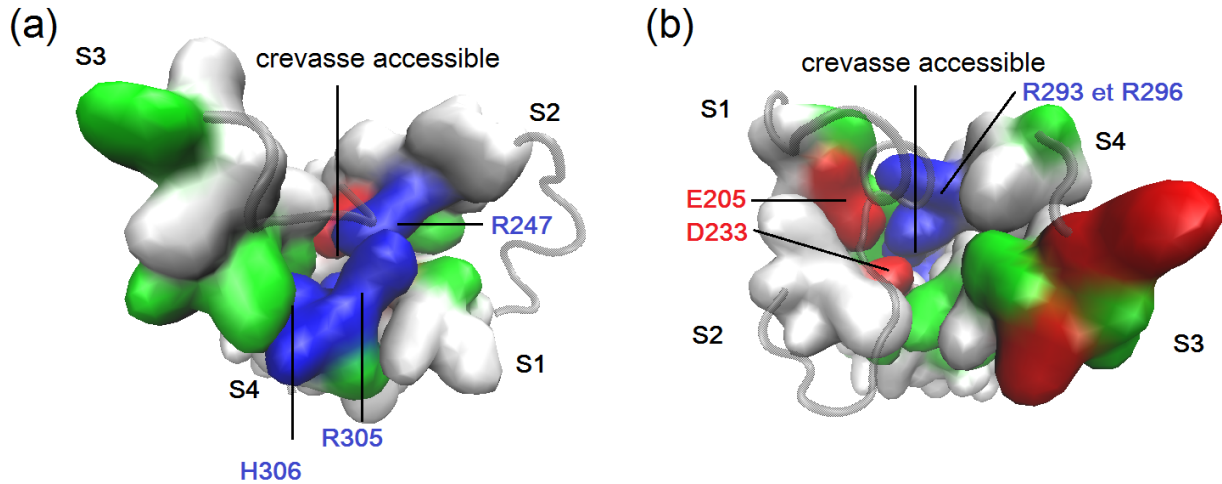


FIGURE 1.9. Crevasses du domaine sensible au voltage accessible au milieu aqueux. En (a), le creux central du VSD, côté extracellulaire, expose les premières arginines de S4, R293 et R296, et leurs résidus stabilisateurs E205 et D233. En (b), le bas du VSD montre une topologie creuse similaire. Produit à partir du modèle d'homologie de $K_V4.2$ de (Heler, Bell et Boland, 2013).

ils transitionnent vers un état conducteur, étape qu'on appelle l'*activation*. Une fois activés, ils peuvent soit se *désactiver*, qui est la transition vers non-conducteur suite à une repolarisation, ou *s'inactiver*, si le potentiel dépolarisant est maintenu assez longtemps. L'inactivation un état dans lequel le canal ne permet plus la conduction de courants ioniques malgré la présence d'un potentiel membranaire dépolarisant. On distingue deux types d'inactivation, selon la cinétique des processus moléculaires à leur origine : une rapide, appelée inactivation de type N et l'autre, lente, appelé de type C. Une quatrième transition se produit lorsque les canaux se *réactive*, i.e qu'il récupère de l'inactivation, lorsque la cellule se repolarise, pour retourner dans un état où ils sont disposés à être activés à nouveau.

Pour une peindre une image claire de la physique impliquée derrière ces transitions et des techniques pour mesurer leurs cinétiques, il convient de d'abord distinguer la notion de seuil d'excitabilité des neurones de l'ouverture sous-jacente des canaux ioniques à sa surface. En effet, il est connu que les potentiels d'action médiés par les canaux Na_V et K_V sont un phénomène tout-ou-rien, qui se produira avec certitude dès que le potentiel membranaire dépasse un certain seuil fixe. Un tel seuil discret n'existe pourtant pas au niveau des *conditions* d'activation (e.g. atteinte d'un certain voltage) des canaux ioniques unitaires. La propriété tout-ou-rien est plutôt le résultat d'une réaction en chaîne d'ordre davantage « macroscopique » qui nécessite des structures anatomiques neuronales particulières. Ce qui est d'ordre tout-ou-rien, par contre, est le *passage* à l'état conducteur (et non le voltage précis qui le déclenche). Autrement dit, il ne faut pas s'imaginer que les canaux peuvent

s'ouvrir partiellement, ou que leur conductance unitaire augmente graduellement à un potentiel donné. Il faut plutôt voir l'activation des canaux ioniques comme un phénomène stochastique dont la probabilité $P(V)$ d'ouverture dépend du voltage de façon *continue*, mais qu'une fois l'activation terminée, le canal est pleinement conducteur. Ces propriétés se reflètent en mesures électrophysiologiques en canal unitaire, où l'on observe, à un voltage donné, des sauts *spontanés* d'un état conducteur à un état non-conducteur (et vice versa), à une fréquence et d'une durée d'autant plus grandes que la cellule est dépolarisée (fig. 1.10. On obtient $P(V)$ en faisant le ratio des durées totales des deux états sur une échelle de temps appropriée (i.e. assez longue pour être représentative, mais assez courte pour éviter les inactivations).



FIGURE 1.10. Courants unitaires typiques des canaux ioniques. Les sauts d'un état fonctionnel à l'autre sont un processus stochastique dont la probabilité dépend continument du voltage. Image libre de droit d'auteur (tirée de Wikipédia, *Patch clamp*).

Cependant, cette approche renseigne très peu sur la dynamique des transitions entre ces états fonctionnels. Pour investiguer la cinétique des changements de conformation requis, une approche fructueuse consiste à penser en termes statistiques et d'observer les populations d'un grand nombre de canaux ioniques. In vitro, ou en conditions de laboratoire de surexpression, deux populations existent simultanément¹ : ceux dans l'état ouvert (O) et ceux dans l'état fermé (F).



La probabilité d'ouverture correspond à la proportion de canaux dans l'état O à l'équilibre :

$$P(V) = \frac{O}{O + F} = \frac{1}{1 + F/O}. \tag{1.2.2}$$

1. Ceci est une première approximation, car, on l'a vu, plusieurs sous-états peuvent exister, mais elle prédit bien la forme mesurée de $P(V)$.

On voit apparaître F/O , le rapport entre les deux populations. Il s'agit du facteur de Boltzmann, lequel se calcule en fonction de la différence d'énergie $E_F - E_O$ entre les deux états :

$$F/O = e^{-(E_F - E_O)/kT}. \quad (1.2.3)$$

Deux sources contribuent à cette différence d'énergie entre les deux états du système : le changement d'énergie conformationnelle $E_{\Delta C}$ et le changement d'énergie potentielle électrique $-zeV$, où e est la charge élémentaire, z , le nombre de charges déplacées (13.6e pour *Shaker*) et V , le changement de potentiel membranaire. Ainsi, avec $E_O - E_F = E_{\Delta C} + (-zeV)$

$$P(V) = \frac{1}{1 + \exp((E_{\Delta C} - zeV)/kT)} = \frac{1}{1 + \exp(-(V - V_{1/2})/\alpha)}, \quad (1.2.4)$$

où l'on a défini le facteur d'inclinaison $\alpha = kT/ze$ et la constante $V_{1/2}$ telle que $P(V_{1/2}) = 0.5$.

Or, ce que l'on mesure en laboratoire, ce sont des courants ioniques, et non des ratios de populations. Le lien de proportionnalité, toutefois, est direct, puisque la conductance unitaire est la même pour tous les canaux. Ainsi, $P(V)$ correspond à la conductance normalisée $G(V)/G_{max}$ à l'équilibre. On voit donc qu'on peut *prédire* la forme sigmoïde de la conductance dépendante du voltage $G(V)$ identifiée par Hodgkin et Huxley. Mais l'avantage, encore une fois, de penser en terme de populations, est qu'il devient beaucoup plus facile de mesurer le taux transition d'un état fonctionnel à l'autre car, bien que les canaux *unitaires* deviennent conducteurs à des moments plus ou moins aléatoires, l'effet combiné d'un grand nombre de canaux rend possible de définir la notion d'équilibre thermodynamique des populations. En électrophysiologie en voltage contrôlé, il est possible de contrôler le potentiel membranaire d'une cellule surexprimant des clones du canal ionique d'intérêt. On peut ainsi créer divers protocoles pour mesurer le temps que prend les canaux pour passer d'un équilibre à l'autre, en forçant divers types de transitions entre les états. La vitesse d'atteinte du nouvel équilibre, mesurée par les cinétiques des courants ioniques macroscopiques – dont la forme est lisse plutôt que discrète –, sont directement liées aux cinétiques des changements de conformation des canaux.

La figure 1.11 montre deux exemples de protocoles forçant des transitions entre différentes populations. En (a) on maintient la cellule à -100mV assez longtemps (1s) pour que quasiment tous les canaux soient fermés, puis on dépolarise soudainement à différents voltages, provoquant l'ouverture d'une partie d'entre eux. Suite à la dépolarisation, les canaux K_V4 s'activent très rapidement (1.11(c)), permettant de mesurer la probabilité d'ouverture $P(V) = G_p(V)/G_pmax = I_p(V)/I_pmax$ via la valeur du courant au pic d'activation (voir courbe

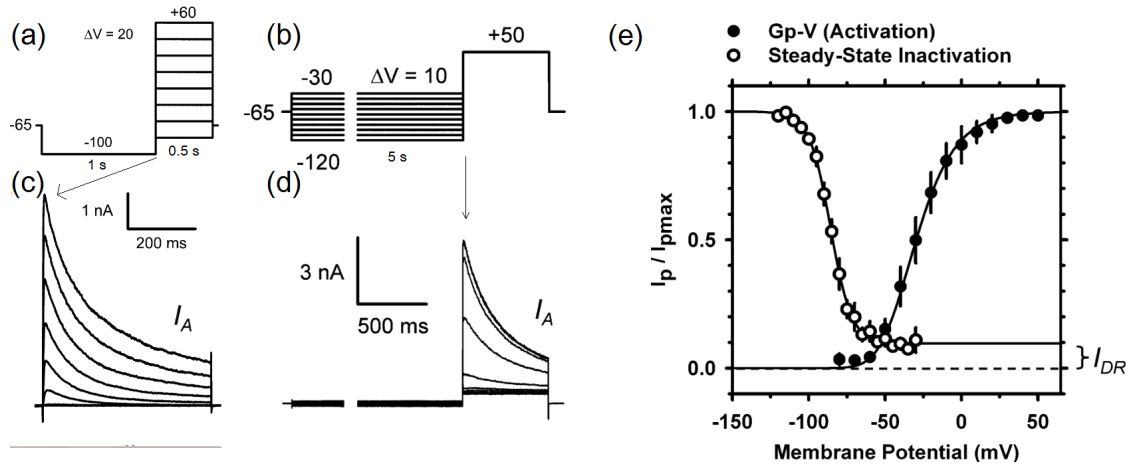


FIGURE 1.11. Voltages et courants pour G(V) et pour l'inactivation à l'équilibre. (a) Protocole de voltage pour G(V). (b) Courants ioniques de type A mesurés dans des ganglions spinaux, suite à la dépolarisation en (a). (c) Protocole de voltage pour l'inactivation à l'équilibre. (d) Courants résultats du protocole en (c). En (e), on voit les courbes G(V) et d'inactivation à l'équilibre résultantes. (Adapté de Phuket et Covarrubias, 2009).

Gp-V fig. 1.11(e)), mais s'inactivent tous également très vite. Le taux auquel le nouvel équilibre est atteint (100% des canaux ouverts s'inactivent) est d'autant plus lent que le potentiel est bas, car il y en a moins à inactiver et la transition (aléatoire) est moins probable. On peut constater ce dernier point, entre autre chose, avec le protocole en (c), où l'on mesure la proportion prête à l'activation après qu'une portion d'entre eux ait été inactivés à l'équilibre (*steady state inactivation*).

1.2.4. Inactivation de K_V4

Les courants potassiques de type A qu'on retrouve médiés par les *Shaker-like* K_V1 à 4 (Birnbaum et al., 2004; Jerng et al., 2004) ont tous la propriété caractéristique de s'activer et s'inactiver rapidement. Cette inactivation rapide (de type N) commune aux quatre isoformes partage plusieurs points communs. En effet, le même modèle de base, appelé le modèle de la balle et de la chaîne, explique les grandes lignes de l'inactivation de type N chez ces canaux. On constate en effet qu'une « balle » faites des 19 premiers acide aminés N-terminaux, reliée au canal par une chaîne, faite des 21 acides aminés suivants, vient s'insérer dans le vestibule hydrophobique interne du pore pour bloquer la conduction ionique (Hoshi et al., 1990; Zagotta et al., 1990; Zhou et al., 2001). Ce blocage du chemin conducteur emprisonne le canal *Shaker* dans un état d'inactivation *ouvert*, ce qu'on peut montrer par les considérables courants de queue (*tail currents*) observés lors d'une repolarisation de la membrane en situation de haute concentration de K^+ externe.

Notons qu'il est possible d'éliminer complètement ce type d'inactivation (type N), en clivant génétiquement ces 40 premiers acides aminés, et de la restaurer en exprimant la chaîne et la balle librement dans le cytoplasme. Cela diffère de l'inactivation de type C, qui est due à une déstabilisation du filtre de sélectivité lors de dépolarisations très prolongées.

Par contre, $K_{V4.2}$ montre des différences avec certaines propriétés classiques de ce modèle. Par exemple, ils ne s'accumulent pas dans un état ouvert-inactivé comme *Shaker* et K_V1 , mais l'occupe seulement très brièvement, pour plutôt s'accumuler dans un état fermé-inactivé (Jerng et al., 1999; Bähring et al., 2001a). Puisque non-conducteur, cet état rend l'inactivation de K_V4 plus difficile à étudier. De plus, la résistance des cinétiques d'inactivation à l'agent bloqueur tetraethylammonium (TEA), qui compétitionne habituellement avec l'inactivation de *Shaker*, et le ralentissement (plutôt que l'accélération) de la récupération de l'inactivation en condition de haute concentration de K^+ extracellulaire, sont deux autres différences qui indiquent que l'inactivation $K_V4.2$ ne repose pas sur les mêmes mécanismes moléculaires que *Shaker* (Jerng and Covarrubias, 1997).

Gebauer et al. (2004) ont investigué ces différences et ont pu montrer que le N-terminal de $K_V4.2$ possède les mêmes propriétés auto-inhibitrices que le modèle classique de la balle et chaîne, mais que les interactions qui prennent place au sein du vestibule sont différentes. Ils ont montré qu'on peut éliminer et restaurer l'inactivation en clivant et en exprimant le peptide dans le cytosol. Mais mieux encore ils ont pu conférer à $K_V2.1$ (fig.1.12) cette inactivation rapide en remplaçant son N-terminal par celui de 4.2, laquelle est normalement beaucoup plus lente en raison de sa résistance aux transitions vers un état inactivé à partir d'un état ouvert conducteur, préférant des transitions à partir des états fermés pré-ouverts.

Par contre, contrairement à *Shaker*, les interactions électrostatiques entre le N-terminal (4 charges positives pour *Shaker*, 1 pour pour 4.2) et le vestibule sont peu importantes pour $K_V4.2$. De plus, les acides aminés 3 et 5 de *Shaker*, identifié par Zhou et al. (2001), jouent un rôle important dans l'affinité entre la balle et le pore, mais une analyse de doubles mutants similaire pour les acides aminés homologues de $K_V4.2$ a révélé que les coefficients d'affinité sont moins importants pour ce dernier, consistant avec la nature transitoire de l'état ouvert-inactivé (Gebauer et al., 2004).

Pour peindre un portrait plus complet des différences moléculaires à l'origine des différences dans l'inactivation, il a fallu attendre l'analyse cinétique et structurale détaillée des états inactivés-fermés de Wollberg et Bähring (2016). Leur recherche était motivée par plusieurs résultats indiquant que la tendance particulière de K_V4 à s'accumuler dans un état *fermé-inactivé* de façon hautement préférentielle était directement reliés aux mécanismes

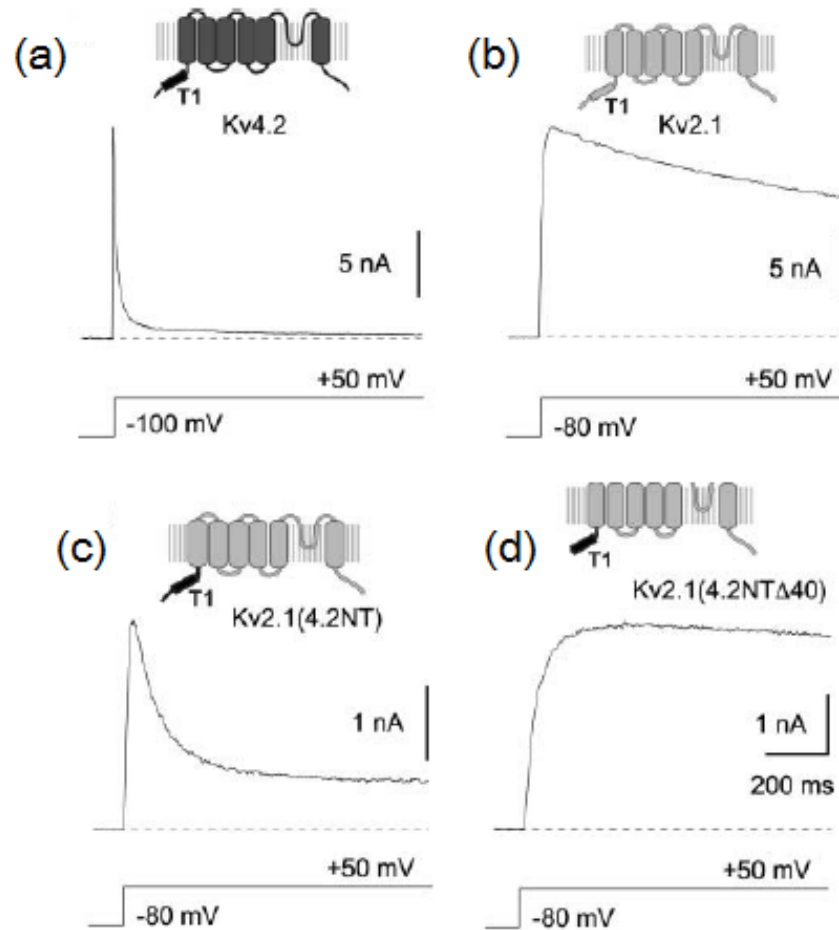


FIGURE 1.12. Comparaison des cinétiques d'inactivation de $K_V4.2$ et $K_V2.1$. En (a), activation et inactivation de $K_V4.2$ *wild type*, exprimé dans HEK 293. Même chose en (b), mais avec $K_V2.1$ wt, qui résiste typiquement à l'inactivation à partir de l'état ouvert, préférant des transitions à partir des états fermés pré-ouverts, ce qui ralentit les cinétiques par rapport à 4.2. En (c), un chimère de $K_V2.1$ avec le N-terminal de 4.2 restaure une partie de la composante rapide de l'inactivation de $K_V4.2$. En (d), l'inactivation de type N disparaît complètement lorsqu'on clive les 40 premiers acides aminés, laissant place à l'inactivation de type C, beaucoup plus lente. (Adapté de Gebauer et al., 2004.)

moléculaires de l'*activation* (Lu, Klem et Ramu, 2002 ; Barghaan et Bähring, 2009 ; Bähring et al., 2012). Avant de résumer leurs travaux, il convient donc de décrire le modèle classique de l'activation pour *Shaker* et de ses différences avec K_V4 .

1.3. COUPLAGE ÉLECTROMÉCANIQUE

Cette section résume le processus par lequel les changements de potentiel membranaire peut changer l'état fonctionnel des canaux K_V . Pour que le canal entre dans un état

conducteur, le *helical bundle crossing* doit s'élargir suite à une dépolarisation. Or, le senseur de voltage étant assez éloigné du bas du pore, un certain couplage, dit électromécanique, doit relier les changements de conformation entre les deux structures. Quels sont les changements conformationnels du VSD induis par les changements de potentiel et comment sont-ils communiqués au pore? La réponse n'est pas évidente, car la structure détaillée n'est forcément qu'un instantané, lequel s'avère capturer les canaux K_V dans leur position ouverte (la différence de potentiel étant évidemment nulle lors d'une cristallisation). Mais même en connaissant la structure dans l'état fermé (plusieurs techniques sont utilisées pour le forcer durant la cristallisation), maints expériences d'électrophysiologie et de fluorométrie en voltage contrôlé, de mutagenèse dirigée, de dynamique moléculaire et d'accessibilité aux agents ont été (et sont encore) nécessaires pour contraindre le modèle *dynamique* exact du passage d'un état à l'autre.

D'abord, ayant vu comment la surface chargée de S4 traverse l'entièreté de la membrane, un premier « défi » à expliquer est le fait que le champ électrique \mathbf{E} dans un environnement aussi peu polarisable que la membrane se communique mal. En effet, si on peut montrer que \mathbf{E} varie linéairement dans une membrane parfaite, il n'est pas évident comment $\mathbf{E}_{\text{local}}$ se comporte près de la protéine. Or, tel qu'introduit à la section précédente, la structure en sablier du VSD forme des cavités étroites et profondes dans la membrane qui sont accessibles au milieu aqueux. Cela diminue l'épaisseur effective de la membrane, ce qui, pour une différence de potentielle membranaire donnée, force le champ électrique à chuter sur une plus petite distance. Conséquemment, autant les résidus chargés de ces crevasses que les résidus chargés enfouis dans le sceau hydrophobique central ressentent un champ local plus prononcé que s'ils étaient enfouis dans une membrane « planaire » (Islas et Sigworth, 2001 ; Asamoah et al., 2003). Cela contribue bien sûr à une sensibilité accrue aux variations du potentiel membranaire. La figure 1.13 illustre comment ces cavités ont l'effet de concentrer $\mathbf{E}_{\text{local}}$ autour du VSD.

Donc, suite à une dépolarisation de la membrane, le senseur de voltage subit un changement de conformation durant lequel S4 passe par plusieurs états fermés avant d'atteindre son état activé (Perozo et al., 1994 ; Zagotta et al., 1994b, Mannuzzu et al., 1996 ; Cha et Bezanilla, 1997). Poussées par l'inversion de polarité de la membrane, les charges de la partie inférieure de S4 brisent leurs ponts salins et, suite à une combinaison de mouvements, vont en former de nouveaux avec les charges négatives de la partie supérieure du VSD. Cette réorganisation des charges de S4 nécessite que l'hélice effectue (1) une inclinaison par rapport à la normale, (2) une rotation sur son axe et (3) de petites translations verticales et radiales (Blunck et Batulan, 2012). Ces mouvement de charges rapides produisent des courant électriques transitoires mesurables, avant même l'ouverture

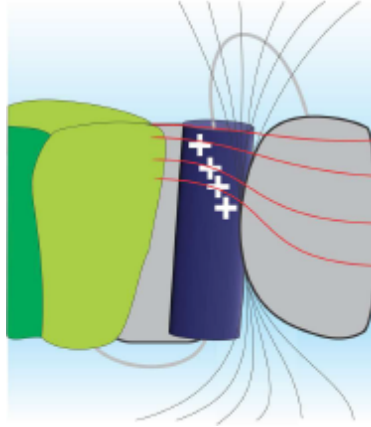


FIGURE 1.13. Comportement local du champ électrique près des VSD. Les lignes noires sont les lignes de champ de \mathbf{E} et les lignes rouges sont les équipotentielles de V . Adapté de (Blunck et Batulan, 2012).

complète du canal, qu'on appelle les courant de *gating* (à distinguer des courants *ioniques* qui suivent l'activation).

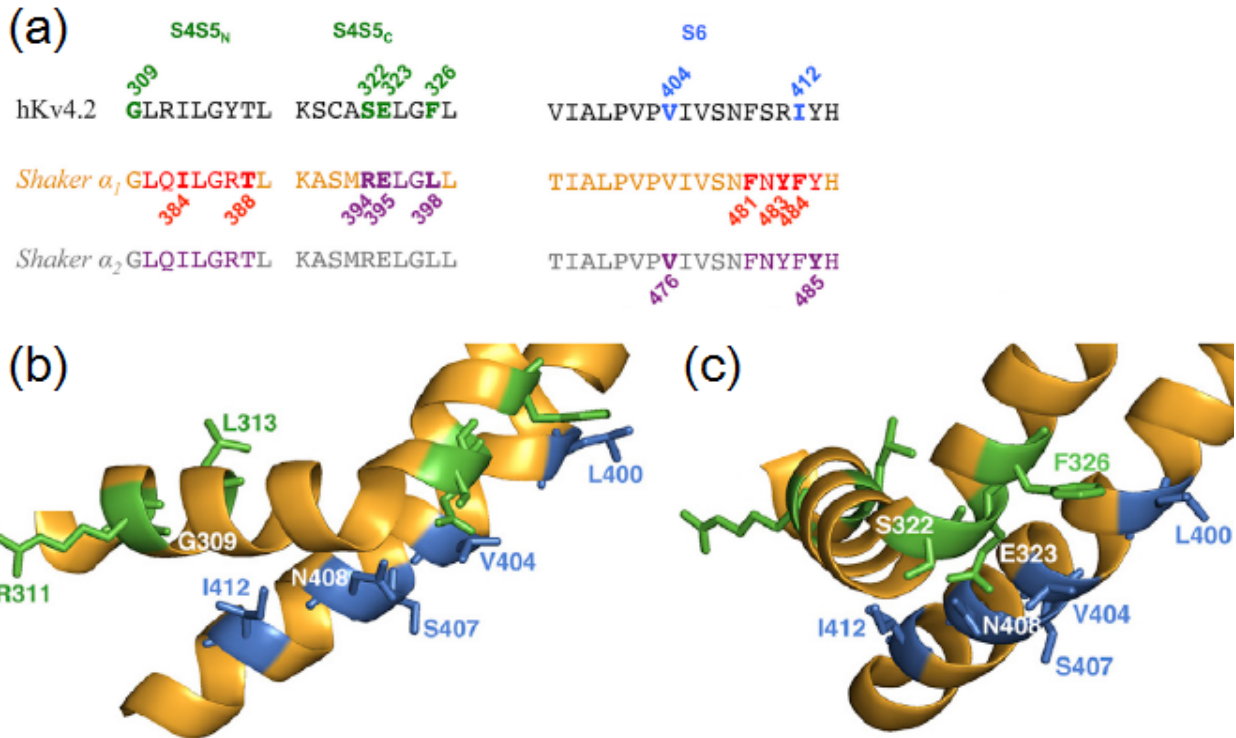
Ces changements conformationnels se produisent indépendamment d'un VSD à l'autre. Mais pour provoquer l'ouverture finale du chemin conducteur, il faut que les quatre VSD soit dans leur état activés simultanément. L'activation du canal se fait donc en deux grandes étapes. Il faut d'abord que les quatre VSD s'activent spontanément – et restent dans cet état activé en attendant les autres – pour ensuite permettre une étape coopérative où les quatre sous-unités collaborent pour élargir le diamètre de l'*helical bundle crossing* bloquant l'accès au pore (Zagotta et al., 1994a; Pathaketal., 2005). Conséquemment, à cause de ce délai, il ne peut pas exister de couplage *direct* entre les transitions de S4 et l'élargissement du *bundle*. Autrement dit, l'énergie mécanique de S4 doit en quelque sorte pouvoir restée emmagasinée jusqu'à l'étape coopérative, où elle sera utilisée pour ouvrir le pore.

Ce transfert d'énergie s'effectue par l'entremise du S4-S5L, le lien covalent liant S4 à S5. Or, étant donné que ce sont les S6 qui forment la partie interne du pore et qui bloquent le passage des ions dans son état fermé, il n'est pas a priori évident de voir comment un mouvement vertical de S4 peut affecter le bas de l'hélice *interne* via une liaison covalente avec l'hélice *externe*. Pour adresser cette question, plusieurs études fonctionnelles sur des mutants ont permis d'identifier une région de contact entre S4-S5L et le bas de S6 (S6_T) essentielle au couplage électromécanique. Cette région est d'ailleurs hautement conservée dans la famille Shaker (K_V1), alors que quelques modifications apparaissent pour les familles K_V2-6.

Pour *Shaker*, il s'agit des séquences 481-FNYFYH-486 de S6_T et 382-LGRTLKAS-392 du S4-S5L (Ding et Horn, 2002, 2003; Hackos et al., 2002; Soler-Llavina et al., 2006; Labro et al., 2008; Batulan et al., 2010; Haddad et Blunck, 2011). Selon Labro et al. (2008) et Haddad et Blunck (2011), les acides aminés F481, Y483 et F484 de S6_T joueraient un rôle primordial en créant un pochette hydrophobique dans laquelle viennent se loger les acides aminés I312 et T316 du S4-S5L de la même sous-unité (fig. 1.4).

Par contre, cette pochette d'interaction n'est pas très bien conservée chez K_V4.2. Entre autre, la phénylalanine F484 de *Shaker* devient l'isoleucine I412 et, plus drastique encore, la tyrosine Y483 devient l'arginine R411. Ces variations sont très probablement à l'origine d'une affinité différente et, donc, d'un couplage différent. Or, Wollberg et Bähring (2016) ont tout de même pu démontrer que I412 (S6_T de K_V4.2) jouait un rôle crucial dans les états inactivés-fermés, surtout par son interaction avec la glycine G309 de la partie N-terminale du S4-S5L, plutôt que l'isoleucine et la thréonine conservés entre les isoformes. En fait, la mutation en alanine de ces acides aminés, dont les homologues sont cruciaux pour le couplage électromécanique de *Shaker*, ont montré des effets négligeables sur les états inactivés-fermés (CSI) de K_V4.2. La même chose est vraie pour le R411, homologue de Y483 : important pour le couplage de *Shaker*, peu d'effet sur les CSI de 4.2 (Wollberg et Bähring, 2016).

Une seconde paire d'interaction essentielle au couplage électromécanique de *Shaker* prend lieu à l'autre extrémité du S4-S5L (i.e. C-terminale). Cette fois, il s'agit par contre d'interactions inter-sous-unité entre la tyrosine Y485 de S6_T et le glutamate E395 et la leucine L399 d'un S4-S5L d'une sous-unité voisine (fig. 1.4) (Lu et al., 2002, Batulan et al., 2010). Pour K_V4.2, on a encore des similarités et des différences. Entre autre, l'homologue au glutamate 395 du S4-S5L de *Shaker*, E323, reste très important pour les CSI de 4.2 via son couplage au S6, mais les interactions importantes pour les CSI s'effectuent au niveau de résidus non conservés chez *Shaker* (Wollberg et Bähring, 2016). Par exemple, la tyrosine 485 du S6 de *Shaker* est conservée (Y413 de K_V4.2), mais n'est pas impliquée dans les CSI (Wollberg et Bähring, 2016). C'est plutôt la valine V404 du S6 qui le sera, comme le montre le fort couplage à plusieurs acides aminés du C-terminal de S4-S5L, dont le glutamate conservé E323, de même que S322 et F326, non-conservés. Cette valine est conservée chez *Shaker* (= V476), où elle joue le rôle stabilisateur du E395 du motif RELY (Batulan, Haddad et Blunck, 2010). Avec R394 et Y485, ces trois résidus sont d'ailleurs appelés la « triade d'activation » (*gating triad*) pour leur rôle dans l'intégrité structurale de l'interface inter-sous-unitaire impliquée dans l'ouverture (Chowdhury, Haegnel et Chanda, 2014).



le S4-S5L (rouge) et le S6_T de deux sous-unités adjacentes sont illustrés, montrant à droite les chaînes latérales de F409 et I412 (S6), qui créent une pochette hydrophobique dans laquelle s'emboîtent les acides aminés I312 et T316 du S4-S5L de la même sous-unité (voir représentation isosurfacique à gauche). En (b), le bas du pore est structuré de façon telle que les S4-S5L entourent S6_T, auxquels ils se lient par un ensemble d'interactions hydrophobiques. À droite de la figure, l'accent est mis sur les interactions inter-sous-unités entre E395 (S4-S5L, magenta) et Y413 (S6, cyan), également présentes sous sa représentation isosurfacique à gauche. Les interactions mis en évidence en (a) n'ont été illustrées qu'une fois, en haut à gauche (rouge et bleu), pour alléger la figure tout en montrant leurs positions relatives. Produit à partir du modèle d'homologie de K_V4.2 de (Heler, Bell et Boland, 2013).

Notons au passage que ce genre d'interactions inter-sous-unités devrait en principe expliquer l'existence de l'étape coopérative, mais il s'avère que les mutations de *Shaker* dans cette région affecte surtout la désactivation, impliquant que ces résidus ne jouent pas de rôle significatif lors de l'activation (Blunck et Batulan, 2012). Par contre, ce type d'interactions *dynamiques* (i.e. contacts seulement dans certaines configurations) entre le S4-S5L et S6 sont cruciales pour l'inactivation particulière via des états fermés (*closed state inactivation*, ou CSI) que présente K_V4 (Barghaan et Bähring, 2009 ; Wollberg et Bähring, 2016). Suivant plusieurs modèles pour les CSI et à partir de leurs résultats, Wollberg et Bähring (2016) suggèrent que la triade d'activation de *Shaker* n'ait pas évolué pour assurer la stabilité de

l'interface inter-sous-unitaire de $K_V4.2$, mais plutôt pour assurer une interaction dynamique en plusieurs sous-étape avant l'activation finale. Donc, le changement conformationnel du VSD durant l'activation entraînerait un découplage temporaire du pore, laissant les canaux dans état fermé et inactivé.

Pour terminer, mentionnons que quelques autres régions de contact ont été proposées comme modulatrices du couplage électromécanique, à différent degré, notamment entre S4 et S5, mais les deux régions décrites ci-haut présentent clairement les effets les plus prononcés (Blunck et Batulan, 2012).

1.4. LES SOUS-UNITÉS BÊTA DE K_V4 ET LES COURANTS I_{SA} ET $I_{to,f}$

La diversité des canaux potassiques et des courants physiologiques auxquels ils participent provient non seulement de la grande quantité d'isoformes α formeuses de pore que l'on retrouve dans la nature, mais aussi de leur association avec différentes sous-unités régulatrices β . Deux courants d'une importance physiologique particulière, le courant transitoire sous-seuil de type A (*subthreshold A-type potassium current*, I_{SA}), retrouvé dans le système nerveux central, et le courant transitoire rapide vers l'extérieur (*fast transient outward current*, $I_{to,f}$), retrouvé les cellules cardiaques, sont des exemples de l'activité de tels complexes protéiques impliquant K_V4 (Baldwin et al., 1991; Johns et al., 1997; Malin and Nerbonne, 2000; Liss et al., 2001; Serodio et al., 1994, 1996; Tkatch et al., 2000; Tsunoda and Salkoff, 1995). Les deux courants s'activent à des potentiels dépolarisants inférieurs à ceux où la plupart des courants potassiques sont activés (*shift* de $G(V)$ vers les potentiels hyperpolarisants) et s'inactive plus rapidement (Connor et Stevens, 1970). Ces propriétés sont spécialement adaptées pour remplir un rôle modulateur. On retrouve I_{SA} impliqué dans la régulation du taux d'activité des neurones rythmiques, dans la suppression des potentiels excitateurs et des *backfiring* des potentiels d'action, dans la plasticité cérébrale, l'émission de neurotransmetteurs (Connor et Stevens, 1971a; Pongs, 1999; Schoppa et Westbrook, 1999; Jerng et al., 2004; Kim, Wei et Hoffman, 2005; Shah, Hammond et Hoffman, 2010). Jerng et Pfaffinger (2014) suggèrent que leur capacité à compartimentaliser l'excitabilité et le stockage d'information au niveau des dendrites ont pu jouer un rôle critique dans l'évolution des systèmes nerveux. Quant à $I_{to,f}$, il module entre autre le profil des potentiels d'action cardiaques, la contractibilité et l'hypertrophie des cardiomyocytes (Sah et al., 2003; Kassiri et al., 2002).

L'identification de l'origine moléculaire de I_{SA} a posé problème car, bien qu'il ait été démontré que K_V4 était impliqué (Serodio et al., 1994), ses cinétiques d'activation et d'inactivation sont trop lentes pour rendre compte de I_{SA} par lui-même. Ce constat a motivé la recherche de protéines auxiliaires modulatrices. Les études de co-immunoprécipitations

ont permis d'identifier deux familles de sous-unité β avec lesquelles K_V4 forment des complexes protéiques in vivo : les *K_V channels interacting proteins*, ou KChIP, et les *dipeptidyl-aminopeptidase-like proteins* (DPLP ou DPP) (Jerng et al., 2004; Birnbaum et al., 2004).

Les premières sous-unités auxiliaires trouvées, les KChIP 1 à 4, sont de petites protéines globulaires sensibles au Ca^{2+} intracellulaire. Par leur interaction avec le domaine de tétramérisation de K_V4 , elles augmentent l'expression surfacique de K_V4 , ralentissent l'inactivation et accélèrent la récupération après l'inactivation (fig. 1.5) (An et al., 2000; Holmqvist et al., 2002; Morohashi et al., 2002). Or, bien que partie intégrante du complexe à l'origine de I_{SA} (Rhodes et al., 2004), la découverte du ralentissement des cinétiques de K_V4 par les KChIP garantissait qu'elles n'étaient pas l'unique médiateur et que le facteur accélérant restait à identifier.

le S4-S5L (rouge) et le S6_T de deux sous-unités adjacentes sont illustrés, montrant à droite les chaînes latérales de F409 et I412 (S6), qui créent une pochette hydrophobique dans laquelle s'emboîtent les acides aminés I312 et T316 du S4-S5L de la même sous-unité (voir représentation isosurfacique à gauche). En (b), le bas du pore est structuré de façon telle que les S4-S5L entourent S6_T, auxquels ils se lient par un ensemble d'interactions hydrophobiques. À droite de la figure, l'accent est mis sur les interactions inter-sous-unités entre E395 (S4-S5L, magenta) et Y413 (S6, cyan), également présentes sous sa représentation isosurfacique à gauche. Les interactions mis en évidence en (a) n'ont été illustrées qu'une fois, en haut à gauche (rouge et bleu), pour alléger la figure tout en montrant leurs positions relatives. Produit à partir du modèle d'homologie de $K_V4.2$ de (Heler, Bell et Boland, 2013).

Une seconde paire d'interaction essentielle au couplage électromécanique de *Shaker* prend lieu à l'autre extrémité du S4-S5L (i.e. C-terminale). Cette fois, il s'agit par contre d'interactions inter-sous-unité entre la tyrosine Y485 de S6_T et le glutamate E395 et la leucine L399 d'un S4-S5L d'une sous-unité voisine (fig. 1.4) (Lu et al., 2002, Batulan et al., 2010). Pour $K_V4.2$, on a encore des similarités et des différences. Entre autre, l'homologue au glutamate 395 du S4-S5L de *Shaker*, E323, reste très important pour les CSI de 4.2 via son couplage au S6, mais les interactions importantes pour les CSI s'effectuent au niveau de résidus non conservés chez *Shaker* (Wollberg et Bähring, 2016). Par exemple, la tyrosine 485 du S6 de *Shaker* est conservée (Y413 de $K_V4.2$), mais n'est pas impliquée dans les CSI (Wollberg et Bähring, 2016). C'est plutôt la valine V404 du S6 qui le sera, comme le montre le fort couplage à plusieurs acides aminés du C-terminal de S4-S5L, dont le glutamate conservé E323, de même que S322 et F326, non-conservés. Cette valine est conservée chez *Shaker* (= V476), où elle joue le rôle stabilisateur du E395 du motif RELY (Batulan,

Haddad et Blunck, 2010). Avec R394 et Y485, ces trois résidus sont d’ailleurs appelés la « triade d’activation » (*gating triad*) pour leur rôle dans l’intégrité structurale de l’interface inter-sous-unitaire impliquée dans l’ouverture (Chowdhury, Haegnel et Chanda, 2014).

Notons au passage que ce genre d’interactions inter-sous-unités devrait en principe expliquer l’existence de l’étape coopérative, mais il s’avère que les mutations de *Shaker* dans cette région affecte surtout la désactivation, impliquant que ces résidus ne jouent pas de rôle significatif lors de l’activation (Blunck et Batulan, 2012). Par contre, ce type d’interactions *dynamiques* (i.e. contacts seulement dans certaines configurations) entre le S4-S5L et S6 sont cruciales pour l’inactivation particulière via des états fermés (*closed state inactivation*, ou CSI) que présente K_V4 (Barghaan et Bähring, 2009 ; Wollberg et Bähring, 2016). Suivant plusieurs modèles pour les CSI et à partir de leurs résultats, Wollberg et Bähring (2016) suggèrent que la triade d’activation de *Shaker* n’ait pas évolué pour assurer la stabilité de l’interface inter-sous-unitaire de $K_V4.2$, mais plutôt pour assurer une interaction dynamique en plusieurs sous-étape avant l’activation finale. Donc, le changement conformationnel du VSD durant l’activation entraînerait un découplage temporaire du pore, laissant les canaux dans état fermé et inactivé.

Pour terminer, mentionnons que quelques autres régions de contact ont été proposées comme modulatrices du couplage électromécanique, à différent degré, notamment entre S4 et S5, mais les deux régions décrites ci-haut présentent clairement les effets les plus prononcés (Blunck et Batulan, 2012).

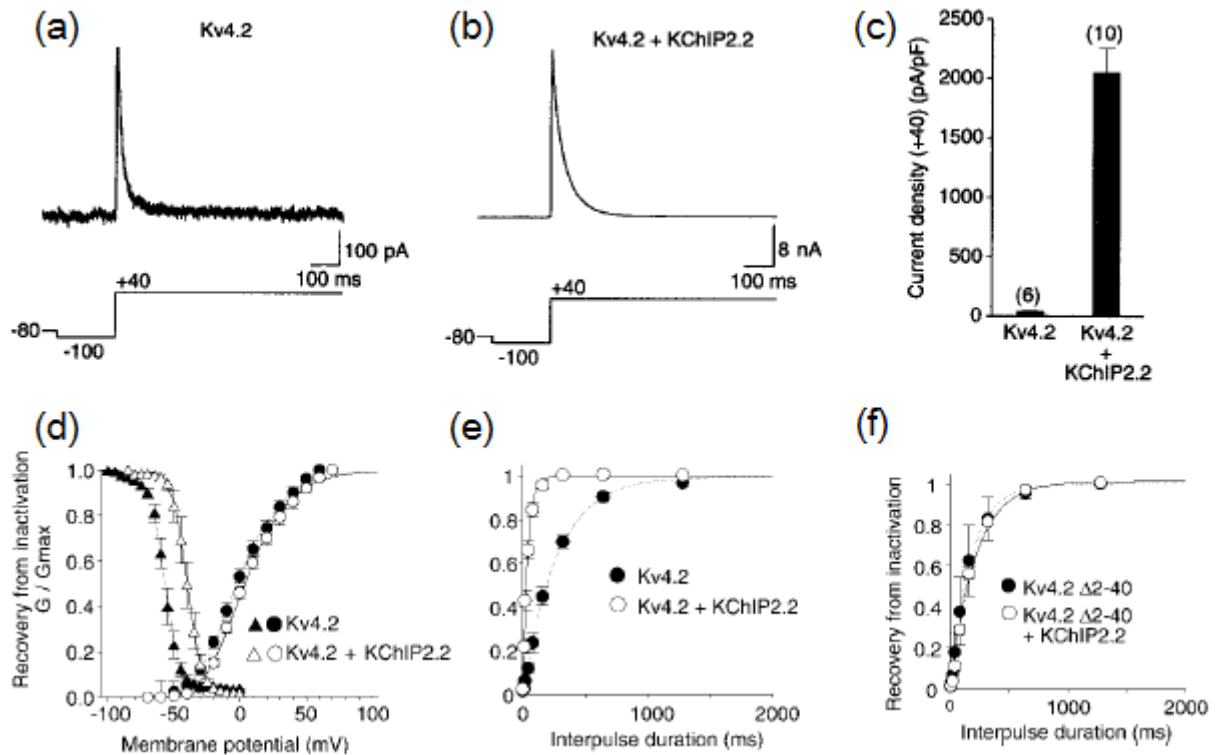
1.5. LES SOUS-UNITÉS BÊTA DE K_V4 ET LES COURANTS I_{SA} ET $I_{to,f}$

La diversité des canaux potassiques et des courants physiologiques auxquels ils participent provient non seulement de la grande quantité d’isoformes α formeuses de pore que l’on retrouve dans la nature, mais aussi de leur association avec différentes sous-unités régulatrices β . Deux courants d’une importance physiologique particulière, le courant transitoire sous-seuil de type A (*subthreshold A-type potassium current*, I_{SA}), retrouvé dans le système nerveux central, et le courant transitoire rapide vers l’extérieur (*fast transient outward current*, $I_{to,f}$), retrouvé les cellules cardiaques, sont des exemples de l’activité de tels complexes protéiques impliquant K_V4 (Baldwin et al., 1991 ; Johns et al., 1997 ; Malin and Nerbonne, 2000 ; Liss et al., 2001 ; Serodio et al., 1994, 1996 ; Tkatch et al., 2000 ; Tsunoda and Salkoff, 1995). Les deux courants s’activent à des potentiels dépolarisants inférieurs à ceux où la plupart des courants potassiques sont activés (*shift* de $G(V)$ vers les potentiels hyperpolarisants) et s’inactive plus rapidement (Connor et Stevens, 1970). Ces propriétés sont spécialement adaptées pour remplir un rôle modulateur. On retrouve I_{SA} impliqué dans la régulation du taux d’activité des neurones rythmiques, dans la suppression des potentiels

excitateurs et des *backfiring* des potentiels d'action, dans la plasticité cérébrale, l'émission de neurotransmetteurs (Connor et Stevens, 1971a; Pongs, 1999; Schoppa et Westbrook, 1999; Jerng et al., 2004; Kim, Wei et Hoffman, 2005; Shah, Hammond et Hoffman, 2010). Jerng et Pfaffinger (2014) suggèrent que leur capacité à compartimentaliser l'excitabilité et le stockage d'information au niveau des dendrites ont pu jouer un rôle critique dans l'évolution des systèmes nerveux. Quant à $I_{to,f}$, il module entre autre le profil des potentiels d'action cardiaques, la contractibilité et l'hypertrophie des cardiomyocytes (Sah et al., 2003; Kassiri et al., 2002).

L'identification de l'origine moléculaire de I_{SA} a posé problème car, bien qu'il ait été démontré que K_V4 était impliqué (Serodio et al., 1994), ses cinétiques d'activation et d'inactivation sont trop lentes pour rendre compte de I_{SA} par lui-même. Ce constat a motivé la recherche de protéines auxiliaires modulatrices. Les études de co-immunoprécipitations ont permis d'identifier deux familles de sous-unité β avec lesquelles K_V4 forment des complexes protéiques in vivo : les *K_V channels interacting proteins*, ou KChIP, et les *dipeptidyl-aminopeptidase-like proteins* (DPLP ou DPP) (Jerng et al., 2004; Birnbaum et al., 2004).

Les premières sous-unités auxiliaires trouvées, les KChIP 1 à 4, sont de petites protéines globulaires sensibles au Ca^{2+} intracellulaire. Par leur interaction avec le domaine de tétramérisation de K_V4 , elles augmentent l'expression surfacique de K_V4 , ralentissent l'inactivation et accélèrent la récupération après l'inactivation (fig. 1.5) (An et al., 2000; Holmqvist et al., 2002; Morohashi et al., 2002). Or, bien que partie intégrante du complexe à l'origine de I_{SA} (Rhodes et al., 2004), la découverte du ralentissement des cinétiques de K_V4 par les KChIP garantissait qu'elles n'étaient pas l'unique médiateur et que le facteur accélérant restait à identifier.



[Modulation de Kv4.2 par KChiP.] Modulation de $K_V4.2$ par KChiP. (a) Courant ionique de type A typique des canaux $K_V4.2$, exprimés seuls dans des cellules CHO. (b) Même courant, mais avec $K_V4.2$ et KChiP2.2 co-exprimés. (c) Densité de courant de $K_V4.2$ comparée à $K_V4.2 + KChiP2.2$. (d) Courbes $G(V)$ et d'inactivation à l'équilibre avec et sans KChiP2.2. (e) Récupération de l'inactivation avec et sans KChiP2.2. (f) Récupération de l'inactivation sans le N-terminal. (Adapté de Bähring et al., 2001.)

Ce sont Nadal et al. (2003) qui ont identifié une protéine de 115kDa, nommée DPP6 (ou DPPX), qui co-immunoprécipite avec $K_V4.2$ dans des neurones montrant des courants I_{SA} . Avec DPP10 (ou DPPY), elle forme la famille des DPP, dont les seules propriétés alors connues étaient sa ressemblance structurelle avec CD26 ($\approx 50\%$ de similarité), une dipeptidyl-aminopeptidase jouant un rôle dans l'adhésion cellulaire. Or, contrairement à CD26, DPP6 n'a pas d'action enzymatique. Ses fonctions restaient inconnues jusqu'en 2003, avant que Nadal et al. reconstituent le complexe $K_V4.2$ -KChiP1-DPP6 dans des ovocytes de *Xenopus laevis* et montrent que DPP6 s'agit bien du facteur accélérant de K_V4 nécessaire pour produire I_{SA} .

La figure 1.15 montre comment DPP6 module les cinétiques autant de l'activation que de l'inactivation. On voit que non seulement la densité de courant est supérieure, mais que le pic du courant est atteint environ trois fois plus vite (c) et que l'activation se fait

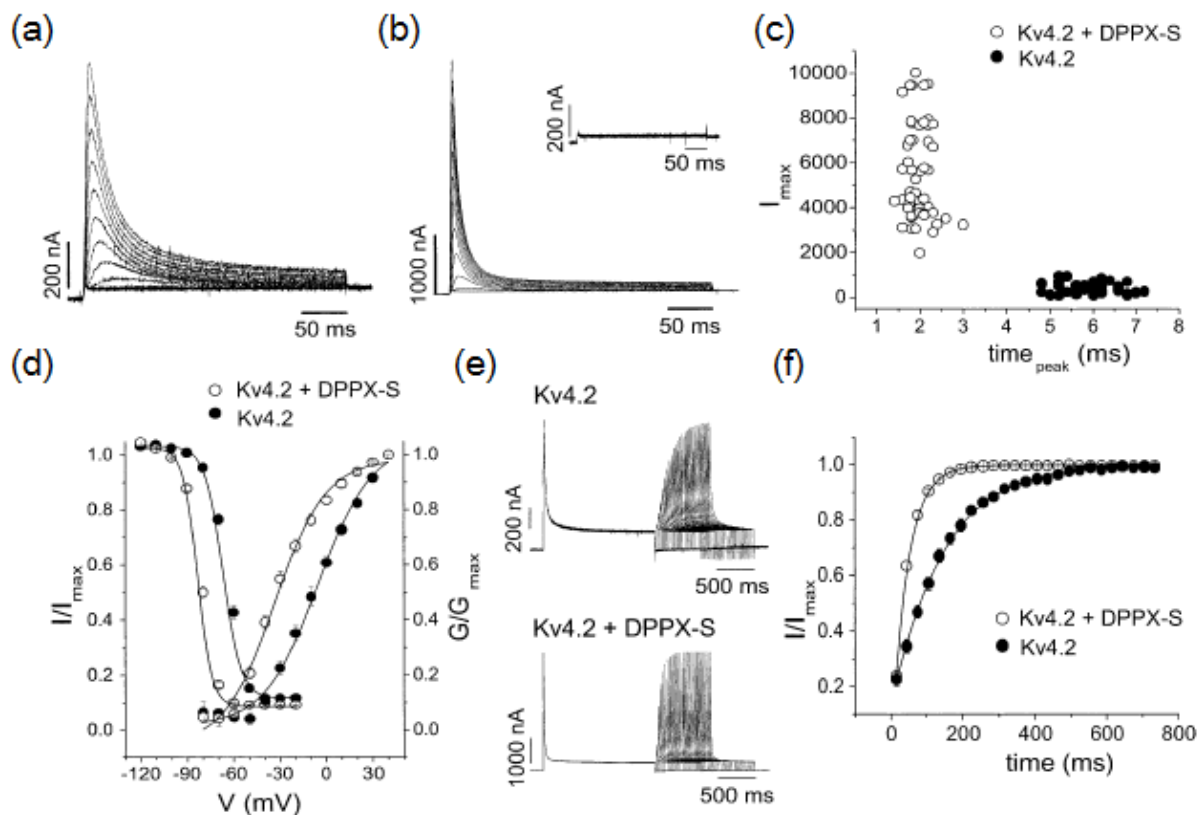


FIGURE 1.15. Modulation de $K_V4.2$ par DPP6. (a) Courants ioniques de type A typiques des canaux $K_V4.2$, exprimés seuls dans des cellules CHO. (b) Mêmes courants, mais avec $K_V4.2$ et DPP6 co-exprimés. La trace du haut est DPP6 injecté seul. (c) Densité de courant de $K_V4.2$ comparée à $K_V4.2$ + KCHIP2.2. (d) Courbes $G(V)$ et d'inactivation à l'équilibre avec et sans DPP6.2. (e) Récupération de l'inactivation avec et sans KCHIP2.2. (f) Récupération de l'inactivation sans le N-terminal. (Adapté de Nadal et al., 2003.)

à des potentiels moins dépolarisants suite à la co-expression avec DPP6 (d). Quant à l'inactivation après l'activation, elle se produit plus rapidement avec DPP6 (a et b). De plus, DPP6 déplace la courbe de l'inactivation à l'équilibre vers des potentiels hyperpolarisants, indiquant qu'à un potentiel donné, il y a moins de canaux inactivés à l'équilibre avec DPP6. Cela est consistant avec le fait que DPP6 accélère la récupération à l'inactivation, comme le montre (e) et (f).

Les mécanismes à l'origine de ces modulations restent pour l'instant inconnus. D'ailleurs, la structure cristalline de DPP6 n'est connue que partiellement. Seul son domaine C-terminal

extracellulaire, constituant la majeure partie de la protéine, a été cristallisé (Strop, Bankovich, Hansen, Garcia et Brunger, 2004). Pour donner une idée de sa taille relative, l'isoforme 6.2 est l'un des trois plus courts, avec 803 acides aminés, et son C-terminal est composé de 739 de ceux-ci. Pour référence, les deux plus longs (DPP6.1 et 6.5) ont une soixantaine d'acides aminés supplémentaires à l'extrémité N-terminal. On connaît relativement peu sur toutes les fonctions accomplies par ce large domaine, mais l'une des plus importantes est son rôle dans le trafic du réticulum endoplasmique vers la membrane. On sait entre autre qu'il est nécessaire pour le trafic de DPP6 vers la membrane et que certains ponts disulfures (C349/C356 et C465/C468) sont cruciaux pour l'augmentation de l'expression surfacique de Kv4.2 (Lin, Long, Hatch et Hoffman, 2014).

Par contre, pour médier l'interaction fonctionnelle avec Kv4 et accélérer ses cinétiques, ces ponts dissulfiques sont inconséquents (Lin, Long, Hatch et Hoffman, 2014). En fait, les propriétés accélératrices de DPP6 ont été pleinement reproduites en l'absence complète du domaine extracellulaire (Lin, Long, Hatch et Hoffman, 2014), indiquant clairement que le mécanisme moléculaire sous-jacent se trouvera au niveau du segment transmembranaire et de l'extrémité N-terminale, mais les détails restent à identifier. Cette connaissance largement incomplète est ce qui a justifiée la présente étude.

Ainsi, les objectifs de cette recherche se décline en trois temps. D'abord, il sera nécessaire de construire des modèles atomiques pour les sous-structures n'ayant pas été cristallisées à ce jour, soit les parties transmembranaire et N-terminale de DPP6.2 *et* Kv4.2. Ensuite, les régions pertinentes (TM-TM et N-terminal-N-terminal) des deux peptides seront soumises à des algorithmes d'amarrage moléculaire pour identifier les sites susceptibles d'interagir. Finalement, on utilisera la dynamique moléculaire pour tenir compte de la flexibilité des ligands et discriminer davantage les différents modèles d'amarrage obtenus sur la base de leur stabilité structurelle relative et la force des interactions inter-moléculaires. Cette approche permettra, on l'espère, de prédire le mode d'interaction le plus probable et de comparer comment les différentes sous-structures fonctionnelle de Kv4.2 se comportent en présence et en l'absence de DPP6.2.

Nous espérons vérifier plusieurs hypothèses. D'abord, nous pensons que les sites actifs dans région d'interaction N-terminale seront les mêmes que ceux identifiés dans la structure cristalline du complexe Kv4.3-KChIP à cause de la compétition fonctionnelle observée en électrophysiologie. Ensuite, nous voulons vérifier une hypothèse de Jerng et Pfaffinger (2014) à l'effet que DPP6 produirait une constriction du pore et que cela serait en partie responsable de la modulation des cinétiques observées. Nous calculerons donc les changements conformationnels prenant place lors de la liaison pour étudier comment la structure est

influencée, ce qui éclairera sur les mécanismes par lesquels l'interaction fonctionnelle prend place.

Chapitre 2

MÉTHODOLOGIE

En général, la capacité de répondre rigoureusement à des questions sur le fonctionnement de la nature repose de façon cruciale sur les méthodes choisies. Cela demeure particulièrement vrai en biophysique numérique, entre autre de par la complexité inhérente aux systèmes étudiés. Malgré le progrès technologique effréné des dernières décennies et malgré l'impressionnant raffinement des modèles bioinformatiques qui s'en est suivi, le principal facteur limitant de l'approche numérique restera toujours la limite des ressources calculatoires. Ces contraintes s'opposent nécessairement au désir des chercheurs d'obtenir des prédictions réalistes, physiologiquement pertinentes et non le moindre, avant leur prochaine demande de subvention. Entre autre, il n'est pas encore possible de procéder à un traitement purement quantique des systèmes d'ordre de grandeur supérieur à une centaine d'atomes. Heureusement, par contre, que la dynamique moléculaire classique (DM) s'avère particulièrement féconde pour investiguer une large gamme de phénomènes biologiques d'un point de vue strictement physique et ce, à un coût beaucoup moindre.

Dans cette première approximation de la réalité, les atomes sont vus comme des charges (partielles) ponctuelles¹ et les liaisons covalentes, comme des ressorts. Cela ne permet pas de capturer le bris de liaisons covalentes ou l'ionisation de molécules comme le permettrait l'approche plus précise de la mécanique quantique, mais les interactions dominant largement la physique des protéines – i.e. les interactions électrostatiques, hydrophobiques et de Van der Waals – sont suffisamment bien modélisées par la DM pour en faire l'approche à succès qu'elle est aujourd'hui.

2.1. LA DYNAMIQUE MOLÉCULAIRE CLASSIQUE

2.1.1. Principes généraux

Les lois de la physique, qu'elles se présentent sous la forme d'équations différentielles ou de principes variationnels, permettent de prédire l'évolution du système qu'elles gouvernent à partir des conditions aux frontières. En mécanique newtonienne, pour un système de N particules de masse m_i , de positions \mathbf{r}_i et soumises à un potentiel V , ce système d'équations différentielles correspondent à la deuxième loi de Newton :

$$m_i \frac{d^2 \mathbf{r}_i}{dt^2} = -\nabla_i V(\mathbf{r}_i(t)) \quad (2.1.1)$$

1. Cette approche n'est pas toujours la plus avantageuse. Le traitement « tout-atome » de certaines molécules s'avère trop raffiné pour résoudre certains problèmes. À cet effet, les champs de force « gros-grain » permettent d'améliorer la rapidité calculatoire en réduisant la résolution des modèles. Par exemple, ils ont initialement été développé pour les phospholipides : plutôt que traiter plus d'une centaine d'atomes individuellement, on peut les regrouper en unités fonctionnelles selon qu'ils font partie de la chaîne hydrophobique ou de la tête polaire. Certains champs de force traitent également le solvant aqueux de façon implicite, encore pour sauver du temps. Ces méthodes ne sont pas utilisées ici et ne sont mentionnées qu'à titre comparatif.

Or, comme le potentiel $V(\mathbf{r}_i(t))$ à la position de la particule i dépend de la position des $N - 1$ autres particules, ce système d'équations du deuxième ordre est couplé, ce qui rend impossible, dans le cas général, sa résolution analytique dès que $N \geq 3$. Pour étudier la physique des systèmes biologiques à partir des interactions atomiques individuelles, la nécessité d'utiliser une approche numérique devient donc incontournable. L'algorithme de Verlet est à cette fin largement utilisé. Source : <https://www.ks.uiuc.edu/Research/namd/2.11/ug/node5.html> en DM pour l'intégration de 2.1.1. L'idée consiste à développer la solution en série de Taylor autour de t :

$$\mathbf{r}_i(t \pm \Delta t) = \mathbf{r}_i(t) \pm \mathbf{r}'_i(t)\Delta t + \frac{\mathbf{r}''_i(t)\Delta t^2}{2} \pm \frac{\mathbf{r}'''_i(t)\Delta t^3}{6} + \mathcal{O}(\Delta t^4). \quad (2.1.2)$$

On trouve donc que

$$\mathbf{r}_i(t + \Delta t) + \mathbf{r}_i(t - \Delta t) = 2\mathbf{r}_i(t) - \mathbf{r}''_i(t)\Delta t^2 + \mathcal{O}(\Delta t^4), \quad (2.1.3)$$

ou encore, avec $\mathbf{a}_i(t) := \mathbf{r}''_i(t)$ et en réarrangeant les termes,

$$\mathbf{r}_i(t + \Delta t) = 2\mathbf{r}_i(t) - \mathbf{a}_i(t)\Delta t^2 + \mathbf{r}_i(t + \Delta t) + \mathcal{O}(\Delta t^4). \quad (2.1.4)$$

C'est cette dernière équation qui permet de faire progresser de $\Delta t > 0$ la trajectoire $\mathbf{r}_i(t)$ de chaque particule à partir des positions précédentes et de l'accélération instantanée. Cette dernière est quant à elle déterminée grâce à la deuxième loi de Newton. Numériquement, cette étape consiste en principe à sommer la contribution des $N - 1$ autres particules au potentiel V et d'en calculer le gradient à la position voulue. Or comme il y a $N(N - 1)/2$ interactions possibles et que N peut facilement dépasser le cap du million, on comprend en quoi il s'agit de l'étape de loin la plus énergivore. C'est également la plus cruciale : une modélisation minutieuse des différents potentiels à l'oeuvre au niveau moléculaire est capitale au réalisme de l'approche et, donc, à son utilité. Ce sujet central sera couvert à la section des champs de force.

Une fois la distribution des positions et des vitesses obtenues, il suffit de les corriger pour respecter les conditions aux frontières et de température et de pression constantes (ex : respecter les conditions pour l'ensemble NVT), au besoin, puis de calculer les quantités physiques d'intérêt avant d'itérer au pas suivant.

Pour terminer la discussion sur les principes généraux de la DM, il est intéressant de noter deux avantages de l'algorithme de Verlet sur l'algorithme d'Euler, lequel est couramment utilisé partout en physique de par la simplicité offerte par l'approximation d'une accélération constante sur Δt . Ici, remarquons que les développements en série initiaux incluent bel et bien la dérivée de $\mathbf{a}_i(t)$. Cela améliore la précision numérique d'un ordre

de grandeur par rapport à Euler, et ce, sans compromettre sa rapidité. Deuxièmement, la vitesse est complètement absente de l'équation discrétisée 2.1.4 obtenue pour $\mathbf{r}(t)$. C'est un avantage considérable sur la méthode d'Euler, car on évite de soit (1) contaminer les erreurs d'ordre $\mathcal{O}(\Delta t^4)$ (Verlet) ou $\mathcal{O}(\Delta t^3)$ (Euler) des positions avec les erreurs d'ordre $\mathcal{O}(\Delta t^2)$ (les deux) des vitesses ou (2) de devoir ralentir les calculs en augmentant la précision des ces dernières. Il est à noter, par contre, que bien que l'algorithme de Verlet permette de reconstruire une solution approximative sans explicitement calculer les vitesses, celles-ci sont bien sûr indispensables à maints calculs thermodynamiques. Les ressources informatiques requises sont donc comparables, mais la précision de Verlet est supérieure.

2.1.2. Dynamique de Langevin et intégrateur BAOAB

L'approche générale présentée à la dernière section est un peu simpliste en ceci qu'elle ne tient pas compte des effets que peuvent avoir la friction (i.e. viscosité des solvants) ou les perturbations aléatoires provenant des rares collisions à haute énergie. Pour pallier à ces degrés de liberté manquants, Langevin a ajouté à l'équation de Newton un terme linéaire en vitesse et un processus stochastique gaussien $\mathbf{R}(t)$:

$$m_i \frac{d^2 \mathbf{r}_i}{dt^2} = -\nabla_i V(\mathbf{r}_i(t)) - \gamma_i \frac{d\mathbf{r}_i}{dt} + \mathbf{R}_i(t), \quad (2.1.5)$$

avec $\langle \mathbf{R}(t) \rangle = 0$ et $\langle \mathbf{R}(t) \mathbf{R}(t') \rangle = 6k_B T \delta(t - t')$, avec δ le delta de Dirac. Cette dernière équation assure un contrôle de la température à la façon d'un thermostat.

Pour intégrer cette équation différentielle non-linéaire, on utilise une méthode numérique appelée *split-step method*. Le principe général consiste à séparer la partie linéaire et la partie non-linéaire, lesquelles sont analytiquement solubles sur des petits intervalles de temps, pour finalement les combiner séquentiellement en une seule solution (numérique) par une transformée de Fourier. La façon précise de décomposer le système en composantes et l'ordre dans lequel on les recombine peut avoir différent effet sur la vitesse de convergence et la précision à un pas donné.

Or, il n'existe pas de règle d'or pour construire un intégrateur pour des systèmes stochastiques tel l'équation 2.1.5 (Leimkuhler et Matthews, 2016). Par contre, Leimkuhler et Matthews ont démontré qu'une méthode d'intégration basée sur le flot géodésique et un *splitting* solvant-soluté intégré dans un ordre particulier (B-A-O-A-B) avait une précision et une stabilité dépassant largement les alternatives connues et ce, pour une grande variété de pas de temps. Par exemple, cette méthode permet d'augmenter d'un facteur deux l'efficacité des échantillonnages de DM sans coût calculatoire supplémentaire (Leimkuhle

et Matthews, 2012). Bien qu'en général, l'augmentation du pas de temps mène à des instabilités, ils ont pu montrer que grâce à cette méthode, ils pouvaient augmenter le pas de temps jusqu'à 8-9 fs sans compromettre la résolution et la stabilité des simulations de longues durées, et ce, avec un coût additionnel « modeste » (Leimkuhler et Matthews, 2016).

Dans cette recherche, des pas de temps de 1 fs et 2 fs ont été utilisés lors de l'étape d'équilibration et de production, respectivement. Or, étant donné que nous étudions les interactions protéine-protéine et que la pose du complexe a été créée manuellement, il est possible que de longues simulations soient nécessaires avant d'atteindre un réel équilibre. Dans ce contexte et pour éviter de devoir réduire l'étape de production à un pas de 1 fs pour diminuer les instabilités sur de longues périodes, nous tirons profit de cette méthode en activant l'option langevinBAOAB après les six étapes d'équilibration par défaut de charmmgui.

2.2. CHAMP DE FORCE

La section précédente a introduit les grandes lignes de la dynamique moléculaire classique. Cela a permis d'esquisser les étapes majeures à suivre pour intégrer les équations du mouvement. On a brièvement vu en quoi l'étape la plus coûteuse, mais la plus importante, est celle du calcul des potentiels d'interaction, aussi appelé *champ de force*.

Plus précisément, un champ de force est un ensemble de fonctions et des paramètres associés qui décrivent les potentiels pour chaque type d'interaction interatomique considérée². Pour la présente recherche, le champ de force additif tout-atome CHARMM36m sera utilisé. Il s'agit du raffinement du champ de force CHARMM36, une version moderne très populaires de CHARMM, le premier champ de force développé dans les années 70 et 80. Ce travail a valu à Martin Karplus, Michael Levitt et Arieh Warshel le Prix Nobel de Chimie en 2013³. Par rapport à la version 36, la version 36m ajoute des paramètres permettant un échantillonnage plus complet des ensembles configurationnels de peptides désorganisés (Huang et al., 2017), ce qui peut peut-être s'avérer utile pour la prédiction structurale de l'extrémité N-terminale de DPP6.

Pour un champ additif, l'énergie potentielle se décompose en énergie liée et non-liée. En condition tout-atome, l'énergie liée est la somme des énergies potentielles associées à chaque

2. Cette définition de « champ de force » est à distinguer de la définition plus courante en physique, qui consiste plutôt en une fonction qui à $x \in \mathbb{R}^3$ associe un vecteur $\mathbf{F}(x)$ représentant la force que subirait un corps s'il était placé en x .

3. En fait, ils ont été récompensés pour le développement de techniques multi-échelles, i.e. non seulement pour le champ de force classique CHARMM (Karplus), mais aussi pour l'implémentation de calculs quantiques sur de plus petites échelles à ces modèles classiques (Warshel et Levitt), une technique hybride connue sous l'acronyme QM/MM (quantum mechanics/molecular mechanics)

degré de liberté des atomes unis par des liaisons covalentes. Ceux-ci peuvent étirer la longueur l des liens entre eux, les angles θ que forment ces liens peuvent varier, ce à quoi on ajoute les potentiels de torsions pour les angles dièdres propres ϕ et impropres φ qui mesurent l'angle entre différents plans formés par quatre atomes consécutifs. Le potentiel lié s'écrit donc

$$V_{\text{lié}} = \sum_{\text{liens}} K_l(l - l_0)^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_0)^2 + \sum_{\text{dièdres impropres}} K_\varphi(\varphi - \varphi_0)^2 + \sum_{\text{dièdres}} \sum_{n=1}^6 K_{\phi,n}(1 + \cos(n\phi - \delta_n)). \quad (2.2.1)$$

Pour le potentiel non-lié, on considère l'interaction électrostatique et les interactions de van der Waals, lesquelles sont suffisamment bien approximées par le potentiel de Coulomb et de Lennard-Jones, respectivement. On obtient donc

$$V_{\text{non-lié}} = \sum_{\text{paires } ij \text{ non-liées}} \frac{q_i q_j}{4\pi D r_{ij}} + \sum_{\text{paires } ij \text{ non-liées}} \epsilon_{ij} \left[\left(\frac{R_{\min,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{\min,ij}}{r_{ij}} \right)^6 \right], \quad (2.2.2)$$

où ϵ_{ij} est la profondeur du puit de potentiel caractéristique de la paire ij et $R_{\min,ij}$ est la distance entre les deux particules. Les paramètres du champ de force, i.e. toutes les constantes de ressorts, la valeur des charges partielles⁴, profondeurs des puits LJ sont déterminés, pour chaque type d'atome, grâce à un mélange de mesures empiriques et de prédictions quantiques.

2.3. MODÈLES ATOMIQUES IN SILICO

Idéalement, le point de départ d'une simulation numérique est la structure cristalline haute résolution du matériau étudié. Or, bien que composant 30% du génome humain, les protéines membranaires ne dépassent pas 1% de toutes les structures cristallines protéiques résolues à ce jour (Jones, 1998; Wallin and von Heijne, 1998). Cela est dû au grand défi technique que présente leur purification. Pour résoudre leur structure, il faut en effet les extraire hors de leur environnement hydrophobe, en assez grand nombre pour détecter un signal et ce, sans trop altérer leur structure. Chacune de ses conditions peut présenter un grand défi. Kv4.2 et DPP6 font partis de ces protéines dont la structure exacte reste à résoudre.

4. En effet, on priorise ici un modèle de charge partielle plutôt que de devoir les contributions multipolaires pour sauver du temps de calcul. Certains champs de force polarisables ont été développés pour combler à cette lacune, quand les effets devient importants. L'interaction majeure étudiée ici étant presque exclusivement dans une membrane lipidique, ignorer la polarisation de l'environnement est justifiée.

2.3.1. Modèle d'homologie de Kv4.2

À défaut d'avoir ces données expérimentales, on a recours à différents outils de modélisation. Une stratégie particulièrement fructueuse est basée sur l'homologie de séquence et l'hypothèse, fortement corroborée par l'expérience, que deux séquences homologues produiront des protéines structurellement similaires. Or, tel que mentionné en introduction, les canaux potassiques se comptent en grand nombre et les sous-structures fonctionnelles sont hautement conservées. Cela facilite bien sûr le réalisme des modèles homologues.

L'idée générale est donc de choisir une (ou plusieurs) protéine cousine dont on connaît la structure et d'enligner leurs séquences primaires. Tous les alignements de cette recherche ont été réalisés grâce à l'algorithme BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). Ensuite, plusieurs méthodes permettent de passer de l'homologie séquentielle à une homologie structurelle (i.e. passé du 1D à 3D). La méthode la plus utilisée est basée sur la satisfaction des contraintes spatiales. Résumé brièvement, cela consiste à construire et optimiser une fonction de densité de probabilité à partir de critères géométriques issus de l'alignement. Autrement dit, les bouts de la séquence primaires qui sont MODELLER est le logiciel de choix pour cette technique (Sali et Blundell, 1993).

Or, il se trouve qu'un groupe a déjà publié un modèle homologue pour la partie transmembranaire de Kv4.2 (Heler, Bell et Boland, 2013). Ils se sont en effet servis de MODELLER 9.5 pour modéliser Kv4.2 grâce à sa similarité avec Kv1.2 et KcsA, dont on possède les structures cristallines. En effet, tout ce qu'il reste à faire, à partir du modèle d'homologie de Kv4.2 et de la structure cristalline de son domaine N-terminal T1, est de « recoller » les morceaux. Je me suis servi de la version la plus à jour, soit 9.19, pour recoller le modèle de Heler et al. avec la structure 1NN7 de la Protein Data Bank (PDB) (Zhou et Choe, 2003). Je me suis aidé de la structure complète de Kv1.2 (3LUT.pdb ; Chen, Ni, Wang et Ma, 2010) pour la modélisation de la chaîne reliant T1 et la partie transmembranaire (TM), pour la distance relative entre ces deux sous-structures, de même que pour modéliser la petite hélice (de Q148 à S159 dans 3LUT.pdb) en N-terminal de S1 (absente du modèle de Heler et al.). J'ai cependant dû couper 7 acides aminés (149-DDADTDT-155) de la chaîne entre T1 et la partie TM car certaines informations structurelles sont absentes pour cette région, faisant en sorte que la « *loop* » était beaucoup trop longue pour être réaliste⁵. Après optimisation de cette région avec la routine DOPE de Modeller, un modèle initial pour Kv4.2 (moins les 41 premiers acides aminés formant la « balle d'inactivation ») a été obtenu.

5. Même après optimisation DOPE (discret optimized protein energy), on pouvait la retrouver entremêlée avec les sous-structures transmembranaires de Kv.

2.4. AMARRAGE MOLÉCULAIRE

Les simulations d’amarrage moléculaire (ou *docking*, en anglais) sont des méthodes qui tentent de prédire la configuration d’un ligand et de son récepteur potentiel. Ces méthodes sont donc au coeur de la présente recherche. Il existe deux approches, une basée sur les simulations de dynamique moléculaire et l’autre sur la complémentarité des topologies. Les deux seront exploitées en combinaison pour notre objectif principal : prédire la configuration du complexe Kv4.2-DPP6.2.

La première approche est la plus sophistiquée en ceci qu’elle tente de simuler le processus d’amarrage de A à Z, tel qu’il se produit *in vivo*. Le ligand est initialement séparé du récepteur et il diffuse aléatoirement, jusqu’à ce qu’il rencontre le récepteur et se lie à son site actif. En principe, suivre chaque étape du processus de liaison dans l’espace configurationnel permet de calculer leur coût énergétique respectif et d’estimer la force d’interaction. En effet, le Saint Graal de ce genre d’études est de pouvoir calculer l’énergie libre de liaison *absolue*.

En pratique, cette approche est très coûteuse. Sans parler de la rareté inhérente à de tels évènements sur les échelles temporelles accessibles à la DM, la principale limitation est la nécessité d’échantillonner des plages énergétiques très vastes. En effet, un seul évènement ne reste que cela : *un* évènement. Bien que riche en informations, il ne s’agit d’un seul point dans l’ensemble statistique. Le calcul de certaines quantités de grand intérêt – le meilleur exemple étant justement l’énergie de liaison – demande d’échantillonner « suffisamment ⁶ » l’ensemble des chemins réactionnels pour converger vers la valeur réelle. Dans le cas général, cela pose un défi méthodologique considérable. Il y a tellement de sources d’erreurs possibles qu’il devient difficile de savoir si une prédiction en accord avec l’expérience est dû la minutie du théoricien ou à l’annulation d’erreurs (Deng et Roux, 2009) Dans cet ordre d’idée, il est de loin préférable d’utiliser une structure cristalline à haute résolution plutôt qu’un modèle d’homologie pour éviter que les petites imprécisions structurelles se propagent en grandes différences dans les résultats.

De surcroît, plus la grosseur et la flexibilité du ligand est grande, plus l’espace configurationnel est long à échantillonner. Pis encore, les ressources calculatoires accessibles aujourd’hui ne permettent pas d’investiguer des processus ayant des cinétiques caractéristiques d’ordre supérieur à $\approx 10 - 100$ ns. Les grands changements conformationnels des protéines et certains mouvements lents des lipides sont deux exemples de contributions à l’énergie de liaison incalculables par les méthodes actuelles (Chipot, 2008). Il va sans dire,

6. La formulation vague est intentionnelle : la question de savoir combien de trajectoires est assez pour assurer une convergence est très difficile à répondre (Chipot, 2008).

alors, que si le problème est complexe pour une petite molécule en solution aqueuse, rien ne s'améliore – et c'est le moins qu'on puisse dire – pour de longs ligands flexibles solubilisés par des composés hydrophobiques diversifiés (cholestérol, différents phospholipides, etc.).

Donc, pour l'interaction protéine-protéine, la deuxième approche aux simulations d'amarrage est souvent privilégiée, de par sa relative simplicité. Plusieurs algorithmes différents existent, mais deux grandes étapes sont communes à tous : l'amarrage et l'évaluation (*docking* et *scoring*).

Dans la première étape, l'algorithme propose des milliards de complexes potentiels basés sur la complémentarité des topologies rigides, à l'image d'une clé dans une serrure. Ensuite, il les classe selon une fonction énergie⁷ - leur « pointage » (ou *score*) - qui mesure le degré d'affinité relative. La forme exacte de cette fonction varie d'un algorithme à l'autre, mais elle comprend habituellement plusieurs contributions qui *estime*, par exemple, l'énergie de solvatation, d'interaction électrostatique, de van der Waals et/ou l'énergie libre de liaison. Il va de soi, après la discussion sur l'approche de DM, que ces fonctions représentent parfois assez grossièrement la réalité, mais puisque tous les modèles sont soumis aux mêmes imprécisions, ces quantités *relatives* demeurent utiles pour un simple classement.

Pour les simulations d'amarrage entre les parties aqueuses des protéines, deux algorithmes ont été utilisés pour avoir une certaine diversité de résultats : les serveur en ligne ClusPro 2.0 (<https://cluspro.org>) (Comeau et al., 2004 ; Kozakov et al. 2006 ; Kozakov et al., 2013 ; Kozakov et al., 2017) et GRAMM-X (<http://vakser.compbio.ku.edu>) (Tovchigrechko et Vakser, 2005 et 2006) ;.

On sépare en effet les parties aqueuses et transmembranaires, car jusqu'à récemment, les algorithmes comme ClusPro ne considéraient pas l'environnement très différents de ces derniers. Du coup, non seulement le bon résultat est diluée parmi les nombreux modèles interdits par les contraintes hydrophobiques, mais l'étape de classement est également compromise, de par la forme invalide de l'énergie de solvatation. En 2016, l'algorithme MemDock est venu palier à ces limites en remplaçant ce terme par une expression qui modélise la solvatation des protéines dans l'environnement amphipatique des phospholipides. Mais avant, justement, d'évaluer le pointage des modèles, MemDock ajoute deux contraintes physiques à l'étape de pré-filtrage pour sélectionner seulement les orientations respectant la symétrie planaire de la membranaire. La première contrainte consiste à limiter les

7. Le terme « énergie » est employé de façon souple car toutes les contributions à cette fonction ne sont pas nécessairement des énergies. La fonction de GRAMM-X, par exemple, comprend un terme pour la conservation évolutive, ce qui, d'après Kaczor et al., contribue significativement à son succès (2013).

translations perpendiculaires à la membrane du centre de masse à ± 8 Å et la seconde, à limiter l'orientation des hélices alpha à ± 0.4 radians (22.91°) par rapport à la normale⁸ (Hurwitz, Schneidman-Duhovny et Wolfson, 2016).

Bref, toutes les simulations d'amarrage comprenant des protéines transmembranaires ont été effectuées grâce à cet algorithme.

2.5. PRÉPARATION DU SYSTÈME, pH ET AUTRES OPTIONS DE SIMULATION

Les algorithmes d'amarrage permettent d'identifier des régions candidates susceptibles d'interagir, mais la pose exacte du ligand à même son site actif est plus difficile à prédire en raison des imprécisions de la fonction de pointage. Ce n'est pas nécessairement le cas pour des ligands de topologie irrégulière, mais si on considère une hélice α , il est facile de comprendre pourquoi comment l'algorithme pourrait avoir du mal à distinguer n'importe quelle rotation autour de l'axe central. À cause de la symétrie cylindrique, la complémentarité des topologies reste intacte sous rotation et la seule différence reste donc dans l'interaction spécifique de quelques acides aminés.

Pour tenter de résoudre la pose exacte du segment TM de DPP6 au sein du site actif prédit par MemDock, on utilisera donc toute la puissance offerte par la dynamique moléculaire. L'idée est d'échantillonner l'ensemble des orientations initiales possibles dans le but de voir quelle configuration est la plus stable, une fois équilibrée. Pour préparer ces différentes conditions initiales, un des modèles MemDock a été choisi arbitrairement. Comme son orientation initiale dans la membrane n'était pas parfaitement verticale, l'hélice a été redressée pour éviter de privilégier le modèle initialement choisi et pour uniformiser l'ensemble des angles initiaux. Ensuite, des rotations de $[0, 360[$ ont été échantillonnées avec un pas de 30, pour un total de douze conditions initiales par région. Pour accélérer considérablement la production de ces douze trajectoires par région, on met à profit la structure tétramérique du canal en préparant quatre orientations différentes sur le même canal Kv4.2, soit une pour chaque réplique du site actif putatif.

Une fois ces modèles d'amarrage obtenus, le logiciel PROPKA 3.0 calcule le pKa local de chaque acide aminé pour un pH de 7.0 et change leur état de protonation en conséquence (Sondergaard et al., 2011 ; Olsson et al., 2011). Ensuite, le fichier .pdb résultant est soumis à l'interface web de CHARMM (<http://www.charmm-gui.org>) pour préparer la bicouche lipidique, les fichiers de structure (.psf), les paramètres et fonctionnelles du champ de

8. Dans un premier temps, des restrictions moins sévères (12 Å et 0.8 rad) sont appliquées, mais le raffinement final correspond aux chiffres ci-haut.

force CHARMM36m et les fichiers de configurations détaillant les options de la simulation. CHARMM-GUI a en effet été développé pour automatiser la très délicate opération que représente la préparation du système et donner accès à la communauté scientifique à des serveurs en ligne pour accélérer la compilation des autrement longs algorithmes (Jo et al., 2017).

Ainsi, la bicouche lipidique symétrique est composée de POPE, POPC et de POPS dans un ratio respectif de 3 :2 :1, lesquels sont distribués selon la méthode d'insertion. Le système bicouche-protéine est complètement hydraté par un solvant explicite, composé de molécules d'eau modélisées par un potentiel TIP3P⁹ et environ¹⁰ 0.15 M de KCl. Les dimensions de la boîte sont déterminées par l'extension de la protéine : l'équivalent d'environ 60-75% de sa largeur/longueur est laissé « libre » (i.e. seulement des lipides) de part et d'autre de celle-ci. Comme les conditions aux frontières sont périodiques, l'objectif était de séparer les protéines virtuelles de deux cellules consécutives par un peu plus que la distance caractéristique du système. Pour la direction normale à la surface (z), une épaisseur de 17.5 Å de solvant est laissée de part et d'autre des extrémités de la protéine.

Le système est ensuite soumis à une étape de *minimisation* (géométrique) de l'énergie. Durant 10 000 itérations, les atomes sont tranquillement relaxés jusqu'à l'atteinte d'un minimum local de l'énergie globale. Cela élimine les mauvais contacts, liaisons et angles qui se seraient formés durant la construction des modèles. Il est important de noter que la trajectoire engendrée pour chaque atome n'a pas de signification physique. Il ne s'agit que d'une minimisation mathématique où les positions sont progressivement relâchées, sous contrainte, vers la direction $-\nabla E(\mathbf{r})$, i.e. vers un minimum local.

Ensuite, on active enfin le champ de force durant 750 ps d'*équibration* séparées en six étapes assurant que le système est tranquillement relaxé vers l'équilibre thermodynamique. Ces trajectoires sont produites grâce au champ de force, mais sous contrainte, pour assurer une relaxation « douce » de la structure cristalline.

Finalement, après ces six étapes, toutes les contraintes sont supprimées, l'algorithme BAOAB est activé et les trajectoires continuent d'évoluer librement pour aussi longtemps que voulu.

9. Pour *transferable intermolecular potential with 3 points*.

10. Environ car le ratio $[K^+]/[Cl^-]$, choisi de façon à neutraliser la charge totale de 15e du système, est légèrement inférieur à 1.

Pour l'ensemble des trajectoires générées, la dynamique de Langevin a été utilisée avec un piston de coefficient d'amortissement $\gamma = 1\text{ps}^{-1}$ qui le couple à un bain thermique de pression 1.01325 bar et de température 303.15 K.

2.6. LOGICIELS ET *hardware*

Toutes les simulations de cette recherche ont été effectuées avec la version 2.12b1 du programme NAMD (Philips et al., 2005), sur le système d'opération Linux-x86_64-multicore-CUDA du superordinateur Hélios de Calcul Québec. C'est un logiciel d'utilisation très répandue, spécialisé dans le calcul parallèle, le rendant approprié pour l'étude de gros systèmes biophysiques comme les complexes protéiques transmembranaires.

Les calculs sont effectués sur 10 processeurs Intel Xeon E5-2670v2 Ivy Bridge (2.5 GHz, 25Mo Cache, 10 coeurs), 1 noeud et 1 noeud physique. Chaque noeud possède 20 coeurs, 8 accélérateurs GPU nVidia K20 (GK110) et 128 Go de mémoire.

L'analyse des trajectoires et la production de figures 3D ont principalement été réalisées grâce au logiciel de visualisation de dynamique moléculaire, VMD1.9.3. Une des grandes utilités de VMD est sa puissante capacité de scriptage dans le langage Tcl/Tk, version 8.5.6.

2.7. MÉTHODES D'ANALYSE

À défaut de pouvoir calculer directement l'énergie de liaison absolue, des méthodes alternatives doivent être utilisées pour identifier la pose la plus probable. Trois approches seront comparées.

La première méthode consiste en une analyse multi-échelle de la déviation de la structure équilibrée par rapport à sa configuration initiale. L'idée générale derrière cette approche est que plus l'hélice doit se déformer et se déplacer pour trouver une position stable, plus sa position initiale était défavorable. Autrement dit, on regarde la stabilité pour inférer une affinité.

La première échelle évidente à considérer est celle des rotations de « corps solide ». Il y a deux rotations pertinentes : l'angle de rotation autour de l'axe central et l'angle d'inclinaison par rapport à la normale. Or, les hélices α n'étant pas des corps solides parfaitement rigides, ces angles sont pour l'instant mal définis. Une façon de procéder consiste à calculer la matrice

4×4 optimale¹¹ pour effectuer un alignement structurel entre le squelette central des segments TM de DPP6 et sa structure initiale. On choisit les 13 (sur 25) acides aminés centraux car les deux premiers tours aux extrémités de l'hélice sont plus fluides et rendent les résultats moins consistants et représentatifs de la réorientation globale. Une fois la matrice 4×4 obtenue, donc, on conserve seulement la sous-matrice de rotation $R = \{r_{ij}\}_{i,j=1,2,3}$, délaissant la partie qui translate les centres de masse, puis on calcule l'axe de rotation $\mathbf{u} = (u_1, u_2, u_3)$ et l'angle de rotation θ grâce aux identités suivantes :

$$u_i = \epsilon_{ijk} r_{kj}$$

$$\text{Tr}(R) = 2 \cos \theta + 1 \quad (2.7.1)$$

où ϵ_{ijk} est le symbole de Levi-Civita, sommé selon la convention d'Einstein. Cette méthode est utile pour avoir une image visuelle de la réorientation des hélices, mais les composantes en x, y et z ne sautent parfois pas aux yeux. On décompose donc R en trois matrices de rotations autour des axes x, y et z de façon à ce que $R(\theta) = R_z(\theta_z)R_y(\theta_y)R_x(\theta_x)$, avec

$$R_x(\theta_x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_x & -\sin \theta_x \\ 0 & \sin \theta_x & \cos \theta_x \end{bmatrix}$$

$$R_y(\theta_y) = \begin{bmatrix} \cos \theta_y & 0 & \sin \theta_y \\ 0 & 1 & 0 \\ -\sin \theta_y & 0 & \cos \theta_y \end{bmatrix}$$

$$R_z(\theta_z) = \begin{bmatrix} \cos \theta_z & -\sin \theta_z & 0 \\ \sin \theta_z & \cos \theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

On peut montrer que les angles $\theta_{x,y,z}$ (souvent appelés les angles de *yaw*, *pitch* et *roll* en aéronautique) satisfont à

$$\theta_x = \text{atan2}(r_{32}, r_{33})$$

11. Optimal signifiant qu'on minimise le RMSD, défini plus bas. Cette matrice est calculée par l'algorithme de Kabsch.

$$\theta_y = \text{atan2}(-r_{31}, \sqrt{r_{32}^2 + r_{33}^2})$$

$$\theta_z = \text{atan2}(r_{21}, r_{11}),$$

où `atan2` est la fonction tangente inverse sensible à la position du quadrant (d'où la nécessité de prendre deux arguments), définie par plusieurs langages de programmation.

Une fois l'orientation finale établie, on peut investiguer les détails à une échelle plus petite en calculant la déviation moyenne d'une sélection quelconque de N atomes $\{\mathbf{r}_i(t)\}_{i=1}^N$ par rapport à leurs positions initiales grâce au RMSD (*root-mean square deviation*), défini par :

$$\text{RMSD}(t) = \sqrt{\frac{1}{N} \sum_{i=1}^N \|\mathbf{r}_i(t) - \mathbf{r}_i(t_0)\|^2}. \quad (2.7.2)$$

Cette approche permet de comparer comment réagissent certains acides aminés ou certaines sous-structures pertinentes, autant de Kv4 que DPP6, en situation liée versus non-liée (contrôle). Les interprétations peuvent rapidement devenir complexes lorsqu'on considère plusieurs acides aminés ou plusieurs échelles différentes, si les données ne convergent pas clairement vers un modèle en particulier, mais, encore, l'idée générale est de détecter les acides aminés et les sous-structures stabilisés (ou non) par le contact.

La deuxième approche est d'évaluer les énergies d'interaction électrostatique et van der Waals grâce au champ de force de CHARMM. NAMD possède l'option `PairInteraction` qui, lorsqu'activée, calcule l'énergie potentielle d'interaction non-liée entre deux sélections d'atomes. Cela permet à NAMD de lire les données dynamiques de la trajectoire et de calculer la contribution à l'énergie potentielle du potentiel de Coulomb et de Lennard-Jones pour toutes les paires provenant de chacune des sélections. Exactement les mêmes paramètres sont utilisés que durant le calcul des dynamiques : mêmes fichiers pour les paramètres du champ, `1-4scaling = 1.0`, `switching on`, `switchdist = 10`, `cutoff = 12`, `pairlistdist = 16`. Donc, plutôt que de passer par la stabilité des configurations initiales avec le RMSD, on a une mesure plus directe de l'affinité d'interaction. C'est essentiellement le même principe que les fonction de pointage qui discrimine entre les modèles d'amarrage, car ces énergies font presque toujours parties de ces fonctions de pointage, mais la différence cruciale est qu'ici, la configuration évaluée est équilibrée et les résultats sont une moyenne autour de cet équilibre plutôt qu'une structure cristalline figée. De plus, les énergies sont calculées avec exactement le même champ de force qui sert pour produire les trajectoires.

Il est peut-être important, à ce point, de distinguer l'énergie libre de liaison et l'énergie d'interaction. L'énergie de liaison est le travail *minimal* à faire pour complètement dissocier le système, alors que l'énergie d'interaction est la contribution à l'énergie totale causée par l'interaction. Pour illustrer la différence, on peut imaginer le cas fréquent d'un ligand qui induit un changement de conformation global important à son récepteur sans que le site actif ne se réorganise. Comme les énergies d'interaction (électrostatique et vdW) dépendent surtout de l'environnement local, elles ne comprendront pas le changement d'énergie conformationnelle. Pourtant, pour dissocier le système et mesurer l'énergie de liaison, on devra certainement travailler contre cette réorganisation (i.e. fournir l'énergie pour la transformation inverse). Ainsi, l'énergie de liaison dépend intimement des énergies d'interaction, mais ces dernières ne sont qu'une des contributions, ce qui sera important de garder en tête lors des analyses.

La troisième approche vise justement à comparer comment les modèles réagissent lorsqu'on *tente* de faire ce travail de dissociation minimal. Le verbe « tenter » est à accentuer, car le système est beaucoup trop gros et inconnu pour espérer appliquer les techniques dans les règles de l'art et obtenir la vraie énergie de liaison en un temps raisonnable. Ayant cette mise en garde en tête, on se permet d'oublier l'exigence de séparer le complexe adiabatiquement et d'échantillonner « complètement » l'espace configurationnel et tout simplement regarder quel modèle reste lié le plus longtemps lorsque soumis à une même force.

2.8. LA VALIDITÉ, LIMITE ET UTILITÉ DES RÉSULTATS NUMÉRIQUES

Avant de plonger au coeur de cette recherche, une discussion s'impose sur la validité des résultats obtenus. Bien sûr, il y a d'abord toute la question de la convergence, de la propagation des erreurs et la robustesse des algorithmes de calcul, lesquels constituent des problèmes centraux de l'analyse numérique. Par soucis de brièveté, il suffit de dire sans le démontrer que les potentiels utilisés sont suffisamment lisses pour assurer une convergence théorique aux algorithmes détaillés plus haut et que les erreurs statistiques et de troncation font autant partie du jeu que les erreurs de mesure expérimentale. Mais plus intéressant encore, pour le physicien, est une question d'ordre épistémologique concernant l'approche numérique en général : même si ces imprécisions étaient magiquement inexistantes, est-ce que les résultats obtenus constituent de « vraies » données expérimentales ? Et si non, pourquoi s'y intéresser ?

La question de la validité au sens strict devrait être évidente à tout scientifique : le seul test pouvant vérifier une hypothèse quelconque demeure l'observation directe du phénomène étudié. L'empirisme restera toujours le fondement philosophique non-négociable de la science. Pourtant, bien que la physique numérique n'était initialement vue que comme un simple outil de calcul permettant aux théoriciens de prédire le comportement

de systèmes non-analytiquement solubles, elle porte aujourd'hui un statut particulier. Plusieurs la considèrent comme une discipline en elle-même, à mi-chemin, en quelque sorte, entre la physique théorique et expérimentale. Les simulations numériques qu'elle développe ne restent bien sûr que des *prédictions* basées sur des théories et modèles, mais elles ont continuellement prouvé la richesse et la puissance de l'approche en prédisant correctement des propriétés de la matière autrement incalculables ou encore, bien avant que les matériaux n'aient été synthétisés et les expériences, conduites. Du coup, il est souvent irrésistible d'étendre notre confiance en nos modèles à des domaines inexplorés et de concevoir les résultats comme des données « empiriques » en elles-mêmes.

De plus, les expériences numériques renferment davantage d'informations sur les systèmes physiques étudiés que les expériences en laboratoire, surtout au niveau de la dynamique microscopique. Par exemple, en biophysique expérimentale, les observables se limitent souvent à un petit ensemble de données macroscopiques (e.g. le courant ou la fluorescence) à partir desquelles on espère inférer les mécanismes moléculaires à l'oeuvre. Dans une expérience numérique, le système est construit au niveau microscopique, donnant accès à une résolution atomique rarement accessible à l'expérience en laboratoire, dont on peut se servir pour calculer en principe n'importe quelle quantité thermodynamique. Cette information microscopique détaillée est souvent très précieuse dans l'élaboration, le raffinement et/ou la discrimination des modèles théoriques, lesquels modèles sont essentiels autant pour l'élaboration d'expériences futures que pour notre compréhension conceptuelle du monde.

Utilisée à des fins exploratoires, la dynamique moléculaire permet en effet d'orienter les décisions concernant les expériences à faire en laboratoire. Par exemple, si les modèles sont suffisamment réalistes, il est beaucoup plus économique et rapide de préfiltrer une large banque de composés pharmaceutiques en estimant numériquement leur degré d'affinité pour leur cible et synthétiser seulement les plus prometteurs en laboratoire. Cela permet aussi de manipuler les conditions expérimentales librement sans trop de conséquences et de difficultés (fixer la température et la pression arbitrairement haute ou petite, par exemple) ou de faire des expériences autrement impossibles, comme observer le comportement d'une protéine lorsqu'on tire sur un seul de ses atomes.

Bref, aujourd'hui, l'approche numérique aux questions scientifiques est plus qu'une simple fabulation académique abstraite. Comme l'a d'ailleurs dit Klaus Schulten, le développeur principal à l'origine de NAMD et VMD : « *If we want to understand health and disease, we need to understand life at the molecular level and to know how all the molecular*

components work together like clockwork. » (Dougherty, 2015).

Chapitre 3

RÉSULTATS

3.1. PRÉDICTION DE LA TOPOLOGIE DE DPP6.2

Avant de pouvoir étudier l'interaction de Kv4.2 et DPP6.2 avec une résolution atomique, il faut bien sûr connaître la structure des deux protéines. La section méthodologie décrit comment le modèle d'homologie pour Kv4.2 a été adapté d'un modèle déjà publié. Pour DPP6, cependant, comme seul son domaine extracellulaire est cristallisé, des méthodes de prédictions structurales *ab initio* sont nécessaires pour construire un modèle complet.

Dans un premier temps, une analyse séquentielle de la partie N-terminale manquante (de 1 à 64) révèle que les acides aminés de 1 à 32 sont probablement du côté cytosolique, en solution, alors que les résidus de 33 à 55 (32-GIAIALLVILVICSLIVTSVILLTP-57) formeraient une hélice α transmembranaire (fig.3.1). À cause de leur environnement profondément différent, les structures des deux domaines seront prédites séparément pour éviter que l'algorithme dépense trop d'effort à suggérer des modèles pour lesquels le N-terminal tente de minimiser l'exposition des résidus hydrophobiques.

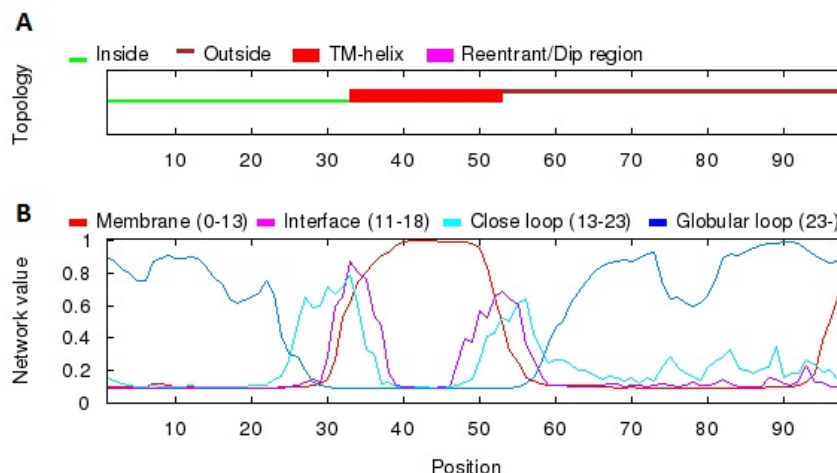


FIGURE 3.1. Analyse séquentielle des cent premiers acides aminés de DPP6.2. (a) Préférence estimée des acides aminés d'être retrouvés à une certaine distance du centre de la membrane (intervalle de distances en Å entre parenthèse). (B) Topologie générale et position relative à la membrane. (Prédiction du serveur OCTOPUS.)

Cette information est pertinente pour contraindre la construction d'un modèle *ab initio*. À cette fin, on utilise le serveur en ligne QUARK, un algorithme qui permet une prédiction structurale à partir de la séquence primaire. Pour le segment transmembranaire, donc, QUARK propose dix hélices alpha de longueur variable, suivies d'une chaîne flexible. Le module MultiSeq de VMD (Roberts et al., 2006) permet d'effectuer le meilleur alignement des C_{α} via l'algorithme STAMP (Russel et Barton, 1992) et de les colorer selon leur degré

de proximité (i.e. *RMSD*) à la moyenne (fig. 3.2). Pour être consistant avec l'analyse de la séquence primaire, seul le modèle ayant la plus longue hélice (sans chaîne flexible) a été retenu. Un bémol pourrait être soulevé concernant ce choix, étant donné que, comme on le voit dans la figure 3.1(b), les résidus à l'interface sont plus incertains et qu'un des modèles possède une courte « chaîne » d'un acide aminé, mais il est fort raisonnable de penser que la dynamique moléculaire corrigera pour ces un ou deux acides aminés peut-être moins rigides. S'en tenir à la topologie générale prédite par 3.1(b) est donc suffisant pour l'instant.

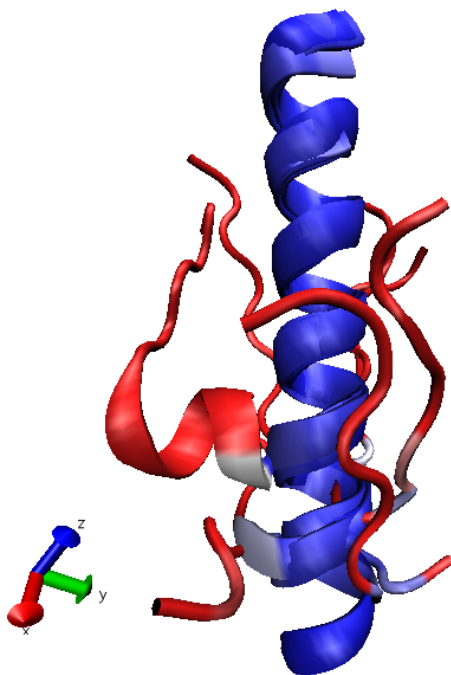


FIGURE 3.2. Alignement structurel des modèles QUARK pour les acides aminés de 32 à 57 de DPP6.2. La coloration indique l'écart à la moyenne, mesuré par le *RMSD* : bleu signifie une proximité et rouge, un écart. On voit que tous les modèles forment une hélice α , suivie d'une chaîne flexible de longueur variable.

Quant à l'extrémité N-terminale (aa. de 1 à 32), QUARK produit également dix modèles. L'alignement structurel STAMP permet d'identifier deux structures communes à tous : une hélice alpha centrale, formée des acides aminés 13-SVQQQEQELVG-23, suivie d'un ensemble désorganisé de boucles en C-terminal (régions 1 et 2 de fig. 3.3). Cette variabilité dans l'orientation des boucles est probablement représentatif du fait que cette région constitue en fait la chaîne qui la relie au TM et doit donc être plus souple. Notons que cette structure est consistante avec l'analyse séquentielle, qui prédisait qu'une transition entre « *close loop* » et « *globular loop* » environ à l'acide aminé 23. (voir flèche sur fig.3.1(a))

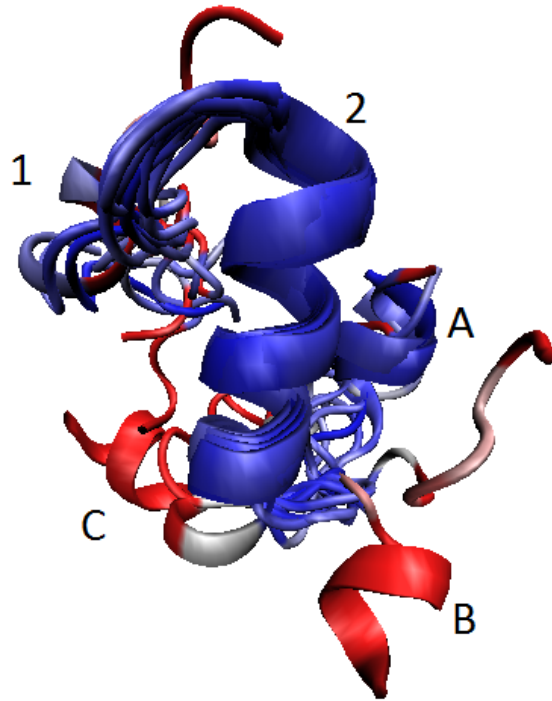


FIGURE 3.3. Alignement structurel des modèles QUARK pour le domaine N-terminal (aa. 1-32). Les modèles présentent deux structures communes : une courte hélice α centrale (2) composée des acides aminés de 13 à 23, suivie, en C-terminal d'un ensemble désordonné de chaînes (1), représentatif de la flexibilité requise pour s'attacher au segment transmembranaire. Les modèles se classifient donc selon l'orientation de l'extrémité N-terminale. La famille A, la plus peuplée, semble composée d'un ensemble désordonné de chaînes orientées plus ou moins parallèlement à l'hélice centrale, dont quelques-unes montrent un tour d'hélice en représentation *NewCartoon*. La famille B possède la même topologie, mais est plutôt orientée perpendiculairement à l'hélice centrale. Le modèle C est le plus différent : son hélice centrale possède un tour supplémentaire, puis est suivie d'une torsion à 90° avant de continuer, toujours sous la forme d'hélice.

Les modèles se classifient donc selon la configuration de l'extrémité N-terminale du peptide (aa. 1 à 13). Basé sur ces résultats, il semble exister trois familles de configurations, libellées A, B et C sur la figure 3.3. La plus peuplée (A) comptent sept modèles dont la boucle N-terminale est plus ou moins parallèle à l'hélice α principale. On voit que certaines de ces boucles semblent s'organiser en courte hélice α , mais il se pourrait que ce soit une fluctuation statistique et un artéfact de la représentation visuelle *NewCartoon* de VMD et qu'en réalité, cette région soit lousse, comme semble l'indiquer le désordre de l'ensemble. La deuxième famille (B) possède la même topologie, mais la boucle est à part, orientée perpendiculairement. La troisième famille (C) sont composés de trois peptides formant une courte hélice dans le plan perpendiculaire à l'hélice principale, orientée dans deux directions

différentes (environ 90° l'une de l'autre).

Or, il est fort probable que le mode d'interaction avec Kv4.2 tel que prédit par les algorithmes d'amarrage classique varie grandement d'une famille à l'autre. En effet, le lecteur se rappellera sans doute de la section sur les méthodes d'amarrage moléculaire (sec.2.4), où il est expliqué que les prédictions de la vaste majorité des algorithmes dépendent intimement de la structure *rigide* des modèles proposés. Or, la rigidité d'une si petite extrémité N-terminale, surtout en solution, est une hypothèse bien forte. La première vérification à faire est donc de s'assurer si la diversité des modèles suggérée par QUARK est bel et bien le reflet de plusieurs états distincts ayant une certaine rigidité ou si l'extrémité est assez flexible pour librement passer d'un état à l'autre en solution. À cette fin, un représentant de chaque famille a été soumis à une simulation DM sous les mêmes conditions. Pour la famille A, un modèle présentant une courte boucle α a été choisi, pour couvrir l'éventualité où celle-ci ne serait pas un artéfact mais interagirait avec l'hélice centrale pour rester dans cette configuration.

D'entrée de jeu, en regardant l'animation des trajectoires, il est clair que deux familles sur trois (A et B) ont des extrémités parfaitement flexibles et sont donc équivalentes. Pour le quantifier, on se place dans le système de référence des axes principaux de l'hélice centrale¹ et on calcule les coordonnées polaires du centre de masse du squelette (N-C $_{\alpha}$ -Carbonyl) du premier acide aminé N-terminal. Les résultats sont résumés à la figure 3.4 pour les trois modèles. On voit que malgré l'atteinte de l'équilibre thermodynamique en environ 10 ns, confirmée par les calculs du RMSD de l'hélice centrale (fig.3.4(a)), les angles polaires du premier acide aminé des modèles A et B ne restent pas près de leurs valeurs initiales comme ce serait le cas s'ils étaient figés dans une structure bien définie, mais fluctuent aléatoirement (fig.3.4(b) et (c)). Surtout pour le modèle B (fig.3.4(c)), ils semblent se stabiliser légèrement vers la fin de la simulation, mais cela est dû au fait que la chaîne s'est beaucoup étirée par rapport à sa position initiale, augmentant son inertie.

Le modèle C, quant à lui, présente une topologie différente. Il y a un tour supplémentaire à l'hélice centrale (a.a. 10-13), suivie d'une torsion de 90° à l'acide aminé 10, suivie d'une autre courte hélice (a.a. 1-9). Sans surprise, la trajectoire suivie lors des simulations montre un comportement qualitatif différent. En visualisant l'animation, on voit que la portion en N-terminale (i.e. après la torsion) vient s'aligner avec le reste de l'hélice, ce qu'on peut aussi constater sur la figure 3.4(d) par la coordonnée r qui devient rapidement stable et par ϕ qui passe d'une valeur d'environ $\pi/2$ à environ π (un peu moins car le centre de

1. On choisit d'enligner l'axe de symétrie de l'hélice avec l'axe z et de prendre son centre de masse pour l'origine. Ainsi, ϕ est l'angle avec l'axe z et θ , l'angle radial.

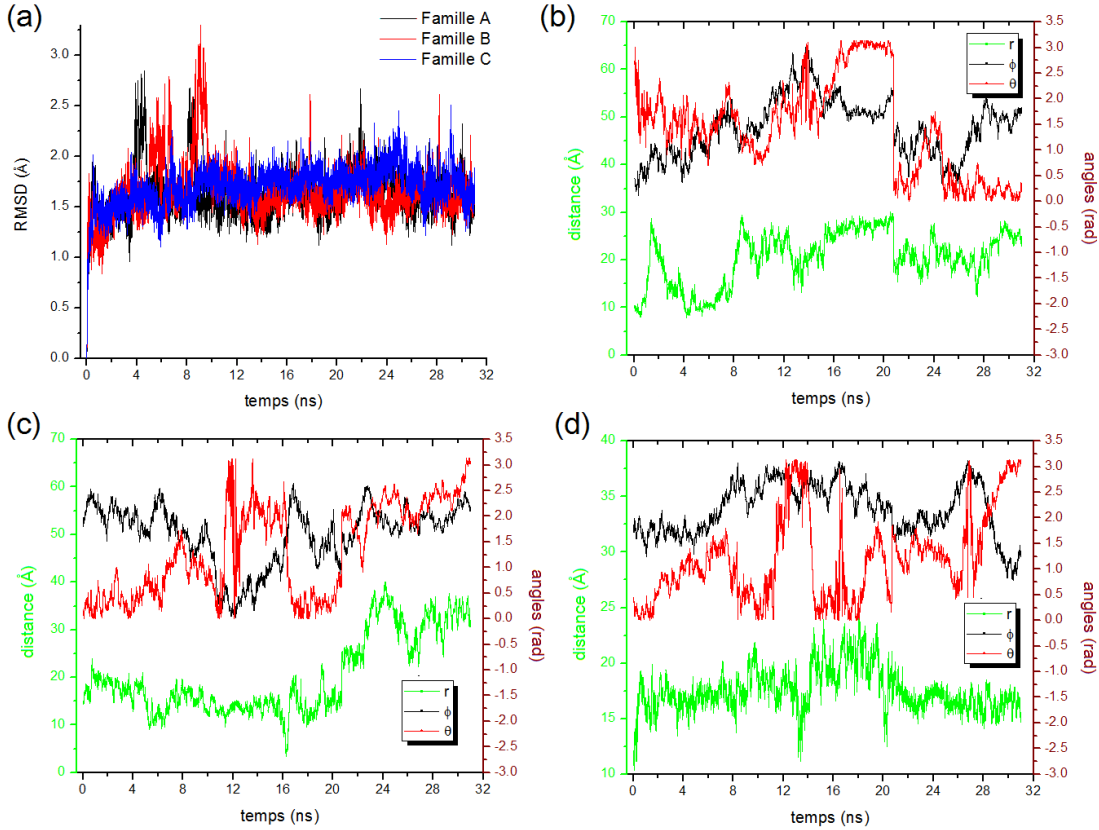


FIGURE 3.4. Dynamique de l'extrémité N-terminale des modèles QUARK pour DPP6.2 en solution. La trajectoire d'un membre de chaque famille a été produite sous les mêmes conditions. (a) L'évolution du RMSD de l'hélice centrale montre qu'elle est stable pour les trois familles et atteint un équilibre thermodynamique un peu avant 10 ns. Par contre, en suivant les coordonnées polaires r (vert), ϕ (noir) et θ (rouge) du centre de masse du premier acide N-terminal dans le système de référence de l'hélice, on voit que les familles A (sous-figure (b)) et B (sous-figure (c)) sont flexibles et n'atteignent pas d'équilibre quelconque. Le cas de la famille C (d) montre davantage de stabilité en (d) Évolution temporelle du RMSD des hélices centrales montre un équilibre atteint autour de 10 ns. (b)

masse n'est pas en ligne directement avec la coordonnée z de l'extrémité). Cependant, cette configuration en longue hélice n'est pas une configuration α archétypale car le point de torsion demeure. Ainsi, la longue hélice ne reste stable que pendant environ 10 ns. Autour de $t = 20$ ns, les deux premiers tours de l'hélice se séparent à nouveau, mais sans pour autant montrer la même flexibilité que dans le cas des familles A et B. Au contraire, ils « pivotent » relativement près de l'hélice centrale, autour du point de torsion initial, tel qu'indiqué par un r constant, mais des angles fluctuants. Cela suggère qu'il existe un état semi-stable dans lequel les acides aminés de 1 à 13 viennent rejoindre le reste de l'hélice α centrale. À

ce moment-ci, on ne peut pas rejeter la pertinence potentielle de cet état transitoire pour les interactions avec T1 ou KChIP, mais il est possible que ce ne soit qu'un artéfact de la présence du tour supplémentaire et de la torsion initialement présente. Il est cependant intéressant de noter que l'hélice centrale des modèles A et B n'a pas spontanément ajouté un tour supplémentaire sur les 30 ns de la simulation, et qu'il est peut probable que la chaîne de 1 à 13 se rétracte pour s'organiser ainsi. De plus, la famille C n'a qu'un seul membre, contrairement à A et B, ce qui indique la faible probabilité d'une telle configuration. Malgré ces deux derniers points, l'hélice plus longue du modèle C sera elle-aussi soumise aux mêmes simulations d'amarrage pour vérifier si elle produit des résultats différents.

Enfin, si ces simulations ne convergent pas vers la conclusion qu'il existe une seule famille d'équivalence, elles permettent du moins de confirmer que l'hélice centrale est une région de grand intérêt pour médier des interactions fonctionnellement intéressantes, vu sa stabilité pour les 3 familles. De plus, en comparant l'alignement BLAST de DPP6.2 avec les autres isoformes (fig. 3.5), on voit que sa séquence est hautement conservée : la fin de sa séquence (ELVG) est commune au cinq isoformes et le début présente un haut degré de similitude pour tous sauf 6.3 (qui n'affecte pas Kv4 de la même façon que les autres, d'ailleurs (Jerng et Pfaffinger, 2012)), notamment en ce qui à trait aux permutations de Q/E/D, de K/R et A/G. Cette conclusion est consistante avec deux études indiquant qu'une structure conservée au sein du N-terminal de DPP6 est en partie responsable des propriétés accélératrices (Jerng et al., 2009 ; Maffie et al, 2009).

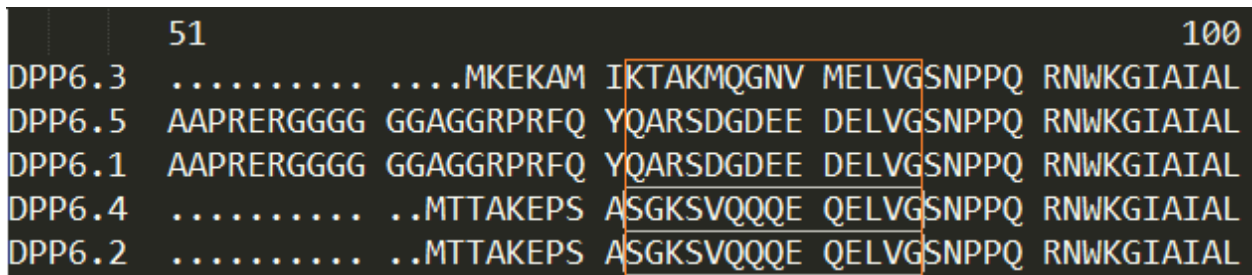


FIGURE 3.5. Alignement BLAST des isoformes de DPP6. Encadrées en blanc sont les séquences des isoformes pour lesquels la région de l'hélice centrale est identiques à DPP6.2 et, en orange, l'ensemble des séquences équivalentes. La numérotation 51 et 100 fait référence aux positions des résidus de DPP6.1 et 6.5, qui ont des N-terminaux plus longs.

La section suivante tente de préciser le mode d'interaction.

3.2. AMARRAGE MOLÉCULAIRE DE DPP6.2 ET Kv4.2

Pour identifier les sous-structures de Kv4.2 susceptibles d'interagir avec les différentes parties de DPP6, les protéines ont été soumises à des simulations d'amarrage moléculaire sur ClusPro2.0, GRAMM-X (partie hydrophile) et MemDock (partie hydrophobe).

3.2.1. Interactions entre le domaine T1 de Kv4.2 et le N-terminal de DPP6.2

3.2.1.1. Prédiction de ClusPro2.0

Étant donné que l'extrémité N-terminale de DPP6 se retrouve en solution, du côté cytosolique, le domaine de tétramérisation T1 de Kv4 semble être le partenaire d'interaction de prédilection. De plus, ce dernier est déjà connu pour interagir avec une autre sous-unité auxiliaire, KChIP, comme l'a montré la structure cristalline du complexe formé du T1 de Kv4.3 et de KChIP1 (Pioletti, Findeisen, Hura et Minor, 2006). La structure cristalline du T1 de Kv4.2 (Nanao, Zhou, Pfaffinger et Choe, 2003) a donc été utilisée telle quelle, sans la partie transmembranaire, pour les simulations avec ClusPro et GRAMM-X.

Au préalable, cependant, les deux queues s'attachant à l'hélice centrale du N-terminal de DPP6.2 ont été coupées pour prévenir la situation où elles nuiraient à l'accessibilité du substrat. Tel que précédemment souligné, leur flexibilité sera mieux prise en compte par la dynamique moléculaire, expérience dont il sera discuté plus loin. Les résultats des simulations d'amarrage sont compilés dans la figure 3.6 et le tableau 3. 1. Pour commencer, on présente les résultats pour l'hélice centrale la plus courte, i.e. celle commune aux trois familles, puis on regarde ensuite si une région différente est identifiée lorsqu'on ajoute les tours supplémentaires de la famille C.

Donc, pour ce qui est de ClusPro, celui-ci regroupe les modèles en six familles, ici numérotée de 1 à 6, peuplées à divers degrés par différents modes d'interaction. Le tableau 3. 1 compile le pointage ClusPro (i.e. la valeur de sa fonction de *scoring*) et le nombre de complexes protéine-protéine appartenant à chaque famille. Pour une représentation visuelle de la topologie des complexes proposés, la figure 3.6 montre une dizaine de configurations caractéristiques par famille. On constate qu'un mode d'interaction sort clairement du lot. En effet, les quatre premières familles sont essentiellement identiques : elles ont le même pointage, approximativement le même nombre de membres, et identifient toutes une unique région d'interaction (voir fig. 3.6 (a) et (b)). La seule différence principale entre les membres d'une même famille (en restant dans les familles 1 à 4) est l'orientation de l'hélice, comme l'indique la position de l'extrémité N-terminale colorée en bleu dans la figure 3.6.

Une deuxième région de contact située au bas de T1 apparaît dans les familles 5 et 6. Cette région est cependant moins probable pour deux raisons. D'abord, ces familles sont environ quatre fois moins peuplées que les familles 1 à 4, signifiant soit qu'il y avait moins de modèles complémentaires lors de l'étape d'amarrage rigide, soit qu'il y en avait moins qui dépassait le seuil pour être retenu pour passer à l'étape de pointage/classification, soit un mélange des deux. Deuxièmement, la fonction de pointage de ClusPro est plus faible pour les familles possédant les modèles situés au bas de l'hélice, suggérant une interaction plus faible.

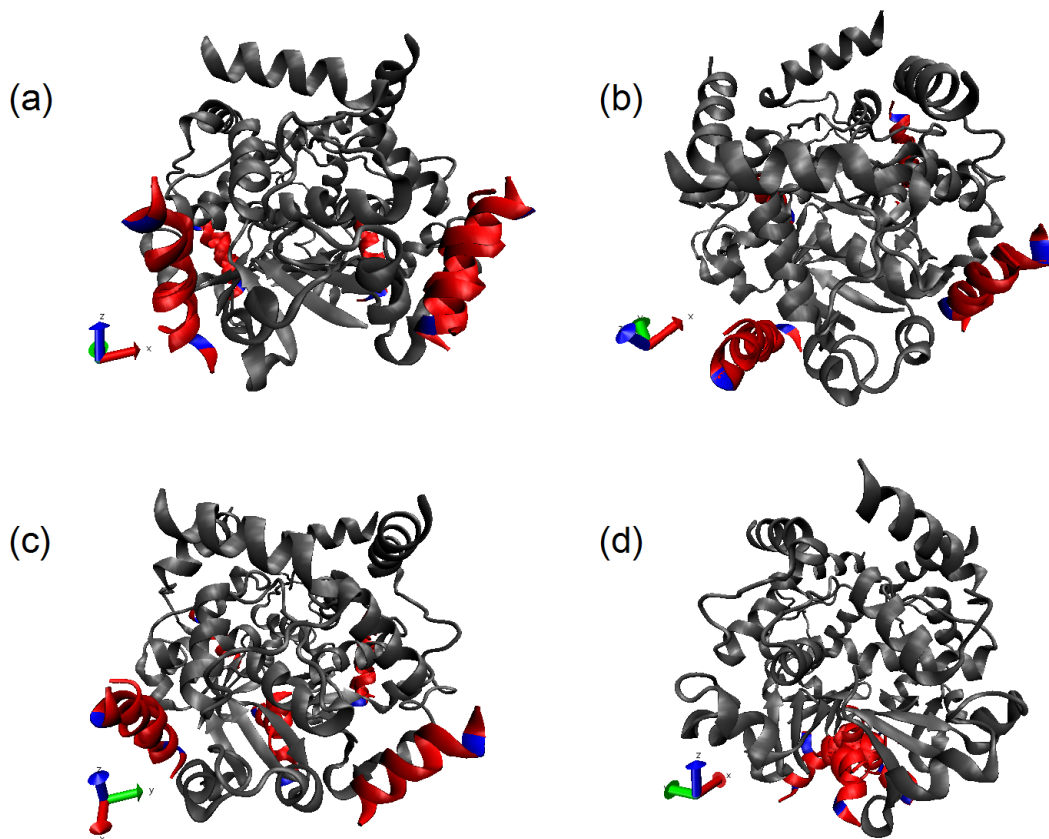


FIGURE 3.6. Interaction de l'hélice centrale du N-terminal de DPP6.2 (rouge et bleu) avec le domaine T1 de Kv4.2 (gris). Illustrés ne sont que quelques membres de la famille 1 (a), la famille 2 (b), la famille 5 (c) et la famille 6 (d). Les familles 3 et 4 sont identiques aux familles 1 et 2 (photos absentes). Le onzième acide aminé (soit le deuxième de l'hélice en partant de son extrémité N-terminale) est coloré bleu pour illustrer l'existence des deux orientations possibles.

Famille	Population	Fonction de pointage ClusPro
1	210	-551.9
2	195	-551.9
3	194	-551.9
4	185	-551.9
5	55	-516.8
6	52	-496.7

TABLEAU 3. 1. Pointage ClusPro et population des six familles identifiées par ClusPro.

Or, malgré les différences de population et de pointage, lesquelles suggèrent toutes deux une plus grande probabilité d'interaction, il n'est pas *a priori* exclu que le bas de T1 puisse interagir avec le N-terminal de DPP6. Cette possibilité est cependant rejetée dès que l'on compare la *distance de contour*² de la chaîne qui devra relier le N-terminal à son segment TM et la distance minimale séparant les modèles de la famille 6 de la membrane. Pour que cette interaction soit possible, il faudrait en effet que cette chaîne de huit acides aminés puisse s'étirer d'au moins³ 52.2 Å, mais la distance de contour typique de huit acides aminés est $8aa \times 3.8\text{Å}/aa = 30.4\text{Å}$ ⁴.

Le même argument permet d'éliminer les modèles pour lesquels l'extrémité N-terminale de l'hélice centrale (colorée en bleue sur la figure 3.6) pointe vers le haut, puisque ces modèles ont un C-terminal pratiquement à la même hauteur que les modèles de la famille 6, ce qui est trop loin pour que la chaîne s'y rende. Au final, ClusPro identifie bien une et une seule région réaliste pour l'hélice centrale principale (aa. 13 à 22).

3.2.1.2. Prédiction de GRAMM-X

Il existe une panoplie d'algorithmes prédisant les interactions protéine-protéine, mais aucun d'entre eux n'est parfait. À titre comparatif, on présente ici les résultats d'un autre exemple basé sur une approche distincte, GRAMM-X (Tovchigrechko et Vakser, 2006).

Contrairement à ClusPro, les résultats ne sont pas classés en familles. Les trois cents modèles sont présentés sans autre classement que la fonction de pointage, le résultat de laquelle n'est pas fournie à l'utilisateur. De plus, on constate d'entrée jeu une bien plus grande diversité de modes d'interaction suggérés. Ces deux derniers points rendent difficile

2. De l'anglais, *contour length*, définie par la longueur d'un polymère lorsqu'étiré jusqu'au maximum permis par la longueur de ses liens covalents.

3. Distance calculée entre la base de T1 et le bas de S1 du modèle d'homologie de K_V4.2.

4. Le chiffre de 3.8Å/aa est commun, mais n'est pas universellement accepté dans la littérature. On retrouve des mesures empiriques variant de 3.4Å/aa à 4.0Å/aa (Ainavarapu et al., 2006), mais même l'estimation la plus généreuse ne change pas la conclusion.

la tâche de les classer en des catégories bien définies objectivement, mais il demeure possible de faire une classification basée sur la fréquence d'apparition car il y a beaucoup de répétitions. Encore une fois, on utilise la récurrence des modèles comme indicateur de la probabilité d'interaction réelle. L'idée est que si des perturbations minimales dans l'organisation atomique du complexe ne perturbent pas sa capacité à franchir les filtres des algorithmes d'amarrage, alors ceux-ci sont de bons candidats pour médier l'interaction réelle. En effet, à cause de la flexibilité des peptides. Autrement dit, on peut le voir en terme de mesure de stabilité du minimum d'énergie : une petite réorganisation (comme celles inhérentes aux fluctuations thermiques) ne changera pas significativement l'affinité. Ou encore, on peut argumenter en terme entropique, plus il y a de micro-états donnant lieu aux mêmes propriétés macroscopiques, plus la configuration globale est probable.

Bref, la première observation est que les modèles de loin les plus fréquents (55%) sont ceux pour lesquels l'hélice N-terminale est placée au bas (comme la famille 6 de ClusPro) et au sommet de T1. Ensuite, la deuxième catégorie de modèles selon leur fréquence d'apparition (20%) est la même région qui ressortait des calculs de ClusPro, à la différence près qu'une minorité d'hélices n'ont pas la même orientation, qui restait constante pour ClusPro. Toujours en fréquence décroissante, une troisième catégorie, absente des prédictions ClusPro mais composant environ 13% des modèles GRAMM-X, suggère quant à elle une zone d'interaction juste au-dessus de la deuxième (en vert sur la fig.3.7 (a)). Enfin, le reste des modèles (12%) ont peu de points en commun, sinon que la plupart semble être mal « collée » à T1 (fig.3.7 (b)).

Dans un premier temps, si on rejette les modèles rendus irréalistes par la contrainte d'être lié au segment TM, la même région d'interaction ressort comme la plus probable pour les deux algorithmes. Notons cependant que les hélices centrales liées au sommet de T1, absentes de ClusPro, ont été comptées dans même catégorie que celles liées au bas. On peut également rejeter ces modèles car en réalité, T1 est lié à la partie TM de Kv4, ne laissant pas assez d'espace pour que l'hélice centrale *et* sa chaîne flexible d'une dizaine de résidus puissent s'y insérer. De plus, non seulement le passage conducteur du canal serait bloqué, mais la littérature indique déjà que certaines interactions impliquent des résidus sur la surface de T1 (Kaulin et al., 2009 ; Jerng et Plfaffinger, 2014).

Concernant les autres configurations absentes des prédictions de ClusPro, elles présentent toutes une complémentarité inférieure au modèle (réaliste) le plus fréquent. Deux possibilités peuvent expliquer ces résultats. La première est que, du moins pour notre système, la fonction de pointage de GRAMM-X est moins efficace que ClusPro pour trier les modèles ayant peu d'affinité. La raison principale corroborant cette intuition est l'absence

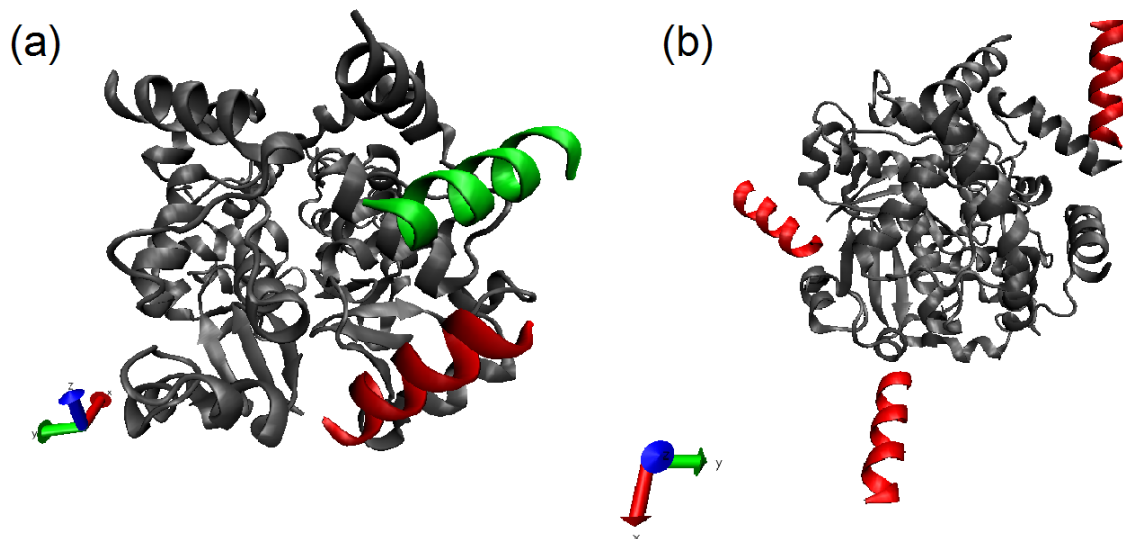


FIGURE 3.7. Catégories GRAMM-X de modèles d’interaction entre T1 et l’hélice N-terminale. (a) Vue de profil du domaine T1, avec deux exemples de modèles d’amarrage. Celui en rouge est le même que celui prédit par ClusPro, qu’on voit apparaître dans environ 20% des trois cent modèles GRAMM-X, alors que celui en vert, juste au dessus, est légèrement moins bien collés et apparaît dans environ 13% des cas. (b) Vue de haut/oblique du domaine T1 pour illustrer trois exemples de modèles parmi les 12 % difficilement classables. On voit qu’il n’existe presque pas de complémentarité entre les surfaces.

des termes pour l’énergie de solvation et l’énergie d’interaction électrostatiques dans sa fonction de pointage (Comeau, Gatchell, Vajda et Camacho, 2014; Kozakov et al., 2017). Elle calcule bien sûr le potentiel de Lennard-Jones, étant l’interaction principale qui unit les complexes protéiques, mais comprend aussi des facteurs de conservation évolutive des surfaces de contact, de préférence statistique aa-aa, et une fonction empirique pour la l’énergie de liaison (Tovchigrechko et Vakser, 2006). Or, il s’avère en effet que la séquence de l’hélice N-terminale comprends plusieurs acides aminés chargés (13-**KSVQQQEQELVG**-23). Ignorer leur contribution potentielle à l’énergie d’interaction et/ou le besoin de les hydrater, au profit de facteurs statistiques, est *probablement* une erreur importante, ici. Il y a peu de doute que ces facteurs jouent un rôle prédictif significatif pour des complexes de plus grandes tailles et/ou plus irréguliers – comme par exemple CAPRI Target 18, dont les auteurs se sont servis pour raffiner leur méthode –, mais ils perdent clairement en importance par rapport aux interactions électrostatique quand 30% du peptide est chargé. Quant à la fonction de pointage de ClusPro, en plus des interactions de van der Waals, cette dernière tient également compte de l’énergie de solvation et des interactions électrostatiques (Comeau, Gatchell, Vajda et Camacho, 2014; Kozakov et al., 2017). Cette résolution supplémentaire au

niveau moléculaire expliquerait peut-être pourquoi il y a moins de variation dans les résultats.

Cela étant dit, il existe une deuxième possibilité qu'on ne peut pas ignorer. En effet, il est possible que GRAMM-X, *un peu* par chance en vue du point précédent, identifie des sites actifs secondaires qui deviennent pertinents si le site actif (putatif) primaire est obstrué. Cette situation pourrait survenir en présence de KChIP, l'autre sous-unité auxiliaire de Kv4 impliquée dans les complexes à l'origine de I_{SA} et $I_{to,f}$. On sait d'ailleurs par la structure cristalline du complexe KChIP1/Kv4.3T1 (2I2R) occupe le site identifié par ClusPro et GRAMM-X. Pour vérifier cette hypothèse, on a soumis l'hélice centrale du N-terminale de DPP6.2 à une simulation d'amarrage sur les serveurs de ClusPro⁵ pour vérifier si les autres configurations identifiées par GRAMM-X étaient reproduites. Les résultats semblent infirmer cette hypothèse puisque aucun modèle ClusPro ne reproduit les modèles additionnels de GRAMM-X. Les positions des hélices centrales sont toutes soit sur KChIP, soit encore au bas de T1.

La prochaine section explore les détails moléculaires sous-jacents à la configuration proposée du complexe. Elle permettra dans un premier temps de tester l'hypothèse que le N-terminal de DPP6.2 interagit avec T1 par un mélange d'interaction de van der Waals et électrostatique et, dans un deuxième temps, de formuler des hypothèse testables en laboratoire sur les paires d'interaction potentielle entre les deux peptides.

3.2.1.3. *Détails moléculaires de l'interaction N-terminale*

Bien que les résultats d'amarrage doivent toujours être interprétés avec précaution, deux algorithmes bien différents convergent vers l'identification d'une même région d'interaction probable, une fois les contraintes du systèmes prises en compte. Cette section vise à en décortiquer la nature moléculaire.

Cette région, donc, consiste en une surface créée par la rencontre de deux sous-unités du domaine T1. Elle est délimitée, d'une part, par la partie du feuillet β formée des aa. 82 à 85, ainsi que des deux hélices formées des aa. 71 à 77 et 112 à 123 d'une sous-unité (en vert foncé, fig. 3.8) et, d'autre part, des parties d'hélices formées des aa. 55 à 62 et 100 à 102 de sa voisine dans le sens horaire, lorsqu'on regarde T1 du haut (en bleu foncé, fig. 3.8). Les acides aminés qui participent à la création de cette surface d'interaction (potentielle) ont été mis en évidence et colorés selon un code de couleur indiquant leur charge électrique.

5. On choisit ClusPro malgré le fait que les autres configurations ont été identifiées par GRAMM-X étant donné qu'on a argumenté qu'il n'est pas judicieux d'ignorer la solvatation et/ou l'interaction des résidus chargés, et pour vérifier si les configurations sont reproductibles en obstruant le site principal.

Les arginines, lysines et histidine (protonée) sont en brun ; les aspartates et glutamates, en orange et finalement, les acides hydrophobiques sont en noir. On observe donc, d'entrée de jeu, la nature amphipatique de la surface : la partie supérieure est surtout constituée d'acides aminés polaires, alors que la partie inférieure est principalement hydrophobe. Pour la suite du texte, les acides aminés individuels seront référés par leur numéros sans distinguer lesquels proviennent d'une sous-unité ou de l'autre, étant donné que l'absence de chevauchement lève toute ambiguïté possible.

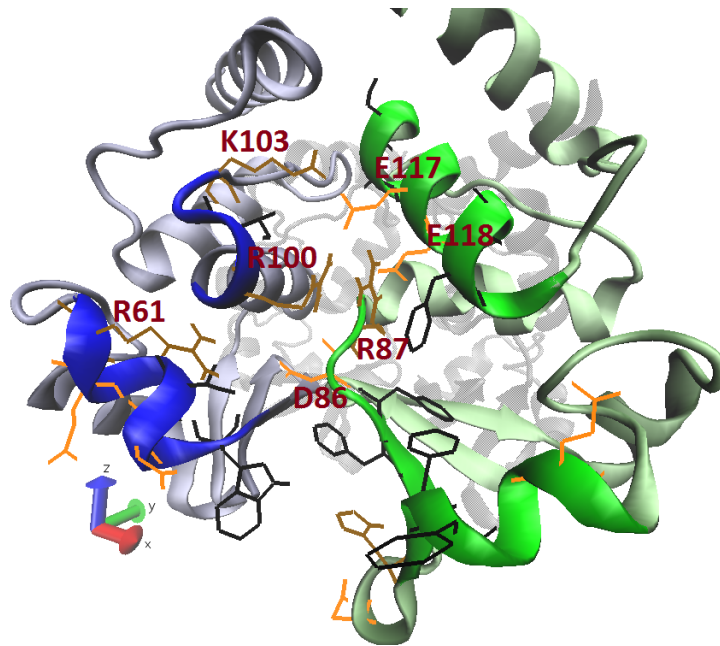


FIGURE 3.8. Détails moléculaires de l'interface putative d'interaction de T1 avec l'hélice N-terminale de DPP6. La surface est créée par la rencontre de deux sous-unités (respectivement en bleu et vert), qui ont toutes les deux des résidus polaires (K, R et H en brun ; D et E en orange) et hydrophobiques (noir).

On remarque que les parties supérieures des sous-unités sont unies par plusieurs interactions électrostatiques, lesquelles sont médiées par les chaînes latérales de R100 et K103 qui projettent vers les chaînes latérales de E117, E118 et R87 (fig. 3.8. Ce groupement de cinq acides aminés crée une région riche en ponts salins, mais dont la charge totale ne semble pas totalement être neutralisée. Il y a en effet trois résidus positifs pour deux négatifs, mais il n'est pas clair que la charge totale est nécessairement e_0 , car juste avant R87, D86 apporte également une contribution négative. Cependant, ce dernier aspartame est légèrement à part de la région concentrée et projette surtout sa chaîne vers R61, qui neutralise en bonne partie sa contribution à la région riche en ponts salins.

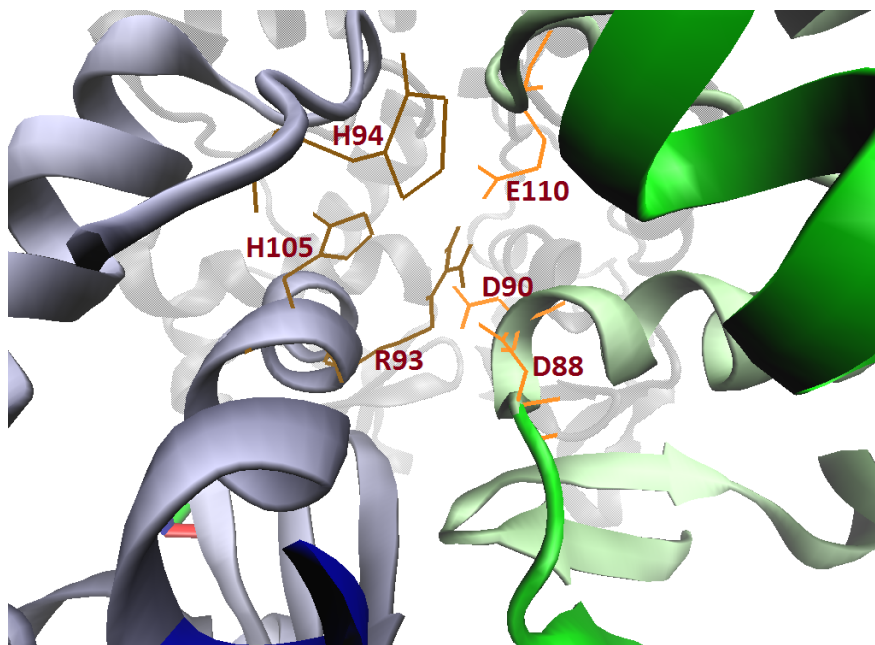


FIGURE 3.9. Interactions électrostatiques inter-sous-unités au coeur du domaine de tétramérisation. Les acides aminés positivement chargés (brun) et négativement chargés (orange) sont en nombre égal et projettent leur chaînes les uns vers les autres.

Par contre, pour s'assurer qu'il existe bel et bien une surface positivement chargée, il importe de s'assurer que les interactions électrostatiques internes s'annulent toutes l'une l'autre. En effet, même si elles ne sont pas accessibles pour interagir avec l'extérieur, elles pourraient quand-même neutraliser une partie des charges superficielles. Il s'avère en effet qu'il existe plusieurs autres interactions électrostatiques près de cette région, située près du coeur du tétramère. Elles sont médiées par les résidus positifs⁶ H94, H105 et R93 d'une sous-unités qui projettent leurs chaînes vers les trois résidus négatifs D88, D90 et E110 de l'autre sous-unité (fig. 3.9). Pour le reste des acides aminés en proximité, ils ne projettent pas leur chaîne dans cette région, ou ne sont simplement pas polaires. Sans trop vouloir simplifier la complexité des ces multiples interactions, il semble par contre qu'il y ait effectivement un équilibre électrostatique interne (3+ et 3-) et que les cinq acides aminés polaires dont les chaînes latérales sont orientées vers la surface ne sont pas neutralisés à l'interne.

Une représentation visuelle ayant moins de résolution, mais plus représentative de l'ensemble de la surface accessible et de sa composition, est la possibilité offerte par VMD de calculer une isosurface via son algorithme « *QuickSurf* » et de la colorer selon la nature polaire de ses résidus (fig. 3.10). Cela permet de bien voir comment les sous-unités

⁶ Les histidines sont bien protonées, dans ce cas. Cela démontre l'importance de bien calculer le pKa avant les simulations !

s'emboîtent l'une dans l'autre et quel type d'interaction (électrostatique ou van der Waals) domine localement. Cette représentation confirme bien que la surface porte une charge positive significative (en bleu), qui couvre non seulement une aire plus grande que sa partie (partiellement) neutralisante (rouge) mais qui est aussi plus concentrée vers le centre. Il y a une petite protubérance rouge au centre, mais celle-ci, bien que certainement accessible, est légèrement plus enfouie. Bref, notons aussi la présence de plusieurs surfaces hydrophobiques (blanc), dont l'une, dans la partie inférieure, est le siège d'une attraction de van der Waals entre les sous-unités alors que le reste est davantage exposé au solvant.

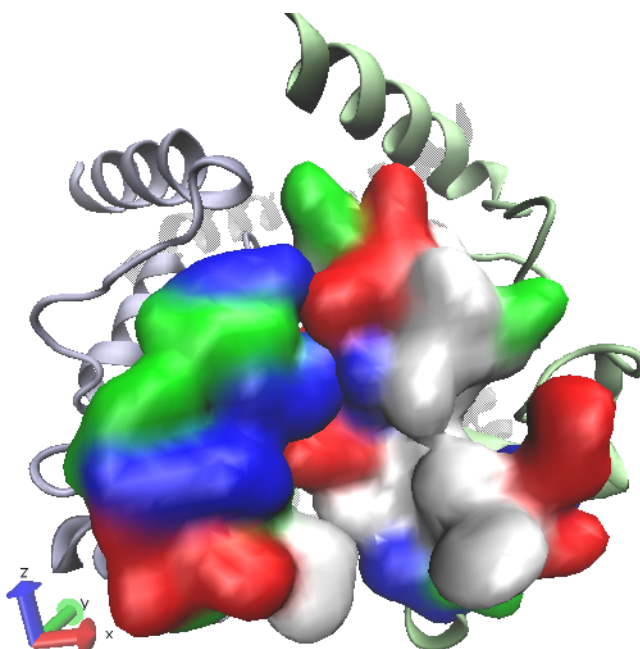


FIGURE 3.10. Représentation « *QuickSurf* » de la région de T1 médiant l'interaction putative avec l'hélice N-terminale, colorée selon que l'acide aminé soit acide (rouge), basique (bleu), non-polaire (blanc) ou polaire non-chargé (vert).

Enfin, la scène est maintenant mise pour étudier la nature des interactions moléculaires entre DPP6 et cette surface selon notre modèle ClusPro/GRAMM-X. La figure 3.11 montrent deux perspectives différentes ((a) et (b)) sur lesquelles six interfaces d'interaction ont été numérotées de (1) à (6) (le même chiffre sur les deux sous-figures représentant la même région). La première observation à mettre en relief est la remarquable complémentarité des surfaces, et ce, malgré le fait que l'hélice n'est qu'un modèle *ab initio* et que ClusPro

tient difficilement compte de la flexibilité des structures.

Commençant par les interactions électrostatiques, on voit d’abord dans la figure 3.11(a) comment le bas de l’hélice positivement chargé (résidu K12) épouse parfaitement la topologie d’une région négativement chargée dans le bas de T1 (D57) pour former un pont salin (région (1)). Ensuite, une partie polaire non-chargée (région (2)) et une partie négativement chargée (région (3)) se logent partiellement dans la crevasse positivement chargée de T1 susmentionnée. Il reste cependant un peu d’espace entre l’hélice et T1, mais il est difficile de résister à la tentation d’imaginer qu’un minime réarrangement conformationnel lors de la vraie liaison pourrait laisser suffisamment de place pour un collage parfait. Notons cependant la présence de la protubérance rouge (D86, fig.3.8) qui semble s’opposer partiellement à cette suggestion (« partiellement » car il a été argumenté que R61, R87 et R100 neutralisait). Seule la dynamique moléculaire (ou l’expérience) pourrait répondre définitivement à ces hypothèses. Plus précisément, si la prédiction de ClusPro est exacte, on devrait pouvoir confirmer l’existence de ces ponts salins par une expérience de type FRET entre les deux régions décrites, soit entre K12 de DPP6.2 et D57 de Kv4.2 et entre le E18 de DPP6.2 et R47 et/ou R100 de Kv4.2.

Ensuite, pour les interactions de van der Waals, surtout visibles sur la figure 3.11(b), on constate que plusieurs protubérances s’emboîtent les unes dans les autres. D’un point de vue topologie, les régions (4) et (6) enlacent la région (5), alors que (7) vient se loger entre (6) et (8). À un niveau moléculaire, la région s’étendant de (4) à (6) sur la figure 3.11(b) correspond aux aa. E20, L21 et V22 de DPP6.2, qui viendraient interagir avec les aa. L66, F74, F75 et F121 de T1. L’interaction entre les régions (6) à (8) de la figure 3.11(b), si elle est juste, devrait pouvoir se confirmer en mesurant les interactions entre V14 de DPP6.2 et W55 et F84 de T1.

Il est intéressant de noter que les acides aminés V14, E18 et E20 de DPP6.2, ici identifiés sur la base de leur complémentarité topologiques et chimiques avec la surface de T1, jouent probablement un rôle fonctionnel dans la modulation des propriétés biophysiques de Kv4. Plus précisément, Kaulin et al. (2009) ont montré que E18 et E20 du motif ELVG joue un rôle dans l’augmentation de la conductance unitaire et Jerng et Pfaffinger (2014) ont montré que l’homologue de V14 et E18 (i.e. M12 et V16 de DPP6.3) jouent un rôle dans la modulation des cinétiques.

À la lumière de ces résultats, on comprend mieux l’origine des résultats différents entre ClusPro et GRAMM-X. Les deux priorisent la complémentarité des surfaces et les interactions de van der Waals et identifient donc la même région, mais comme celle-ci est le berceau de multiples interactions électrostatiques, mieux capturées par ClusPro,

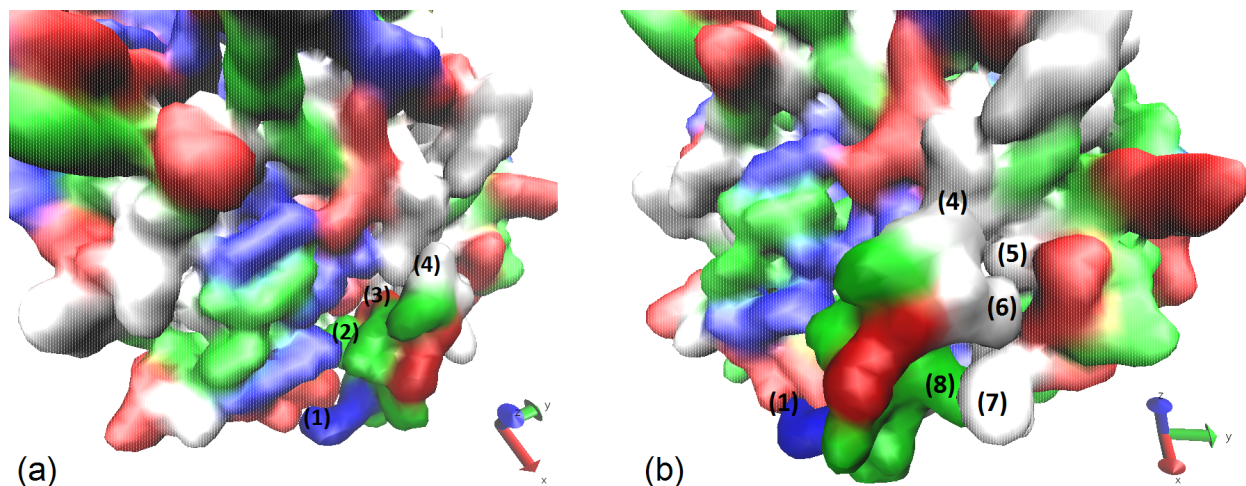


FIGURE 3.11. Interfaces d'interactions putatives entre l'hélice N-terminale (couleurs solides) et T1 (couleurs diffuses). Les surfaces sont hautement complémentaires et montrent un mélange d'interaction polaires (1 à 3), visibles surtout sur la perspective (a) et non-polaires (4 à 8), mieux visibles en (b).

GRAMM-X est moins discriminatoire.

Enfin, pour raffiner davantage la précision du modèle (équilibration, changements de conformation durant la liaison, ajout des chaînes de part de d'autre de l'hélice, etc.) dans le but de prédire exactement les interactions à un niveau moléculaire, il est préférable d'attendre les résultats de dynamique moléculaire. Une expérience que cette recherche espérait mettre à terme, une fois la pose du segment TM obtenue, était de capturer la réaction de liaison entre le N-terminal de DPP6, laissé libre dans le bain alors que son TM est lié à Kv4, et le complexe T1-KChIP. Idéalement, le mieux aurait été d'inclure les parties TM dans cette réaction 'libre lié', mais l'augmentation considérable du nombre de degrés de liberté, l'environnement visqueux de la membrane et la rareté inhérente à de tel événement rendait cette expérience plutôt ambitieuse pour une maîtrise.

La première étape, donc, à identifier le mode d'interaction des segments TM est de recommencer ces simulations d'amarrage avec les peptides d'intérêt.

3.2.2. Amarrage des parties transmembranaires

Comme l'explique la section 2.4, la majorité des algorithmes d'amarrage sont mal adaptés au besoin des chercheurs étudiant les protéines membranaires. Par exemple, ClusPro et GRAMM-X suggèrent tous les deux des orientations relatives entre les segments TM de DPP6 et Kv4 inconsistantes avec leur solvatation par une bicouche de phospholipides

commune. Les modèles ClusPro suggèrent que le TM de DPP6 est pratiquement horizontal, configuration également proposée par GRAMM-X, mais ce dernier place plutôt la majorité des modèles partiellement dans la cavité interne du pore de Kv4 (fig. 3.12). Il est évident que ces résultats n'ont aucune assises réalistes.

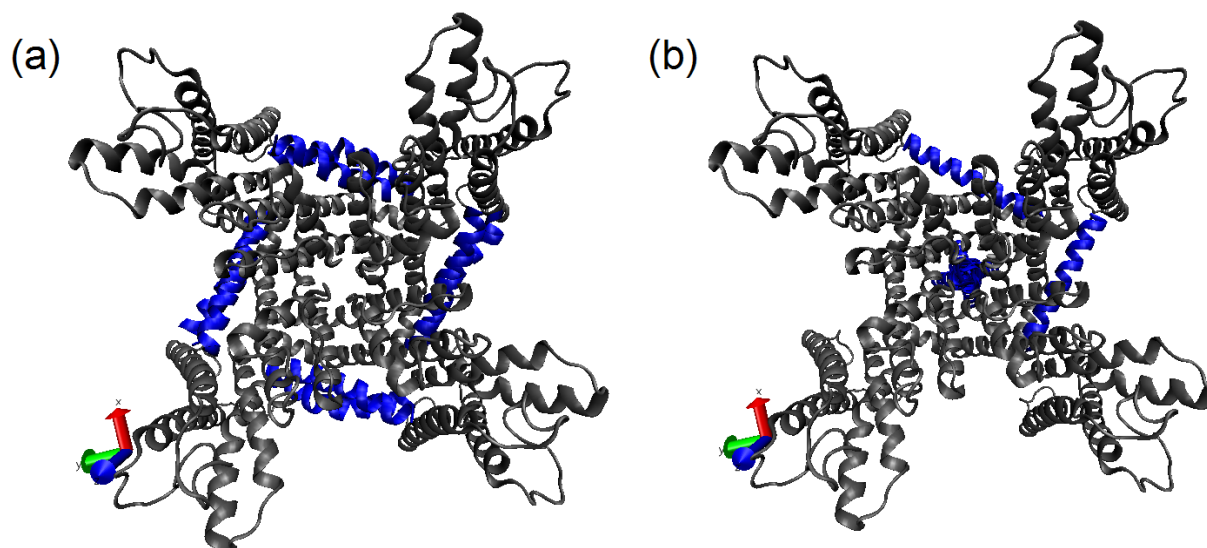


FIGURE 3.12. Amarrage moléculaire des segments transmembranaires de K_V et DPP6 par deux algorithmes « classiques » : (a) ClusPro et (b) GRAMM-X. K_V4 vu de haut.

MemDock corrige le tire en imposant deux critères de sélection géométriques à l'étape de création des configurations complémentaires : une contrainte sur le déplacement vertical et une sur l'orientation par rapport à la normale (voir section 2.4). Les résultats, rapportés à la figure 3.13, sont beaucoup plus convaincants, de prime abord. L'algorithme identifie une seule région potentielle d'interaction, soit la pochette délimitée principalement par le segment S1 du VSD et l'hélice externe du pore (S5). Ce résultat vient corroborer certaines données préliminaires suggérant que les DPP interagissent avec une portion du S1 et du S2 (Ren et al., 2003). Cependant, malgré la région unique identifiée, l'orientation relative des modèles varie beaucoup. Pour le voir, la figure 3.13(b) illustre l'ensemble des hélices en (a) une fois leur centre de masse superposés et leur orientation corrigée pour le fait qu'ils ne proviennent pas toutes du même quadrant (i.e. après une rotation soit 0° , 90° , 180° ou 270° , selon leur quadrant d'origine). Les hélices ont été rendues transparentes pour mettre en évidence le même acide aminé central, coloré en bleu. Constatant que ce même acide aminé couvre le cercle unitaire complet, il devient difficile d'identifier une surface de contact privilégiée entre DPP6 et Kv4 à partir de MemDock seulement. Notons aussi l'ambiguïté concernant l'orientation verticale des hélices dans la membrane, aussi visible sur 3.13(b) .

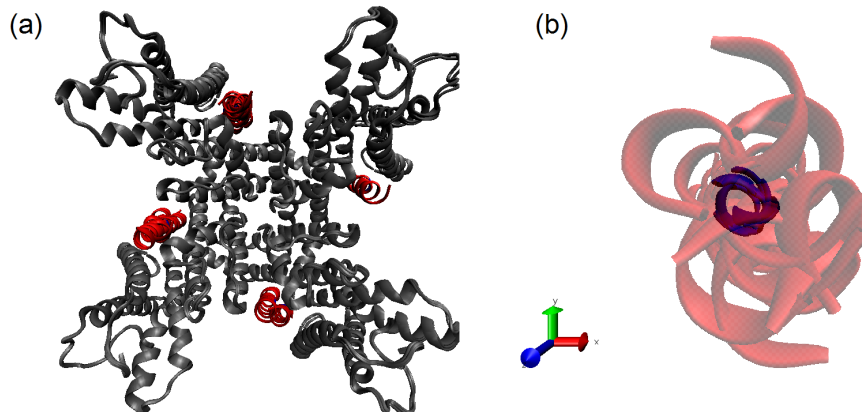


FIGURE 3.13. (a) Les simulations d’amarrage moléculaire des segments transmembranaires de $K_V4.2$ et DPP6.2 par MemDock suggère une pochette d’interaction unique ($K_V4.2$ vu de haut). (b) L’orientation relative des TM de DPP6 (i.e. une fois corrigée pour les rotation autour de l’axe du pore) montre que MemDock discrimine mal la configuration exacte, au sein de cette pochette.

Une option pour choisir l’option la plus probable est de se fier à la fonction de pointage de MemDock (tableau 3. 2. MemDock en calcule deux, en fait : la fonction énergie *MemDock* qui calcule la combinaison linéaire de plusieurs contributions à l’affinité d’interaction (environnement, ponts H, liaisons π , vdW, électrostatique, etc.) et elle fournit le pointage de *FiberDock*, un algorithme qui raffine les complexes d’amarrage en tenant compte de la flexibilité du squelette des peptides (le « *backbone* »)

Famille	Pointage <i>MemDock</i>	Pointage <i>FiberDock</i>
1	366.85	81.89
2	355.35	73.06
3	350.35	62.21
4	346.92	69.24
5	341.61	64.81
6	335.90	47.07
7	335.81	70.52
8	333.79	65.28
9	331.86	61.34
10	319.05	43.67

TABLEAU 3. 2. Pointages de MemDock en ordre décroissant et facteur de raffinement FiberDock.

Donc, il y bel et bien une façon d’identifier une orientation préférable aux autres, surtout avec ce classement beaucoup plus diversifié que ClusPro, qui étaient presque identiques en

raison du fait que les calculs convergeaient vers un seul modèle avec une seule orientation⁷. Cependant, comme l'a montré la comparaison entre ClusPro et GRAMM-X, les résultats de l'approche de l'amarrage moléculaire dépendent intimement de la nature imparfaite des fonctions de pointage. Et bien que cette méthode soit très utile pour identifier des interfaces d'interaction probables, espérer identifier avec confiance un mécanisme moléculaire précis derrière la modulation de Kv4 par DPP6 requiert une approche plus robuste, ayant une résolution atomique. C'est qu'ici qu'entre en jeu la dynamique moléculaire.

3.3. STABILITÉ RELATIVE DES CONFIGURATIONS MEMDOCK

3.3.1. Équilibre thermodynamique

L'objectif général de cette recherche est d'identifier les mécanismes par lesquels DPP6 module Kv. Puisque cette interaction module tous les aspects des cinétiques de Kv4 (activation, désactivation, récupération de l'inactivation) et que celles-ci sont principalement contrôlées par le VSD et son couplage au pore, connaître la pose exacte du TM⁸ par rapport à ces structure devient alors d'une importance primordiale, qu'on ne peut laisser sujette aux imprécisions des algorithmes d'amarrage. C'est à cette fin qu'on a produit trois complexes protéine-protéine modèles dans lesquels quatre répliques du TM ont été placées à la verticale (pour uniformiser les conditions initiales), dans la pochette identifiée par MemDock, à des angles différents les uns par rapport aux autres. Les douze TM résultant permettent d'échantillonner le cercle unitaire avec un pas de 30° et d'étudier la stabilité relative pour chacune de ces orientation. Une expérience contrôle, dans laquelle les quatre TM sont placés à environ 30 Å de la surface de Kv4, a également été produite à des fins comparatives.

Avant toute chose, il importe de vérifier si les systèmes simulés ont bel et bien atteint l'équilibre thermodynamique. Une des façons de le constater est de suivre l'évolution du RMSD du système complet. Les figures 3.14 (a), (b) et (c) comparent comment les douze TM individuels s'éloignent de leur structure cristalline respective ($t = 0$) et en (d), alors qu'en (d), on compare l'évolution des trois complexes « complets » (i.e. les parties transmembranaires de Kv⁹ et ses quatre DPP6). On peut remarquer aux figures 3.14(a) à

7. Rappelons que la position de l'extrémité N-terminale de l'hélice variait pour toutes les familles, mais on parle ici de la direction vers laquelle l'hélice pointe, et non du sens.

8. Si le début du mémoire utilisait TM comme simple raccourci à « transmembranaire », l'abréviation est exclusivement utilisée pour désigner le segment transmembranaire de DPP6 pour la suite de ce mémoire.

9. Bien que domaine intracellulaire T1 ait été simulé par soucis de réalisme, il n'est pas inclus dans les données de la figure 3.14(d) car sa diffusion « libre » ajoute une composante considérable qui pollue l'information voulue sur l'équilibration lorsqu'on tente d'aligner les parties transmembranaire et cytosolique comme un tout rigide.

(c) que la plupart des TM adoptent une configuration relativement stable à un RMSD entre environ 1.5 Å et 2.0 Å. Sans surprise, la vitesse de convergence diffère pour chaque modèle, mais on constate qu'après $t = 40$ ns, soit près de la moitié de la durée totale de simulation, très peu d'entre eux montre une instabilité structurelle majeure. Les exceptions les plus flagrantes sont les cas de $\theta = \{30^\circ, 150^\circ, 210^\circ$, pour lesquels il faudrait davantage de temps de simulation. En effet, on le voit à la figure 3.14(d) – la référence pour réellement mesurer l'équilibration du système complet – l'équilibre n'est pas parfait, surtout pour le modèle comprenant les angles de 120° à 210° .

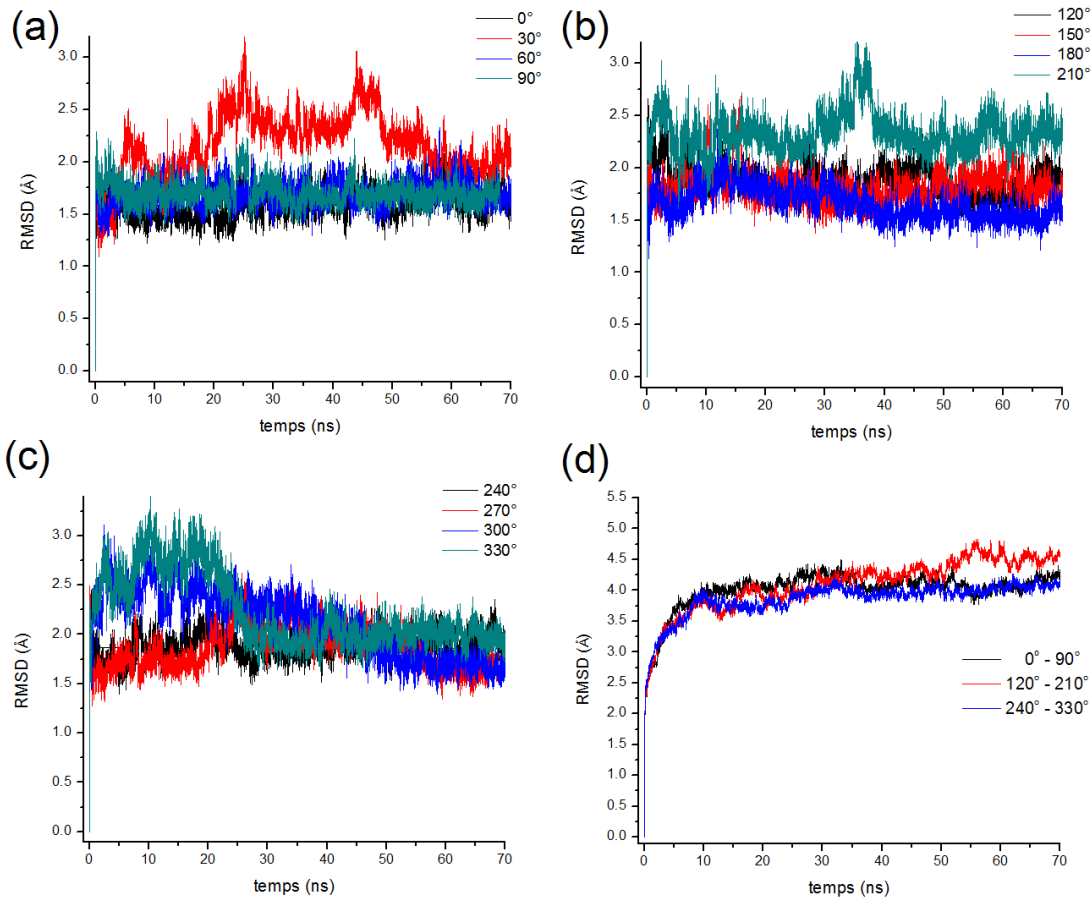


FIGURE 3.14. Évolution du RMSD des TM individuels (sauf les hydrogènes), orientés de (a) 0° à 90° , (b) 120° à 210° et (c) 240° à 330° . En (d), on suit le RMSD des systèmes complets (Kv + 4 DPP6). On constate que les TM sont pour la plupart dans un état d'équilibre stable, à l'exception de quelques-uns comme $\theta = \{30^\circ, 150^\circ, 210^\circ$. Bien sûr, leur présence affecte aussi le RMSD total en (d).

Pour l'instant, bien qu'il est important de garder ce constat en tête, notons tout de même que la composante majeure de l'équilibration se développant sur $t \in [0, 30]$ ns reste

assez loin derrière, à $t = 70$ ns, pour pouvoir analyser les grandes tendances et, espérons-le, classifier les modèles. En principe, comme c'est souvent le cas en dynamique moléculaire de systèmes biologiques, il est vrai qu'il faudrait plus de temps pour atteindre un meilleur équilibre. Mais en pratique, il est possible que pour un complexe d'interaction de cette taille et avec des configurations initiales artificiellement créées (i.e. orientation purement verticale et rotation à intervalle fixe dans le but, rappelons-le, d'uniformiser l'effet des conditions initiales sans prioriser un modèle d'amarrage donné a priori), un équilibre parfait n'existe pas. En effet, il n'est pas impossible qu'on puisse artificiellement emprisonner un TM donné dans un état qui cherche à se délier plutôt que se réorienter, un processus rare en soi, et d'autant plus difficile à capturer spontanément que le ligand est complexe.

Si un tel cas devait se présenter, on observerait des sauts aléatoires dans le RMSD dans cet « équilibre local » et ceux-ci n'empêcheraient pas la capacité à tirer des conclusions sur la nature de cet équilibre local relativement aux autres TM. Ce n'est probablement pas ce qu'on observe ici. Le plus probable est sûrement l'effet combinés de quelques acides aminés n'ayant pas de position stable et/ou qui oscillent entre plusieurs états malgré une orientation globale stable assurée par le reste des acides aminés. Bref, dans un cas comme dans l'autre, le point demeure : il ne faut pas nécessairement s'attendre à un RMSD distribué de façon purement gaussienne à l'équilibre. Très rarement dans la littérature observe-t-on des droites constantes, et ce bémol ne fait que s'accroître avec la grosseur des ligands. La valeur réelle de notre approche repose surtout dans un échantillonnage adéquat et un équilibre suffisamment avancé pour pouvoir *comparer* les modèles. Nous savions dès le départ que celle-ci demeure indirecte et que ce compromis pragmatique rendraient les analyses relatives plutôt qu'absolues. Donc, pour la suite, on parlera tout de même du système *équilibré* en se rappelant sans la mentionner de la nuance ici soulevée.

3.3.2. Orientations finales ϕ_f

Sachant maintenant que le système est assez stable, il est possible de calculer les quantités physiques désirées dans cet état équilibré. Une première quantité d'intérêt est de vérifier si l'orientation initiale des TM est conservée tout au long de la trajectoire. En effet, avec un pas de 30° (équivalent à un pas de $1/3$ d'acide aminé puisque qu'on compte $100^\circ/3$), il est très probable de voir les TM se réorienter vers un angle plus favorable durant l'équilibration, surtout étant donné que les dix orientations de MemDock couvre plus ou moins le cercle entier. Cette vérification est d'autant plus importante que l'information sur les rotations est absente des analyses basées sur le RMSD puisque le calcul de celui-ci est précédé par un alignement structurel minimisant le RMSD. Conséquemment, les translations

du CM et les rotations de type « corps rigide » ne contribuent pas au RMSD¹⁰.

Ainsi, la figure 3.15 compile l'évolution temporelle de l'orientation relative des différents modèles. À la sous-figure (a) est illustrée l'évolution des angles de rotation autour de l'axe centrale des douze TM. On y voit que malgré l'équidistribution initiale, les modèles se réorientent rapidement (entre 5 et 20 ns) vers angle souvent différent. On observe la présence de plusieurs foyer de convergence, ou classe d'équivalence, notamment autour de $\phi_f = 0^\circ = 360^\circ$, où se rejoignent $\phi_i \in \{0^\circ, 300^\circ, 330^\circ\}$; de $\phi_f = 60^\circ$, où se rejoignent $\phi_i \in \{30^\circ, 60^\circ, 90^\circ\}$ et finalement $\phi_f = 250^\circ$, où se rejoignent $\phi_i \in \{210^\circ, 240^\circ, 270^\circ\}$. Les configurations $\phi_i = 120^\circ, 150^\circ, 180^\circ$ restent quant à elles isolées. Pour ce qui est des sous-figures (b) à (d), on peut y voir l'évolution des trois angles d'Euler $\theta_{x,y,z}$ tout au long des trajectoires pour trois exemples de TM, soit $\phi_i = 30^\circ, 60^\circ$ et 90° respectivement. Rappelons que ces angles ont été définis à la section 2.7 comme intervenant dans la décomposition de la matrice de rotation $R(\theta) = R_z(\theta_z)R_y(\theta_y)R_x(\theta_x)$ nécessaire pour amener la configuration actuelle du TM sur le modèle de départ, une fois les CM superposés. Ainsi, s'il faut une rotation de θ_z pour ramener ϕ_f sur ϕ_i , alors $\phi_f = \phi_i - \theta_z$.

Un point important à soulever pour les figures 3.15(b) à (d) est que les modèles choisis convergent vers la même orientation finale ϕ_f (voir 3.15(a)), mais les orientations verticales $\theta_{x,y}$ se comportent assez différemment les unes des autres¹¹. Il faut donc éviter de penser que si deux ϕ_i convergent vers des ϕ_f équivalents, les mêmes interactions ont nécessairement lieu, car l'inclinaison verticale peut changer l'exposition de certains résidus. De plus, d'autres effets entre en ligne de compte : le centre de masse n'est peut-être pas à la même hauteur dans la membrane et l'hélice peut être tordue (légèrement) à différent degré. C'est d'ailleurs le cas de $\phi_i = 60$, qui est un « étage » plus bas que les autres, exposant un tour de plus à la partie cytosolique. On verra plus tard que cela se répercute dans une énergie d'interaction beaucoup plus faible, mais on en voit déjà les effets indirects ici, de par le θ_y qui dépasse parfois la limite permise d'environ 23° pour l'inclinaison des hélices transmembranaires (voir section 2.4 dans les méthodes).

Finalement, il convient d'adresser l'instabilité des angles telle que visible aux figures 3.15(b) à (d) et sa signification pour la « qualité » de l'équilibre thermodynamique. Entre autre, il faut garder en tête qu'un changement dans les orientations verticales ($\theta_{x,y}$) peut simplement être dû à une réorientation complémentaire du VSD. Mais le point principal, surtout pour θ_z , est l'échelle impliquée. Il faut se rappeler qu'il faut une rotation de 100°

10. Si cet alignement est ignoré, le RMSD est dominé par la diffusion libre et la rotation globale. Il devient alors difficile de distiller l'information pertinente à l'échelle moléculaire.

11. Notez que les données ne corrigent pas pour le quadrant d'origine (i.e. $\theta_{x,y} \rightarrow \pm\theta_{y,x}$ selon le cas), mais il reste facile de voir que la forme et la valeur finale ne sont pas les mêmes.

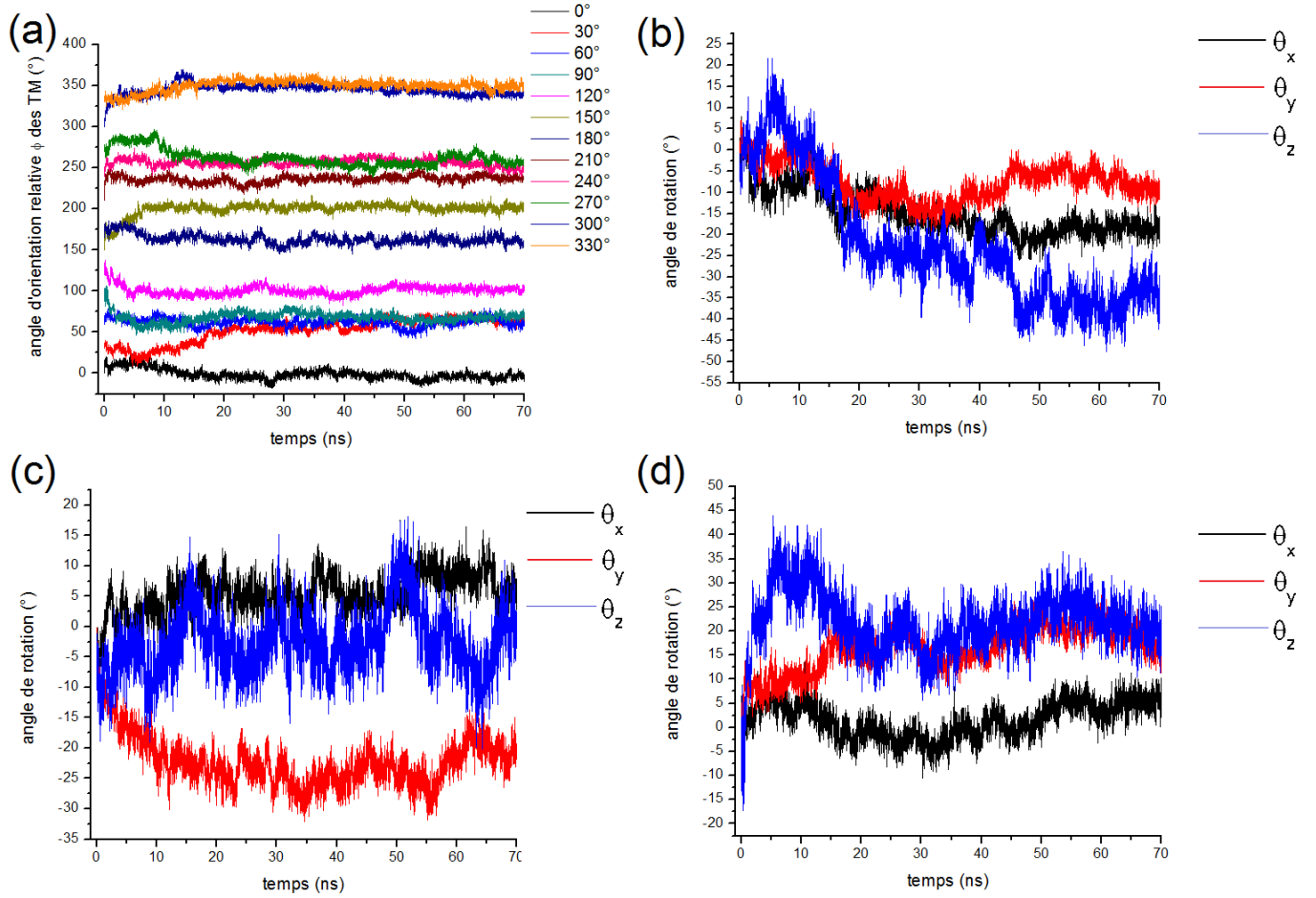


FIGURE 3.15. Évolution des angles de rotations $\theta_{x,y,z}$ pour trois configurations initiales représentatives : (a) $\phi_i = 0^\circ$ (b) $\phi_i = 120^\circ$ (c) $\phi_i = 240^\circ$.

pour passer d'un acide aminé au prochain de la séquence. Ainsi, des sauts comme de 10-15° de θ_z comme celui visible à 45ns en (b) n'ont pas autant d'impact que la figure le suggère.

Enfin, tout cela étant dit, on compile maintenant le tableau 3.3.2 des statistiques à l'équilibre (i.e. sur la dernière ns) de toutes les valeurs pertinentes à l'orientation finale de tous les TM, c'est-à-dire la direction \mathbf{u} de la meilleure rotation θ entre l'équilibre et la structure cristalline, sa décomposition en $\theta_{x,y,z}$ et l'angle relatif final ϕ_f .

Tel que remarqué plus tôt, plusieurs configurations ϕ_i convergent vers certains mêmes ϕ_f . Comme ces cas *peuvent* (mais pas nécessairement) représenter le même minimum local d'énergie, on aimerait les regrouper dans une même classe d'équivalence, et ce, même si les intervalles de confiance définis par la déviation standard ne sont pas assez larges pour se chevaucher. Par exemple, $\phi_i = 0^\circ$ et $\phi_i = 300^\circ$ terminent respectivement à $\phi_f = (-2 \pm 3)^\circ$ et

ϕ_i (°)	θ_x (°)	θ_y (°)	θ_z (°)	\mathbf{u}	θ (°)	$\phi_f = \phi_i - \theta_z$ (°)
0	17 ± 2	-1 ± 3	2 ± 3	(0.97, -0.07, 0.12)	17.3	-2 ± 3
30	-18 ± 2	-10 ± 3	-34 ± 3	(-0.49, -0.10, -0.86)	40.7	64 ± 3
60	6 ± 2	-21 ± 3	0 ± 4	(0.28, -0.94, 0.04)	22.7	64 ± 4
90	5 ± 2	16 ± 3	18 ± 3	(0.10, 0.67, 0.72)	24.2	69 ± 3
120	-11 ± 1	-9 ± 3	18 ± 3	(-0.43, -0.46, 0.76)	22.1	102 ± 3
150	1 ± 2	18 ± 3	-51 ± 3	(0.17, 0.31, -0.94)	54.6	201 ± 3
180	3 ± 2	-12 ± 4	19 ± 4	(0.20, -0.50, 0.83)	23.2	161 ± 4
210	-14 ± 2	-2 ± 2	-28 ± 2	(-0.46, 0.05, -0.89)	31.9	238 ± 2
240	-6 ± 1	1 ± 3	-9 ± 3	(-0.57, 0.15, -0.77)	11.4	249 ± 3
270	9 ± 2	8 ± 3	14 ± 3	(0.44, 0.49, 0.74)	18.4	256 ± 3
300	-5 ± 2	8 ± 2	-39 ± 2	(-0.05, 0.23, -0.97)	40.1	339 ± 2
330	1 ± 2	7 ± 3	-21 ± 3	(0.08, 0.29, -0.95)	22.2	351 ± 3

TABLEAU 3. 3. Rotations de corps rigide du squelette des TM de DPP6 de θ autour de \mathbf{u} par rapport à leurs orientations initiales respectives et décomposition de cette rotation en x, y et z .

$\phi_f = (351 \pm 3)^\circ = (-9 \pm 3)^\circ$, mais la distance entre leur intervalle de confiance est non-nulle : $(-2 - 3) - (-9 + 3) = 1$. Considérant que chaque acide aminé d'une hélice α balaie un angle de 100° et considérant que des petites différences dans θ_x et θ_y peuvent influencer le résultat d'un θ_z donné (et donc influencer la réelle orientation relative des ϕ_f), on aimerait se permettre une certaine souplesse dans la définition des classes d'équivalence. On dira donc que deux ϕ_i sont équivalents si la distance entre leur intervalle de confiance est inférieur à 20° .

Ainsi, en plus de la classe $[\phi_f] = 0^\circ$ déjà identifiée, on peut regrouper $\phi_i \in 30^\circ, 60^\circ, 90^\circ$ dans la classe $[\phi_f] = 60^\circ$ et $\phi_i \in 210^\circ, 240^\circ, 270^\circ$ dans la classe $[\phi_f] = 250^\circ$ ¹². Finalement, les angles $\phi_i = 120^\circ, 150^\circ$ et 180° sont seuls dans leur famille, avec des $[\phi_f]$ respectifs de $\approx 100^\circ, 200^\circ$ et 160° .

Le lecteur aura peut-être remarqué certaines curiosités qui méritent d'être adressées. Par exemple, il peut sembler étrange que $\phi_i = 180^\circ$ termine sa course à $\phi_f = 161^\circ$ alors que $\phi_i = 150^\circ$ termine à $\phi_f = 202^\circ$, passant « par-dessus » ce qui semble être un minimum local à $\phi_f = 159^\circ$. Ou encore, si $\phi_f = 102^\circ$ est l'état stable final de $\phi_i = 120^\circ$, pourquoi $\phi_i = 90^\circ$ n'y converge pas, mais s'oriente plutôt vers $\phi_f = 69^\circ$? Pour y répondre, il est important de rappeler que l'espace des configurations est extrêmement plus complexe que le veut la présente description, qui se limite uniquement à la meilleure rotation rigide du squelette central de l'hélice. L'angle de rotation obtenu ne peut en aucun cas paramétriser adéquatement les chemins réactionnels par lesquels les acides aminés individuels doivent passer, dans le paysage énergétique multidimensionnel des peptides, pour effectuer cette

12. On choisit $[\phi_f]$ en fonction du ϕ_f de la configuration ϕ_i se déplaçant le moins, puis on arrondit.

réorganisation globale. Un corollaire de cette observation pouvant expliquer les points soulevés est que si plusieurs minimums locaux sont accessibles à un ϕ_i donné, une seule simulation n'est pas suffisante pour échantillonner l'ensemble des configurations finales et les multiples façons de s'y rendre. Cela demande prudence lors de l'interprétation.

3.3.3. Énergies d'interaction

À la section précédente, on a identifié plusieurs familles d'équivalence en terme d'orientation de la surface d'interaction du TM, soit la première mesure d'existence de minimum locaux d'énergie, mais on ne sait encore rien de la stabilité relative de ces familles, ni des autres types de mouvements internes des hélices qui pourraient révéler des différences dans l'affinité des peptides (les torsions locales du squelette, mouvements des chaînes latérales, position verticale du CM, etc.). Les calculs des énergies d'interaction et de RMSD sont plus appropriés pour cette analyse.

Dans cette section, on calcule l'énergie d'interaction non-liée entre TM et K_V grâce au champ de force CHARMM (voir méthodologie : 2.7). Cette énergie est la somme de l'interaction électrostatique (potentiel de Coulomb du champ de force) et de l'interaction de van der Waals (potentiel de Lennard-Jones). Le tableau 3. 4 compile le résultat de ces calculs pour chaque ϕ_i , moyennés sur les deux dernières nanosecondes.

ϕ_i (°)	$[\phi_f]$ (°)	$E_{\text{non-lié}}$ (kcal/mol)
0	0	-86 ± 13
300	0	-201 ± 15
330	0	-135 ± 12
30	60	-153 ± 13
60	60	-64 ± 8
90	60	-124 ± 15
210	260	-143 ± 18
240	260	-155 ± 11
270	260	-127 ± 15
120	100	-172 ± 15
150	200	-152 ± 14
180	160	-111 ± 12

TABLEAU 3. 4. Énergie d'interaction non-liée entre le segment transmembranaire de DPP6.2 et $K_V4.2$ à l'équilibre pour chaque configuration initiale ϕ_i .

Plusieurs remarques s'imposent sur la distribution des résultats. Plusieurs orbitent très près autour d'une même valeur intermédiaire d'environ -140 ou -150 kcal/mol, mais ceux-ci proviennent de classes d'équivalence variées ($[\phi_f] = 0, 60, 200, 260$). Quant aux

classes $[\phi_f] = 0^\circ$ et 60° , dont on aurait pu espérer qu'elles consistaient en des minimums stables d'énergie vu la propension de plusieurs modèles à y converger malgré une distance initiale considérable, elles montrent quant à elles une distribution étrange : elles ont toutes les deux des membres ayant des énergies d'interaction inférieures *et* supérieures à la valeur intermédiaire de -140kcal/mol . Cela rend impossible l'argument qu'il s'agit du vraie équivalence entre les minimums d'énergie. Au contraire, ce résultat montre que la variable « macroscopique » qu'est l'angle de rotation rigide θ_z ne semble pas jouer un rôle aussi important qu'on avait pensé. Dans certains cas, vu la grande disparité entre, par exemple, $E(\phi_i = 0) = (-86 \pm 13) \text{ kcal/mol}$ et $E(\phi_i = 330) = (-201 \pm 15) \text{ kcal/mol}$, il faut regarder les détails pour comprendre davantage.

À cet effet, la figure 3.16 illustre un des facteurs dans cette différence. On compare la conformation du complexe entre les paires $\phi_i \in \{0^\circ, 300^\circ\}$ (fig. 3.16 (a) et (b)) et $\phi_i \in \{210^\circ, 240^\circ\}$ (fig. 3.16 (c) et (d)), lesquelles appartiennent à des classes d'équivalence ayant des énergies d'interaction qui diffèrent grandement ((a) et (b)), ou sont les mêmes ((c) et (d)). On voit que le modèle moins fortement lié en (b) épouse moins bien la forme du VSD qu'en (a) à cause du bas de l'hélice en retrait. En (c) et (d), on voit les modèles $\phi_i = 210^\circ$ et $\phi_i = 240^\circ$, respectivement. Ces modèles appartiennent à la classe dont les énergies d'interaction non-liée sont essentiellement les mêmes, et leur conformation se ressemble davantage. Il va de soi qu'un TM plus rapproché aura des interactions les plus fortes et il s'avère même que ce facteur est visible à l'oeil nu dans le cas de la classe $[\phi_f] = 0^\circ$. Cette coïncidence est d'appoint, car le fait que les deux modèles ayant les énergies extrêmes converge vers le mêmes ϕ_f demandait des explications.

Finalement, on constate que notre hypothèse de départ à l'effet qu'il est possible d'identifier la pose la plus probable en dénombrant la population des classes d'équivalence rotationnelle n'est pas supportée. En effet, force est d'admettre qu'il n'y a pas de réelle équivalence physiquement pertinente : la rotation axiale, bien qu'essentielle à calculer pour bien comparer les modèles, manque de raffinement pour prédire la force d'interaction. A priori, l'orientation axiale semblerait être la variable qui véhicule le plus de poids dans la force d'interaction puisqu'elle est plus libre et un changement dans celle-ci a le potentiel de complètement changer la surface d'interaction. Mais la figure 3.16 et les valeurs d'énergie montrent comment un degré de liberté beaucoup plus contraint, l'orientation verticale, ne converge pas nécessairement vers la même valeur dès que l'angle de rotation axiale ϕ est fixé. Cette section montre que les variations possibles des autres degrés de liberté de DPP6.2 entraînent souvent des variations énergétiques plus importantes que de changer complètement la partie de surface exposée à Kv. Il est donc impossible de lire une probabilité d'occupation quelconque dans la population

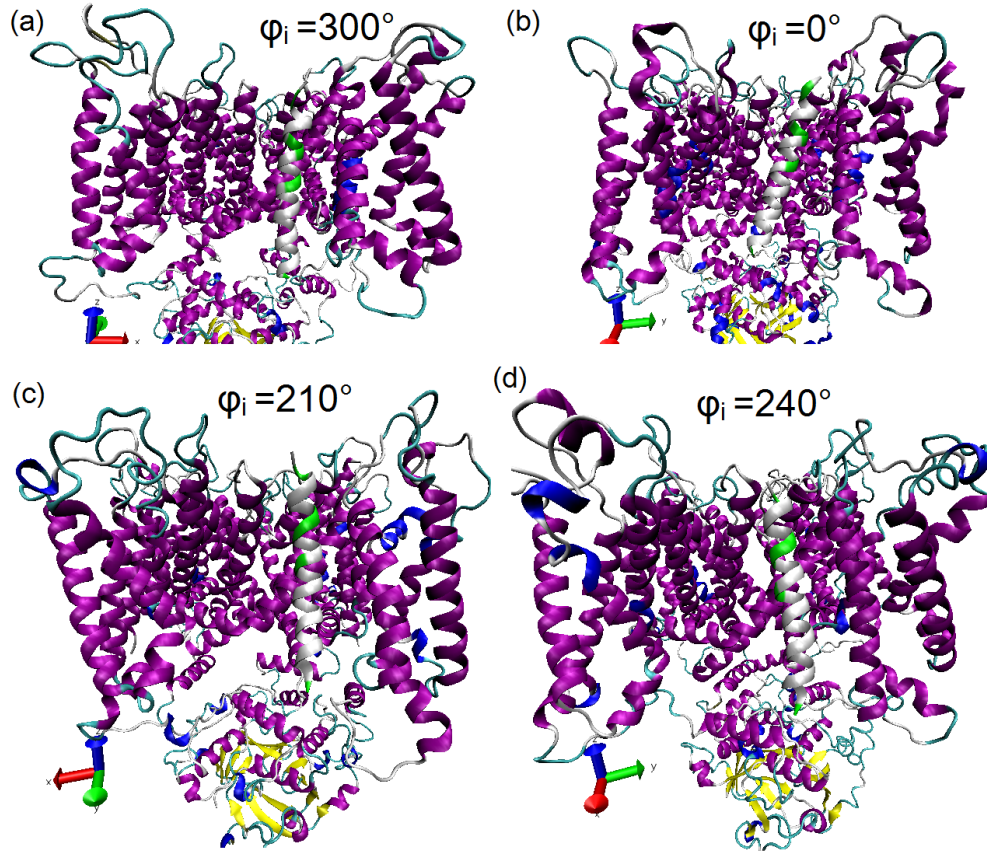


FIGURE 3.16. Comparaison de deux paires de TM d'une même classe d'équivalence rotationnelle. Dans la partie supérieure, on voit illustrés en blanc et vert les modèles $\phi_i = 300^\circ$ (a) et $\phi_i = 0^\circ$ (b), soient les modèles ayant respectivement la plus forte et la plus faible force d'interaction au sein de la classe $[\phi_f] = 0^\circ$. On constate que bien qu'ils aient un ϕ_f similaire (voir les résidus polaires colorés en vert), le bas du TM en (a) est en retrait alors que celui en (b) est davantage complémentaire et s'insère mieux dans la pochette créée par S1, S4S5L et S4 et S5. En (c) et (d), les modèles $\phi_i = 210^\circ$ et $\phi_i = 240^\circ$, membres de la classe ayant des énergies similaires, ont des orientations verticales plus rapprochées.

relative des $[\phi_f]$. Pour cela, bien qu'ils pouvaient sembler secondaires a priori, il faudrait inclure θ_x , θ_y et la position du CM dans la formation des classes d'équivalence ; or, une lecture du tableau 3.3.2 montre qu'aucune paire a une orientation 3D similaire, mais surtout, le fait est qu'il faudrait produire beaucoup plus qu'une simulation par ϕ_i pour avoir un échantillon représentatif. Cette approche serait non seulement très longue, mais son plus grand défaut serait d'ignorer les degrés de liberté à l'échelle moléculaire. Plutôt que de considérer tous les degrés de liberté moléculaires séparément, on utilisera une approche indirecte en analysant leurs effets concertés sur la stabilité structurelle à la section suivante.

3.3.4. Analyse de la stabilité par le RMSD

Jusqu'à maintenant, nous avons établi que la majorité des TM sont relativement bien équilibré, nous avons calculé les orientations et les énergies d'interaction finales. Nous avons vu que plusieurs modèles convergent vers la même orientation axiale, le facteur majeur pour déterminer la surface d'interaction, mais nous n'avons pas pu confirmer ou infirmer l'hypothèse que le nombre de modèles présentant la même surface d'interaction à Kv était indicatif de sa probabilité d'occupation. Pour l'instant, donc, nous n'avons que les énergies d'interaction pour discriminer les modèles.

Par contre, comme on l'a rappelé plusieurs fois, ce n'est pas nécessairement le cas que le modèle ayant l'énergie d'interaction la plus forte est le plus probable d'être tous car on doit tenir compte de l'énergie conformationnelle et de solvation. Ces énergies présentent des barrières énergétiques supplémentaires dans le processus de liaison moléculaire. Or, comme l'énergie libre de liaison peut difficilement se calculer pour notre système, on prend donc une approche indirecte basée sur la stabilité structurelle, laquelle repose sur l'hypothèse suivante. On supposera que le modèle le plus structurellement stable est celui qui est le plus probable d'être le mode d'interaction réel. En effet, la pose d'un ligand, i.e. sa conformation spatiale relative au récepteur, est unique. L'hypothèse revient donc à dire deux choses. D'abord, qu'une liaison plus forte réduit les fluctuations thermiques en contraignant davantage les chaînes latérales du ligand lié, et ensuite, que la nature est économe, c'est-à-dire que s'il existe un chemin dans l'espace configurationnel vers un état de plus petite énergie, le ligand va éventuellement l'emprunter et c'est ce minimum énergétique absolu qui correspond à la vraie pose fonctionnelle, et non les minimums locaux transitoires. Donc en observant le modèle le plus stable, on a de meilleures chances de tomber soit sur le minimum absolu, soit sur un état transitoire.

Pour mesurer la stabilité structurelle des complexes, on calcule les statistiques du RMSD des régions pertinentes. La première région naturelle à étudier est la surface de contact entre les peptides. Le langage Tcl/Tk permet à cet effet de définir une sélection variable d'atomes en fonction d'une distance fixe d'une autre sélection. La figure 3.17 montre la région de Kv4.2 résultante pour deux distances d'un TM donné. Si on couvre surtout les régions en contact direct avec $d = 10\text{\AA}$ (S1, S5, bas du S4, parties de S6 et du P-loop), on remarque qu'à $d = 12\text{\AA}$ (non illustré) et surtout $d = 14\text{\AA}$, certains acides aminés isolés commencent à être inclus dans la sélection. Plus importants encore sont le S4-S5L, le S4 et le S6_T, qui sont beaucoup plus complets dans la sélection à $d = 14\text{\AA}$, alors que, par exemple, seul le dernier tour du S4-S5L est inclus dans $d = 10\text{\AA}$. De plus, dans les deux cas, on remarque que la chaîne entre le segment transmembranaire et T1 est partiellement incluse.

Pour éviter que ces sélections viennent contaminer les RMSD, on les exclut de la sélection ((within "d" of "TM") and not "loops"). Il faut en effet une certaine prudence dans ce choix car l'ajout d'autres sous-structures n'ajoutent pas seulement leurs contributions individuelles, mais influence en principe tous les autres calculs en changeant le choix de la meilleure matrice d'alignement. Malgré tout, comme le choix d'un d reste relativement arbitraire, on présente les résultats pour trois valeurs, soit $d = 10\text{\AA}$, 12\AA , 14\AA

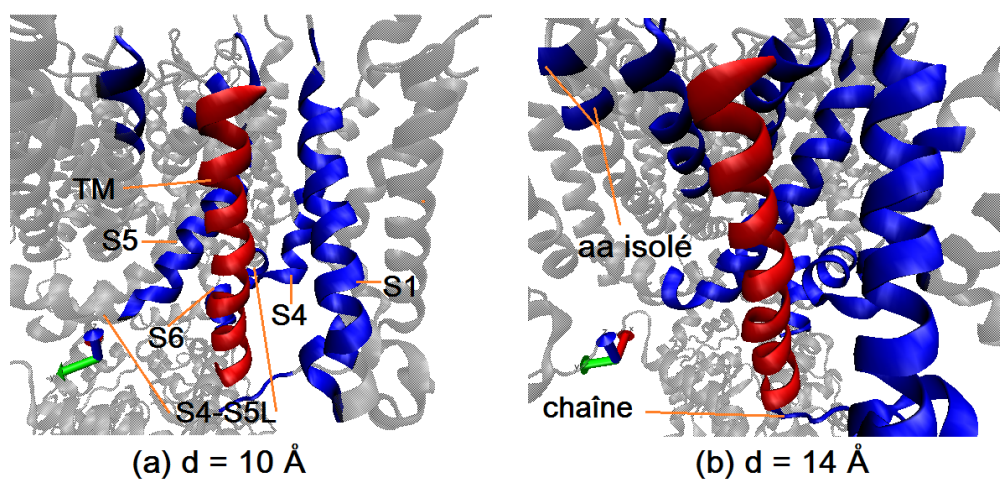


FIGURE 3.17. Sous-structures de $K_V4.2$ (bleues) à d Å de TM (rouge), avec $d = 10$ (a) et $d = 14$ (b).

Le tableau 3.3.4 compile les calculs des RMSD moyens et de leur écart-type à l'équilibre pour les régions d'interaction définies par tous les atomes (sauf les hydrogènes et les boucles pour diminuer le bruit) à une distance d du centre de masse de TM, avec $d = 10\text{\AA}$, 12\AA , 14\AA . Pour interpréter ces données, il vaut la peine d'accentuer ce que mesure vraiment le RMSD. Il a été défini à la section 2.7 comme étant la racine de la distance au carré moyenne entre un groupe d'atomes et leurs positions dans la structure initiale. Or, il faut faire attention car plusieurs auteurs dans le domaine utilisent le RMSD directement pour mesure la stabilité des complexes simulés. Ceci n'est valide que pour de très petits ligands n'ayant pas ou très peu de flexibilité. En général, on s'intéresse à la stabilité à l'équilibre, soit l'écart-type, et non pas le degré auquel on s'éloigne de la structure cristalline. Dans le cas d'un ligand simple dont la structure est fixée par ses liens covalents et rien d'autre, le RMSD moyen sera toujours près de 0 et utiliser le RMSD revient au même. Dans le cas des protéines, les interactions non-covalentes fondamentales à leurs structures offrent plus de liberté. Donc, bien qu'il pourrait être tentant d'avancer que plus le RMSD est petit, moins de réorganisation est nécessaire et donc plus le modèle initial était stable tel quel, le fait est qu'une réorganisation n'est pas nécessairement signe d'incompatibilité entre ligand et récepteur. En général, un changement conformationnel peut avoir un effet positif ou

négatif sur l'énergie d'interaction, mais le lien entre la stabilité moyenne des résidus à l'origine des liaisons non-covalentes et la force de ces dernières est beaucoup plus unidirectionnel.

ϕ_i (°)	$[\phi_f]$ (°)	RMSD($d = 10$) (Å)	RMSD($d = 12$) (Å)	RMSD($d = 14$) (Å)
0	0	3.91 ± 0.10	4.58 ± 0.15	5.12 ± 0.15
30	60	3.88 ± 0.09	4.70 ± 0.13	5.02 ± 0.16
60	60	3.46 ± 0.07	4.17 ± 0.15	4.66 ± 0.19
90	60	4.07 ± 0.13	5.15 ± 0.14	5.52 ± 0.14
120	100	4.13 ± 0.12	6.84 ± 0.20	7.62 ± 0.21
150	200	5.26 ± 0.15	5.53 ± 0.16	5.67 ± 0.18
180	160	2.97 ± 0.10	3.58 ± 0.13	4.05 ± 0.13
210	260	3.78 ± 0.09	5.75 ± 0.26	6.17 ± 0.28
240	260	3.43 ± 0.09	3.96 ± 0.11	4.50 ± 0.11
270	260	4.26 ± 0.14	5.16 ± 0.18	5.94 ± 0.18
300	0	3.17 ± 0.05	3.49 ± 0.05	3.81 ± 0.09
330	0	4.01 ± 0.10	5.91 ± 0.27	6.77 ± 0.27

TABLEAU 3. 5. RMSD moyen et déviation standard pour les régions de contact définies par une distance d fixe du TM pour chaque configuration initiale ϕ_i .

Bref, en observant les écarts-types du tableau 3.3.4, on constate qu'un modèle se distingue clairement des autres peu importe la distance. La région d'interaction correspondant à $\phi_i = 300^\circ$ montre la plus grande stabilité pour les trois d et ce, de façon significative. En effet, pour $d = 10$ Å et $d = 12$ Å, l'écart-type de la région d'interaction associée au TM $\phi_i = 300^\circ$ est systématiquement au moins la moitié des autres modèles. La seule exception Les 0.09 ne sont pas des exceptions car ils ont été arrondi vers le bas, alors que le 0.5 est en fait un 0.0467 arrondi vers le haut. est le 0.07 de $\phi_i = 60^\circ$, Pour $d = 12$ Å, on constate que l'écart-type de $\phi_i = 300^\circ$ a augmenté plus rapidement que tous les autres, mais il demeure inférieur à toutes les autres valeurs. Ce résultat sera mieux compris plus loin, à la section où l'analyse des écarts-type est séparée selon les sous-structures fonctionnelles de Kv. Pour l'instant, il convient de remarquer que le modèle le dont la structure est la plus stable est $\phi_i = 300^\circ$. Basée l'hypothèse introduite en début de section, on conclut que c'est le meilleur candidat pour déterminer la pose réelle de l'interaction protéine-protéine. Ainsi, deux analyses distinctes pointent vers le même candidat, c'était ce modèle qui montrait l'énergie d'interaction non-liée entre Kv4.2 et DPP6.2 la plus forte. La figure 3.18 montre son emplacement relatif exact et la complémentarité des surfaces.

Maintenant, avec ces données, on peut quantifier à quel point est vérifiée l'hypothèse avancée en début de section sur la corrélation entre la stabilité et l'énergie d'interaction, et comparer la corrélation pour chaque distance. La figure 3.19 présente le résultat de ces

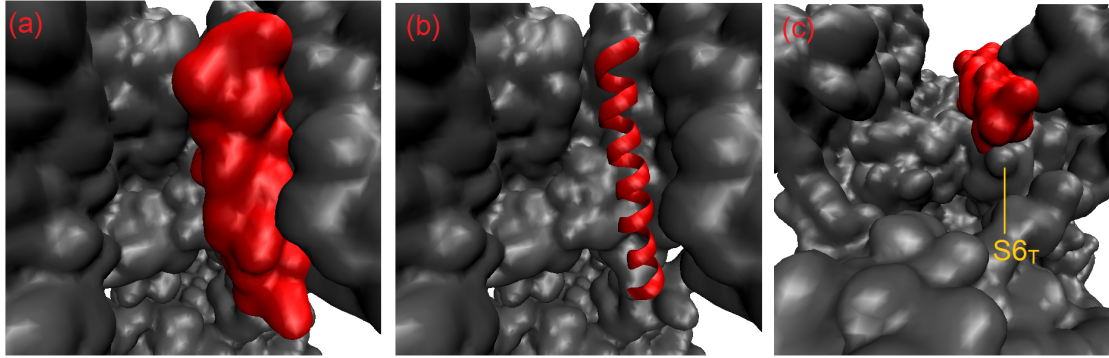


FIGURE 3.18. Pose du segment transmembranaire de DPP6.2 (rouge) dans la pochette créée par S1 et S5 de Kv4.2 (surface grise). (a) Vue de côté, représentation isosurfacique pour montrer la complémentarité. (b) Même vue qu'en (a), avec TM représenté sous forme NewCartoon pour montrer le relief de Kv4 caché (a). (c) Vue du bas, pour montrer l'interaction avec le bas de S6.

analyses. On constate qu'il a bel et bien une tendance générale à l'effet qu'une interaction plus forte (énergie d'interaction plus négative) augmente la stabilité structurale. Par contre, les corrélations sont assez faibles pour chaque distance, avec des facteurs de corrélation R valant 0.087, 0.17 et 0.142 pour $d = 10\text{\AA}$, $d = 12\text{\AA}$ et $d = 14\text{\AA}$, respectivement. Or, il est évident que le fait que certains modèles ne soient pas encore parfaitement équilibrés augmente l'écart-type et influence négativement la qualité du facteur de corrélation. Il y a donc quelques points aberrants. On verra aussi plus loin que certaines sous-structures se font déstabiliser, alors que d'autres se font stabiliser, par la présence de DPP6.2, même pour $\phi_i = 300^\circ$, et nous pensons que ce résultat a une signification fonctionnelle pour l'interaction. Inclure ces différentes sous-structures dans un même ensemble produit donc un effet de moyenne.

Cela étant dit, il est intéressant de noter que la corrélation augmente de près du double entre $d = 10\text{\AA}$ et $d = 12\text{\AA}$, pour diminuer légèrement à $d = 14\text{\AA}$. Ce résultat est particulier, car étant donné que les interactions électrostatiques et de van der Waals opèrent à courte portée, il aurait été naturel de s'attendre à ce que les énergies d'interactions soient plus fortement corrélées à l'écart-type à courte distance puisque les interactions sont médiées localement, et que le bruit provenant des structures lointaines, moins affectées, contamine progressivement cette corrélation à mesure que la distance augmente; or, on observe l'effet inverse. Comme l'avait montré la figure 3.17, on sait aussi que $d = 10\text{\AA}$ devrait être suffisant pour capturer cette corrélation puisque non seulement DPP6 est toujours inclus dans ces distances, mais on inclut complètement la surface d'interaction directe. À $d = 12\text{\AA}$ et $d = 14\text{\AA}$, on commence à capturer des régions indirectes. Notons d'ailleurs qu'on observe bel et bien la dilution attendue à plus grande distance. Le facteur R^2 -ajusté permet de mesurer

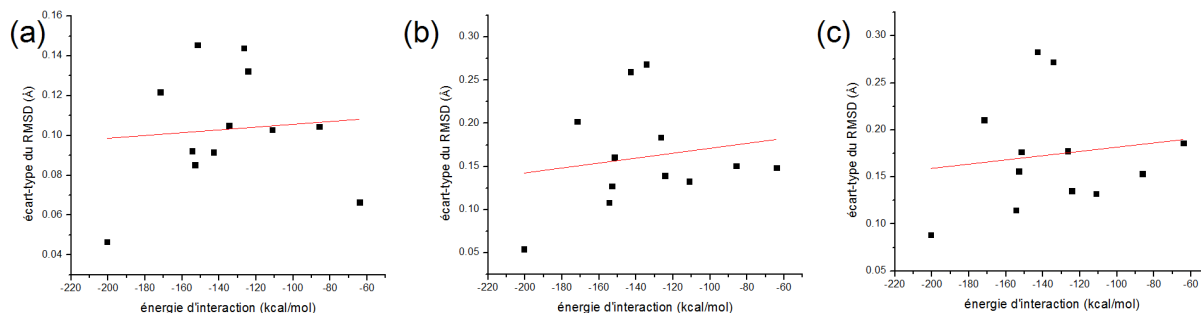


FIGURE 3.19. Relation entre l'énergie d'interaction non-liée et l'écart-type des RMSD des structures à $d \text{ \AA}$ du centre de masse de DPP6.2 (a) Pour $d = 10 \text{ \AA}$, la régression linéaire (droite rouge) a comme coefficient de corrélation $R = 0.087$ et un R^2 -ajusté = -0.0917 (b) Pour $d = 12 \text{ \AA}$, $R = 0.17$ et R^2 ajusté = -0.0667 (c) Pour $d = 14 \text{ \AA}$, $R = 0.142$ et R^2 -ajusté = -0.07795 .

la signification statistique de l'ajout de variables dans une régression. Son augmentation entre $d = 10 \text{ \AA}$ et $d = 12 \text{ \AA}$ (et entre $d = 10 \text{ \AA}$ et $d = 14 \text{ \AA}$) suggère que l'effet n'est pas dû à une fluctuation statistique. Pour le passage à $d = 14 \text{ \AA}$, il y a une légère baisse dans les coefficients R et R^2 -ajusté par rapport à $d = 12 \text{ \AA}$, mais on ne peut pas dire s'il s'agit *exclusivement* d'un effet de dilution (lequel est certes présent comme le confirme la baisse du R^2 -ajusté) ou si on capture également la déstabilisation fonctionnellement importante d'une sous-structure éloignée. La section suivante visera à décortiquer l'analyse selon les sous-structures.

Conceptuellement parlant, notons que ce résultat n'est pas sans rappeler l'exemple typique de la liaison d'un petit ligand induisant une réorganisation minimale de son site actif, mais des changements conformationnels importants ailleurs dans le récepteur. Par exemple, on peut penser au cas fréquent de deux antagonistes qui compétitionnent pour un même site (ayant donc des interactions locales similaires), mais qui induisent des changements conformationnels correspondant à des états fonctionnels distincts. On sait d'ailleurs qu'il doit exister ce genre « de couplage à distance » dans notre système puisque la présence de DPP6 à la surface de $K_V4.2$ doit (entre autre) accélérer les transitions de l'état ouvert vers les états inactivés-fermés caractéristiques de K_V4 , lesquels sont gérés par des liaisons dynamiques prenant place à l'interne, surtout entre $S4$ - $S5_L$ et $S6_T$. Dans la prochaine section, on regardera donc l'origine des changements conformationnels de $K_V4.2$ en présence de DPP6.2.

3.3.5. Changements conformationnels lors de la formation du complexe et mécanisme d'action.

La section précédente a permis de supporter le choix du modèle ayant la plus forte énergie d'interaction comme étant le plus probable en analysant les données structurales de la région proximale d'interaction. Comme l'énergie interne (conformationnelle, thermique, etc.)

et l'énergie de solvation sont des facteurs importants mais ignorés du calcul des énergies d'interaction, l'analyse de la stabilité structurelle permet de voir leur effet indirectement. Si l'on admet l'hypothèse qu'une plus grande stabilité implique une mode d'interaction plus probable, cette analyse permet de palier partiellement à l'information incomplète, à défaut d'avoir les ressources calculatoires nécessaires aux calculs d'énergies libres de liaison. Cette hypothèse fait du sens intuitivement puisqu'il faut que plusieurs résidus exposés à l'environnement lipidique de la membrane passe d'un état de fluctuations thermiques « libres » à un état de lié. Curieusement, cependant, l'analyse a montré que la stabilité accrue due à l'interaction prenait plus d'importance avec la distance, et diminuait par la suite. Cette section vise donc à identifier comment les sous-structures fonctionnelles sont impactées et quelle est l'origine du résultat de la section précédente. À cette fin, on compare les sous-structures pertinentes du complexe protéique à une simulation contrôle dans laquelle les peptides ne sont pas liés. On calculera le degré auquel une sous-structure subit un changement conformationnel de deux façons. La première est de comparer le changement relatif du RMSD d'un état à l'autre (lié vs non-lié) en calculant ΔR , qu'on définit par

$$\Delta R := \frac{\langle \text{RMSD}^{\text{lié}} \rangle - \langle \text{RMSD}^{\text{non-lié}} \rangle}{\langle \text{RMSD}^{\text{non-lié}} \rangle}, \quad (3.3.1)$$

où le $\text{RMSD}^{\text{non-lié}}$ est la moyenne des quatre sous-unités de l'expérience contrôle. Le tableau 3.3.5 compile les valeurs des RMSD, de leurs écarts-types et le ΔR pour chaque sous-structure du modèle $\phi_i = 300^\circ$, soit le plus stable et le plus fortement lié. Consistant avec la section précédente, on constate bel et bien que la réorganisation des sous-structures n'est pas uniforme. L'hélice externe du pore S5 ne se réorganise pas significativement ($\Delta R = 6.1\%$, soit dans les marges d'erreur), contrairement par exemple à S1 du VSD ($\Delta R = 35.0\%$), malgré un contact direct similaire entre ces deux structures et DPP6.2. On voit également que le S4S5L et le pore S6_T subissent des réorganisations importantes. Ces résultats sont consistants avec notre hypothèse que seules les sous-structures responsables de l'activation et l'inactivation seront affectées. Nous convenons qu'en pratique, cela ne laisse que le S5, le haut de S6 et le filtre, mais il n'était pas évident que notre méthode de liaison manuelle et forcée n'allait pas perturber l'ensemble des structures. Au contraire, on constate même que S4S5L reste plus prêt de sa structure cristalline dans l'état lié qu'à l'état non-lié.

On observe que ces résultats expliquent une partie¹³ des variations dans les coefficients de corrélation R avec la distance de la section précédente. La région d'interaction à $d = 10\text{\AA}$

13. L'autre partie étant le bruit attendu des acides aminés isolés et autres structures capturées mais n'ayant peu d'impact fonctionnel sur le couplage électromécanique et l'activation, comme le P-loop ou le haut du S6. Tel qu'expliqué, les variations du R^2 -ajusté supporte cette affirmation.

	S1	S4	S4S5L	S5	S6 _T
RMSD ^{lié} (Å)	2.641 ± 0.071	2.995 ± 0.055	1.551 ± 0.064	1.638 ± 0.113	2.951 ± 0.115
RMSD ^{non-lié} (Å)	1.956 ± 0.021	2.821 ± 0.043	1.827 ± 0.040	1.813 ± 0.074	1.918 ± 0.030
ΔR (%)	35.0 %	6.1%	-15.1%	-9.6%	53.9%

TABLEAU 3. 6. Comparaison des RMSD moyens des sous-structures fonctionnelles à l'état lié et non-lié par ΔR .

3.17 n'incluait qu'un acide aminé du S4S5 et trois du S6_T. Comme S4S5L est perpendiculaire et S6_T est incliné par rapport au TM de DPP6, on capture plus rapidement le S4S5L avec la distance, ce qui fait augmenté la corrélation car, comme le montre le tableau 3.3.5, celui-ci est plus *contraint* vers la structure cristalline, jusqu'à ce qu'on capture assez du S6 pour la faire chuter légèrement.

Bref, la stabilité supérieure de $\phi_i = 300^\circ$ semble surtout provenir de cette région clé pour le couplage électromécanique, où interagissent S4S5L et S6_T. Étant donné que S4S5L se voit contraint de rester plus prêt de sa structure cristalline et que S6_T s'en éloigne davantage lors de l'équilibration en présence de DPP6.2, nous suggérons que cela *favorise* la liaison dynamique entre ces structures après l'activation (rappelons que le canal est dans l'état ouvert dans ces simulations). Comme les états fermés-inactivés atteints dans cette liaison dynamique est à l'origine de l'inactivation rapide, cela rendrait compte de l'accélération des cinétiques de l'inactivation. Pour vérifier cette hypothèse numériquement, il faudrait cependant recommencer les simulations seulement avec le modèle $\phi_i = 300^\circ$ en évitant les contaminations des autres angles car les états fermés-inactivés dépendent autant des liaisons intra-sous-unités que inter-sous-unités. Nous avons tout de même calculer l'énergie d'interaction entre le S4S5L et le S6_T d'une même sous-unité avec et sans la présence de DPP6.2 (soit le modèle $\phi_i = 300^\circ$ vs simulation contrôle), et le résultat est prometteur. L'énergie d'interaction avec DPP6 était de -147 ± 20 kcal/mol, alors que sans le TM, on obtient -109 ± 10 kcal/mol. Cela vient corroborer une hypothèse de Jerng et Pfaffinger (2014) a l'effet que DPP6 change la structure du pore, qui se voit davantage contraint, favorisant l'interaction des résidus à l'origine des états fermés-inactivés.

Concernant l'activation plus rapide que cause DPP6, il faut rendre compte de deux composantes, soit le décalage de la courbe de conductance vers les potentiels hyperpolarisés et l'atteinte plus rapide du courant maximum à un potentiel donné. Or les présents résultats suggèrent que S1 doit se réorganiser en présence de DPP6.2. Étant donné qu'il possède deux ensembles de charges stabilisant la surface positive de S4 (un en haut et l'autre en bas), nous faisons l'hypothèse que cette réorganisation viendra abaisser la barrière d'énergie nécessaire pour que les charges de S4 fasse le « saut » d'un « étage » de charge à l'autre.

Conséquemment, si les S4 sont moins « figés » en présence de DPP6, les interactions stabilisatrices opérant en son sein seraient plus facilement brisées et, conséquemment, la probabilité d'ouverture sera augmentée (ce qui équivaut de dire que la courbe de conductance est décalée vers les potentiels hyperpolarisants). Aussi, si le couplage électromécanique est renforcé par un effet de constriction du pore, l'activation plus rapide du VSD entraînera l'ouverture plus rapide du canal, et donc l'atteinte du courant maximal plus rapide. Tout ceci ne reste cependant que des hypothèses à vérifier. Il demeure des plus cavaliers d'espérer inférer des mécanismes *dynamiques* avec précision à partir de calculs moyennés indirects dans un état d'*équilibre*. D'un point de vue numérique, pour renforcer la confiance en cette hypothèse, il faudrait calculer la valeur du champ électrique local au sein du VSD réorganisé en présence de DPP6.

3.3.6. Discussion

En résumé, on a dans un premier temps prédit les structures transmembranaire et N-terminale de DPP6.2 avant de les soumettre à des simulations d'amarrage moléculaire approprié en fonction de leur environnement. Pour l'extrémité N-terminale, nos résultats suggèrent que la liaison s'effectue surtout au niveau d'une hélice α constitué d'une dizaine d'acide aminé (résidus 12 à 23 pour DPP6.2), laquelle viendrait se lier à l'intersection de deux sous-unités du domaine T1 grâce à une combinaison d'interactions hydrophobiques et électrostatiques. Entre autre, les acides aminés L66, F74, F75 et F121 de T1 créent une pochette hydrophobique qui épouse la topologie des résidus E20, L21 et V22 de DPP6.2. Une autre interaction hydrophobique contribue à la liaison, soit la valine V14 de DPP6.2 qui vient s'insérer dans une pochette créée par W55 et F84 de T1. Quant interactions électrostatiques, nos simulations suggèrent l'existence de ponts salins entre K12 de DPP6.2 et D57 de Kv4.2 et entre le E18 de DPP6.2 et R47 et/ou R100 de Kv4.2.

Il se trouve que plusieurs de ces acides aminés ont été caractérisés d'un point de vue fonctionnel. On savait que les acides aminés E18 et E20 du motif conservé ELVG joue un rôle dans l'augmentation de la conductance unitaire (Kaulin et al., 2009) et que l'homologue de V14 et E18 (i.e. M12 et V16 de DPP6.3) jouent un rôle dans la modulation des cinétiques (Jerng et Pfaffinger, 2014). Cette recherche permet de prédire leur partenaire d'interaction et d'identifier la pose plus précisément. De plus, elle a permis d'établir que l'extrémité N-terminal composée de aa. 1 à 13 était libre de toute topologie rigide et disponible pour interagir davantage. Nous avançons qu'il est possible qu'il y ait interaction avec la chaîne et la balle libre de Kv4. Il faudrait pour cela construire un modèle bioinformatique de cette structure, ou mieux encore, effectuer des études fonctionnelles et/ou de FRET en laboratoire. Notons aussi que nos résultats expliquent pourquoi certains isoformes de DPP

compétitionnent avec KChIP, puisque cette région est la même que le site 2 identifié dans la structure cristalline du complexe Kv4.3-KChIP1 (Pioletti, Findeisen, Hura et Minor, 2006). D'ailleurs la pochette dont fait partie F73, F74 et F119 est à proximité des aa. E70, F73 et D78 (conservés entre Kv4.2 et Kv4.3), dont on sait qu'ils influencent la modulation des cinétiques de Kv4.3 par KChIP (Scannevin et al., 2004; Callsen et al., 2005).

Pour ce qui est de la partie transmembranaire de l'interaction, aucune information était connue avec une résolution moléculaire. Des indices indiquaient une interaction partielle avec S1 et S2, mais rien sur les acides aminés en soi. Or, nos simulations d'amarrage pointaient vers la pochette créée par S1, S5 et une partie du S4-S5L, mais l'orientation de la surface des modèles était ambiguë puisque l'orientation des modèles était équidistribuée sur le cercle unitaire. Pour discriminer entre ces modèles et tenir compte de la flexibilité de l'hélice α , on a soumis douze modèles séparés d'une rotation axiale de 30° degré les uns des autres (de 0° à 330°) à la dynamique moléculaire. Après 70 ns, la plupart des modèles étaient équilibrés et on les a comparés sur la base de l'énergie d'interaction non-liée et leur stabilité structurelle. Dans les deux cas, $\phi_i = 300^\circ$ était le candidat qui se distinguait des autres par sa plus grande affinité et sa stabilité. Cela a ensuite permis deux choses : formuler des prédictions sur les interactions stabilisatrices et étudier comment la conformation de Kv4 s'adapte en présence de DPP6. En regardant les structures les plus affectées, on a constaté que S1, S4-S5L et S6_T devaient se réorganiser davantage.

On a suggéré que les changements structurels de S1, dont on sait qu'il joue entre autre le rôle de stabiliser la surface chargée de S4 par la présence de charge négative, déstabilisent le VSD et abaissent la barrière d'énergie nécessaire à l'activation des S4. Pour le confirmer, cependant, il faudrait calculer comment cette réorganisation du S1 influence le champ électrique local. Par contre, à supposer que le S4 ressent bel et bien un champ plus puissant à un potentiel donné, l'instabilité a un impact dans les deux sens : sans un couplage électromécanique plus fort, cette déstabilisation du VSD voudrait aussi dire que les S4 peuvent se désactiver plus rapidement sans provoquer l'ouverture du pore. Or nos calculs suggéraient que S4S5L était plus contraint vers sa structure cristalline en présence de DPP6.2 qu'en son absence, alors que l'inverse était vrai pour S6_T, probablement pour compenser. On a donc regardé ce que cela impliquait d'un point de vue énergétique, pour constater que l'énergie d'interaction entre S4S5L et S6_T était environ 50% plus grande en présence de DPP6.2. Cela vient démontrer (du moins numériquement) une hypothèse qu'avait formulée Jerng et Pfaffinger (2014) à l'effet que l'ouverture du pore était contraint par DPP6 et que cela justifierait le fait que les états fermés-inactivés étaient favorisés en sa présence.

Cependant, nos expériences et analyses ont plusieurs limites évidentes. Il y a bien sûr le fait que les modèles structurels ne sont que des prédictions, *ab initio* pour DPP6 et d'homologie pour Kv4.2. Comme le modèle d'homologie est basé sur Kv1, Kv2 et Shaker, on n'a pas la structure exacte des unités responsables du couplage électromécanique. Or, plusieurs différences importantes existent. À défaut d'avoir ces structures, une amélioration évidente des présents travaux seraient d'éviter les contaminations entre les angles en mettant plusieurs configurations différentes sur le même canal. Nous l'avons fait par soucis d'économie des ressources en espérant que l'impact ne sera pas trop grand. Nous avons confiance que nous avons gagné ce pari puisque les grandes lignes sont consistantes avec la littérature (dans la mesure du possible avec le peu d'information) et les résultats semblent pointer vers un mécanisme d'action sensé. Mais pour un calcul plus raffiné à l'échelle des acides aminés clés dans les interactions inter- et intra- sous-unités au niveau du pore pour le couplage électromécanique, il faudrait utiliser la vraie orientation aux quatre sites actifs et avoir la vraie structure du pore. Pour l'avenir, nous suggérons d'utiliser ces résultats et les prédictions qui en découlent pour explorer la nature des interactions en laboratoire. Cette information contraindra davantage des simulations et améliorer leur réalisme.

Chapitre 4

CONCLUSION

Les canaux potassiques voltage-dépendants jouent plusieurs rôles physiologiques fondamentaux dans le système nerveux de toutes les espèces vivantes qui en sont doté. Si certains isoformes sont impliqués dans la transduction du potentiel d'action lui-même – ce signal électrique voyageant le long des axones des neurones pour relayer de l'information entre ceux-ci –, plusieurs autres produisent des courants qui viennent moduler les propriétés du potentiel d'action. C'est le cas des courants I_{SA} et $I_{to,f}$, qui sont tous les deux médiés par des complexes protéiques constitués de la sous-unité α formeuse de pore K_V4 et des sous-unités β auxiliaires DPP et KChIP. Ces sous-unités sont nécessaires pour accélérer les cinétiques de K_V4 , une propriété essentielle à sa capacité de réguler les autres courants de type A comme ceux médiés par la famille *Shaker*. Grâce à l'interaction avec DPP6 et 10, K_V4 voit son activation, son inactivation et sa récupération de l'inactivation s'accélérer considérablement. Par contre, le mécanisme d'action exact derrière le facteur accélérant des DPP reste majoritairement inconnu.

Cette recherche avait donc pour objectif de mieux définir l'origine structurelle permettant à DPP6.2 d'accélérer les cinétiques de $K_V4.2$, surtout au niveau des segments transmembranaires, soit la partie la plus importante. Pour ce faire, il fallait dans un premier temps construire des modèles théoriques. En effet, seuls le domaine intracellulaire T1 de $K_V4.2$ et le large domaine extracellulaire de DPP6.2 ont été cristallisés. Or il se trouve que ce dernier ne contribue seulement qu'à l'augmentation du niveau d'expression de $K_V4.2$, et non à la modulation de ses cinétiques. Il fallait donc avoir recours à différents outils de modélisation pour compléter les structures des protéines, avant de construire un modèle « complet » pour les interactions. Pour $K_V4.2$, MODELLER a utilisé pour recoller la structure cristalline de T1 au modèle d'homologie de la partie transmembranaire, construit par Heler, Bell, et Boland (2013) à partir de la structure cristalline de $K_V1.2-2.1$ et KcsA. Pour DPP6.2, l'analyse séquentielle a permis d'identifier la partie de la séquence primaire située dans la membrane et quelle partie était au niveau cytosolique. L'algorithme de prédiction structurelle ab initio QUARK a ensuite permis de prédire un ensemble de modèle pour les deux parties.

Pour le segment TM, il suffisait de prendre le modèle correspondant à l'hélice α complète. Pour l'extrémité N-terminale, après avoir classé les modèles selon leur topologie, on a dû faire recours à la dynamique moléculaire pour démontrer l'équivalence de deux familles sur les trois suggérées. L'analyse des trajectoires a permis de confirmer la flexibilité des deux chaînes attachées à leur hélice centrale, laquelle est commune et stable pour toutes les familles. Comme cette hélice présente une séquence hautement conservée parmi les isoformes de DPP6, elle est un bon candidat pour médier des interactions fonctionnelles. Pour la troisième famille, la DM a montré que la torsion du modèle QUARK n'était pas stable et

que celle-ci est rapidement détendue, donnant naissance à une hélice centrale plus longue de cinq acides aminés. Par contre, l'ajout de ces cinq résidus supplémentaires n'influencent pas les résultats des simulations d'amarrage, qui identifient (après élimination des modèles irréalistes) l'interface des sous-unités du domaine T1 comme la région d'interaction la plus probable dans les deux cas.

Ces surfaces d'interactions, en plus d'être hautement complémentaires, sont composées d'un mélange d'interactions hydrophobiques et électrostatiques. En particulier, la lysine K12 de DPP6.2 semble établir un pont salin avec l'aspartame D57 du domaine T1, alors que le glutamate E18 de DPP6.2 est probablement attiré, lors de la formation, par la surface négative créée par les arginines R47 et R100 de T1. Ce E18 (ainsi que E20 du motif ELVG conservé, voir plus loin) a été identifié comme jouant un rôle déterminant pour l'augmentation de la conductance unitaire (Kaulin et al., 2009). Pour ce qui est des interactions hydrophobiques, elles semblent surtout confinées dans deux régions : d'abord, une surface créée par les acides aminés L66, F74, F75 et F121 de T1, sur laquelle viennent se déposer les résidus E20 (du moti 'ELVG'), L21 et V22 de DPP6.2 ; puis enfin, la valine V14 de DPP6.2 semblent s'insérer dans une pochette créée par W55 et F84 de T1. L'alignement séquentiel des isoformes de DPP6 révèle que la valine V14 et le glutamate E18 de DPP6.2 sont respectivement homologues à la méthionine M12 et la valine V16 de DPP6.3, lesquels sont deux résidus ayant été identifiés comme jouant un rôle crucial dans la modulation des cinétiques de K_V4 (Jerng et Pfaffinger, 2014).

Il s'avère que l'homologue $K_V4.3$ de la région d'interaction identifié par les simulations d'amarrage est déjà connu pour interagir avec KChIP1, tel que le montre la structure cristalline du complexe (Pioletti, Findeisen, Hura et Minor, 2006) (voir figure 4.1). La surface hydrophobique homologue (dont F73, F74 et F119 font partie) s'agit du site 2 identifiés par les auteurs. En particulier, la structure cristalline vient confirmer une interaction directe avec les résidus E70, F73 et D78 (conservés chez $K_V4.2$), cruciaux pour la modulation des cinétiques de $K_V4.3$ par KChIP (Scannevin et al., 2004 ; Callsen et al., 2005). Étant donné que notre modèle suggère que la boucle de $K_V4.2$ comprenant les résidus homologues E71 et F74 est en contact direct avec DPP6.2, cela pourrait expliquer la compétition entre DPP6.2 et KChIP observée dans les courants en situation de co-expression (fig. 4.2). On sait d'ailleurs que cette trace est mieux reproduite par deux constantes de temps (Jerng, Dougherty, Covarrubias et Pfaffinger, 2009).

Comme le montre la figure 4.2, différents isoformes peuvent avoir divers effets sur les cinétiques de K_V4 , particulièrement à cause des variantes des N-terminaux. Selon l'intensité de l'interaction, il est y avoir différent effet sur les modulations en présence de coexpression

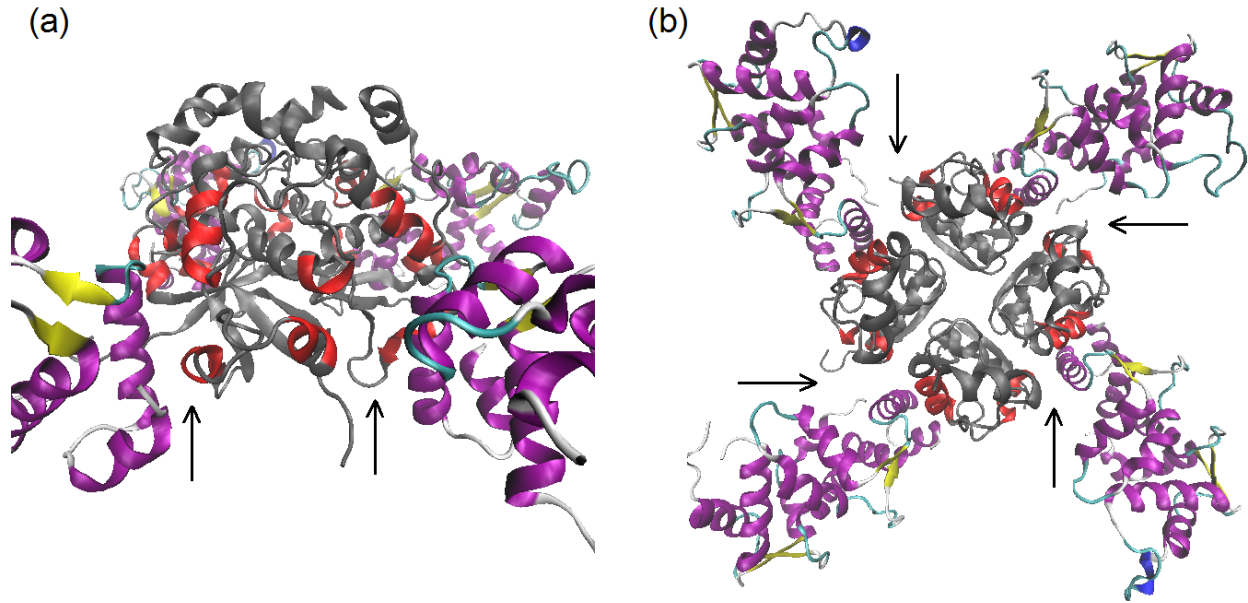


FIGURE 4.1. Structure cristalline du domaine T1 de $K_V4.3$ (en gris et rouge) et KChIP1 (en mauve et jaune). Les acides aminés de l'interface entre deux sous-unité de T1 (en rouge) délimite la région d'interaction avec DPP6.2 identifiée par la présente recherche, laquelle correspond à un des sites d'interaction avec KChIP (site 2 identifié par Pioletti, Findeisen, Hura et Minor, 2006, voir flèches).

des sous-unités β . Dans la figure 4.2, on voit que les effets de DPP10 dominant ceux de KChIP1 (de même que DPP6.3 et DPP6a, alors que l'effet inverse se produit pour DPP6.2, par exemple (Jerng et Pfaffinger, 2011)), mais il est probable qu'il s'agisse quand-même d'un assemblage des trois protéines (Jerng, Hunjilwar et Pfaffinger, 2005). Cependant, c'est surtout le domaine TM des DPP qui confère la majorité des effets fonctionnelles de cette interaction (Jerng et Pfaffinger, 2014). Pour mieux l'étudier, un modèle de l'hélice α a été construit grâce à QUARK pour le soumettre à l'algorithme d'amarrage MemDock. Il a identifié une seule région d'interaction, en contact avec S1, S5 et des parties de S4 et S4-S5L. Ceci est partiellement en accord avec certains résultats préliminaires indiquant des interactions avec des portions de S1 et S2, sans pourtant identifier le site exact (Ren et al., 2003). Par contre, l'orientation axiale (et, dans une moindre mesure, verticale) des modèles suggérés par MemDock était hautement variable. Pour vérifier laquelle de ces orientations est la plus probable de médier les interactions de DPP6 et pour tenir compte de la flexibilité de l'hélice, on a construit douze modèles séparés par une rotation axiale de 30° qu'on a équilibré avec la dynamique moléculaire. Plusieurs méthodes ont permis d'identifier le candidat principal.

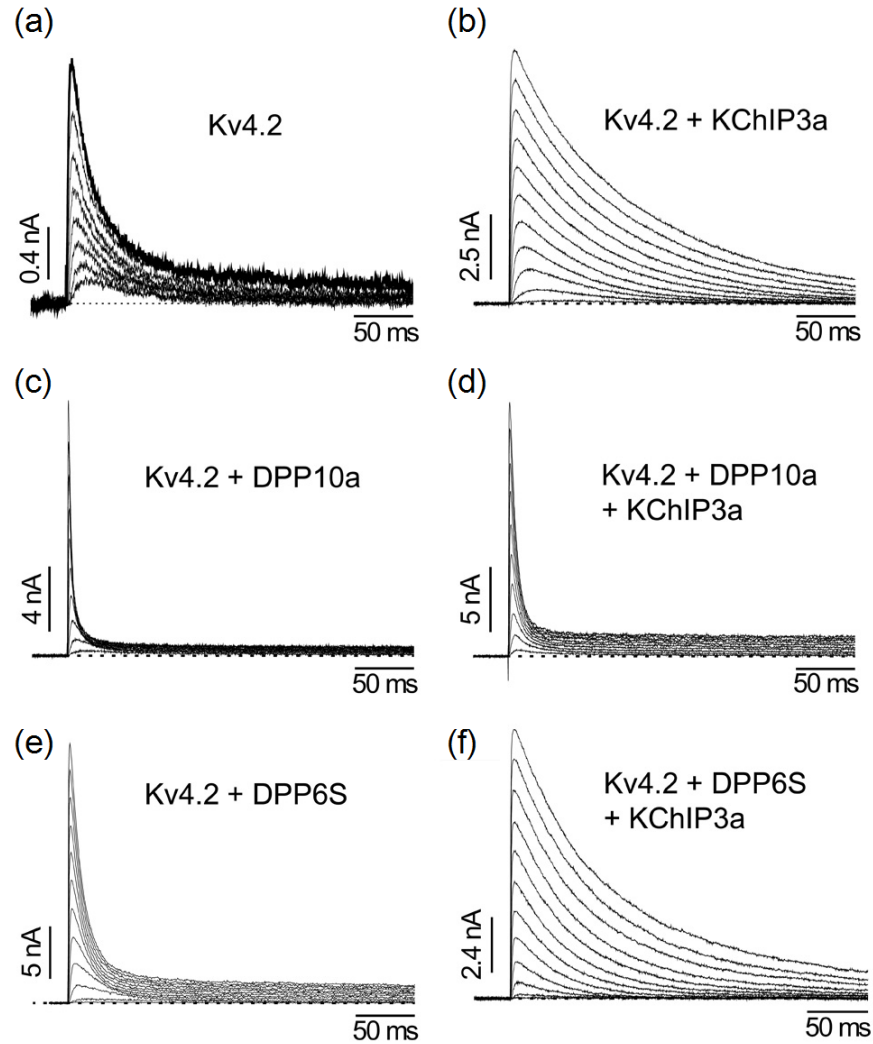


FIGURE 4.2. Effets des différentes sous-unités β de Kv4.2 sur ces cinétiques. (a) Courants de $K_V4.2$ isolé suite à une dépolarisation de cellules CHO. (b) La coexpression avec KChIP3a augmente la densité de courant mais ralentit les cinétiques d'inactivation. (c) La coexpression avec DPP10a accélère les cinétiques d'inactivation. (d) En coexprimant $K_V4.2$ avec DPP10 et KChIP3, on observe une accélération similaire à $K_V4.2$ +DPP10, montrant que KChIP3 interfère peu avec le mécanisme accélérant. Ce n'est pas le cas pour DPP6S (= DPP6.2), qui accélère bel et bien les cinétiques de $K_V4.2$ lorsqu'exprimé seul avec $K_V4.2$ (e), mais qui ne produit pas d'effet lorsqu'il est coexprimé avec KChIP3a (f). (adapté de Jerng, Hunjilwar et Pfaffinger, 2005)

Dans un premier temps, étant donné que l'angle de rotation axiale est le degré de liberté qui influence le plus la nature chimique de la surface de contact, nous avons formulé l'hypothèse que les modèles allaient converger vers quelques orientations finales communes, et que celles-ci correspondraient à des minimums locaux de l'énergie d'autant plus probables qu'il y a de modèles adoptant cette orientation. Nous avons pu vérifier la première partie

de l'hypothèse en montrant qu'il existait trois orientations (ou classes d'équivalence, soit $\phi_f = 0^\circ, 60^\circ, 260^\circ$) favorables à plus d'un modèle, accumulant chacun le quart des TM équilibrés. Cependant, la deuxième partie de l'hypothèse a été infirmée par le calcul des énergies d'interaction non-liée. Ces dernières étaient en effet très variables pour deux des trois classes. La classe $\phi_f = 0^\circ$, par exemple, comprenait le modèle ayant une des interactions les plus faibles ($\phi_i = 0^\circ$ à (-86 ± 13) kcal/mol) et celui ayant l'interaction la plus forte ($\phi_i = 300^\circ$ à (-201 ± 15) kcal/mol). Nous avons montré qu'un facteur important pour expliquer ces grandes différences malgré une même orientation axiale était l'orientation verticale, et notamment à quel point le TM épouse bien la topologie de son environnement. À ce moment, nous étions aux prises avec deux limitations majeures de cette approche basée sur la population des classes d'équivalence et l'énergie d'interaction non-liée : 1) nous n'avons pas assez d'échantillons pour augmenter le nombre de degrés de liberté considérés dans la classification des modèles et 2) l'énergie d'interaction ne considère pas (entre autre) l'énergie conformationnelle et l'énergie de solvation, deux contributions importantes l'énergie libre de liaison. Pour combler à ces limites, nous nous avons regardé la structure du complexe et sa stabilité, une mesure indirecte mais qui inclut l'ensemble des contributions.

Nous avons voulu tester l'hypothèse à l'effet que le modèle ayant l'énergie d'interaction la plus forte était aussi celui dont la région d'interaction était la plus stable. Celle-ci a été confirmée avec une grande certitude par le calcul de la déviation du RMSD des régions d'interaction pour les douze modèles. En effet, le TM $\phi_i = 300^\circ$ était au moins deux fois plus stable que (presque) tous les autres modèles pour la région directe d'interaction. Cette information structurale a permis de corroborer notre choix basée uniquement sur l'énergie d'interaction non-liée.

Cependant, lorsqu'on s'éloignait du centre de masse du TM pour inclure des régions indirectes, la corrélation entre la stabilité moyenne et l'énergie d'interaction augmentait vers un maximum et rediminuait légèrement, mais toujours à un niveau supérieur que la région de contact direct. Pour comprendre ce résultat, on a donc analysé les changements conformationnels et la stabilité de chaque sous-structure fonctionnelle séparément. Nous avons montré que S1 doit se réorganiser et que le S4-S5L et le S6_T sont plus contraints les uns sur les autres. Basé sur ces résultats, on suggère dans un premier temps que la réorganisation déstabilise le VSD et réduit la barrière d'énergie pour l'activation des S4. Cette déstabilisation peut expliquer le décalage vers les potentiels hyperpolarisés de la courbe de conductance et l'atteinte plus rapide du courant maximal en permettant aux S4 individuels de briser les liaisons électrostatiques stabilisatrices plus aisément, diminuant donc le temps nécessaire pour que les quatre soient activés simultanément. Cependant, tout ceci n'est possible que si l'étape concertée suivant l'activation des quatre VSD est également plus rapide/sensible. On a donc vérifié l'implication énergétique de cette réorganisation

du pore pour conclure que l'interaction entre S4S5L et S6_T était bel et bien plus intense à cause de DPP6. Ce résultat sur le S4-S5L et le S6_T vient appuyer et étendre une hypothèse déjà postulée dans la littérature à l'effet que le segment transmembranaire des DPP ont un impact sur la constriction¹ du pore (Jergn et Pfaffinger, 2014). Les auteurs se basent sur le fait les DPP ont des effets sur la conductance unitaire et la liaison avec des toxines, deux propriétés qui dépendent de la structure du pore. Ils suggèrent que ces effets de constriction restreignent les états intermédiaires accessibles au canal durant le processus de *gating*, accélérant du coup les cinétiques des transitions (Jergn et Pfaffinger, 2014). Mettant le tout ensemble, on voit bien comment l'accélération a lieu : les S4 s'activent plus rapidement, le couplage plus grand catalyse l'atteinte du courant maximal *et* la liaison dynamique des résidus du pore responsable des états fermés-inactivés, accélérant également l'inactivation.

Enfin, plusieurs limites à cette étude demandent prudence lors de la formation de tels modèles pour le mécanisme d'action. On a déjà parlé du fait que quelques acides aminés clés peuvent parfois faire toute la différence. C'est le cas pour plusieurs mécanismes moléculaires, dont le couplage électromécanique et les états fermés-inactivés, qui ont des bases moléculaires très similaires . Or, toutes les simulations pour la partie transmembranaire reposent sur un modèle d'homologie basé sur la structure cristalline du chimère K_v1.2-2.1, qui présente des différences structurelles au niveau des résidus responsables du couplage électromécanique. Par exemple, dans la structure "cristalline" du modèle, les aa I312 et T316 (homologues de I384 et T388 de Shaker) s'insèrent très bien dans la pochette de F481 et F484, mais les données expérimentales indique que ces interactions n'influent pas les états fermés-inactivés, que G309, plutôt que I et T, vient fortement interagir alors qu'ils ne font pas contact dans le cristal. Idéalement, donc, il aurait fallu une structure cristalline haute résolution plutôt qu'un modèle d'homologie.

Pour améliorer la modélisation, davantage de données expérimentales sont nécessaires pour contraindre les modèles. On espère que ces travaux aideront un peu dans le design des futures recherches nécessaires. De cette façon, il serait possible d'utiliser le module de variable collectives (*colvars*) de NAMD pour simuler, en des temps raisonnables, le processus d'amarrage dès le départ, sans forcer des positions initiales déjà liées. Pour confirmer le mécanisme proposé par ces travaux, il serait intéressant de suivre la *dynamique* de chaque sous-structure fonctionnelle lors du processus de séparation plutôt que comparer des expériences indépendante dans un état d'équilibre.

1. traduction de *packing*

BIBLIOGRAPHIE

Ainavarapu, S. R. K., Brujić, J., Huang, H. H., Wiita, A. P., Lu, H., Li, L., ... Fernandez, J. M. (2007). Contour Length and Refolding Rate of a Small Protein Controlled by Engineered Disulfide Bonds. *Biophysical Journal*, 92(1), 225–233.

An, W. F., Bowlby, M. R., Betty, M., Cao, J., Ling, H. P., Mendoza, G., et al. (2000) Modulation of A-type potassium channels by a family of calcium sensors. *Nature*, 403, 553–6.

Batulan, Z., G. A. Haddad, and R. Blunck. (2010). An intersubunit interaction between S4-S5 linker and S6 is responsible for the slow off gating component in Shaker K⁺ channels. *J. Biol. Chem.*, 285 :14005–14019.

Blunck, R., Batulan, Z. (2012). Mechanism of Electromechanical Coupling in Voltage-Gated Potassium Channels. *Front Pharmacol*, 3(166).

Bahring, R., J. Barghaan, J., Westermeier, R., Wollberg, J. (2012). Voltage sensor inactivation in potassium channels. *Front Pharmacol*, 3(100).

Barghaan, J., and R. Bahring. (2009). Dynamic coupling of voltage sensor and gate involved in closed-state inactivation of Kv4.2 channels. *J. Gen. Physiol.*, 133 :205–224.

Birnbaum, S. G., Varga, A. W., Yuan, L. L., Anderson, A. E., Sweatt, J. D., Schrader, L. A. (2004). Structure and function of Kv4-family transient potassium channels. *Physiol Rev*, 84, 803–33.

Callsen, B., Isbrandt, D., Sauter, K., Hartmann, L. S., Pongs, O. and Bahring, R. (2005) Contribution of N- and C-terminal channel domains to Kv4.2 domains to KChIP interaction. *J. Physiol.*, 568(2), 397-412.

Comeau, S. R, Gatchell, D. W., Vajda, S., Camacho, C. J. (2004). ClusPro : an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics*, 20(1), 45-50.

Comeau, S. R, Gatchell, D. W., Vajda, S., Camacho, C. J. (2004). ClusPro : a fully automated algorithm for protein-protein docking. *Nucleic Acids Research*, 32(Web Server issue), W96-9.

Chipot, C. (2008). Free energy calculations applied to membrane proteins. *Methods in Molecular Biology*, 443, 121-144.

Chowdhury, S., B. M. Haehnel, and B. Chanda. (2014). Interfacial gating triad is crucial for electromechanical transduction in voltage-activated potassium channels. *J. Gen. Physiol.*, 144, 457–467.

Deng, Y., Roux, B. (2009) Computations of Standard Binding Free Energies with Molecular Dynamics Simulations. *J Phys Chem B*, 113(8).

Dougherty, E. (2015). Computing Cellular Clockworks : Klaus Schulten. *SBGrid Consortium*. President and Fellows of Harvard College.

Duhovny, D., Nussinov, R. and Wolfson, H. J. (2002). Efficient unbound docking of rigid molecules. *Workshop on Algorithms in Bioinformatics*, 2452, 185-200.

E. Mashiach, R. Nussinov and H. J. Wolfson. (2009). FiberDock : Flexible induced-fit backbone refinement in molecular docking. *Proteins*, 78(6) :1503-1519.

Fiser, A., Do, R. K. et Sali, A. (2000) Modeling of loops in protein structures, *Protein Science*, 9, 1753-1773.

Haddad, G. A. and R. Blunck. (2011). Mode shift of the voltage sensors in Shaker K⁺ channels is caused by energetic coupling to the pore domain. *J. Gen. Physiol.* 137, 455–472.

Heler, R., Bell, J. K. et Boland, L. M. (2013). Homology model and targeted mutagenesis identify critical residues for arachidonic acid inhibition of Kv4 channels. *Channels*, 7(2), 72-84.

Holmqvist, M. H., Cao, J., Hernandez-Pineda, R., Jacobson, M. D., Carroll, K. I., Sung, M. A., et al. (2002). Elimination of fast inactivation in Kv4 A-type potassium channels by an auxiliary subunit domain. *Proc. Natl. Acad. Sci. U.S.A.*, 99, 1035–1040. doi :10.1073/pnas.022509299

Huang, J., Rauscher, S., Nawrocki, G., Ran T., Feig, M., de Groot, B. L., Grubmüller, H., MacKerell Jr, A. D. (2017). CHARMM36m : an improved force field for folded and intrinsically disordered proteins. *Nature Methods*, 14, 71–73.

Hurwitz, N., Schneidman-Duhovny, D., Wolfson, H. J. (2016). Memdock : an α -helical membrane protein docking algorithm. *Bioinformatics*. 32(16), 2444–2450

Jerng, H. H., Pfaffinger, P. J., Covarrubias, M. (2004). Molecular physiology and modulation of somatodendritic A-type potassium channels. *Mol Cell Neurosci*, 27, 343–69.

Jerng, H. H., Dougherty, K., Covarrubias, M., and Pfaffinger, P. J. (2009). A novel N-terminal motif of dipeptidyl peptidase-like proteins produces rapid inactivation of Kv4.2 channels by a pore-blocking mechanism. *Channels*, 3, 448–461. doi : 10.4161/chan.3.6.10216

Jerng, H. H., Kunjilwar, K., and Pfaffinger, P. J. (2005). Multiprotein assembly of Kv4.2, KChIP3 and DPP10 produces ternary channel complexes with ISA-like properties. *J. Physiol*. 568, 767–788. doi :10.1113/jphysiol.2005.087858

Jerng, H. H., Lauver, A. D., and Pfaffinger, P. J. (2007). DPP10 splice variants are localized in distinct neuronal populations and act to differentially regulate the inactivation properties of Kv4-based ion channels. *Mol. Cell. Neurosci*. 35, 604–624. doi :10.1016/j.mcn.2007.03.008

Jerng, H. H., and Pfaffinger, P. J. (2008). Multiple Kv channel-interacting proteins contain an N-terminal transmembrane domain that regulates Kv4 channel trafficking and gating. *J. Biol.Chem*. 283, 36046–36059. doi :10.1074/jbc.M806852200

Jerng, H. H., and Pfaffinger, P. J. (2011). The cytoplasmic N-terminal domain of DPP6K disrupts KChIP modulation of Kv4 channels. *Biophys. J.*, 100, 98a. doi : 10.1016/j.bpj.2010.12.742

Jerng, H. H., and Pfaffinger, P. J. (2012). Incorporation of DPP6a and DPP6K variants into ternary Kv4 channel complex reconstitutes properties of A-type K current in rat cerebellar granule cells. *PLoS ONE* 7 :e38205. doi :10.1371/journal.pone.0038205

Jerng, H. H., Pfaffinger, P. J., and Covarrubias, M. (2004a). Molecular physiology and modulation of somatodendritic A-type potassium channels. *Mol. Cell. Neurosci.*, 27, 343–369. doi :10.1016/j.mcn.2004.06.011

Jerng, H. H., Qian, Y., and Pfaffinger, P. J. (2004b). Modulation of Kv4.2 channel expression and gating by dipeptidylpeptidase10(DPP10). *Biophys. J.*, 87, 2380– 2396.

doi :10.1529/biophysj.104.042358

Jo, S., Cheng X, Lee, J., Kim, S., Park, S. J., Patel, D. S., Beaven, A. H., Lee, K. I., Rui, H., Park, S., Lee, H. S., Roux, B., MacKerell, A. D. Jr, Klauda, J. B., Qi, Y., Im, W. (2017) CHARMM-GUI 10 years for biomolecular modeling and simulation. *J Comput Chem*, 38(15), 1114-1124.

Jones, D. T. (1998). Do transmembrane protein superfolds exist? *FEBS Lett.*, 27(423), 281 – 285.

Kaczor, A. A., Selent, J., Sanz, F., Pastor, M. (2013). Modeling Complexes of Transmembrane Proteins : Systematic Analysis of Protein-Protein Docking Tools. *Mol. Inform.*, 32(8), 717-733.

Kassiri, Z., Zobel, C., Nguyen, T. T., Molkentin, J. D., Backx, P. H. (2002) Reduction of I(to) causes hypertrophy in neonatal rat ventricular myocytes. *Circulation research*, 90(5) :578–85.

Kaulin, Y. A., De Santiago-Castillo, J. A., Rocha, C. A., Nadal, M. S., Ruby, B., Covarrubias, M. (2009). The dipeptidyl-peptidase-like protein DPP6 determines the unitary conductance of neuronal Kv4.2 channels. *The Journal of Neuroscience*, 29(10).

Kim, J., Wei, D. S. et Hoffman, D. A. (2005). Kv4 potassium channel subunits control action potential repolarization and frequency-dependent broadening in rat hippocampal CA1 pyramidal neurones. *J Physiol.*, 569, 41–57.

Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, Beglov D, Vajda S. (2017). The ClusPro web server for protein-protein docking. *Nature Protocols*, 12(2), 255-278.

Kozakov, D., Beglov, D., Bohnuud, T., Mottarellam S., Xia B, Hall D., R., Vajda, S. (2013). How good is automated protein docking? *Proteins*, 81(12), 2159-2166.

Kozakov, D., Brenke, R., Comeau, S. R., Vajda, S. (2006) PIPER : An FFT-based protein docking program with pairwise potentials. *Proteins*.

Leimkuhler, B., Matthews, C. (2016) Robust and efficient configurational molecular sampling via Langevin dynamics. *The Journal of Chemical Physics*, 138(174102).

Leimkuhler B, Matthews C. (2016) Efficient molecular dynamics using geodesic integration and solvent–solute splitting. *PRSA*, 472(20160138).

Lin, L., Long, L., Hatch, M. et Hoffman D. (2014). DPP6 Domains Responsible for Its Localization and Function. *The Journal of Biological Chemistry*, 289(46), 32153–32165.

Lu, Z., Klem, A. M. and Ramu, Y. (2002). Coupling between voltage sensors and activation gate in voltage-gated K⁺ channels. *J. Gen. Physiol.* 120 :663–676.

Long, S. B., Campbell, E. B. and R. MacKinnon, R. (2005). Voltage sensor of Kv1.2 : structural basis of electromechanical coupling. *Science*, 309 :903–908.

Long, S. B., Campbell, E. B., and MacKinnon, R. (2005). Crystal structure of a mammalian voltage-dependent Shaker family K⁺ channel. *Science*, 309 :897–903.

Madej, T., Lanczycki, C. J., Zhang, D., Thiessen, P. A., Geer, R. C., Marchler-Bauer, A., Bryant, S. H. (2014). MMDB and VAST+ : tracking structural similarities between macromolecular complexes. *Nucleic Acids Res.*, 42(Database issue), 297-303.

Marti-Renom, M. A., Stuart, A., Fiser, A., Sánchez, R., Melo, F., Sali, A.. (2000) Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* 29, 291-325.

Morohashi, Y., Hatano, N., Ohya, S., Takikawa, R., Watabiki, T., Takasugi, N. et al. (2002). Molecular cloning and characterization of CALP/KChIP4, a novel EF-hand protein interacting with presenilin 2 and voltage-gated potassium channel subunit Kv4. *J. Biol. Chem.*, 277, 14965–14975. doi :10.1074/jbc.M200897200.

Nadal, M.S., Ozaita, A., Amarillo, Y., Vega-Saenz de Miera, E., Ma, Y., Mo, W., et al. (2003). The CD26-related dipeptidyl aminopeptidase-like protein DPPX is a critical component of neuronal A-type K⁺ channels. *Neuron*, 37, 449–61.

Nanao, M. H., Zhou, W., Pfaffinger, P. J., Choe, S. (2003). Determining the basis of channel-tetramerization specificity by x-ray crystallography and a sequence-comparison algorithm : Family Values (FamVal). *Proc. Natl. Acad. Sci. U.S.A.*, 100(15), 8670-8675.

Olsson, M. H. M., Chresten, R., Sondergaard, C. R., Rostkowski, M. and Jensen, J. H. (2011). PROPKA3 : consistent treatment of internal and surface residues in empirical pKa

predictions. *Journal of Chemical Theory and Computation*, 7(2), 525-537.

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L., Schulten, K. (2005). Scalable molecular dynamics with NAMD. 26(16), 1781-1802.

Pongs O. (1999). Voltage-gated potassium channels : from hyperexcitability to excitement. *FEBS Lett*, 452, 31-5.

Pioletti, M., Findeisen, F., Hura, G. L., et Minor, D. L. (2006). Three-dimensional structure of the KChIP1-Kv4.3 T1 complex reveals a cross-shaped octamer. *Nature Structural and Molecular Biology*, 13(11), 987-995.

Roberts, E., Eargle, J., Wright, D. et Luthey-Schulten, Z. (2006). MultiSeq : unifying sequence and structure data for evolutionary analysis. *BMC Bioinformatics*, 7(382).

Russell, R. B., Barton, G. J. (1992). Multiple protein sequence alignment from tertiary structure comparison : assignment of global and residue confidence levels. *Proteins*, 14(2), 309-323.

Sah, R., Ramirez, R. J., Oudit, G. Y., Gidrewicz, D., Trivieri, M.G., Zobel, C., (2003). Regulation of cardiac excitation-contraction coupling by action potential repolarization : role of the transient outward potassium current (I_{to}). *The Journal of Physiology*, 546(Pt 1) :5-18.

Sondergaard, C. R., Olsson, M. H. M., Rostkowski, M. and Jensen, J. H. (2011). Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pKa Values. *Journal of Chemical Theory and Computation* 7(7), 2284-2295.

Tovchigrechko, A. Ilya, Vakser, A. (2006). GRAMM-X public web server for protein-protein docking, *Nucleic Acids Research*, 34(2), 310-W314.

Tovchigrechko, A., Vakser, I.A. (2005). Development and testing of an automated approach to protein docking. *Proteins*, 60(2), 296-301.

Viklund, H. et Elofsson, A. (2008). Improving topology prediction by two-track ANN-based preference scores and an extended topological grammar. *Bioinformatics*

Scannevin, R. H., et al. (2004). Two N-terminal domains of Kv4 K(+) channels regulate binding to and modulation by KChIP1. *Neuron*, 41, 587–598.

Sali, A. et Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779-815.

Serodio, P., Kentros, C., and Rudy, B. (1994). Identification of molecular components of A-type channels activating at subthreshold potentials. *J. Neurophysiol.* 72, 1516–1529.

Wallin E. et von Heijne G. (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Science*, 7, 1029–1038.

Webb, B. et Sali, A.. (2014) Comparative Protein Structure Modeling Using Modeller. *Current Protocols in Bioinformatics*, John Wiley Sons, Inc., 5.6.1-5.6.32

Yu, F. H. (2005). Overview of Molecular Relationships in the Voltage-Gated Ion Channel Superfamily. *Pharmacological Reviews*, 57(4) : 387-395.

Zagotta, W. N., Hoshi, T., and Aldrich, R. W. (1994a). Shaker potassium channel gating. III : evaluation of kinetic models for activation. *J. Gen. Physiol.*, 103, 321–362.

