

Molecular Dynamic - LIVRE

7.1 Introduction

La dynamique moléculaire (MD) se base sur l'intégration des lois de Newton (inertie, $f=ma$, action-réaction). Ces simulations fournissent la position et la vitesse des particules qui composent le système en fonction du temps. La trajectoire est résolue en résolvant la deuxième loi de Newton :

$$\frac{d^2 x_i}{dt^2} = \frac{F_{x_i}}{m_i} \quad (1)$$

7.2 Dynamique Moléculaire : modèles simples

Premier modèle de MD apparaît en 1957 en utilisant le modèle de sphère-dure. Dans ce modèle, les sphères se déplacent à vitesse constante et les collisions sont parfaitement élastiques. Ce modèle fonctionne comme suit :

- identifier la future paire de sphères qui vont se rencontrer, et déterminer quand cela arrivera
- calculer la position des sphères au moment de la collision
- déterminer les vitesses des sphères après la collision
- répéter 1

Les vitesses sont calculées sur base du principe de la conservation du moment linéaire.

7.3 Potentiels Continus

Dans des modèles plus réalistes d'interactions moléculaire, la force qui s'exerce sur chaque particule change à chaque changement de position. L'influence continue du potentiel sur les mouvements fait surgir de nombreux problèmes et rend la résolution analytique impossible. La solution du problème consiste à séparer le temps de simulation en une série de pas de temps δt . On calcule alors la force qui s'exerce sur chaque particule à chaque instant δt . La force résultante permet de mesurer l'accélération et d'ensuite déterminer la position des particules après un pas de temps. Il existe de nombreux algorithmes qui mesurent cela, et tous assument que cette mesure peut être approximée par une série de Taylor. Parmi ces algorithmes, l'algorithme de Verlet est celui qui est le plus utilisé en MD. Il utilise les positions et les accélérations mesurées à l'étape précédente $r(t - \delta t)$ pour calculer les nouvelles positions au temps $t + \delta t$ à la position $r(t + \delta t)$.

$$r(t + \delta t) = r(t) + \delta t v(t) + \frac{1}{2} \delta t^2 a(t) + \dots \quad (2)$$

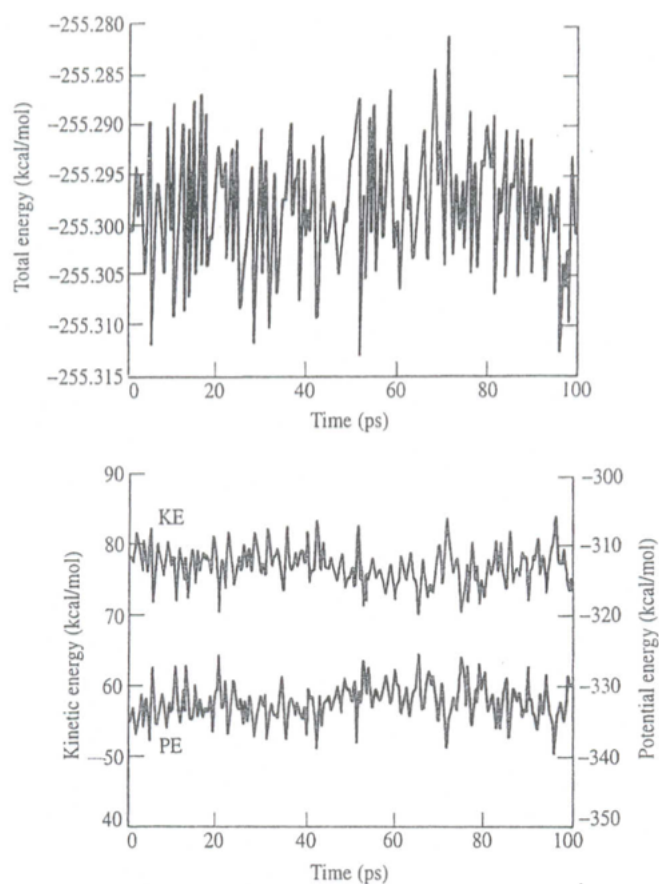
$$r(t - \delta t) = r(t) - \delta t v(t) + \frac{1}{2} \delta t^2 a(t) - \dots \quad (3)$$

Ces deux équations donnent :

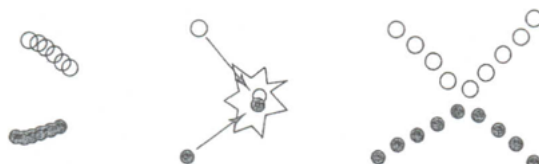
$$r(t + \delta t) = 2r(t) - r(t - \delta t) + \delta t^2 a(t) \quad (4)$$

La vitesse peut être estimée en divisant la différence de position par le temps, c'est à dire par $2\delta t$. Cet algorithme présente quelques désavantages. Premièrement, les positions $r(t + \delta t)$ sont obtenues en ajoutant un petit terme ($\delta t^2 a(t)$) à la différence de deux termes plus grands, $2r(t)$ et $r(t - \delta t)$, ce qui induit une perte de précision. Un autre désavantage, est que les vitesses d'un instant t ne peuvent être mesurées qu'après avoir calculé l'étape suivante. Aussi, cet algorithme

ne donne pas les positions de départ puisqu'il requiert les positions en $t - \delta t$. On peut alors utiliser une série de Taylor tronquée au deuxième terme pour estimer la position de départ. Il existe plusieurs variantes de l'algorithme de Verlet. Le "leap-frog" par exemple, inclue une mesure directe des vitesses et ne nécessite plus le calcul de la différence entre deux valeurs plus importantes. Il est dès lors possible de calculer l'énergie cinétique en même temps que les positions. Quel algorithme utiliser ? En simulation MD, on choisira celui qui présente le calcul le plus rapide de la force de chaque particule composant le système. Il faudrait également qu'il soit réversible dans le temps, et permette d'utiliser des grands δt . Aussi, il faudrait que l'algorithme donne les mêmes résultats qu'une trajectoire résolue analytiquement. Outre le coût de calcul propre à chaque algorithme, un autre facteur déterminant est la conservation d'énergie. Il est attendu que l'énergie cinétique et que l'énergie potentielle fluctuent de façon opposée.



Le choix du pas de temps est évidemment important. Un pas de temps trop petit implique une trajectoire dans l'espace des phases plus petite. Inversement, un pas d'intégration trop long peut faire surgir des instabilités dues à des énergies de recouvrement entre atomes trop proches.



Les vibrations les plus importantes dans les molécules flexibles sont dues aux liens "étirables" tels que les liens H. A la page 362 du livre, il y a un tableau qui reprend les pas de temps les plus appropriés en fonction des système que l'on étudie en MD. Il est évident que les pas de temps devraient être d'un ordre de grandeur plus petit que les plus petits mouvements du système.

Etant donné que certains petits mouvements n'ont pas de rôles majeures dans le système, il est courant de freezer ceux-ci en contraignant (ou en mettant des restrictions) ces liens par leur valeur à l'équilibre en permettant tout de même quelques degrés de liberté.

7.4 Setting Up and Running Simulations

Dans un premier temps, il faut établir les conditions initiales du système (sur des coordonnées x , y , z par exemple). Ce peut être à partir de données expérimentales ou à partir d'un modèle théorique (cfr article) ou par une combinaison des deux. Pour initialiser les vitesses des atomes, on peut utiliser l'équation de Maxwell Boltzmann qui donne la proba de vitesse d'un atome i d'une masse m dont les probas varient en fonction de la T . Les vitesses peuvent aussi être déterminées à partir d'une distribution uniforme ou Gaussienne. On initialise généralement les vitesses de sorte que le moment général du système soit égal à zéro (mesure des v sur x , y , z , on divise ces vitesses par la masse totale de l'atome et on soustrait cette valeur par la vitesse de l'atome).

A chaque pas de temps, la force s'appliquant sur chaque atome doit être calculée. Elle doit inclure les contributions les plus importantes, telles que les liens atomiques, les angles, les torsions, les interactions non liées. On utilise généralement le potentiel de Lennard-Jones pour mesurer la force d'interaction entre deux atomes. (rappel : premier terme : attractif et domine à grande distance (= interaction de VDW) et second terme : répulsif qui domine à courte distance (= répulsion des nuages électroniques mutuels).

Ensuite, on applique ce qu'on appelle une phase d'équilibrage dont le but est de faire atteindre l'équilibre au système à partir des conditions initiales. En MD, il faut que la phase d'équilibrage soit plus longue que la phase de relaxation, de sorte que le système perde toute mémoire des conditions initiales.

Suit alors la phase dite de production où l'on mesure l'énergie cinétique, potentielle, les vitesses, la température et la pression du système.

7.5 Contraintes Dynamiques

Le modèle le plus simple en MD est de considérer le système comme étant rigide sans liberté de conformation moléculaire. Dans ce cas-ci, la dynamique peut être vue par les modifications des centres de masses et par les rotations. De plus, comme dit plus haut, considérer un corps rigide permet d'utiliser des pas de temps plus grand puisqu'on néglige certains mouvement qui n'influencent pas significativement la conformation du système. C'est dans ce but que l'on utilise des contraintes. Il est nécessaire de différencier contrainte et restriction. Contrainte = forçage du système vs restriction = guide, comprenant une certaine flexibilité. On peut contraindre des angles ou des liens à adopter certaines valeurs par exemple tandis qu'une restriction va "encourager" le système à adopter ces valeurs. En général, lorsque l'on ajoute une restriction, on fonctionnera avec un système de pénalité lorsque l'on dévie de la valeur cherchée (voir 9.10).

7.6 Propriétés dépendants du temps

Pouvoir mesurer les propriétés qui dépendent du temps est l'un des grands avantages de la dynamique moléculaire comparée à la méthode Monte Carlo (= échantillonnage aléatoire de l'espace des conformations, méthode probabiliste). Ces propriétés dépendantes du temps sont souvent calculées par le coefficient de temps de corrélation.

Imaginons que nous ayons deux jeux de données (x et y), et que nous voulions évaluer la corrélation qui pourrait exister entre eux. Par exemple, une simulation d'un fluide dans un

capillaire, quelle serait la corrélation entre la vitesse x d'un atome et sa distance y par rapport à la surface du tube. Une manière d'évaluer cela serait de porter ces deux jeux de données sur un graphe. Une fonction de corrélation, est la fonction qui fournit une valeur numérique englobant les données et qui quantifie la force de corrélation. Une fonction de corrélation commune pour M valeurs de x et de y :

$$C_{xy} = \frac{1}{M} \sum_{i=1}^M x_i y_i \equiv (x_i, y_i) \quad (5)$$

En général, on va normaliser la corrélation de sorte qu'une valeur de 0 indique qu'il n'y a pas de corrélation, et qu'une valeur de 1 indique une forte corrélation.

Une simulation MD fournit des données à des temps spécifiques. Ceci permet de corréler des valeurs mesurées à un certain temps t avec ces mêmes valeurs (ou d'autres) d'un autre temps t . Ces valeurs sont appelées *coefficient de temps de corrélation*. Lorsque les deux jeux x et y sont différents, on appelle la fonction de corrélation *cross correlation function* tandis que lorsque x et y sont les mêmes, on l'appelle *autocorrelation function*. Cette dernière indique dans quelle mesure le système garde en mémoire les valeurs mesurées à différents t . En général, le coefficient d'autocorrélation de vitesse est initialisé à 1. Le temps pris pour que sa valeur soit de zéro est appelé temps de corrélation ou temps de relaxation. Il est donc logique que le temps de simulation soit plus grand que le temps de relaxation. Une valeur négative du coef de corr de vitesse signifie que la particule se déplace en direction opposée à celle en temps $t = 0$. Dans la même idée que le coefficient d'autocorrélation de la vitesse, on peut mesurer l'autocorrélation du dipole qui s'applique à toute la molécule entre deux temps t (cfr page 368 si vous voulez voir la formule). Dans le cas du dipole, on peut comparer les résultats obtenus en MD avec des résultats obtenus expérimentalement en utilisant une transformée de Fourier qui permet de comparer les spectres expérimentaux avec ceux des simulations.

7.7 MD à Température et Pression Constante

Généralement, les simulations MD s'opèrent à température constantes (+ énergie et volume constant), néanmoins, il y a plusieurs raisons pour lesquelles il est intéressant de travailler à température constante. Il est assez commun de moduler la température durant la phase d'équilibre. Il est souvent utile de travailler à T cste si l'on veut déterminer si le système change avec T . Comme par exemple dans le cas du repliement d'une protéine. La température du système est liée à l'énergie cinétique.

Deux méthodes sont généralement utilisées pour contrôler la température du système. La première possibilité consiste à jouer sur les vitesses. Si on a la température à un temps, $T(t)$, et que l'on multiplie les vitesses par un facteur λ . On peut déterminer le changement de température par cette expression (démonstration page 383) :

$$\lambda = \sqrt{T_{new}/T(t)} \quad (6)$$

Pour contrôler la température il suffit donc de multiplier les vitesses par le facteur $\lambda = \sqrt{T_{req}/T_{curr}}$ où T_{curr} est la température actuelle et T_{req} est la température désirée. L'autre façon utilisée pour contrôler la température du système est de le coupler avec un bain thermostatique dont la température est fixée à la température désirée. Le bain sert de source d'énergie et ajoute ou soustrait la chaleur au système.

$$\frac{dT(t)}{dt} = \frac{1}{\tau} (T_{bath} - T(t)) \quad (7)$$

τ est le paramètre qui détermine à quel point le bain et le système son lié. Si τ est large, alors le lien est faible, inversement s'il est petit. L'avantage de cette méthode est qu'elle permet au système de fluctuer autour de la température désirée. La page 384 et 385 discute d'autres méthodes stochastiques. Je ne pense pas que le prof en ait parlé et que ce soit important. On y parle également d'un degré de liberté s ajouté au système qui comprend la température.

Tout comme pour la température, ce peut être utile de travailler à pression constante, ou du moins en contrôlant la pression. Cela permet par exemple, d'étudier la sensibilité du système aux transitions de phase dues à la pression. Travailler à P cste peut également être utile lorsque l'on travaille avec des systèmes dont le nombre de particules du système change. La pression, au contraire de la température, fluctue davantage que l'énergie totale du système. D'un point de vue macroscopique, on module la pression avec le volume. La compressibilité isothermale κ (= kappa) est liée aux fluctuations de volume :

$$\kappa = -\frac{1}{V}\left(\frac{\delta V}{\delta P}\right)_T \quad (8)$$

Une substance facilement compressible présentera un κ grand, et donc des fluctuations plus importantes de volumes pour une pression fixée que pour des valeurs de κ plus petites. Et pour un volume fixé, une substance moins compressible présentera des fluctuations de pression plus importantes. La plupart des méthodes permettant de travailler à pression constante sont analogues à celles pour la température. La pression peut être maintenue en contrôlant le volume. Ou alors, on peut coupler le système à un "bain de pression". Comme pour la température, un degré de liberté traitant de la pression est ajouté dans les modèles plus récents.

7.9 Changements conformationnels de simulations MD

La MD peut fournir des informations sur les changements de conformation du système. En effet, on peut porter l'évolution des données des positions des particules qui évoluent en fonction de temps sur un film. Les mouvement sont souvent chaotiques, et ce chaos est souvent dues aux mouvements hautes fréquences qui ne jouent pas de rôles importants dans l'étude du système. Les mouvements basses fréquences sont souvent plus intéressant et significatifs. On peut alors utiliser des analyses de Fourier pour filtrer les hautes fréquences non désirées. Une transformée de Fourier permet de convertir des variations de quantités en fonction du temps en fonction de fréquence et inversement. Donc, en utilisant la fréquence des mouvements, on peut filtrer les mouvements que l'on veut garder (observer). A chaque étape d'une analyse de Fourier d'une simulation MD, la variation des coordonnées cartésiennes en fonction du temps est convertie en fonction de fréquence. On filtre alors le spectre de fréquence en enlevant les hautes fréquences non désirée.

9.10 Résolution de Structures Protéiques à l'aide de Restrictions

Une application particulièrement imortante de la MD, est le raffinement de données expérimentales X-ray et NMR pour déterminer la structure 3D de molécules larges telles que des protéines. Le but est donc de déterminer la conformation qui explique le mieux les données expérimentales. On utilise en général une variante à la MD, qui s'appelle *restrained molecular dynamics*. Cette méthode inclue des nouveaux termes tels que des fonctions de pénalités qui s'ajoutent aux fonctions de potentiels d'énergie. Ces fonctions "pénalisent" les conformations qui ne s'accordent pas bien avec les données expérimentales. Pour explorer l'espace des phases et trouver la meilleur conformation, on utilise souvent la méthode *simulated annealing* qui consiste à jouer sur la température.

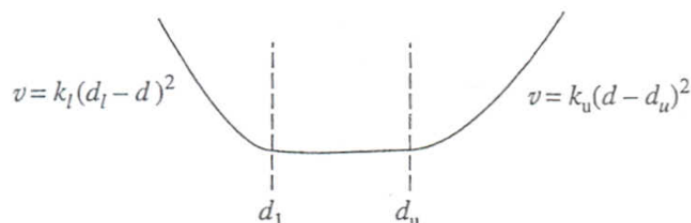
X-ray cristallographie - C'est une technique puissante qui permet de résoudre la structure de protéines. Je ne rappelle pas comment la technique fonctionne, il y a un rappel à la page

484 si vous voulez. La MD est utilisée pour raffiner les structures calculées de sortes à ce qu'elles correspondent le plus avec les données expérimentales.

$$R = \frac{\sum ||F_{obs}| - |F_{calc}||}{\sum |F_{obs}|} \quad (9)$$

Le facteur R en cristallographie permet d'évaluer si la structure observée (expérimentale) correspond bien avec la la structure calculée. La cristallographie par rayon X fourni la densité électronique de la molécule étudiée. Sur base de cette densité, l'objectif est de déduire la structure. Cette déduction peut induire des erreurs, l'idée sera d'utiliser des restrictions pour limiter celles-ci (je suis pas certain d'où viennent les erreurs, mais l'idée de restrictions pour les limiter est bonne). En restrained MD, on utilise l'énergie potentielle pour introduire des restrictions. Dans le livre, il est indiqué comment on introduit cette restriction, je ne le détaille pas car le cas NMR est plus simple et le prof s'était davantage exprimé dessus. L'espace des conformations est balayée par des simulations *annealing* où l'on utilise des températures élevées pour permettre au système de varier. On diminue alors graduellement la température de sorte que la structure adopte une structure basse énergie avec un petit facteur R (grâce aux restrictions).

Raffinement de données NMR par MD - La façon la plus commune pour raffiner des données NMR est d'ajouter un terme harmonique de restriction $v(d) = k (d - d_0)^2$ où d est la distance entre les atomes et d_0 est la distance désirée obtenue à partir du spectre NMR. k détermine dans quelle mesure la restriction s'applique. On peut procéder de la même façon avec les angles de torsions ou avec d'autres grandeurs. Il d'autres façons pour introduire des restrictions, par exemple, en introduisant des pénalités uniquement lorsque la distance d excède d_0 . Il existe des fonctions plus sophistiquées qui tiennent compte des imprécisions expérimentales.



Sur la figure ci-dessus, on observe que deux restrictions s'appliquent. Si la distance d est inférieure d_l ou si la distance d est supérieure à d_u . Entre ces deux valeurs, la restriction $v(d) = 0$.

Time-averaged NMR Refinement - Si la molécule peut changer de conformation sur une échelle de temps plus petite que le "chemical shift" (voir thème 2), alors le spectre NMR obtenu ne montrera qu'une moyenne de toutes ces conformations. Si les deux conformations (voir figure ci-dessous) sont capables de former des interactions différentes, alors on peut trouver deux restrictions de distances différentes. Un raffinement standard ne tiendrait pas compte de ces deux restrictions et ne permettrait pas de trouver les deux conformations. Une solution serait d'utiliser une autre restriction : *the Time-averaged restraints*. J'avoue ne pas trop avoir compris ce passage (page 488), mais je crois sincèrement pas que ce soit important. Je ne pense pas que le prof l'ait mentionner au cours, mais il ne l'a pas barré.

Recent Progress in the study of G-protein-coupled receptors with molecular dynamics computer simulations - REVIEW

Abstract

Les récepteurs couplés aux protéines G (GPCRs) appartiennent à une famille de protéines importantes. Leurs conformations varient en fonction des ligands susceptibles de s'y lier. La dynamique moléculaire est devenue un facteur déterminant dans la biophysique des GPCRs, permettant de mieux comprendre les interactions lipides-protéines, les mécanismes d'activation, ainsi que l'hydratation interne.

Introduction

La famille des GPCRs est la plus grande du génome humain, et sont exceptionnellement importante d'un point de vue biomédicale. Il a été estimé que plus d'un quart des nouvelles drogues développées ciblent les GPCRs. Mode d'activation agoniste et antagoniste des GPCR. Les GPCRs sont des protéines de membranes, et sont donc difficiles à manipuler expérimentalement (nécessite un environnement similaire à celui d'une membrane pour garder sa conformation). En 2000, l'étude des GPCRs a fait un bond en avant avec la cristallisation de la rhodopsine de bovin, une GPCR ; offrant un point de vue atomique sur la structure des GPCRs. En 2007, après d'autres résolutions de structures, celle de récepteur β_2 -adrenergic (B2AR), suivi du récepteur A_2A -adenosine (A2A) et du récepteur β_1 -adrenergic (B1AR). Plus récemment, l'opsin (forme de la protéine avec son ligand) liée ou non à une protéine G.

Ces différentes résolutions de structure ont été le point de départ de nombreuses simulations de dynamiques moléculaires. Cette review résume l'étendue du travail effectué dans ce domaine, en concentrant le review sur les méthodes quantitatives de MD.

Rhodopsine

La rhodopsine est une GPCR photorecepteur trouvée dans le système de vision des mammifères. Il s'agit d'un pigment protéique photosensible, formée d'une protéine transmembranaire, l'opsine, sur laquelle se fixe le rétinal (vitamine A). La rhodopsine existe sous plusieurs états, la forme *Dark State*, *Meta-I*, et *Meta-II* (sa forme active). Le passage de la forme DS à Meta-I sur une échelle de temps de l'ordre de la microseconde, tandis que le passage de l'état Meta-I à l'état Meta-2 s'opère sur une échelle de temps de l'ordre de la milliseconde. Elle est relativement tolérante aux ratios élevés lipides/protéines. Cela a permis de réaliser de nombreux travaux de recherche sur cette protéine (NMR, fluorescence, spectro IR,...). La rhodopsine a également été beaucoup étudiée en MD, plus que toutes les autres GPCRs.

Interactions Lipides - Protéines dans l'état sombre (DS)

L'environnement entourant les biomolécules est primordiale pour leur comportement. Pour les protéines transmembranaires, la composition de la bicouche (importance des types de lipides et de la composition de la membrane) influence la stabilité et l'efficacité des protéines qui s'y intègrent. La rhodopsine en est un bon exemple, en effet, la membrane cellulaire des cellules du ROS (rod outer segment) dans lesquels on retrouve cette protéine présente une composition lipidique particulière. Celle-ci est riche en acide gras ω -3 polyinsaturés, la concentration en cholestérol est élevée mais pas uniforme. Les ω -3 désordonnent la membrane et amplifie l'action de la rhodopsine en l'activant tandis que le cholestérol a un effet opposé sur la rhodopsine et stabilise la membrane. Est-ce que ces effets sont dus aux modifications des propriétés de la membrane, ou sont dus aux interactions de ces molécules directement avec la rhodopsine. Dans le cas des lipides polyinsaturés, des simulations ont été effectuées et il semblerait que les lipides

interagissent directement avec la protéine. Seulement le temps des simulations MD était relativement court (12.5 ns). D'autres simulations MD similaires ont été faites, cette fois avec 10 simulations de 100ns, plus réalistes. Les résultats sont consistant avec les précédents, suggérant des interactions spécifiques dirigées entropiquement. Certaines interactions sont trop longues pour être simulées en all-atom MD, il y a donc un équilibre entre le temps caractéristique du mouvement étudié, et le temps de la simulation.

Il a été montré que beaucoup de GPCRs sont capables de se dimériser in vitro avec elle-même (homodimères) ou avec d'autres protéines (hétérodimères). Mais rien n'a encore été montré sur l'intérêt physiologique de la dimérisation possible des GPCRs. C'est pourquoi la plupart des simulations ont été effectuées en considérant l'état monomérique des GPCRs. Est-ce le bon choix ? L'état d'avancement des recherches actuelles ne permettent pas encore de l'affirmer. Si l'on travaille sur des simulations MD de dimérisation, plusieurs challenges se présentent : augmentation considérable des temps de calculs (implique diminution des trajectoires), augmentation de la taille du système protéique, et le plus difficile, quel dimère utilisé (pas de structures cristallisée, seulement des modèles) ? Il semblerait que des simulations de l'ordre de la microseconde ne soit pas suffisantes pour caractériser correctement la plupart des phénomènes d'intérêt. L'échantillonnage statistique (déjà difficile à calculer pour des systèmes monomériques) des simulations MD est lié au temps de simulation (qui augmente avec la taille des systèmes).

Puisque les résolutions expérimentales d'oligodimères manquent, il faut compter sur des simulations. Une approche développée est la simulation "*Gros Grains*" (coarse-graining CG) qui offre des temps de simulations permettant l'étude de la formation et de la dissociation de dimères. Certaines études sur la dimérisation de la rhodopsine ont été faites avec des simulations CG pour différentes épaisseurs de couche lipidique. A noter que cette approche fait intervenir des contraintes et des approximations importantes.

Etude du mécanisme d'activation

Quelle est la nature des changements conformationnels qui s'opèrent lorsque les GPCRs s'activent ? Le modèle le plus commun d'un GPCR - protéine G est celui où le ligand d'origine extracellulaire induit une série de changements conformationnels permettant à la protéine G (rappel : les protéines G sont trimériques) d'échanger un GDP contre un GTP. Cet échange provoque la dissociation de G en sous-unité G_α et en $G_{\beta\gamma}$ qui induit alors une cascade de signalisation.

Pour le moment, aucun cristal de GPCR sous forme active n'a pu être obtenu (obtenir une forme active en cristallographie est extrêmement difficile). C'est pourquoi les chercheurs se sont tournés vers les simulations MD. Plusieurs simulations ont été réalisées, tantôt sur des temps courts (10ns) mais avec des calculs de trajectoires différents, tantôt sur des temps plus longs (150ns), observations de rupture de ponts salins, de déformations de coudes. Les résultats de ces recherches sont consistants avec la littérature et les recherches parallèles. D'autres simulations MD ont inclus des distances atomiques déterminées expérimentalement comme contraintes énergétiques, les chercheurs ont utilisé ce qu'ils appellent le "*Magic Angle Spinning NMR*" (MAS NMR, voir le petit paragraphe page 1871 s'il faut en parler, c'est souligné en rouge).

Il y a tout un sujet exploré par les auteurs sur un pont salin sur la base protonée de Schiff de la rhodopsine où deux modèles s'opposaient. La résolution du problème est passée par la collaboration entre simulations et approches expérimentales (comme lorsque les chercheurs ont utilisé le MAS NMR pour d'autres recherches) en incluant des nouvelles contraintes/trajectoires aux simulations. Cela montre l'importance d'utiliser les résultats expérimentaux pour améliorer les simulations, plutôt que de les utiliser uniquement pour évaluer les simulations.

Les interactions entre les protéines et l'eau sont cruciales sur la structure, la dynamique, et la fonction. L'effet hydrophobe est un élément clé dans le repliement des protéines. Pour une

protéine déjà repliée, l'influence de changements d'hydratation sur la fonction n'est pas évidente. Néanmoins, des expériences menées sur la rhodopsine ont montré l'importance de son hydratation malgré que celle-ci soit transmembranaire. Il semblerait que l'état Meta-II de la rhodopsine soit plus hydraté que l'état Meta-I. Les structures cristallographiques révèlent aussi la présence de molécules d'eau localisée près de zones fonctionnelles. Des simulations MS de 1500ns ont également révélé une augmentation de l'hydratation interne importante entre l'état DS et Meta-I. Pour contrôler si leurs simulations étaient justes, les chercheurs ont reproduit des simulations de longueurs de temps similaires pour vérifier si l'hydratation du DS était constante (pour rejeter l'hypothèse d'erreurs dans la construction, ou de défaut des fonctions d'énergie qui seraient alors à l'origine de ces changements d'hydratation). L'hydratation est restée stable, comme attendu. Les molécules d'eau réagissent de façon spécifique au champ magnétique, et une augmentation de l'hydratation mène à une atténuation de la résonance des lipides (pas bien compris). Est-ce que ce comportement d'hydratation s'observe également chez les autres GPCRs ? Oui, des simulations sur d'autres rhodopsines et d'autres GPCRs le montre. Beaucoup de recherches indiquent que l'hydratation des GPCRs est probablement centrale dans la modulation des fonctions de celle-ci.

Simulations des interactions GPCR - protéines G

Une autre étape critique dans les simulations des GPCRs est de caractériser les interactions avec la protéine G suivie de la dissociation du trimère. Il n'y a pas de résolution expérimentales précises qui traite du sujet. Aujourd'hui, on pense qu'une protéine G ne s'associe à un récepteur GPCR que lorsque le GPCR est sous sa forme active. Il faudrait donc utiliser cette structure-ci en point de départ. Or, on ne dispose pas de bonnes résolutions de ces structures. Aussi, simuler l'évolution d'un complexe GPCR-G impose un coût important de simulation. Malgré ces difficultés, certains chercheurs sont parvenus à étudier la dynamique d'un complexe GPCR-G en combinant la "modélisation par homologie", le "docking" (thème 3) et la MD. L'article est très ambitieux, mais les résultats sont tout de même critiquables. Premièrement, leur structure de départ n'est pas basée sur une structure expérimentalement résolue (comme dit plus haut) et deuxièmement, le temps de leur simulation est très court, seulement 10 ns (plus court que le temps de corrélation sans considérer de relaxation de la structure de départ, j'imagine que statistiquement ça pose des soucis). En 2010, un autre article présentait une simulation d'une μ s. Leurs simulations révèlent la nature de la dynamique des interactions entre la rhodopsine et la sous-unité α de la protéine G, et explorent les détails de la dynamique intra et intermoléculaires des ponts salins. À noter que leur structure de départ n'est pas une structure cristallographique, mais est issue d'une modélisation docking. L'interface entre les deux protéines, lors de la simulation, présente un ensemble lâche de contacts. Ceci pourrait signifier que le complexe est dynamique ou bien que la structure de départ présente des erreurs. Pour prouver qu'il n'y a pas d'erreur il faudrait faire des simulations deux ou trois fois plus longues. Ces résultats sont tout de même très prometteurs.

Simulation d'autres GPCRs

La B2AR, une autre GPCR a été résolue et cristallisée expérimentalement. Des simulations à grandes échelles ont été réalisées avec cette protéine. Leurs simulations ont été effectuées sur une B2AR chimérique pour laquelle un lysozyme T4 a été inséré entre l'hélice 5 et 6 (je sais pas pourquoi, le prof l'a souligné c'est pourquoi je le note dans le résumé) en présence d'un ligand non naturel, le carosol. Il ont effectué la même chose avec le ligand naturel, l'adrénaline et ont observé que l'hydratation interne de la protéine augmente avec l'adrénaline. Ce qui est consistant avec ce qui a été décrit plus haut (augmentation de l'hydratation lorsque l'on active la GPCR).

D'autres études (sur B2AR et B1AR) ont porté sur l'étude du blocage ionique ("ionick lock"),

qui correspond à un pont salin hautement conservé entre Arg (arginine de l'hélice 3) et Glu (acide glutamique de l'hélice 6). Ce blocage est crucial dans le maintien de la forme inactive du GPCR. Les simulations MD ont permis d'étudier ce phénomène et de mettre en évidence certains acides aminés jouant un rôle majeur. Ils ont également découvert que le blocage ionique pouvait exister sous trois différents états : ouvert, fermé, et lié à l'eau (où les deux hélices (3 et 6) sont attachées ensembles via la même molécule d'eau). L'état du blocage ionique peut changer dans des temps de l'ordre de la nanoseconde. Il y a plus de détails dans le point 4 de l'article, mais je ne trouve pas cela super intéressant à relever. L'idée étant de voir un peu les champs d'études possibles avec la MD.

Évaluation des erreurs statistiques

Est-ce que les statistiques sont suffisamment bonnes pour tirer des conclusions ? Toutes les questions statistiques que l'on pourrait se poser sur les résultats que l'on peut obtenir de simulations MD sont intimement liées à une grandeur : le temps de corrélation du système (combiné avec la longueur de trajectoire) qui détermine le nombre de points du système. Les statistiques de jeux de données indépendants et non corrélés sont relativement simples.

Les auteurs du review présentent plusieurs formules de grandeurs statistiques. Je ne pense pas que ça soit important, néanmoins, toutes les grandeurs développées par les auteurs sont facilement trouvable dans l'article, et les éléments qui les composent sont définies juste en dessous. Ils parlent notamment de l'évaluation de l'incertitude statistique, comment l'estimer, notamment en utilisant le "block averaging". Mais cette méthode présente des inconvénients, notamment face au bruit et au fait que cette mesure ne fournit pas directement le temps de corrélation. Il en ressort qu'il est nécessaire de trouver un autre moyen d'évaluer certaines statistiques. Peut-être en examinant directement le système d'intérêt sur une plus grande échelle en incluant seulement les termes les plus importants. Par exemple, en ne considérant pas le dépliage spontané qui peut s'effectuer. Mr Gilis a émis un commentaire en disant que ce que les auteurs du review avancent n'est pas bien explicité et trop bref.

La "covariance overlap" est une autre grandeur qui a été développée par certains auteurs. Elle a permis de mettre en évidence que le temps de simulations de 100ns (qui à l'époque était très long) de certaines simulations s'avéraient trop courts. Il semblerait que même les simulations récentes sur la rhodopsine et B2AR de l'ordre de la microseconde ne soient de bons échantillons statistiques. Notamment en évaluant le RMSD de plusieurs simulations. Des méthodes de clustering ont été développées et permettent d'évaluer les vitesses les plus lentes offrant la possibilité de se concentrer sur les mouvements à large échelle. Les auteurs pensent que ces évaluations deviendront un standard en dynamique moléculaire car pour le moment, selon la littérature, il semblerait que l'on fasse trop confiance à l'égard de l'échantillonnage statistiques. (pas bien compris)

Il est aussi à noter que d'autres modèles de réseaux (qui ne sont pas basés en "all-atome") sont bien plus simples, permettent d'éviter beaucoup de problèmes, et fournissent des résultats prometteurs, notamment sur des GPCRs. Le coût de calculs de cette méthode est drastiquement réduit (de l'ordre de la minute vs semaines ou mois). Il est dès lors évident que ces modèles doivent être considérés lors de la réalisation de projet de modélisation.

Conclusion

Les simulations MD ont grandement contribué dans de nombreux domaines, incluant les interactions protéine-lipide, l'hydratation interne des protéines, et sur les changements conformationnels qui s'opèrent lors de l'activation des GPCRs. Les collaborations rapprochées entre simulations et études expérimentales (particulièrement le "solid state NMR") ont été fructueuses et ont offert des nouveaux aspects des fonctions des GPCRs.