

Projet court : prédiction des segments transmembranaires

MARINE BAILLIF
N° étudiant : 22200055
M2-BI

Introduction :

Les protéines transmembranaires sont une catégorie de protéines qui sont ancrées dans la bicouche lipidique. Elles sont impliquées dans de nombreuses fonctions comme la signalisation ou encore le transport de molécules à travers la membrane ¹. En plus de leurs rôles physiologiques, ce sont d'importantes cibles thérapeutiques. Lors de la résolution des structures protéiques par des méthodes de cristallographie, les lipides de la membrane sont dissous ce qui rend difficile l'identification des segments membranaires d'une protéine. De nombreuses méthodes ont été développées au fil du temps afin d'arriver à prédire avec une plus ou moins grande efficacité les parties transmembranaires d'une protéine. Néanmoins, cela reste un réel challenge. Parmi les nombreuses méthodes, les travaux de Tusnády et al.² ont permis en se basant sur des critères d'hydrophobicité et de géométrie de prédire les segments transmembranaire de certaines protéines. En s'appuyant sur cette méthode, l'objectif de ce projet est de prédire le meilleur plan de membrane de protéine transmembranaire en se basant sur l'hydrophobicité.

Matériels et Méthodes

Afin d'utiliser le programme développé en python, un fichier de structure 3D obtenu de la PDB (Protein Data Bank) contenant la chaîne de la protéine d'intérêt est utilisé. Dans ce cas, les protéines prises en compte sont celles utilisées dans les travaux de Tusnády et al et d'autres protéines de la base de données OPM (Orientations of Proteins in Membranes).

À partir de la structure 3D l'accessibilité au solvant est calculée grâce au logiciel DSSP³. Un acide aminé est considéré comme exposé (accessible) quand sa valeur relative d'accessibilité au solvant est supérieure à 30 %, il est sinon considéré comme enfoui. En connaissant la nature hydrophobe et hydrophile des acides aminés ce paramètre permet d'estimer les résidus exposés et hydrophobes qui sont ainsi susceptibles d'être exposés à la bicouche lipidique.

Afin de trouver la meilleure position de la membrane au niveau de la protéine, le centre de masse de la protéine est calculé. Il est utilisé afin de construire des droites partant du centre et permettant de quadriller le plus de directions possible. Deux points étant nécessaires à la génération d'une droite, l'algorithme de Saaf et Kuijlaars⁴ a été utilisé afin d'obtenir la distribution de points à la surface d'une demi-sphère. Le nombre de points à placer sur la demi-sphère est précisé par l'utilisateur du programme. Les droites seront ainsi calculées en prenant le centre de masse et chaque point de la demi-sphère.

Afin de "modéliser" la membrane sur chaque droite, deux plans sont utilisés. Les plans sont espacés de 15 Å. Chaque membrane (une par axe) est dans un premier temps glissée le long de

la droite (dans un sens puis l'autre) afin d'estimer pour chaque direction, la tranche avec l'hydrophobicité relative la plus élevée. Cette hydrophobicité relative est déterminé en prenant en compte le nombre de résidus exposés polaires hors membrane, le nombre de résidus exposés polaires, le nombre de résidus hydrophobes dans la membrane et le nombre de résidus hydrophobes :

$$\text{hydrophobicité relative} = \frac{\text{Nb res exposés polaires hors membrane}}{\text{Nb res exposés polaires}} + \frac{\text{Nb exposés res hydrophobes dans la membrane}}{\text{Nb res hydrophobes}}$$

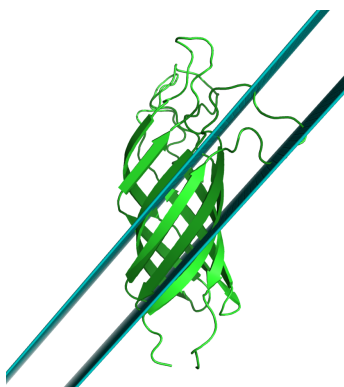
Les membranes sont déplacées par pas de 1 Å et à chaque pas l'hydrophobicité relative est calculée. Sur chaque droite, les membranes sont déplacées tant que des résidus exposés sont présents dans la tranche. Cette première étape permet de sélectionner le plan de membrane maximisant la valeur d'hydrophobicité relative. Il est possible que pour plusieurs membranes la même valeur soit trouvée, dans ce cas-là elles sont toutes conservées pour la suite.

Pour la ou les membranes gardées, l'épaisseur de la membrane est augmentée de 0.5 Å à chaque pas. Lors de chaque augmentation la valeur d'hydrophobicité est calculée. L'épaisseur membranaire est augmentée tant que la valeur d'hydrophobicité augmente et si elle diminue trop ou ne change pas pendant un nombre d'itérations limite l'élargissement est stoppé. Ce nombre, limite d'itération, est fixé par l'utilisateur.

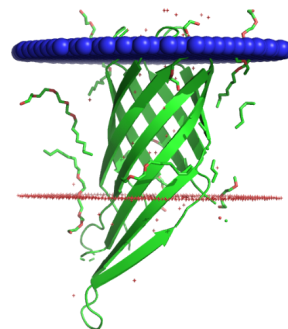
Après avoir déterminé le plan de membrane maximisant l'hydrophobicité relative, les équations des plans correspondant sont utilisés pour générer une représentation 3D de la membrane. Afin que le plan soit centré sur la protéine, les coordonnées maximales et minimales des x et z de la protéine sont utilisées. En sortie, les coordonnées atomiques des atomes du plan sont stockées dans un pdb et peuvent donc ainsi être superposées à la protéine dans un logiciel de visualisation tel que Pymol.

Résultats et Discussion :

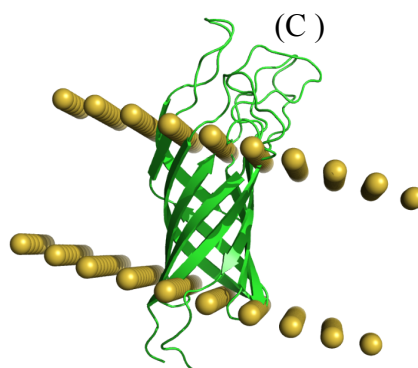
Le programme qui permet de rechercher des plans de membranes pour des protéines transmembranaires à ainsi pu être testé sur plusieurs pdb. Afin d'estimer l'efficacité du programme les résultats ont pu être comparés avec ceux d'autres méthodes de prédiction de segment transmembranaire.



(A)



(B)



(C)

Figure 1 : représentation tridimensionnelle pour la protéine de la membrane externe A (code PDB 1g90) avec des plans de membrane : (A) d'après le programme du projet (B) d'après la base de données OPM (C) d'après OREMPRO

Pour la protéine 1g90 (figure 1) les plans proposés par le programme du projet sont différents de ceux trouvés par les autres méthodes. En effet, l'axe de la membrane est, dans un sens, différent. Globalement parmi toutes les protéines tester très peu ont été correctement bien prédites. Cela est sûrement dû au fait que l'hydrophobicité relative peut difficilement à elle seule prédire les segments transmembranaires d'une protéine. Les autres méthodes combinent d'autres paramètres. Les travaux de Tusnády et al se basent d'ailleurs sur des facteurs structuraux non exploités dans ce programme.

Conclusion :

Le programme n'a permis que dans peu de cas de prédire correctement les segments transmembranaires des protéines. D'autres paramètres seraient à ajouter au programme afin d'augmenter les chances de trouver le bon axe pour la membrane.

1. Tsirigos, K. D. *et al.* Topology of membrane proteins—predictions, limitations and variations. *Curr. Opin. Struct. Biol.* **50**, 9–17 (2018).
2. Tusnády, G. E., Dosztányi, Z. & Simon, I. Transmembrane proteins in the Protein Data Bank: identification and classification. *Bioinformatics* **20**, 2964–2972 (2004).
3. Touw, W. G. *et al.* A series of PDB-related databanks for everyday needs. *Nucleic Acids Res.* **43**, D364–D368 (2015).
4. Saff, E. B. & Kuijlaars, A. B. J. Distributing many points on a sphere. *Math. Intell.* **19**, 5–11 (1997).