

# Détection et Visualisation des membranes dans les Structures Protéiques à l'Aide de PyMOL : Une Approche Basée sur les Plans

GUEDOUAR Sana

## Résumé

Cette note présente une approche pour détecter et visualiser les membranes dans les structures protéiques en utilisant un pipeline bioinformatique intégré avec PyMOL. Nous avons développé des outils pour l'analyse des plans de membrane et la visualisation interactive, facilitant ainsi l'étude des interactions protéines-membranes. Les résultats démontrent que notre méthode est efficace pour identifier les orientations des membranes et générer des visualisations intuitives, bien que certaines limitations et axes d'amélioration soient notés.

## Introduction

Les protéines transmembranaires jouent un rôle essentiel dans divers processus cellulaires et sont cruciales pour la découverte de médicaments et la compréhension des maladies. Cependant, l'identification et la classification de ces protéines sont complexes car la structure de la membrane n'est pas directement résolue. Dans cette note, nous présentons une méthode innovante pour détecter et visualiser les régions membranaires dans les protéines en utilisant des plans géométriques et PyMOL. Notre approche intègre une analyse basée sur des plans avec des outils de visualisation interactifs pour améliorer l'étude des interactions protéines-membrane.

## Méthodologie

### Vue d'ensemble de l'algorithme

Notre algorithme vise à identifier le positionnement optimal des plans de membrane au sein de la structure 3D d'une protéine. L'objectif principal est de maximiser l'hydrophobicité relative des résidus exposés au solvant entre deux plans représentant les limites de la membrane. Voici les étapes clés de notre méthode :

#### 1. Prétraitement des Données :

- **Analyse du Fichier PDB** : Nous chargeons et analysons les fichiers Protein Data Bank (PDB) pour initialiser les structures protéiques et encapsuler les données des résidus, y compris l'hydrophobicité et l'accessibilité au solvant en utilisant l'algorithme DSSP [2].

#### 2. Analyse de la Structure :

- **Génération des Plans** : Nous calculons le centre de masse de la protéine et générons des vecteurs vers des points sur une hémisphère en utilisant l'algorithme de Saff et Kuijlaars [3]. Pour chaque vecteur, nous créons deux plans orthogonaux séparés par un écart par défaut. Ces plans sont analysés alors qu'ils glissent le long de l'axe, en évaluant l'hydrophobicité relative des résidus entre eux.
- 3. **Calcul de l'Hydrophobicité** :
  - **Évaluation** : L'hydrophobicité est calculée en fonction de la proportion de résidus hydrophobes par rapport aux résidus polaires entre les plans. Nous itérons sur diverses positions de plans pour déterminer la configuration avec la plus haute hydrophobicité relative.
- 4. **Visualisation** :
  - **Création d'Objets CGO** : Nous convertissons les plans identifiés en objets CGO pour la visualisation dans PyMOL, permettant une exploration interactive des régions membranaires.

## Programmation Orientée Objet (POO)

Pour assurer une meilleure modularité et maintenabilité, nous avons utilisé le paradigme de Programmation Orientée Objet (POO) dans le développement de notre outil. La POO nous a permis de structurer le code de manière plus claire et efficace :

- **Encapsulation des Données** : Nous avons défini des classes pour encapsuler les informations relatives aux protéines, tels que les objets Plane pour les plans et Axis pour les axes de détection des plans.
- **Modularité** : Le code est organisé en modules distincts, facilitant ainsi l'ajout de nouvelles fonctionnalités et la maintenance.
- **Réutilisabilité** : Les objets et méthodes créés peuvent être réutilisés dans différents contextes ou pour d'autres projets similaires.

## Entrées et Sorties

- **Entrées** : L'outil est conçu pour les protéines transmembranaires à chaîne unique. Les utilisateurs peuvent personnaliser des paramètres tels que le nombre de points initiaux, la largeur de la membrane, et la taille de la fenêtre de glissement pour s'adapter à différentes structures protéiques.
- **Sorties** : L'outil génère une session PyMOL [4] avec la structure protéique et les plans de membrane identifiés. Les résultats sont sauvegardés dans un fichier `.pse`, permettant une visualisation et une exploration ultérieure dans PyMOL.

## Résultats

Nous avons appliqué notre méthode à plusieurs structures protéiques pour évaluer son efficacité. Par exemple, en analysant la structure cristalline de la channelrhodopsine activée par lumière rouge Chrimson. (PDB : 5zih), notre méthode a réussi à identifier les plans de membrane, bien qu'il y ait eu des différences par rapport aux données de référence de la

base de données OPM. Les écarts dans la largeur de la membrane et le positionnement des plans mettent en évidence des domaines d'amélioration.

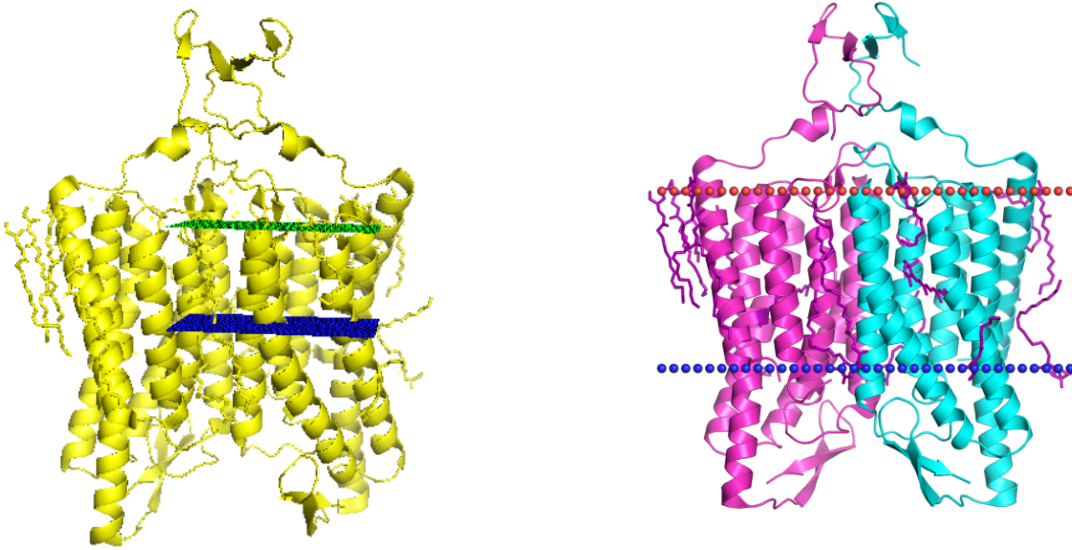


Figure 1 : Comparaison entre la sortie de notre programme (à gauche) et le fichier PDB donné par la base de données OPM [5] (à droite).

## Discussion

Notre méthode offre un outil prometteur pour la détection et la visualisation des membranes, mais révèle également plusieurs limitations :

1. **Contraintes Géométriques** : L'approche actuelle repose sur des critères géométriques qui peuvent ne pas capturer pleinement la complexité des structures membranaires. Les régions membranaires peuvent avoir des formes et des orientations variées qui ne sont pas toujours bien représentées par des plans géométriques simples.
2. **Résolution et Échelle** : La résolution des fichiers PDB d'entrée et l'échelle utilisée pour la visualisation des plans peuvent affecter la précision des plans détectés. Une échelle uniforme peut ne pas être adaptée à toutes les structures protéiques, impactant potentiellement la précision de la représentation des plans.
3. **Limitations de l'Algorithme** : Les algorithmes utilisés pour identifier les axes et les plans peuvent nécessiter des améliorations. Des perfectionnements pourraient inclure l'intégration de caractéristiques géométriques supplémentaires ou l'adaptation de l'algorithme pour mieux correspondre aux structures membranaires déterminées expérimentalement.

Les travaux futurs devraient se concentrer sur l'amélioration de l'algorithme pour mieux gérer diverses structures protéiques et améliorer l'alignement avec les plans membranaires connus. De plus, l'intégration de techniques géométriques plus sophistiquées ou d'apprentissage automatique pourrait améliorer la précision de la détection des membranes. Des études comparatives avec d'autres méthodes et une validation supplémentaire contre

les données expérimentales sont recommandées pour adresser ces limitations et améliorer la performance globale.

## Conclusion

Notre approche fournit un outil utile pour la détection et la visualisation des membranes dans les structures protéiques, facilitant l'analyse des interactions protéines-membranes. Bien que la méthode offre des résultats prometteurs, elle présente des limitations qui nécessitent des améliorations, notamment dans l'optimisation des algorithmes et des paramètres de visualisation. Les futures améliorations viseront à perfectionner ces aspects pour une détection plus précise et une meilleure alignement avec les données expérimentales.

## Références

- [1] Tusnady GE, Dosztanyi Z, Simon I. Transmembrane proteins in the Protein Data Bank: identification and classification. Bioinformatics. doi:10.1093/bioinformatics/bth340 2004;20(17):2964-2972.
- [2] DSSP algorithm: <https://swift.cmbi.umcn.nl/gv/dssp/>
- [3] Saff B, Kuijlaar A.B.J. , Distributing Many Points on a Sphere, 9 Springer Verlag New York, Volume 19. Number 1, 1997
- [4] PyMol Software: <https://pymol.org/2/>
- [5] OPM databank : <https://opm.phar.umich.edu/>