

Profil d'hydrophobicité d'une protéine

I/ Introduction

Il existe en biologie de multiples interactions de natures différentes qui régissent le vivant, notamment les liaisons covalentes et non covalentes. Dans le cas des protéines, ce sont les interactions non covalentes de type hydrophobe qui nous intéressent. L'interaction hydrophobe est un phénomène spontané. En effet, en solution aqueuse, les substances non polaires vont tendre à s'agréger entre elles, ce qui va induire l'exclusion des molécules d'eau. Ainsi, la réalisation de ce regroupement conduit à la réduction de la surface de contact entre les molécules hydrophobes et le solvant.

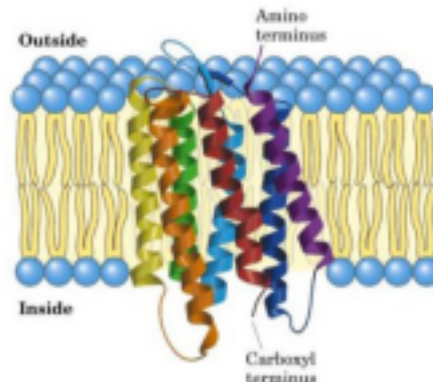
L'hydrophobicité d'une protéine représente sa capacité à interagir avec des molécules d'eau: au plus une molécule est hydrophobe, au plus elle limite ses interactions, à l'inverse d'une molécule hydrophile. Dans le cas des protéines, l'hydrophobicité peut être variable en fonction des différents acides aminés et de ses structures II, III et IV. Ainsi, réaliser le profil d'hydrophobicité ou profil d'hydrophobie consiste à prédire les régions hydrophobes et hydrophiles de protéines en utilisant leurs séquences d'acides aminés. En effet, chaque acide aminé possède une chaîne dont la nature chimique détermine son affinité pour l'eau. Ainsi, certains acides aminés, comme la leucine ou la phénylalanine, ont des chaînes latérales apolaires, ce qui les rend hydrophobes. D'autres, comme la lysine ou l'aspartate, portent des charges ou des groupes polaires, ce qui les rend hydrophiles. La valeur d'hydrophobicité attribuée à un acide aminé indique donc si celui-ci est hydrophobe ou hydrophile. Ces valeurs sont essentielles pour prédire la structure et la localisation des différentes régions d'une protéine, notamment dans le repliement ou l'ancrage membranaire.

Les résultats obtenus sont particulièrement intéressants car ils peuvent permettre de prédire les régions transmembranaires ; certaines échelles permettent de prédire des zones antigéniques potentielles (Hopp-Woods), d'autres les hélices alpha (Cornette) [5], et d'autres usages.

Concernant les protéines transmembranaires, les profils d'hydrophobicité sont utiles pour localiser les segments qui traversent la bicouche lipidique. En effet, leurs rôles, qu'il s'agisse de transport, de signalisation ou de catalyse, dépendent directement de leur bonne insertion dans la bicouche lipidique. Puisque ces régions sont généralement composées d'acides aminés fortement hydrophobes et qui sont organisés en hélices alpha capables de s'insérer entre les queues hydrophobes des phospholipides qui composent la membrane. Ainsi, en analysant la répartition de l'hydrophobicité tout le long de la séquence de la protéine, on peut prédire la présence et la position des domaines transmembranaires de la protéine. La mise en œuvre de ce procédé est essentielle pour la compréhension de la structure et de la fonction des protéines membranaires qui jouent des rôles centraux dans le vivant. Il convient toutefois de noter que les résultats obtenus ne sont que des prédictions et ne tiennent pas compte des facteurs comme la forme dans l'espace de la protéine puisqu'ils sont seulement basés sur la structure primaire. Des analyses plus poussées peuvent être nécessaires pour confirmer les régions transmembranaires par exemple.

II/ Exemple d'intérêt

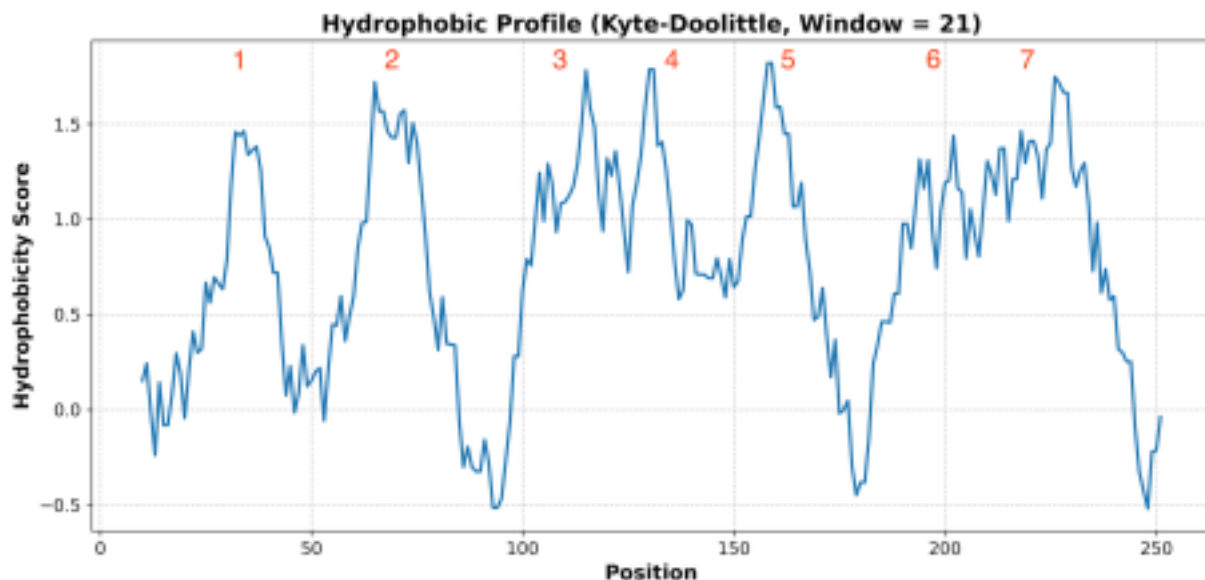
L'exemple d'intérêt est la bactériorhodopsine. C'est une protéine activée par la lumière, présente dans la membrane de certaines archées halophiles telles que *Halobacterium salinarum* [2]. Celle-ci vit dans des environnements extrêmes comme les marais salants ou les lacs hypersalés. La protéine est au cœur d'un système de conversion de l'énergie lumineuse en énergie chimique, permettant à ces micro-organismes de produire de l'ATP en l'absence de photosynthèse classique. Son rôle est d'assurer le transport actif de protons vers l'extérieur de la cellule afin de permettre la création d'un gradient électrochimique, par la suite exploité pour la synthèse d'ATP [1].



La bactériorhodopsine est composée de sept hélices alpha transmembranaires formées principalement d'acides aminés hydrophobes. C'est une protéine idéale pour réaliser un profil d'hydrophobicité. Les segments traversent la bicouche lipidique et assurent l'ancrage stable de la protéine dans la membrane. Quant aux zones situées à la surface extracellulaire, elles sont principalement composées d'acides aminés hydrophiles, ce qui est courant chez les protéines membranaires. Cette distribution nous permet de visualiser clairement les segments transmembranaires lors d'un profil d'hydrophobicité. Ce sont les pics hydrophobes qui correspondent aux hélices insérées dans la membrane.

La méthode standard pour calculer le profil d'hydrophobicité consiste à définir une échelle de manière expérimentale, assignant à chaque acide aminé une valeur correspondant à son caractère hydrophobe. Ensuite, pour chaque acide aminé de la séquence protéique, on réalise une moyenne de l'hydrophobicité des acides aminés alentour, l'acide courant étant au centre. La moyenne peut s'effectuer sur un nombre variable d'acides aminés appelé fenêtre. Différentes fenêtres peuvent mettre en évidence différentes propriétés : l'échelle Kyte Doolittle prédit les régions exposées ou non de la protéine avec de petites fenêtres (5-7), et prédit les régions transmembranaires avec des fenêtres plus grandes (19-21) [3]. Enfin, on réalise un graphe présentant les différentes valeurs calculées, ce qui nous donne le profil d'hydrophobicité ou d'hydropathie.

Pour réaliser le profil, nous récupérons en fichier FASTA ou PDB la séquence protéique. Pour cet exemple, nous utilisons l'échelle d'hydrophobicité de Kyte et Doolittle [3].



Le profil obtenu nous permet d'identifier clairement les 7 hélices alpha hydrophobes, puisque nous obtenons 7 pics distincts avec un score d'hydrophobicité élevé. Cela signifie que les acides aminés à ces positions sont plutôt hydrophobes. Ainsi, le résultat obtenu est cohérent avec les résultats connus sur la bactériorhodopsine, ce qui conforte l'utilité de cette technique [1].

Mais cette technique a des limites : le profil ne permet pas, par exemple, de déterminer la structure complète ; il ne montre ni l'orientation des segments, ni les boucles extra- et intracellulaires.

Aussi, selon l'échelle utilisée et la technique de calcul, les résultats ne sont pas les mêmes et des régions hydrophobes de petite taille peuvent ne pas apparaître. Si on change la taille de la fenêtre, il peut être difficile de déterminer les régions transmembranaires. Pour confirmer les résultats, il est donc envisageable de réaliser de la cristallographie [4].

III/ Programme Python

Un programme permettant de générer un profil d'hydrophobicité a été réalisé. Il est téléchargeable depuis le repository [GitHub](#) [6]. La procédure d'installation est spécifiée dans le fichier README.md. Afin d'utiliser l'outil pour générer un profil d'hydrophobicité, exécutez la commande: `python3 cli.py gui`, ce qui lancera l'interface graphique du programme. On peut aussi utiliser l'outil depuis le terminal en exécutant: `python3 cli.py hydrophob -h`. -h retourne une liste des options possibles.

Le programme est divisé en plusieurs parties, celles qui nous sont importantes sont: le fichier `cli.py`, `scales.py`, `Main.py` et le dossier `pages/`, le dossier `hydrophob/`.

cli.py :

Ce script est la porte d'entrée du programme et permet d'exécuter différentes commandes. Selon ce qui est saisi par l'utilisateur, `cli.py` redirige vers tel ou tel outil.

scales.py :

Ce script contient un ensemble de fonctions pour charger les différentes échelles sauvegardées dans le dossier `data/` et manipuler de manière générale ces échelles.

Main.py et le dossier pages/ :

Pour l'interface, nous avons utilisé le package `streamlit` qui permet de réaliser rapidement et facilement des interfaces web prototypes et plutôt modernes. Le fichier `Main.py` est la porte d'entrée pour accéder aux différents outils en mode GUI.

Dans le dossier `pages/` on trouve les outils en mode GUI: `Scales.py` permet de gérer les échelles (ajout, modification, suppression) et `Hydrophob.py` permet de générer des profils d'hydrophobicité et de les télécharger.

Le dossier hydrophob/ :

Le dossier `hydrophob/` contient un fichier `__init__.py` et un fichier `utils.py`. Le fichier `__init__.py` indique à python que le dossier est un package, il contient aussi une fonction `run(<args>)` qui permet à `cli.py` de générer des graphes en mode CLI (ligne de commande). Le fichier `utils.py` contient une seule fonction: `compute_profile(seq, scale, window)` permettant de calculer le profil d'hydrophobicité sur une séquence donnée, selon une échelle et une taille de fenêtre.

IV/ Références

1. National Center for Biotechnology Information (NCBI). (n.d.). UEB91126.1.
<https://www.ncbi.nlm.nih.gov/protein/UEB91126.1>
2. Darling, D. (n.d.). Bacteriorhodopsin.
<https://www.daviddarling.info/encyclopedia/B/bacteriorhodopsin.html>
3. Kyte, J. and Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J Mol Biol.*
[https://doi.org/10.1016/0022-2836\(82\)90515-0](https://doi.org/10.1016/0022-2836(82)90515-0)
4. Engelman, D. M., Steitz, T. A., and Goldman, A. (1986). Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Annu Rev Biophys Chem.*,
<https://doi.org/10.1146/annurev.bb.15.060186.001541>
5. Cornette, J. L., Cease, K. B., Margalit, H., Spouge, J. L., Berzofsky, J. A., and DeLisi, C. (1987). Hydrophobicity scales and computational techniques for detecting amphipathic structures in proteins. *J Mol Biol.*
[https://doi.org/10.1016/0022-2836\(87\)90189-6](https://doi.org/10.1016/0022-2836(87)90189-6)
6. BioUtils: A small set of tools for bioinformatics
<https://github.com/Egeyae/BioUtils>