Assignation and detection of protein transmembrane regions

Iphigenie Gonnet

September 14, 2022

Abstract

As transmembrane proteins make prime drug targets, it is important to develop the right tools to study them. As their biological activity is correlated with their position in the membrane, we present here a python tool to assign transmembrane regions for proteins and visualize them.

1 Introduction

A big potion of all proteins are embedded in the membrane. Because those membrane proteins play a central role in a plethora of diverse cellular function, they make prime drug targets. [3] We know that the conformation and biological activity of these proteins depends highly on the way they are placed in the membrane.[1] Being able to compute the position of the membrane given a protein PDB structure would help answering those problematics.

2 Approach

Our approach was loosely based on the article by Tusnády, Gábor E *et al.*[4]
To start computation, we retrieve the accessible alpha carbons of the protein chains and center them on a Cartesian coordinate system.
We then use a Fibonacci's lattice sphere algorithm to generate uniform points on a hemisphere. With a 100 sample points we obtain the distribution shown Figure1. These points support vectors that indicate the direction perpendicular to a membrane plane. Looping through all generated points, we obtain membranes uniformly distributed across all possible directions. For each membrane plane generated this way, we create a membrane object of a width selected by the user (by default : 14 Å). Looping through all alpha carbons retrieved, we keep the ones localized in the membrane object and compute the relative hydrophobicity of the corresponding collection of residues.
Hence why we only need a hemisphere of these points as we scan through the entire protein for each selected membrane plane.
After all iterations we keep the membrane position that yielded the best hydrophobicity score overall.

3 Exemple

The following exemple is guided through on the git repository of this project. Found on the OPM database [2], we work with the transmembrane protein Structure of Sarcolipin (PDB 1JDM) show Figure 2. Our Algorithm doesn't yield results but loops infinitely instead. Given the results obtained by the other student with this project we could have reasonably expected to find a membrane placement close to the one in Figure 2 as it is a simple protein and we probably would have had more trouble with more complex structures.

4 Discussion

Given that our code doesn't yield results one can conclude that I didn't acquire the project managment skills required by this unit yet. I would love to try and understaind better where I went wrong what parts of my process should be improved on. It is likely that even though I started coding quite carefully

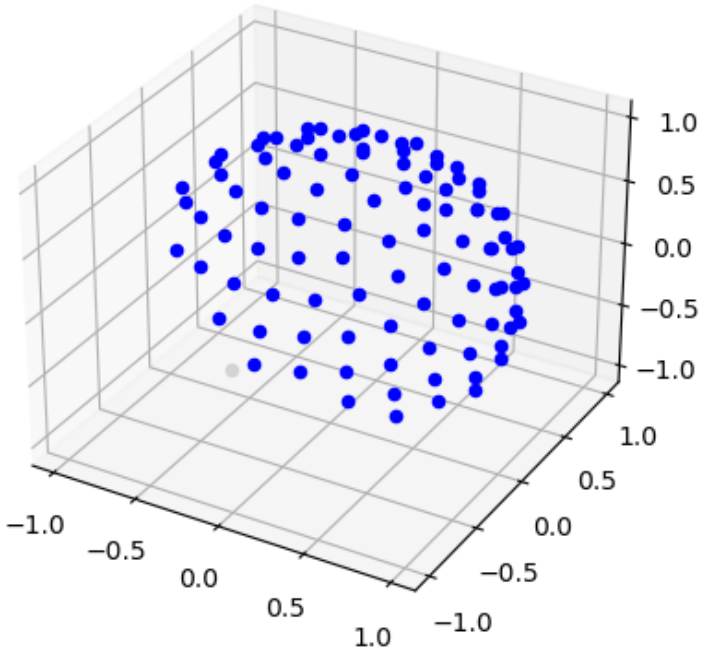

Figure 1: 100 Fibonacci sphere points

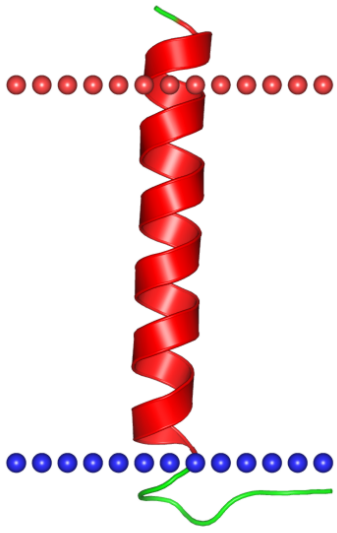

Figure 2: Sacopilin structure with membrane position registered in the OPM database

testing each separate element as I went, I got more careless as the deadline got closer. Once I arrived to a point where my code is all together but doesn't work with no way of knowing where the problem comes from, debugging all scripts to fin the critical error was an overwhelming task in the remaining time. In retrospect, I should not have tried OOP style coding as I am not yet totally at ease with it and I probably lost some valuable time because of it.

However I have a ferm grasp on the logical aspects of the problem and were I to start implementing it again I feel confident that I would succeed.

References

[1] Andrei L Lomize, Irina D Pogozheva, Mikhail A Lomize, and Henry I Mosberg. The role of hydrophobic interactions in positioning of peripheral proteins in membranes. *BMC Struct. Biol.*, 7(1):44, June 2007. doi:10.1186/1472-6807-7-44.

[2] M. A. Lomize, A. L. Lomize, I. D. Pogozheva, and H. I. Mosberg. OPM: Orientations of proteins in membranes database. *Bioinformatics*, 22(5):623–625, January 2006. doi:10.1093/bioinformatics/btk023.

[3] Hwee Tong Tan Sandra Tan and Maxey C. M. Chung. Membrane proteins and membrane proteomics. *PROTEOMICS*, 8(19):3Z24–3932, 2008. doi:https://doi.org/10.1002/pmic.200800597.

[4] Gábor E Tusnády, Zsuzsanna Dosztányi, and István Simon. Transmembrane proteins in the protein data bank: identification and classification. *Bioinformatics*, 20(17):2964–2972, November 2004. doi:10.1093/bioinformatics/bth340.