PROJECT MSI 1: MEMBRANE ADDITION IN PDB INTERMEMBRANE PROTEIN

Joan Francesc Gilabert & Andreu Bofill
April 16

USAGE: membrane.py name.pdb

The function of this python module is to obtain a HTMD molecule with a PDB intermembrane protein plus a network of atoms representing the membrane leaflet.

MATRIX ROTATION

First of all, we need to load the pdb file and download and save the molecule to a variable.

With this object, we apply a function called create_planes(), that allows us to rotate the molecule parallel to the z axis, so it will be perpendicular to the membrane (which will be located in the x,y plane). We know that, theoretically, the aromatic rings formed from every aromatic amino-acids that are anchored to the membrane are parallel to it, so knowing the mean of all perpendicular vector of each aromatic-ring, we can obtain a vector that is indicative of the orientation of the protein, to rotate it into the z axis.
To do that, we can obtain the coordinates of all the Ca (alpha carbons) and two other Carbon atoms from the aromatic ring and create two vectors with them.

Then, we need to calculate the rotation angle between the vector calculated and the raw initial z-axis and the axis around where the rotation will take place. This axis is simply the cross product between the mean vector and the z-axis. With the coordinates of the rotation axis and the rotation angle (in radiants), we can rotate the molecule.

At this point, we got the molecule with the orientation that will allows us to put the membrane atoms in the good plane.

MEMBRANE LOCATION

Now, we need to know the z-coordinate of the membrane plane. To go trough this, we developed another function called histog(), that will return a set of plots about the distribution of hydrophobic and charged residues through a set of z-axis segments. Hypothetically, we will find more hydrophobic residues in the intermembrane-residues than in the outside residues. Knowing that, we expect to find two clear boundaries on the distribution of hydrophobic residues through the z-axis.

In order to obtain a better representation of the distribution, we calculate the "density" of hydrophobic residues. This density is basically the ratio between the number of hydrophobic residues and the total number of residues in a given segment of z. This allows us to better differentiate the regions with higher local concentration of hydrophobic residues, which are supposed to be the intramembrane region. Moreover, we apply a windowing to this density plot, where for each z segment we calculate the mean between its density, the density of the successor and the density of the predecessor, this windowing allows for a smoother curve and facilitates the prediction of the membrane position.

Next, we need to read this information, and to do that we develop the find_z() function which returns a membrane start and end positions. To predict these positions, we assume that the mean gives a good enough threshold for the prediction, i.e. the membrane will be the z-segment that has a windowed hydrophobic frequency greater than the mean. The functions return the left and right extremes of that segment as the positions of the membrane.

When we locate the places that supposedly the membrane should be, we need to create the membrane with a slab of dummy atoms. We need to know the minimum and the maximum coordinates of the molecule, to create a right size atom matrix. To create this matrix, we create the dummy_leaflet() function. In addition, we also need to give to the function the start and end positions in the z-axis, where we will put the dummy atom matrix.

PDB EXAMPLES

We test this program with the following PDBs: 1dxr, 1m0l, 1qjp, 1r3j, 2maf; and we compare with the same in OPM to be sure that we add the membrane in a quite good position. The pdb file with our predictions can be found in the folder "pdb" of our git hub repository (https://github.com/cescgina/Membrane_MSI) with the filename PDBcode_mem.pdb. The plots with the hydrophobic density and other information can be found in the folder "plots".

Certainly, this program does not work with all possible intermembrane proteins PDBs. It may be a problem of our molecule orientation approach. To orientate the molecule, we calculate the mean perpendicular vector of the aromatic rings of all aromatic amino-acids of the molecule. This approach does not so well, because we use all the aromatic amino-acids and not only these ones that are closer to the membrane. In addition, other approaches should be considered, since base it in only this aromatic rings method seems to be insufficient.