

## 1. Lipid bilayer and protein (from OPM database)

Option 5 and 6 of AMBAT applies to protein insertion. In this example we insert a medium size polypeptide (CXCR4) in a 242 POPC lipid bilayer. CXCR4 PDB (4rws) is downloaded from OPM database hence, its already in the right orientation with respect to the bilayer. AMBAT is fully compatible with PDB's downloaded from OPM database.

(a) Choose Option '5' after executing AMBAT code.

```
tarunkhanna@Taruns-MacBook-Pro: ~/Desktop/Examples$ tclsh AMBAT_MB.tcl
*****
THIS IS AMBER BASED LIPID BILAYER BUILDER FOR
LIPID BILAYER AND TRANSMEMBRANE PROTEIN
SIMULATIONS SETUP
*****

DEVELOPED BY TARUN KHANNA AND DR. IAN GOULD
IMPERIAL COLLEGE LONDON, U. K.

### ENTER THE TASK YOU WANT TO PERFORM

CHOOSE ANY OF THE BELOW OPTIONS
0 = BUILD A PURE LIPID BILAYER
1 = INSERT A MOLECULE INSIDE AND BUILD A BILAYER
2 = INSERT A MOLECULE OUTSIDE AND BUILD A BILAYER
3 = INSERT A MOLECULE INSIDE A PREBUILD BILAYER
4 = INSERT A MOLECULE OUTSIDE A PREBUILD BILAYER
5 = INSERT A PROTEIN AND BUILD A BILAYER AROUND IT
6 = INSERT A PROTEIN IN A PREBUILD BILAYER
7 = TO INITIATE THE VESICLE BUILDER
8 = TO INITIATE THE MEMBRANE BUILDER FROM A INPUT FILE AND A LIPID
PDB FILE
9 = JUST SOLVATE THE MEMEMBRANE SYSTEM
IMP NOTE: PRE BUILD BILAYER SHOULD BE IN THE AMBER FORMAT AND IN T
HE SAME FOLDER WHERE THE CODE IS EXECUTED
```

(b) Give the composition of the upper and the lower leaflet. For more information regarding this input please refer pure bilayer example.

```
### ENTER THE NUMBER OF POPC LIPIDS IN UPPER LEAFLET
121
### ENTER THE NUMBER OF DIFFERENT LIPIDS IN LOWER LEAFLET
1
(FOLLOW LIPID 14 NOMENCLATURE)
ENTER THE HEAD GROUP OF LIPID 1
PC
ENTER FIRST TAIL GROUP OF LIPID 1
OL
ENTER SECOND TAIL GROUP OF LIPID 1
PA
### ENTER THE NUMBER OF POPC LIPIDS IN LOWER LEAFLET
121
```

```

9 = JUST SOLVATE THE MEMEBRANE SYSTEM
IMP NOTE: PRE BUILD BILAYER SHOULD BE IN THE AMBER FORMAT AND I
N THE SAME FOLDER WHERE THE CODE IS EXECUTED
5
ET
1
PC
OL
PA
121
    ### ENTER THE NUMBER OF DIFFERENT LIPIDS IN UPPER LEAFL
    ( FOLLOW LIPID 14 NOMENCLATURE)
    ENTER THE HEAD GROUP OF LIPID 1
    ENTER FIRST TAIL GROUP OF LIPID 1
    ENTER SECOND TAIL GROUP OF LIPID 1
    ### ENTER THE NUMBER OF POPC LIPIDS IN UPPER LEAFLET

```

- (c) AMBAT will then enter the protein insertion suite of AMBAT. Input the name of the protein PDB. Type in '4rws.pdb'

```

.....
THIS CODE HELPS TO ORIENT THE MEMBRANE PROTEIN IN
RIGHT ORIENTATI ON
.....

DEVELOPED BY TARUN KHANNA AND DR. IAN GOULD
IMPERIAL COLLEGE LONDON, U.K.

    ### ENTER THE NAME OF THE PDB ###
    ( MAKE SURE TO REMOVE THE REMARK SECTION OF
THE PDB)

```

- (d) Inbuild AMBAT\_PI suite allows for the manual movement of the polypeptide (which is really cumbersome). Hence, we recommend the PDB to be downloaded from OPM database in right oreintaton with respect to the bilayer. So, type in 'N' for manual insertion.

```

    ### DO YOU WANT TO MANUALLY INSERT THE
PROTEIN BASED ON THE HYDROPHOBIC AND HYDROPHYLIC AMINO ACIDS? (Y/N)
n
    ***** NOTE : BASED ON THE INPUTS YOU GIVE BELOW
THE INPUT PROTEIN PDB WILL BE TRANSLATED AND ROTATED ACCORDING TO THES
E VALUES *****

    ### ENTER THE VALUE FOR Z ROTATION (NOTE: PROT
EIN IS INVARIANT UNDER Z ROTATION, THIS VALUE ONLY CONTROLS THE SIMULAT
ION SETUP) ###

```

- (e) As mentioned at the end, protein is invariant under z rotation. So, type in **0.0** for this input.

```

##### ENTER THE VALUE FOR Z ROTATION (NOTE: PROT
EIN IS INVARIANT UNDER Z ROTATION, THIS VALUE ONLY CONTROLS THE SIMULAT
ION SETUP) #####
0.0
6.435 ***
1.329 ***
7.509 ***

*** THE SPREAD OF THE PROTEIN ALONG X AXIS IS 4
*** THE SPREAD OF THE PROTEIN ALONG Y AXIS IS 4
*** THE SPREAD OF THE PROTEIN ALONG Z AXIS IS 9

DETERMINING THE PROTEIN PROFILE ***
(x,y,z) = (0,1,2)
-59.039 ALONG DIRECTION 2
-58.039 ALONG DIRECTION 2
-57.039 ALONG DIRECTION 2
-56.039 ALONG DIRECTION 2
-55.039 ALONG DIRECTION 2
-54.039 ALONG DIRECTION 2

```

- (f) There are few important files which will be generated at the end of this step “pro\_ori.pdb” which shows the orientation of the polypeptide with respect to the bilayer (only placement along z dimension is important) , “ac\_nature”; which contains 6 columns for the density of various classes of amino acids along the bilayer normal. Column 2 and 3 represents the most important class of hydrophobic and hydrophilic amino acids respectively ( (-36.0 to 0) is the bilayer interior) and “ref1\_no.pdb” which will be used further in the code for insertion into the bilayer

For version, as the bilayer is symmetric select 3.0 and for non-random grid type ‘N’ (non-random grid is explained further in pure bilayer example).

```

***** CHECK CAREFULLY THE FILES GENERATED BEFORE CONTINUING *****
WHICH VERSION DO YOU WANT TO EXECUTE? (
VERSION 3.0 IS THE LATEST ONE)

##### BOTH LIPID BILAYER AND INSERTION
PART OF THE CODE WILL BE EXECUTED #####
##### DO YOU WANT TO FORM A NON-RANDOM G
RID BASED ON SIMPLE == AND != RULES? #####

```

- (g) Next, AMBAT will form a 242 lipids POPC bilayer and insert 4rws.pdb inside it.

```

***** FORMING A LIPID BILAYER *****
***** BUILDING THE UPPER LAYER ACCORDING TO METH
OD1 *****
***** BUILDING THE LOWER LAYER ACCORDING TO METH
OD1 *****
***** PUTTING THE LIPIDS INSIDE *****
***** PDB :: POPC_A.pdb *****
***** USING VERSION 3.0 *****

```

```

***** INSERTING A PROTEIN WITH PDB NAME
ref1_no.pdb INSIDE THE LIPID BILAYER *****
***** CHECKING RESIDUE 1 *****
***** CHECKING RESIDUE 2 *****
***** CHECKING RESIDUE 3 *****
***** CHECKING RESIDUE 4 *****

```

- (h) After forming a pure bilayer and inserting a polypeptide, AMBAT will print out the final bilayer composition. And ask the user for solvation. Type in 'Y' to solvate the system.

```

***** CHECKING THE FINAL COMPOSITION OF LIPIDS *
***** CHECK CAREFULLY *****
***** LIPID COMPOSITION BEFORE AND AFTER PROTEIN
INSERTION *****
***** DO YOU WANT TO SOLVATE THE SYSTEM
? (REQUIRES AMBERTOOLS) (Y/N) *****

```

LIPID	INITIAL(UL)	INITIAL(LL)	FINAL(UL)	FINAL(LL)
PC	121	103	121	103

- (i) Type in 'n' for the next input regarding more than 100,000 residues. And type in the parameter files. As this contains only protein and lipids, ff99SB/ff14SB and lipid 14 will suffice.

```

***** DOES THE SYSTEM CONTAINS MORE THAN
100,000 RESIDUES (EXCLUDING WATER)? (Y/N) *****
n
***** ENTER THE NAME OF THE PARAMETER FI
LES (EACH SEPARATED BY A SPACE) *****
ff99SB lipid14
***** IS THERE ANY ADDITIONAL PARAMETER
FILE? (frcmod file) (y/n) *****
n
***** IS THERE ANY .LIB OR .OFF STRUCTUR
E FILES YOU WANT TO ADD? (Y/N) *****
n
***** IS THERE ANY .prepin STRUCTURE FIL
ES YOU WANT TO ADD? (Y/N) *****
n

```

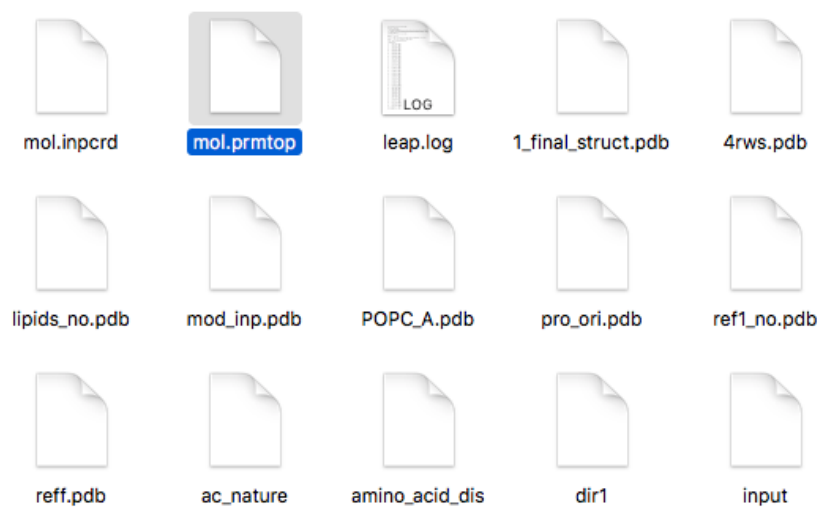


(j) Next, AMBAT will ask for water thickness. Type in '**10.0**'.

```
##### ENTER THE THICKNESS OF THE WATER L
AYER IN ANGSTOMS YOU WANT TO ADD #####
```

(k) Finally, AMBAT will print out the description of various files formed. And prmtop and inpcrd files for running MD simulation.

```
##### DESCRIPTION OF THE FILES #####
(NOTE: SOME FILES MAY NOT BE GENERATED
DEPENDING UPON THE OPTION YOU HAVE CHOSEN)
lipids_no.pdb : UNSOLVATED PDB OF PURE LIPID BILAYER
1_final_struct.pdb : UNSOLVATED PDB OF LIPID BILAYER WITH INSERTED COMP
ONENT
lip.pdb: SOLVATED PURE LIPID BILAYER
reff.pdb: FINAL SOLVATED PDB OF THE SYSTEM
mol.prmtop: FINAL AMBER PRMTOP FILE
mol.inpcrd: FINAL AMBER COORDINATE FILE
dir1: PROFILE OF THE PROTEIN INSIDE THE BILAYER
lipids_wo.pdb: FINAL DRY VESICLE OR DIBs
sol_lipid.pdb: FINAL SOLVATED VESICLE OR DIBs
```



## 2. Protein not in OPM database (AMBAT\_PI example)

In this example we would like to show the functioning of AMBAT\_PI code for the proteins which are not there in OPM database or in some rare cases like 2n28 (virus protein U) wrongly identified.

**NOTE:** AMBAT\_PI does an extensive search on the regions around the bilayer and hence is very slow at this point.

AMBAT\_PI - Stands for AMBAT'S protein insertion code. It runs as a separate script and have its own interactive command line interface.

Enter the name of the PDB you want to orient. For this example, we use **2n28.pdb** (downloaded from RCSB database, with remarks section removed).

```
tarunkhanna@Taruns-MacBook-Pro: ~/Desktop/Examples$ tclsh AMBAT_PI.tcl
.....
THIS CODE HELPS TO ORIENT THE MEMBRANE PROTEIN IN
RIGHT ORIENTATION
.....
DEVELOPED BY TARUN KHANNA AND DR. IAN GOULD
IMPERIAL COLLEGE LONDON, U. K.
##### ENTER THE NAME OF THE PDB #####
(MAKE SURE TO REMOVE THE REMARK SECTION OF
THE PDB)
```

Next, if you have an option to either just put the protein at the centre of the box (N) or use AMBAT\_PI to position it (Y). Type 'Y' to try AMBAT\_PI

```
NOTE: AMBAT_PI does an extensive search on the regions around the bilayer and hence is very slow at this point.
AMBAT_PI stands for AMBAT'S protein insertion code. It runs as a separate script and have its own interactive command line interface.
##### ENTER THE NAME OF THE PDB #####
(MAKE SURE TO REMOVE THE REMARK SECTION OF
THE PDB)
2n28.pdb
##### DO YOU WANT AMBAT PI TO TRY TO FIND THE BEST POSITION FOR THE PROTEIN? (Y/N) #####
```

As the speed is a real problem with this code, (after a wait of around 15-20 min) you will have the final structure 'ref1\_no.pdb' and its representative orientation ("pro\_ori.pdb") in the execution folder.

