PepSuMD survival guide

v0.91

GENERAL INFO:

pepSuMD [1] is a more general version of SuMD [2,3] able to consider also peptide as ligand. In any case, They share the same supervision algorithm and all the functions of SuMD are included within pepSuMD. The algorithm is deeply explained in the 2016 reference.

INSTALLATION:

Requirements:
-pepSuMD is a python code (python 2.7) and requires two libraries:
numpy
prody
you can install them as you want, for example with *pip*
(e.g. sudo pip install numpy; sudo pip instalProDy==1.9)
-pepSuMD is based on AceMD MD engine, that can be obtained from acellera.com

Installation:
The code is an executable, just unpack the tar.gz and place the folder some in your machine.
For example:
#extract the pepSuMD folder:
tar -xvzf pepSuMD_0.91.tar.gz

#mv the folder wherever you prefer according to your user permissions:
mv pepSuMD_0.91/ /usr/local/

#to run digit the entire path of the executable:
/usr/local/pepSuMD_0.91/suMD

CONFIGURATION:
In the pepSuMD_0.91 folder, you will find a hidden file named *.suMDrc* in which the user has to declare the executable name (if executable from anywhere) of Acemd. The default setting should work for most of acemd installation (e.g. Acemd=acemd).

RUNNING SuMD:

pepsumd simulations can be performed starting from a pre-equilibrated system in which the ligand is placed far from the protein. Consider that the distance separating the protein from the ligand directly affects the system size and, therefore, the speed of the simulation. Nonetheless, take care of the PME threshold (9 Å), the long-range component of attractive forces. So the general idea is to place the ligand at a distance, at least, bigger than the PME cut-off from any protein atom. The distance should be set even depending on the hydrodynamic properties of the ligand and the complexity of the binding event. As rule of thumb, we suggest to place the ligand in a random conformation, in a range of 30-70 Å.

pepSuMD requires a configuration file (here named "*selection.dat*") organized in three major sections containing information about (i) the system, (ii) the supervision procedure, and (iii) the simulation settings. In the system settings section, the following details about the molecular system need to be provided: (i) the *PDB* file name containing the starting coordinates, the *VMD selection* style of the ligand, the *residues* describing the target binding site. In the supervision settings section (ii), the following values are declared: the *slope threshold* (default value: 0) and the number of maximum consecutive *failed steps* (default value: 31) to stop the simulation. In the simulation settings section, the following details must be specified: the *force field* to use, the *parameter* file, the SuMD-number of steps, the MD time step, and the *GPU* device ID to which the calculation will be addressed. In this section, a Boolean operator manages the introduction of a randomization step that changes the ligand positionthrough 600 ps (Section 2.5.2) of nonsupervised MD simulation. In the same directory from which SuMD is launched, a file containing the cell dimension as a parameter file as well (prmtop/psf with the same name of the PDB) must be provided.

###SYSTEM SETTING
structure=file.pdb #pdb file name assigned to the last frame of equilibration (psf or prmtop; xsc should have the same root name)
ligand=chain B #vmd selection style of ligand/peptide
ligand_cm=chain B #vmd selection style of the ligand/peptide region desired for mass centroid calculation (it may coincide with the previous selection)
residues=380 353 207 189 #id (vmd resid) binding site residues by which mass centroid is calculated (supervision target)
###SUPERVISION SETTINGS
randomize=no #perform randomization step (classical non-supervised MD) before run : yes/no
n_device=0 # GPU device ID-number (similar to *acemd --device* command)
constrain=no #experimental, leave unchanged
timestep=2 #MD time step
slope=0.00 #unchanged
MaxFailed=31 #maximum number of consecutive failed attempts
###SIMULATION SETTINGS
n_steps=300000 #classical MD steps for each supervised one
ForceField=CHARMM #force field, AMBER and CHARMM
parameters=par_all27_prot_lipid.prm #FF parameters for CHARMM FF
opt_dist=2 #v. paper
bound_dist=5 #v. paper
meta_dist=9 #v. paper
non_prot_AA=SAR # OPZ. report here non-standard residues (this line can be present several times)

Command to start a pepSuMD simulation:
/usr/local/pepSuMD_0.91/suMD selection.dat

Note:
There is a selection example in the folder

REFERENCE

(1) Salmaso, V.; Sturlese, M.; Cuzzolin, A.; Moro, S. Exploring Protein-Peptide Recognition Pathways Using a Supervised Molecular Dynamics Approach. *Structure* 2017, *25*, 655–662.e2.

(2) Cuzzolin, A.; Sturlese, M.; Deganutti, G.; Salmaso, V.; Sabbadin, D.; Ciancetta, A.; Moro, S. Deciphering the Complexity of Ligand-Protein Recognition Pathways Using Supervised Molecular Dynamics (SuMD) Simulations. *J. Chem. Inf. Model.* 2016, *56*, 687–705.

(3) Sabbadin, D.; Moro, S. Supervised Molecular Dynamics (SuMD) as a Helpful Tool to Depict GPCR-Ligand Recognition Pathway in a Nanosecond Time Scale. *J. Chem. Inf. Model.* 2014, *54*, 372–376.