Rosalind J. Xu 18’ Spring 2018
呢

Cyanylation on protein + water + ions (the entire system) — using cyanylator_system.py

Good for: Inserting the MSCN probe and continuing a simulation from a frame in MD trajectory, with the waters / ions around the protein.

Structural Editing:

1. If necessary, rewrite the residue and / or hydrogen names in .pdb files so that they match the force field .rtp entries.
    1. Change the residue names so that they reflect the correct protonation states.
    2. Use **switchHydrogenName-charmm.awk** to edit the hydrogen names (for charmm force field)
    3. Check to make sure that all arginines have NH1 cis to CD, and NH2 trans to CD. If not, switch the names of NH1, NH2 and attached hydrogens.

2. Replace the respective residues in each representative structure with the SCN probe:
    1. Edit **cyanylator-system.py**, which does cyanylation on a simulation system (includes protein, water, ions, etc). Run in Chimera to generate cyanylated .pdb files.
        1. Selection of rotamers: criteria: c (clash cores) - h (h-bonds) - p (probability in rotamer library)
    2. Open each .pdb file in Chimera. Adjust CA-CB-S-C dihedrals.
        1. Type in command line “select #0:[MSCN residue number]@SD z<7”, and Actions -> Atoms/Bonds -> Show, to see nearby waters / residues
        2. Adjust CA-CB-S-C dihedrals to 180 degree
        3. Further adjustment in case of clashes (definition of clashes: for non-water: < 3 A; for water: < 2.5 A)
    3. Transfer the headings from the original .pdb files to the newly generated *-initlal pdb files, for the simulation box size

Gromacs Preparation:

1. Modify 6-md.mdp to the correct length and saving parameters (and set gen-vel = yes and continuation = no)
2. Run:

```
protein=[name of protein]
length=[length of simulation, or other identifiers]
# site of probe insertion
for site in l105c m109c m145c; do
```

```
# sampling point identifier (from where the initial frame is taken)
for spoint in n.1 n.2 n.3 n.4 n.5 n.6; do
	mkdir $protein-$site-$spoint-$length-setup
	mv $protein-$site-$spoint-$length-initial.pdb $protein-$site-$spoint-$length-setup
	cp 6-md.mdp $protein-$site-$spoint-$length-setup
	cd $protein-$site-$spoint-$length-setup
	gmx pdb2gmx -f $protein-$site-$spoint-$length-initial.pdb -o $protein-$site-$spoint-$length-initial.gro -p $protein-$site.top -water tip3p -ff charmm36MSCN
	gmx editconf -f $protein-$site-$spoint-$length-initial.gro -o test-initial.pdb
	gmx grompp -f 6-md.mdp -c $protein-$site-$spoint-$length-initial.gro -p $protein-$site.top -o $protein-$site-$spoint-$length.tpr -maxwarn 2
	gmx editconf -f $protein-$site-$spoint-$length.tpr -o $protein-$site-$spoint-$length-solefp.gro
	cd ..
	mkdir $protein-$site-$spoint-$length
	mv $protein-$site-$spoint-$length-setup $protein-$site-$spoint-$length
done
done
```

3. Check test-initial.pdb in Pymol, make sure that the structure is correct; compare the box dimensions to the wild type .pdb, make sure that they are identical.

The output directory (containing the run .tpr file) is now ready for MD simulation. The output MD trajectory needs further processing before being used for SolEFP / QMMM calculations. The $protein-$site-$spoint-$length-solefp.gro should be used as the “top” (topology file) in SolEFP simulations.