TXS Simulation Documentation

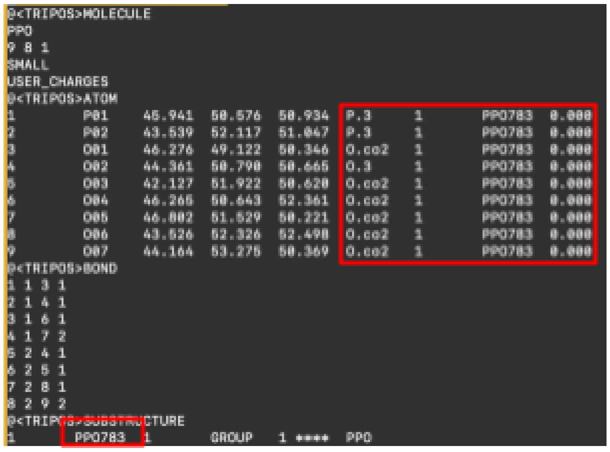
TXS Simulation Workflow

1. Prepare PDBs
    a. Prepare a PDB file for the entire TXS-ligand complex: **TXS-*lignd*/TXS-*lignd*_start.pdb**
        i. Any mutations should be done to the corresponding wild-type TXS-*lignd*_start.pdb file
            1. Mutations can be done in PyMol Wizard → mutagenesis → protein
            2. IMPORTANT: if you are mutating a residue to a histidine or you are mutating a residue next to a histidine (spatially, not only sequentially) you should re-check the histidine protonation states using H++
        ii. Ensure ligand residue names match 3 letter code (i.e., *LIG*)
            1. E.g. in the TXS_txn4_start.pdb we have TXN for all atoms corresponding to taxen-4-yl

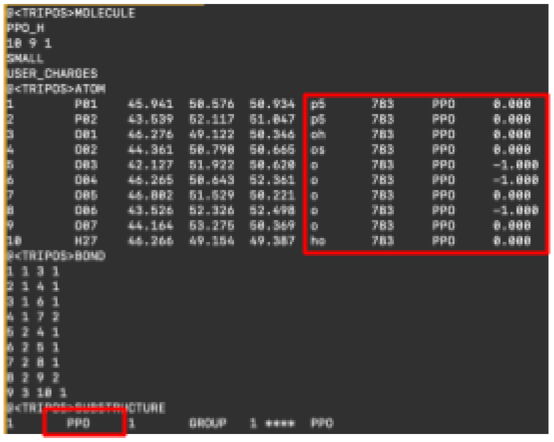
    b. IF RE-DOCKING: Prepare a separate PDB file for any ligand(s) in the complex: **TXS-*lignd*/ligand files/*ligand*.pdb**
        i. Create the ligand file by copying the associated atoms from TXS-*lignd*_start.pdb into a file called *ligand*.pdb
        ii. Ensure ligand residue names match 3 letter code (i.e., *LIG*)
        iii. Ensure HETATM and atom positions match that in corresponding TXS-*lignd*_start.pdb file
        iv. Remove any CONECT records at the end of the PDB file
        v. REMEMBER: All TXS-ligand complexes except GGPP and taxadiene have PPO^{4-} as well as their designated intermediate ligand. TXS-taxadiene complexes also have a PPO^{3-} ligand. Repeat steps 1b and 2 for BOTH

2. IF RE-DOCKING: Prepare ligand mol2 files: **TXS-*lignd*/ligand files/*LIG*.mol2**
    a. Create the mol2 file
        i. Open the ligand(s) in PyMol
        ii. Using the 'Builder' tool add any double bonds and atom charges
        iii. Save the molecule as a mol2 file

    b. Correct the mol2 file
        i. Open TXS-*lignd*/ligand files/*LIG*_key.mol2
        ii. Change the atom specifier, atom charges, residue name, and residue number in *LIG*.mol2 to matchD



```
@<TRIPOS>MOLECULE
PPO
9 8 1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
1        P01    45.941  50.576  50.934  P.3     1       PPO783  0.000
2        P02    43.539  52.117  51.047  P.3     1       PPO783  0.000
3        O01    46.276  49.122  50.346  O.co2   1       PPO783  0.000
4        O02    44.361  50.790  50.665  O.3     1       PPO783  0.000
5        O03    42.127  51.922  50.620  O.co2   1       PPO783  0.000
6        O04    46.265  50.643  52.361  O.co2   1       PPO783  0.000
7        O05    46.802  51.529  50.221  O.co2   1       PPO783  0.000
8        O06    43.526  52.326  52.498  O.co2   1       PPO783  0.000
9        O07    44.164  53.275  50.369  O.co2   1       PPO783  0.000
@<TRIPOS>BOND
1 1 3 1
2 1 4 1
3 1 6 1
4 1 7 2
5 2 4 1
6 2 5 1
7 2 8 1
8 2 9 2
@<TRIPOS>SUBSTRUCTURE
1        PPO783  1        GROUP    1 ****   PPO
```



```
@<TRIPOS>MOLECULE
PPO_H
10 9 1
SMALL
USER_CHARGES
@<TRIPOS>ATOM
1        P01    45.941  50.576  50.934  p5      783     PPO     0.000
2        P02    43.539  52.117  51.047  p5      783     PPO     0.000
3        O01    46.276  49.122  50.346  oh      783     PPO     0.000
4        O02    44.361  50.790  50.665  os      783     PPO     0.000
5        O03    42.127  51.922  50.620  o       783     PPO     -1.000
6        O04    46.265  50.643  52.361  o       783     PPO     -1.000
7        O05    46.802  51.529  50.221  o       783     PPO     0.000
8        O06    43.526  52.326  52.498  o       783     PPO     -1.000
9        O07    44.164  53.275  50.369  o       783     PPO     0.000
10       H27    46.266  49.154  49.387  ho      783     PPO     0.000
@<TRIPOS>BOND
1 1 3 1
2 1 4 1
3 1 6 1
4 1 7 2
5 2 4 1
6 2 5 1
7 2 8 1
8 2 9 2
9 3 10 1
@<TRIPOS>SUBSTRUCTURE
1        PPO     1        GROUP    1 ****   PPO
```

3. Push TXS-ligand complex pdb (and ligand mol2 files if re-docking) to Compute Canada
   a. git commit TXS-*lignd*_start.pdb LIG.mol2 PPO.mol2 -m "*message*"

4. Prepare cell-like environment for simulations
   a. If running locally:
      i. ./prep_envr.sh /Path_to_amber22

   b. If running on Compute Canada
      i. sbatch --account=def-codruta prep_env_remote.sh

   c. Outputs:
      i. TXS-*lignd*.prmtop, TXS-*lignd*.inpcrd
         1. Topology and trajectory files for the complex to be used for relaxation and QM/MM MD simulations
      ii. TXS-*lignd*_solvated.pdb
         1. Unrelaxed protein complex solvated

5. Relax the structure
   a. If relaxing locally (very slow):
      i. ./relax.sh /Path_to_amber22

   b. If relaxing on Compute Canada:
      i. sbatch --account=def-codruta relax_remote.sh

   c. Outputs:

        i. Relaxation/TXS-*lignd*_relaxed.rst7
            1. Restart file for TXS-ligang trajectory
        ii. TXS-*lignd*_relaxed.pdb
            1. PDB file for relaxed TXS complex
        iii. Other files in Relaxation/
            1. Read .out files for steps where relaxation may have had issues

6. Run QM/MM MD simulations
    a. If simulating on Compute Canada:
        i. sbatch --account=def-codruta run_qmmm.sh
    b. Outputs:
        i. QM_MM/TXS-*lignd*_qmmm.mdcrd
            1. Can be used in conjunction with TXS-*lignd*.prmtop to display simulation as a movie in Chimera
        ii. TXS-*lignd*_qmmm.pdb
            1. PDB file for TXS complex following QM/MM MD simulation

7. Perform umbrella sampling
    a. Define reaction coordinates using distances between specific atoms (given in WT table) using Chimera

        i. Example for WT:

Step	**Reaction coordinate**	**Start dist.**	**End dist.**
GGPP ← gerger (*7-gerger-GGPP*)	O01 (PPO) – C02	C2_O1ini =	C2_O1fin =
gerger ← cmbrn15 (*6-cmbrn15-gerger*)	C02 – C01	C2_C1ini =	C2_C1fin =
cmbrn15 ← vert12 (*5-vert12-cmbrn15*)	C11 – C15	C11_C15ini = 1.605	C11_C15fin =4.966
vert12 → vert8 (*1-vert12-vert8*)	C11– H14	C11_H14ini =1.097	C11_H14fin = 4.312
	H14 – C7	C7_H14ini = 2.487	C7_H14fin = 1.097
vert8 → txn4 (*2-vert8-txn4*)	C03 – C08	C3_C8ini =	C3_C8ini =
txn4 → taxa-4-5 (*3-txn4-taxa-4-5*)	C05 – H05		
	H05 – O03 (PPO)		
txn4 → taxa-4-20 (*4-txn4-taxa-4-20*)	C20 – H31		
	H31 – O03 (PPO)		

b.

Package File Structure

- ***TXS/Simulations***
    - ***WT***
        - ***TXS-GGPP***
            - ***Ligand_files***
                - *GGP.mol2*
                    - GGPP mol2 file
                - *GGP_key.mol2*
                    - GGPP mol2 file with correct atom identifiers
                - *GGPP.pdb*
            - ***Prep_environment***
                - *GGPP.in, complex.in*
                    - Leap input scripts for preparing force fields for GGPP and TXS-GGPP complex
            - ***Relaxation***
                - *TXS-GGPP_relaxed.rst7*
                    - Final relaxation restart file (end state of complex)
                - *1min.in – 9md.in*
                    - Input parameter files for 8-step relaxation process
                - *to_pdb.in*
                    - File for converting .prmtop and .inpcrd files into *TXS-GGPP_relaxed.pdb*
            - ***QM_MM***
                - *TXS-GGPP_qmmm.rst7*
                    - Final QM/MM MD restart file (end state of complex)
                - *TXS-GGPP_qmmm.mdcrd*
                    - File to be used with *TXS-GGPP.prmtop* to generate MD movie
                - *qmmm.in*
                    - Input parameter file for QM/MM MD simulation
                - *To_pdb.in*
                    - File for converting .prmtop and .inpcrd files into *TXS-GGPP_qmmm.pdb*
            - *Prep_env.sh*
                - Script for prepping TXS-GGPP locally
            - *Prep_env_remote.sh*
                - Script for prepping TXS-GGPP remotely
            - *Relax.sh*
                - Script for relaxing TXS-GGPP locally

    - *Relax_remote.sh*
      - Script for relaxing TXS-GGPP remotely
    - *Run_qmmm.sh*
      - Script for running QM/MM MD simulation remotely
    - *TXS-GGPP_start.pdb*
      - Initial active TXS docked with GGPP
    - *TXS-GGPP_solvated.pdb*
      - TXS-GGPP with water and NaCl added, no relaxation
    - *TXS-GGPP_relaxed.pdb*
      - TXS-GGPP complex following 8-step relaxation process
    - *TXS-GGPP_qmmm.pdb*
      - TXS-GGPP complex following QM/MM MD simulation
    - *TXS-GGPP.prmtop*
      - Topology file for TXS-GGPP
  - ***TXS-cmbrn15***
    - . . .
  - ***TXS-gerger***
    - . . .
  - ***TXS-vert12***
    - . . .
  - ***TXS-vert8***
    - . . .
  - ***TXS-txn4***
    - . . .
  - ***TXS-taxa-4-5***
    - . . .
  - ***TXS-taxa-4-20***
    - . . .
- ***Umbrella_sampling***
  - ***1-vert12-vert8***
    - ***MD***
      - *md#.in, md#,out*
        - Input and output files for each umbrella run associated with constrained atom distance(s) of #
      - *md#.mdcrd*
        - Trajectory file for run associated with md#.in
    - ***dat***
      - *distances#.dat*
        - Distance file for constraint atom distance(s) of #
    - *umbrella.sh*

        - Script for making sequential md.in and distance.dat files and using them to perform sequential umbrella runs

- ***2-vert8-txn4***
- ***3-txn4-taxa-4-5***
- ***4-txn4-taxa-4-20***
- ***5-vert12-cmbrn15***
- ***6-cmbrn15-gerger***
- ***7-gerger-GGPP***

Starting a GitHub Repo

1. In the **local** master folder you would like to connect to a repo type:

   ```
   git init
   ```

2. Add all files to the repo

   ```
   git add *
   ```

3. Commit files

   ```
   git commit -m “message”
   ```

4. Go to github.com and create new repo
5. Connect to repo

   ```
   git add origin main git@github.com:Username/repo
   ```

6. Generate an SSH key
7. Add SSH key to your github account
8. Connect to compute canada and clone the repository

   ```
   git clone git@github.com:Username/repo.git
   ```

Naming Conventions

	Name ***'ligand'***	**Abbreviation** ***'lignd'***	**3 Letter Code** ***'LIG'***	**Charge**
	Geranylgeranyl diphosphate	GGPP	GGP	–3
A	Geranylgeranyl cation	gerger	GGC	+1
B	Cembren-15-yl cation	cmbrn15	CMB	+1
C	Verticillen-12-yl cation	vert12	VRT	+1
D	Verticillen-8-yl cation	vert8	VRT	+1
E	Taxen-4-yl cation	txn4	TXN	+1
	Taxa-4,5-11,12-diene	taxa-4-5	TAX	0
	Taxa-4,20-11,12-diene	taxa-4-20	TAX	0
	Diphosphate	PPO	PPO	–4
	Diphosphate (H)	PPO_H	PPO	–3