Mucin Runs Data

128.192.9.183/eln/lachele/2019/11/30/mucin-runs-data

November 30, 2019

This post describes the data relevant to the mucin runs with a focus on quantity of data.

Contents [hide]

Raw Data

Location: /home/lachele/RESEARCH/DLIVE_MD_FINAL_STATE

Total storage: 42 GB

Contents:

- The scripts used to:
  - Set up the simulations. These are in a "moved_for_safety" subdirectory (because if they were run a second time they could cause trouble).
  - Perform initial analyses.
- The complete set of force field parameters used in the simulations
- The raw data that came straight out of the simulations. This accounts for most of the storage.
  - 13G d4g – The GalNac
  - 14G d4m – The Mannose
  - 16G protein – The un-glycosylated protein
- Some initial data analysis. This data is inside the d4g and d4m directories and accounts for about 0.3 GB of storage, total.
  - It contains mostly J-coupling and RMSD calculations.
  - I suspect this analysis has been superseded by the treated data (see below).

Note: I have moved the implicit solvent runs to a separate directory (see below).

The Treated Data

This is the main location of all the analysis. I separated it from the original data because of the need to remove corrupted frames from the trajectories.

Location: /home/lachele/RESEARCH/DLIVE

Total Storage: 71 GB

Contents:

- Most of the scripts from the MD_FINAL_STATE directory plus many more analysis programs and scripts.
- Force field parameters used in the simulation (redundant, but not large).
- The scripts and data files associated with the process of removing the corrupted data frames from the trajectories.
- Many analysis files
- Here are pertinent directories & subdirectories and their contents. For large directories, storage amounts are indicated.
    - d4g : the GalNAc runs (36 GB)
        - 1 – 32 : the directories for each of the 32 runs. (~7 GB)
        - ANALYSIS : most of the analysis for the d4g data (28 GB)
        - FIX : Files associated with converting the trajectories to a more convenient format for me so that I could remove the corrupted frames.
        - input : the initial input files for the 32 runs
        - original_pdbs : the original PDB files from which the simulation input files were made
        - VIDEO : files associated with making videos
    - d4m (17 GB)
    - protein (11 GB)

Note: I have moved the implicit solvent runs to a separate directory (see below).

Implicit Solvation Runs

I think these runs will not be part of this paper. Their purpose was to see if the mucin would behave differently if there were no explicit waters available for forming any sort of bridging or networking hydrogen bonds.

Location: /home/lachele/RESEARCH/DLIVE_Implicit_Solvent

Total storage: 81 GB

Contents: Three different sets of implicit solvent runs. I will not take time, at this time, to figure out more about them.

What I remember: The implicit solvation methods were not well enough behaved for me to be comfortable making any conclusions from them. I don't recall the results of any analysis beyond basic validation.