Mucin: 0. Set up run environment

For safety reasons, the analysis was typically performed in a pared-down copy of the directory structure used for data collection. That is, the analysis directory contained only the simulation output files needed for analysis. The attached script refresh_run_data.bash was used to copy the necessary simulation output files into the analysis directory structure.

In these notes, comments begin with '#'.

The main simulations directory structure:

DLIVE
- scripts # custom scripts and executables
  - examples # scratch space for testing and creating sample procedures
  - legacy # old versions
  - moved_for_safety # scripts that, if accidentally run, could cause significant data loss
  - src # source code for compiled programs
    - 2D_JCOUPLING # 2D J-coupling calculations
    - extract_data # some data manipulations used during the analysis phase
    - fix_bad_data # remove trajectory frames corrupted by the hardware failure
    - statistics # perform statistical analyses on data and/or results
- parms # force field parameters used in the simulations
- d4g # N-acetylgalactosylated runs
- d4m # mannosylated runs
- protein # unglycosylated runs

The three run directories, d4g, d4m and protein, contain equivalent directory structures.

For data collection:
- original_pdb # the initial NMR structures
- input # parmtop & inpcrd files and related information
- [numbered] # directories numbered 1 to N containing information relevant to each of the individual runs

The separate directories used for analysis also contained:
- FIX # information about the data curation
- ANALYSIS # data analyses
  - 2D_J-Coupling # files related to calculating the 2D backbone J-couplings
  - ANGLES # angular data extracted from the trajectories
  - DIST # distance data extracted from the trajectories
  - HBOND # hydrogen bonding data extracted from the trajectories
  - JCOUP # files related to calculating the standard J-couplings
  - MANUAL # any analyses handled manually (not written into scripts)
  - PTRAJ_INPUTS # ptraj and cpptraj input files (for angles, etc.)
  - RMSD # RMSD data extracted from the trajectories
  - SUMMARIES # certain data summaries
- VIDEO # any videos generated

Created: 23 Apr 2013 16:02:06 GMT , Updated: 23 Apr 2013 19:01:04 GMT

Created: 23 Apr 2013 15:57:43 GMT , Updated: 23 Apr 2013 15:57:57 GMT

The rest of the directory and file structure was created automatically using the script setup_runs.sh.

The script does the following things. For programming reasons, in the script, some tasks happen concurrently or do not happen in quite the same order as they appear in the list below.

1. Reads in a file containing setup parameters specific to each set of runs. These files are attached below.
2. If any files or directories that it generates are present, it removes them. For this reason, after it was used to generate the structure that was finally used, it was moved to a directory called "moved_for_safety".
3. Makes directories for each available NMR structure. These directories are named by the number assigned to them in the NMR-generated pdb files. For example, for the GalNAc structures, there were 32 directories, named with numbers 1-32.
4. Generates a run control script for each numbered directory. This script will instruct the computer to start each portion of the simulation (minimization, heating, equilibration and production).
5. Generate a top-level script that will start the jobs in each numbered directory.

Created: 19 Apr 2013 18:17:36 GMT , Updated: 23 Apr 2013 15:57:42 GMT

setup_runs.sh

This is the script used to generate the initial run environment.
Created: 19 Apr 2013 18:19:59 GMT

setup1.sh

This is the first setup file for the d4g (GalNAc) runs.
Created: 19 Apr 2013 18:55:20 GMT

setup2.sh

This is the second setup file (for the second set of runs) for the d4g (GalNAc) runs.
Created: 19 Apr 2013 18:55:51 GMT

setup.sh

This is the setup file for the d4m (mannose) runs.
Created: 19 Apr 2013 18:56:31 GMT

setup.sh

This is the setup file for the unglycosylated peptide runs.
Created: 19 Apr 2013 18:56:57 GMT