Chemistry 438/550: Computational Chemistry, Spring 2022

Homework #8, due 22 April

The goal of this homework project is to have you work through a free energy calculations related to pKa's of amino acid side chains in proteins; details can be found at

ambermd.org/tutorials/advanced/tutorial6/index.php

This uses a thermodynamic cycle that is related to the one we discussed in class, but instead of comparing gas to liquid, this compares a model compound in solution to a similar Asp side chain in a protein.

As noted in earlier homework, you should use *tleap* in place of *xleap*. Also, these calcualtions are longer ones, so you need to create a PBS submission script. Copy the one you used for Gaussian jobs. Inside that script, there will be several non-trivial changes: first, in your queue submission script, have this line, which had "-n 1" it the Gaussian version:

```
#SBATCH -n 2      # asks to use two cores
```

Second, you need to change the "module load gaussian" line to load the modules you need for amber/20. Also, you won't need to make the GAUSS_SCRDIR environment variable.

Third, you will change the "srun g16" command with a sander command; take the top line from the *run_model.x* script in the tutorial, and replace "`mpirun -np 2`" with "`srun -mpi=pmi2`", e.g. in the you would have a line like this:

```
srun --mpi=pmi2 sander.MPI -ng 2 -groupfile model_step2.group
```

(You won't need the "$AMBERHOME/exe" part, since *sander.MPI* will be in your PATH. This is also true for *tleap*.)

1. Work through sections 1 and 2.1 of the tutorial. Do just the first λ value, (that is, the first line of the *run_model_x* script), but **set λ to 0.1 instead of 0.0**. You can use whatever plotting program you like (even Excel) to look at the data file you create. Note that the details of the fluctuations will be different than those shown in the examples; this would be true even if you weren't using a different value of λ: the details of an MD simulation run on today's codes and machines will be different from one run 15 years ago, but the averages should be the nearly the same. [Note also that on amarel, the calculation took me about 450 sec for a single λ value, compared to the value of 4400/5 = 880 sec listed in the tutorial for 2006. The real jump in speed for recent MD simluations has come from using GPUs, and shows up more on bigger systems.]

2. Compute $< dV/d\lambda >$ for $\lambda = 0.1$; also compute the standard deviation of this value about its mean. Put these numbers into a text file called lastname.txt.

3. "Hand this in" by doing the following:

   ```
   cp lastname.txt /projects/community/classes/ccb550/homework8
   ```

   (Again this is a pretty trivial thing to "hand in". The point is to make sure you are actually learning how to prepare systems in *tleap*, and to run and analyze thremodynamic integration calculations.)

4. If you wish, you can go on the section 3 of the tutorial, and see how to set up the same calculations in the protein environment. But this is optional, and is not part of the homework assignment.