BMB 173 – Winter 2017

Homework Set 9.2 (80 points) – Assigned 3/9/17. Due 3/14/17

TA: Emily Wyatt, ewyatt@caltech.edu

Office Hours:

Friday 3/10/17 from 2-3pm in Spalding B123 or by appointment
Monday 3/13/17 from 3-4pm in Spalding B123 or by appointment

1. **Course-grained Molecular Dynamics** (30 points)
   a. (5 points) In your current research, what might be a potential use of MD simulation?

      Many possible answers.

   b. (10 points) Estimate the number of atoms in your system.

      Many possible answers.

   c. (5 points) Is this system size tractable, or would you need to perform course-grained molecular dynamics? What length of time is needed to answer your question?

      Tractable if system is no larger than 10^9 atoms (for nanoseconds).

   d. (10 points) Describe one potential strategy for course graining your system if you wanted to expand your system size or simulation time. For example, how might you cluster your atoms?

      Many possible answers – e.g. clusters of ~10 atoms are substituted by a single CG bead, several water molecules become one CG bead, etc.

2. **Structural Genomics** (20 points)
   a. (10 points) One challenge in structural genomics is developing technology to automate the process. Thinking about the structural genomics pipeline, what steps do you think were the most difficult to automate? Why?

      Many possible answers.

      Example: One step that must have been very difficult to automate was initially identifying protein crystals in growth trays. More specifically, it is very difficult to identify true protein crystals, distinct from precipitated protein or salt crystals that will also form for various growth conditions. Another step in the structural genomic pipeline that must have been extremely difficult to automate is the indexing of the diffraction spots in a given collected pattern. From what was described in class, for a while there was still some level of expert touch required to correctly index the data. Finally, I imagine there was significant computational challenge in determining all of the phases for the diffraction data (Harker construction, etc.) to be able to calculate the density function from the structure factor data.

b. (10 points) If you were to write a grant, how would you pick the targets for a structural genomics project?

Many possible answers.

Example: If I was to write a grant, I would pick the targets for a structural genomics project based on a number of criteria. First, it's probably important to try to pick targets that have new folds, such that the structural information can be representative of a large structural family for which structural information is currently absent. This would maximize the impact for a given investment in the characterization. Similarly, I'd probably select targets from different protein families, again to better fill the "fold space". Another consideration would be the protein stability. It might be advantageous, for example, to choose a homolog of a desired target that is more stable to various stresses or one that is smaller in size to increase the likelihood of being able to solve the structure effectively. A final consideration might be whether other groups are already targeting this protein or a similar one, and to assess the status of their structural genomics project and whether it is worth the investment.

3. **The Future of Structural Biology** (30 points)

a. (15 points) Please read the following review on X-ray Free-Electron Lasers (XFELs):
http://www.ncbi.nlm.nih.gov/pubmed/24914150

i. (5 points) Briefly explain how XFELs work.

XFELs produce extremely short (tens of femtosecond in duration), but intense X-ray pulses (1012 X-ray photons/pulse that can be focused down to a sub μm focal spot). This radiation will cause the molecule to explode (highly destructive), but you can collect diffraction patterns (from single molecules or microcrystals only a few unit cells across) before this destruction occurs to solve the crystal structure. The review discusses several membrane proteins for which feasibility of using XFELs was demonstrated, despite the challenges of merging diffraction data from many randomly oriented microcrystals of varying shape and size.

ii. (10 points) Describe two advantages to XFELs in comparison to traditional x-ray crystallography.

One advantage of XFELs in comparison to traditional X-ray crystallography is that structural information can be extracted from thousands of microcrystals that are too small to be used by traditional X-ray crystallography. In our X-ray crystallography lectures, one key point was the difficulty in growing large, ordered crystals. Therefore, we could bypass this difficulty by using XFELs to collect diffraction data from microcrystals. A second advantage of XFELs in comparison to traditional X-ray crystallography is the ability to perform XFEL-based time resolved structural studies. More specifically, because the pulses are so short, you can get very high temporal resolution, opening up the possibility

for ultrafast time-resolved diffraction studies of biomolecules, a capability not accessible using traditional X-ray crystallography.

b. (15 points) Of the ~125,000 structures in the PDB, 90% have been solved by traditional x-ray crystallography, another 9% by NMR, and the last 1% have been determined using other strategies.

Explain the relative advantages and disadvantages of three up and coming techniques to solve structures of proteins in the future – cryo-EM, XFELs, and MicroED.

Many possible answers.