Final Exam Practice Solutions

CS/BIOE/CME/BIOPHYS/BIOMEDIN 279

Fall 2017

Instructions

These are practice questions in the style of questions you might expect on the exam. Each question should be answerable in a few sentences (that is, youre not required to provide a great deal of detail). This practice exam is not representative of the length of the final exam.

Question 1: Compare and contrast the energy functions used for molecular dynamics simulations and those used for ab initio protein structure prediction.

> **Solution:** *Molecular dynamics simulations use a molecular mechanics force field. This is an approximate potential energy function — that is, it specifies the energy of each precise arrangement of atoms, and thus allows one to compute the force on each atom. The terms in most molecular mechanics force fields are physics-based. Ab intio protein structure prediction generally uses an (approximate) free energy function, which specifies the energy associated with a set of atomic arrangements, and thus allows one to pose proteins structure prediction as a minimization problem. It is usually knowledge-based, at least in part.*

Question 2: Describe, at a high level, how Rosetta predicts protein structure (in particular, for ab initio modeling).

> **Solution:** *It attempts to minimize a free energy function that is largely knowledge-based (i.e., determined in large part based on statistics of the PDB). To help address the computational complexity of this minimization process, it uses a fragment-based Monte Carlo search.*

Question 3: How would you go about estimating how long it would take to run an MD simulation? What information would you need to consider?

> **Solution:** *You would need to consider:*
>
> - *Number of time steps (which depends on total time to be simulated)*
> - *Total number of atoms in the system being simulation.*
> - *The average number of non-bonded interactions to be computed for each atom at each time step. (Non-bonded interactions dominate the overall computation.). This will depend on the algorithm thats being used to compute the non-bonded interactions.*
> - *Time to compute each non-bonded interaction.*

Roughly speaking, youd multiply the four quantities above.

Question 4: We would like to estimate how tightly a particular drug candidate binds to a particular target protein.

(a) Provide a quantitative definition of binding affinity (i.e., binding strength). That is, what does it mean for one ligand to bind more tightly than another?

> **Solution:** *Ligand A binds more tightly than ligand B another (that is, binds with higher affinity) ifwhen either is present at some reference concentrationligand A is bound a larger fraction of the time. Affinity can be defined either as a difference in free energy between the bound and unbound state, or as the concentration of free ligand such that half the target molecules have a ligand bound and the other half do not.*

(b) Suppose we have a single molecular dynamics simulation in which the drug candidate binds to the target and stays bound for the remainder of the simulation. Can we accurately estimate the binding affinity from that simulation? Why or why not?

> **Solution:** *No. You would need to see the ligand binding and unbinding many times in order to estimate accurately the fraction of time it will remain bound.*

Question 5: Describe one common approximation made by ligand docking methods, and explain why it helps simplify the problem to be solved.

> **Solution:** *A common approximation is that the protein is rigid. This simplifies the problem to be solved because one doesnt need to consider all possible arrangements of the protein atoms (and weight the binding energy/score across them).*

Question 6: X-ray crystallography and single-particle electron microscopy are both techniques for determining the structure of a molecule or molecular complex. Why is single-particle electron microscopy typically used for larger molecules or complexes than x-ray crystallography?

> **Solution:** *Two reasons: (1) its generally harder to form crystals of large molecules/complexes than of small ones, (2) in single-particle electron microscopy, its usually harder to solve the computational reconstruction problem for smaller molecules.*

Question 7: Discuss the trade offs of a stochastic particle-based reaction-diffusion simulation versus a continuum approach (in which concentrations of each type of molecule are represented in each voxel).

> **Solution:** *The continuum approach is often much faster, particularly for large numbers of molecules, but its less accurate, particularly for small numbers of molecules.*

Question 8: There is a very efficient algorithm for computing the Fourier Transform known as the Fast Fourier Transform (FFT). Describe how this algorithm is useful for one of the methods covered in this course.

> **Solution:**

- *The FFT allows one to compute convolutions more efficiently. This is useful, for example, when low-pass filtering an image to denoise it.*
- *In x-ray crystallography, the FFT allows one to calculate efficiently the diffraction pattern that would be associated with a hyphothesized electron density.*