Übung: revision – lattices and coarse grain proteins

SS 2009 Übung zu Struktur und Simulation für 14 July 2009

1. Overview

This is a revision Übung which requires no computer. The questions are typical of what will be in the final exam.

2. Questions

1. I am simulating with a protein in a box of water with periodic boundary conditions using conventional molecular dynamics and with a cutoff for interactions of 15 Å. In "big O" notation, what is the running time. Does the worst case running time change if I do not use cutoffs ? Does the running time change in practice ?
2. Exam examples only – no need for discussion
    - Write down the energy due to electrostatics – use any reasonable choice of nomenclature.
    - Write down the energy due to electrostatics if we use a distance dependent dielectric constant.
    - Write down the force acting on a charge due electrostatics if we use a distance dependent dielectric constant. Write it simply as a function of the distance between two particles (no need for vectors). Use nomenclature consistent with the first two parts of this question.
3. Describe a water model based on solvent accessible surface area. Explain in terms of simple equations why this will lead to a force acting on particles such as oxygen or nitrogen atoms. Describe an effect of water which will not be represented by this model (there are several sensible answers).

4. I am simulating a protein in a box with an explicit SPC water model (3 point charges, 3 mass sites and 1 Lennard-Jones site). An argon atom approaches the water. Later, a charged sodium ion approaches the water. From the points of view of the argon atom and sodium ion, is the water spherical (like a football / billiards ball) or some different shape ?

5. Exam example question with many possible answers:
   Describe a low-resolution protein model which can

   - keep adjacent (successive) amino acids at the correct distance
   - form regular α-helices and β-sheets
   - has some mechanism to maintain roughly the correct density within the middle of the protein

   Explain what you choose as interaction sites. Explain which features of your model are necessary for each feature.

6. I have built a score function for proteins based on the frequency with which certain pairs of amino acids are seen at certain distances. It allows me to give a score to any conformation of a protein by looking up scores in tables.

   - I am not able to use it for molecular dynamics simulations. Why ?
   - I try to make a continuous form of the functions by fitting the data to a polynomial function like $score(r_{ij}) = k_1 + k_2 x + k_3 x^2 + k_4 x^3 + \ldots$ Why would a sane person do this ? Where may it break down ?

7. I have built a Boltzmann / knowledge-based score function for proteins using the methodology based on potentials of mean force. It is based on C^{α}-C^{α} distances. I do not distinguish between amino acids which are separated by one residue (i,i+2) and those separated by many residues. Why will this be a very bad approximation ?

8. I am working with a lattice model for a protein. The model is simple so I can computationally visit all conformations. Describe how I would use the Boltzmann relation to work out the absolute probability for a certain configuration of points.

9. I would like to investigate protein conformational switches using a lattice model. These are proteins which can adopt more than one conformation. The hamburger Abendblatt has claimed that the protein switches are more likely to be found in proteins with less hydrophobic residues.

Describe a set of steps to see if this is plausible. Describe calculations for a protein of length 18 in the HP model.

10. A popular view of protein folding is that as a protein folds, its potential energy decreases, but the entropy also decreases. How would you see this be reflected in a simple lattice based model and how could you check if it is true ?

11. I look at sequence alignments and measure the variability of residues at different sites in order to see how "important" each residue is. Why can I not compare mutation rates at different sites ? Could I correct for this (in my simple model) ? Where would this property show up in Prof Kurtz's or Frau Willhoeft's lectures ?