1. (10 pts) Name three types of terms in typical potential energy functions for proteins.

Possible answers: bond length, bond angle, dihedral angle, Lennard-Jones (non-bonded), improper angles, electrostatic, Urey-Bradley

2. (10 pts) What is the densest packing fraction for disordered collections of monodisperse hard spheres? What is the packing fraction of all-atom hard-sphere (i.e. explicit hydrogen) representations of residues in protein cores?

Monodisperse spheres: $ϕ=0.64.$

Residues with explicit hydrogens in protein cores: $ϕ=0.56$

3. (10 pts) Plot the purely repulsive linear spring potential (V/ε versus rij/σij) given by

for rij ≥ σij and V(rij) = 0 for rij < σij, where rij is the center-to-center separation between atoms, σij = (σi+σj)/2 is the average diameter of spherical atoms i and j, and the strength of the potential ε is a constant.

Answer: see plot below



4. (10 pts) Calculate the x-, y-, and z-components of the force on atom i from neighboring atom j, $\vec{F}\_{ij}=-\frac{dV}{dr\_{ij}}\hat{r}\_{ij}$ from the purely repulsive linear spring potential in question #3.

Answer: $\frac{∂V}{∂r\_{ij}}=\frac{ϵ}{σ\_{ij}}\left(1-\frac{r\_{ij}}{σ\_{ij}}\right)$, so  $\vec{F}\_{ij}=-\frac{ϵ}{σ\_{ij}}\left(1-\frac{r\_{ij}}{σ\_{ij}}\right) \hat{r\_{ij}}$ where $\hat{r\_{ij}}$ is a unit vector pointing from atom i to atom j

5. (10 pts) What is the difference between a deterministic and stochastic model? Which term applies to Ordinary Differential Equation (ODE) models?

No randomness is involved in a deterministic model. Given model structure, parameter values, and initial conditions, there is no variation in output. In stochastic models, the next state of is not fully determined by the previous state – probability is involved.

6. (10 pts) What does it mean for a model parameter to be identifiable?

Identifiable parameters are those which effect the value of the data and can be estimated with some degree of certainty.

7. (10 pts) Below, a Ramachandran plot based on the original theoretical calculations is shown. Label the regions of the plot that correspond to α-helix and β-sheet backbone conformations and describe roughly what is the difference between the yellow and red regions of the plot.



Answer: Top left/bottom left region are beta sheets, region at left at phi = -60 is a right-handed alpha helix, small region at psi = 60 and phi = 60 is left-handed alpha helix.

Red regions are the “normally allowed” region, where standard (i.e. strict) contact distances between hard sphere atoms are assumed, and the yellow regions are the “outer” regions, where shorter (i.e. lenient) contact distances between hard sphere atoms are assume.

8. (10 pts) Name two common features of intrinsically disordered proteins.

Answer: No well-defined tertiary structure/dynamic conformational changes, can liquid-liquid phase separate/aggregate, tend to be mostly composed of charged amino acids

9. (10 pts) In the context of mathematical modeling, what is an F test used for? What does it mean when F >> 1?

The F test can be used to compare two model fits. It is often applied to determine whether a parameter (or variable) should be added or deleted from a model. One compares the more complex model (with the parameter) to a simpler model (without the parameter). F >> 1 indicates that the more complex model is better (i.e., the inclusion of the parameter leads to a significantly better fit). The F test can be used as part of a forward selection or backward elimination strategy in model building.

10. (10 pts) The SIR model can be used to predict the course of an epidemic in a large population. In the model equations given below, describe the meaning of each state variable (S, I and R) and each parameter ($β$ and $μ$):



S is the population of susceptible individuals

I is the population of infectious individuals

R is the population of individuals who were infected, but have now recovered (also referred to as removed)

$β$ is the infection rate

$μ$ is the recovery rate

11. (10 pts) How does the radius of gyration Rg ~ Nν grow the number of residues N for a protein that behaves as a a) a fully extended chain, b) random walk, and c) self-avoiding random walk?

Answer: a) $ν=1$, b) $ν=1/2$, c) $=3/4$