Due Date: Sunday, April 21, 2019 (23:59)

CENG 465
Introduction to Bioinformatics
Spring 2018-2019

Assignment #3

Written Assignment on Protein Structures

Problem 1 (30 Points):

(a) Find a protein structure in the Protein Data Bank (https://www.rcsb.org/) which contains 4 alpha helices, 8 beta sheets and is composed of 4 chains. Write down the PDB ID of the protein structure you have found.
(b) Which SCOP category/categories (class, fold, superfamily, and family) does this protein (or its subsequent chains/domains) belong to?
(c) How many polar amino acids does each chain contain?

Problem 2 (30 Points):

How many different (self-avoiding) folds a 5 amino acid length polypeptide sequence can exhibit on a 90 degree 2D lattice in *ab initio* modeling? In a self-avoiding fold, each lattice point can be occupied by at most one amino acid. On a 90 degree lattice, only 90 degree moves are allowed.

Problem 3 (40 Points):

Consider two polypeptide structures each of which contains 6 amino acids. Find an upper bound on the minimum RMSD between these two protein structures. In other words, find two (very different) protein structures of length 6 each, so that the minimum RMSD between them is as large as possible when they are superimposed. Assume that each amino acid is represented as a single point in 3D space, i.e., the position of the alpha-carbon, and the distance between two consecutive amino acids is fixed at 4 angstroms. In other words, each protein structure is modeled a sequence of 6 three-dimensional points, in which the distance between each consecutive pair of points is 4 angstroms. Do not consider any biological or chemical constraints. Solve this problem as a geometric problem using only the geometric constraints stated above.

Submission

Submit your answers to the problems above as a single PDF document via ODTU-Class before the deadline (handwritten and scanned PDF documents are accepted as well). Late submission is -20 points per day.