1. **(20 pts) Acid-base, reactivity, and artistic properties of lysine.**

   One of the active site lysines (324) in the enzyme fumarase has a $pK_a$ of 7.1, surprisingly low.

   (a; 3 pts) Complete the positively charged lysine side chain on the sketch below and write down the chemical equilibrium corresponding to its deprotonation.

   ![Lysine Side Chain Sketch](image)

   (b; 3 pts) It turns out that the active site has a second lysine next to Lys324. The second one has a normal $pK_a$. How does this help explain the unusual $pK_a$ of Lys324?
(c; 5 pts) Use the Henderson-Hasselbach equation to calculate the ratio of deprotonated to protonated lysine at both pH 5.9 and at pH 8.1 for Lys324, with its $pK_a$ of 7.1.

(d; 5 pts) Because Lys324 has such a low $pK_a$, it is a much better nucleophile than free Lys. Draw the product of acetylating Lys324. Why is acetyl-lysine important in gene regulation?

(e; 4 pts) Write down the name of a molecular visualization program and describe how to select and emphasize one residue using the program.
2. **(20 pts) Rite of Passage:**

Draw the structure of the peptide Ile-P-Tyr-Met-Asp, including the correct stereochemistry at Cα’s and all ionizable groups in their correct protonation states at pH 7. P-tyr = phosphotyrosine, which has pKₐ’s of ~2 and ~5.8, so its charge at pH 7 is –2. If you need more space, the sequence is also on the last page.

3. **(25 pts) Bioinformatics:**

(a; 12 pts) Briefly describe a four steps in a typical bioinformatics/biochemistry “workflow” that a bench biochemist might perform in learning what she can about the likely structure and function of a protein sequence that she has just connected to a function of interest. Don’t forget the last and most important step!
(b; 13 pts) Contemplation of the BLOSUM matrix can provide much insight into protein and amino acid properties. Referring to the pictorial version of the top half of the substitution matrix, answer the questions below, whose numbers correspond to the indicated areas on the matrix:

(1) Why are the scores for identical W, C, P, and H residues higher than the scores for other residues? (One answer for all three).

(2) Why are I and V more similar to each other than either is to L, given that L and I are isomers of each other? In general, why do the bulky hydrophobic residues appear to substitute for each other quite readily?
(3) Replacing E with K has a positive similarity score. Why is this initially surprising? Considering where the residues are likely to be located in the protein structure, explain why they do in fact often substitute for each other.

(4) What are the special features of C, G, and P that make each of them different from all other amino acids (one feature each).

4. (35 pts) Secondary, tertiary, and miscellaneous structure:
   (a; 8 pts) In general, how do chaperones use the free energy available from ATP hydrolysis to improve the success rate of protein folding and avoid aggregation?
(b; 9 pts) The two sequences below are known to be amphipathic. Which one is part of a beta sheet and which one is an alpha helix? Briefly explain your reasoning? For your convenience seven-pointed stars are sketched below in case you need them.

Sequence A:  L E D K V E E L S S K N Y H L E N E V A R L

1 2 3 4 5 6 7 8 9 10 12 14 16 18 20 22

Sequence B:  E I K N G I D L T L K A H M T F K I S F K W

(c; 4 pts) Under what conditions is an exothermic ordering reaction thermodynamically favorable? Give an example of an exothermic ordering reaction.
(d; 9 pts) Explain why molecule (i) below forms a micelle, molecule (ii) forms a bilayer, and molecule (iii) forms a globule. Briefly describe two biological functions for membranes.
(e; 5 pts) How does the steric zipper model provide a quite general and yet individualized failure mode for proteins?