

**Problem 1.** (12 points total, 4 points each)

The molecular weight of an unspecified protein, at physiological conditions, is 70,000 Dalton, as determined by sedimentation equilibrium measurements and by gel filtration chromatography. The SDS-polyacrylamide gel electrophoresis (SDS PAGE) of the protein yields a single band corresponding to molecular weight of 70,000 Dalton. However, in the presence of the reducing agent,  $\beta$ -mercaptoethanol, the SDS PAGE shows two bands, corresponding to molecular weights of 30,000 and 20,000 Dalton.

(a) From these data, describe the native protein in terms of the number of subunits present, their molecular weight, stoichiometry of subunits, and the kinds of bonding (covalent, noncovalent) existing between the subunits.

The protein consists of 3 subunits with molecular weights: 30,000 Da (one subunit) and 20,000 Da (2 subunits), connected with each other by disulfide bonds.

(b) You treat your protein with protease (e.g. trypsin), run the product of the proteolytic cleavage on the SDS PAGE, and discover that the SDS gel in the absence of  $\beta$ -mercaptoethanol still shows one band, but shows 4 bands in the presence of  $\beta$ -mercaptoethanol. How would you explain this result?

Trypsin cleavage sites (Lys, Arg) are located in the polypeptide sequences of the subunits of this protein between the sites involved in disulfide bonding, so that no peptides are released after the cleavage: the cleaved parts of the sequence are still held together by disulfide bonds.

(c) How would your answers to problem (a) change if the molecular weight of the protein was 100,000 Dalton?

$$100,000 = 2 \times 30,000 + 2 \times 20,000$$

This protein consists of 4 subunits, 2 with molecular weight of 30,000 Da and 2 with 20,000 Da.

**Problem 2.** (total 20 points, 4 points each)

You have to determine the amino acid sequence of a peptide. You perform the following steps using enzyme cleavage of your peptide (see table below):

*Step 1.* Treatment with trypsin yields three fragments with the following sequences (in the order of their length): **WGA, AGTK, YLDR**

*Step 2.* Treatment with chymotrypsin gave the following three peptide fragments: **GA, LDRW, AGTKY.**

(a) What is the sequence of your peptide?

**A G T K Y L D R W G A**

(b) Explain why step 1 alone was not sufficient to unambiguously determine the sequence of your peptide.

From the results of step 1 it is clear that **WGA** is the C-terminal part of the sequence: no R or K on the C-terminus of this tripeptide, so it cannot be a product of trypsin cleavage. However, these data alone do not indicate the correct order of the remaining two tetrapeptides, i.e. we are unable to distinguish between **AGTKYLDR** and **YLDRAGTK**.

Assume you cannot perform step 2 as described above because your lab ran out of chymotrypsin supply. You are desperate to have the answer and you are looking for an alternative method to use in step 2. Consider the following possibilities:

(c) Can you use any of the other three cleavage agents listed in the table in order to unambiguously determine the sequence? Explain your answer.

Yes. You can use pepsin: it will cleave exactly at the same sites as trypsin, but this will be sufficient to solve the puzzle: because **YLDR** has a **Y** at the N-terminus and **AGTK** does not have an aromatic residue at the N-terminus, it is clear that **AGTK** precedes **YLDR**.

Alternatively, you can use Asp-N protease: you will get **AGTKYL** and **DRWGA**, and the order of these peptide fragments will become obvious.

(d) Would peptide hydrolysis with 6M HCl help you? Explain your answer.

No. Hydrolysis will give only the amino acid composition, not the sequence.

(e) Would Saenger's reagent (1-fluoro-2,4-dinitrobenzene) or dansyl chloride help you? Explain your answer.

Yes, because treatment with these compounds will identify Ala as the N-terminal residue.

Treatment	Cleavage points	Cleavage location at
Trypsin	Lys, Arg	carboxyl side of these residues
Chymotrypsin	Phe, Tyr, Trp	carboxyl side of these residues
Asp-N protease	Asp, Glu	amino side of these residues
Pepsin	Phe, Tyr, Trp	amino side of these residues
Cyanogen bromide	Met	carboxyl side of this residue

**Problem 3.** (20 points total, 4 points each)

Suppose you have a mixture of five proteins listed in the table below.

#	Protein	pI	Mol.Weight, kDa
A	Ovalbumin	4.6	45
B	Myoglobin	7	16.7
C	Serum albumin	4.9	68.5
D	Ubiquitin	6.4	8.5
E	Cytochrome c	10.6	13

(a) Indicate the order in which these proteins will elute from a gel-filtration column (starting with the one that elutes first). You can use letters A-E (see table) for simplicity.

**C, A, B, E, D**

(b) You load this mixture on a cation exchange column (i.e. column that bears negatively charged groups). The buffer you use for this column is acetate buffer, pH 4.76.

-- List proteins that will appear in the flow-through (i.e. will not bind to the column).

**A, C** (In the latter case, the difference between pH 4.76 and pI 4.9 is too small, so serum albumin will also be in the flow-through).

-- In order to elute those proteins that are immobilized on the column, you then apply a linear salt gradient, with NaCl concentration gradually increasing from 0 to 1 M. Indicate the order in which the proteins bound to the column will elute as the salt concentration increases.

**D, B, E**

(c) You repeat the same procedure as in (b) but now you use anion exchange column (that bears positively charged groups), and the buffer is TRIS (pH 8.0). Answer the same questions as in (b).

Flow through: **E**

Bound proteins elute as: **B, D, C, A**

(d) You load this protein mixture and run it on the SDS PAGE. Predict the order they will migrate on the SDS gel, starting with the fastest.

**D, E, B, A, C**

(e) You are interested in further studies of cytochrome c. Based on the results of your experiments above, devise a reliable procedure for its purification from this mixture of proteins.

Most straightforward: use anion exchange column (see problem (c)): cytochrome c will be in the flow-through

Other feasible procedures, e.g., a combination of size-exclusion chromatography and ion exchange were also given the credits they deserved.

**Problem 4.** (12 total, 1 point each)

Indicate which of the following statements are true and which are false.

**Note: To discourage random guessing in this problem, each wrong answer will give you -0.5 point.**

True	False	
	X	$\alpha$ -helix is held by hydrogen bonds between amino acid side chains in the positions $i$ and $i+4$
X		Ramachandran plot is a graphical representation of the sterically allowed conformations of peptide planes
X		The extended, $\beta$ -conformation, is characterized by a zigzag backbone geometry, where the backbone NH bonds belonging to residues in the positions $i$ and $i+2$ are located on the same side of the $\beta$ -strand and are almost parallel to each other
X		Hydrophobic effect is the main driving force in protein folding
X		The backbone NH bonds of the neighboring residues in the $\beta$ -strand are located on the opposite sides of the strand and are almost orthogonal (perpendicular) to the direction of the strand.
	X	The <i>cis</i> isomer is the highly favorable isomer for most peptide bonds except those preceding proline.
	X	The geometry of the $\beta$ -sheet formation requires that the participating residues be nearby on the polypeptide chain
X		The term secondary structure refers to local conformation of the protein backbone, disregarding the conformation of side chains
X		The dihedral angle $\phi$ characterizes the orientation of the peptide plane at the amino end of the residue
	X	Protein tertiary structure is the three-dimensional arrangement of its 2 <sup>o</sup> structural elements, $\alpha$ -helices and $\beta$ -sheets, held together by hydrogen bonds between them.
X		Nuclear magnetic resonance (NMR) provides structural information in terms of interatomic distances and bond orientations, but not the actual atom coordinates.
X		Quarternary structure refers to the three-dimensional arrangement of polypeptide subunits in a protein consisting of two or more polypeptide chains.

**Problem 5.** (24 Points total, 4 points each).

A specific protein has an 18-residue long  $\alpha$ -helix with the following sequence:

**PENQWKQELDTRYRNALQ**

(a) How many full turns are in this  $\alpha$ -helix?

It's 3.6 residues per turn, so 18 residues make  $18/3.6 = 5$  turns

(b) What is the length of the helix (in Angströms) in the direction of the helix axis?

One turn is 5.4A in length, therefore 5 turns make  $5 \times 5.4 \text{ A} = 27 \text{ A}$

(c) How many hydrogen bonds between the backbone atoms are in this helix? Explain your reasoning.

Consider H-bonds between backbone atoms. Each residue is capable of making two H-bonds with other residues, via amide N (donor) and carbonyl O (acceptor). For 18 residues this will make up to  $18 \times 2 = 36$  backbone atoms capable of H-bonding. Not all these atoms are actually H-bonded. The 4 N-terminal and 4 C-terminal residues in the helix are missing a H-bonding partner from one side; this reduces the number of H-bonded atoms to  $36 - 8 = 28$ . Since a H-bond involves two atoms, the number of hydrogen bonds is  $28 / 2 = 14$ .

(d) In the peptide sequence below, circle all residues that are most likely to be packed against the rest of the protein and, therefore, will be involved in the formation of the hydrophobic core of the protein in aqueous environment.

PENQWKQELDTRYRNALQ

(e) In the sequence below, circle those residues that would favor contacts with water.

PENQWKQELDTRYRNALQ

(f) In addition to hydrogen bonding, what other interactions could contribute to the stabilization of this helix? Specify the interactions and use lines to connect residues involved in these interactions.

PENQWKQELDTRYRNALQ

Ion pairs (salt bridges) between opposite-charged side chains: these residues are 4 residues apart in the sequence and therefore are close to each other in space in the  $\alpha$ -helix.