

Problem 1. (total 30 points)

You have to determine the amino acid sequence of a peptide. You performed the following steps using enzyme cleavage of your peptide (see table on the front page) combined with the subsequent sequencing:

Step 1. Treatment with trypsin yielded three fragments with the following sequences:

YLDR, GSAK, WGSM

Step 2. Treatment with pepsin gave the same three peptide fragments:

YLDR, GSAK, WGSM.

(a) (10 pts) From these data determine the correct sequence of your peptide.

GSAK YLDR WGSM

(b) (5 pts) Explain why neither of these steps alone is sufficient to unambiguously determine the sequence of your peptide.

Since each of these steps resulted in three peptide fragments, there are two possible solutions for each step:

Step 1: **GSAK YLDR WGSM** or **YLDR GSAK WGSM**

Step 2: **GSAK YLDR WGSM** or **GSAK WGSM YLDR**

Neither of the two steps *alone* can discriminate between these solutions

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

(c) (5 pts) Can you use any of the other three cleavage agents listed in the table in order to unambiguously determine the sequence? Explain your answer.

-- Cleavage with **chymotrypsin** will give: **GSAKY, LDRW**, and **GSM** – combined with the results of step 1, this will unambiguously determine the right sequence;

-- Cleavage with Asp-N-protease will result in two fragments: **GSAKYL** and **DRWGSM**, which will unambiguously determine the right sequence.

-- Cyanogen bromide will be useless in this situation, because there is no Met in the sequence.

(d) (5 pts) Can you use peptide hydrolysis with 6M HCl to solve the problem? Explain your answer.

No, this will give us amino acid *composition*, not the sequence.

(e) (5 pts) Can you use Saenger's reagent (1-fluoro-2,4-dinitrobenzene) or dansyl chloride to solve the problem? Explain your answer.

Yes, this will identify the N-terminal residue as Gly, which will unambiguously discriminate between the two possible peptides after Step 1.

Problem 2. (20 total)

Answer the following questions about protein structure (use back of the page if you need more space):

(1) (5 pts) What does the Ramachandran plot represent? What are the parameters shown on the Ramachandran plot and what do they characterize?

Ramachandran plot represents sterically allowed conformations of the backbone. Shown on the plot are dihedral angles Φ and Ψ . Φ characterizes the conformation of the peptide plane preceding the C_α atom (at the amino-end of C_α atom) – this angle is associated with the plane rotation about the N-- C_α bond. Ψ characterizes the conformation of the peptide plane at the carboxyl end of the C_α atom – this angle is associated with the peptide plane rotation about the C_α --C' bond.

(2) (5 pts) What groups are connected by the hydrogen bonds in the α -helix and in the β -sheet? Compare the characteristic hydrogen bonding patterns (i.e. what residues are connected) and the orientation of the hydrogen bonds in these two elements of secondary structure.

Backbone NH and CO groups.

In the α -helix the H-bonds are parallel to the helix axis, and connect carbonyl oxygen of residue i with amide N of residue $i+4$. These bonds are within the same fragment of the polypeptide chain.

In the β -sheet, the H-bonds are oriented perpendicular to the direction of the β -strand. These bonds are between carbonyl oxygen of one strand and amide nitrogen of an adjacent (in 3D space) strand. These bonds connect different pieces of a polypeptide chain that could be far apart in the sequence.

(3) (10 pts) Explain the meaning of terms “primary”, “secondary”, “supersecondary”, “tertiary”, and “quarternary” structure.

Primary structure refers to the amino acid sequence of a protein (including the covalent bonds that link amino acids).

Secondary structure refers to stable arrangements of amino acid residues giving rise to recurring structural patterns.

Supersecondary structure (or motif) refers to particularly stable arrangements of several elements of secondary structure and the connections between them.

Tertiary structure includes all aspects of the three-dimensional structure of a protein

Quarternary refers to the three-dimensional arrangement of two or more polypeptide subunits.

Problem 3. (30 Points total, 6 points each).

An unspecified protein contains a 15-residue long α -helix with the following sequence:
WEANIKQLSTEYKQ

(a) How many full turns are in this α -helix?

One turn contains 3.6 residues; therefore 15 residues form 4 full turns ($15/3.6 = 4.2$).

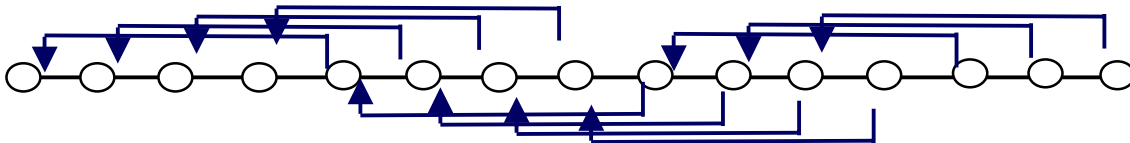
(b) What is the length of the helix (in Angströms) in the direction of the helix axis?

The length of one turn is 5.4 Å, so a 15-residue α -helix will be $15/3.6 \times 5.4 \text{ Å} = 22.5 \text{ Å}$ long

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

You can figure this out using the following simple considerations. 1 residue can be involved in max 2 H-bonds, therefore 15 residues can make up to $2 \times 15 = 30$ H-bonds. In the α -helix, 4 residues at the N-terminus and 4 at the C-terminus make only 1 bond pre residue. This makes the total number of H-bonds $30 - 2 \times 4 = 22$. When calculating this number, each H-bond was counted twice: one time for the donor residue and one time for the acceptor. The real number of H-bonds is then $22/2 = 11$.

You can also draw the H-bond connections (between res. i and $i+4$) and count them. An example is shown here, where the acceptors (CO) are indicated by an arrow.



(d) In the protein, this helix is packed against a β -sheet, thus forming the hydrophobic core. Predict what side of the helix is facing the β -sheet: circle those residues involved in the hydrophobic contacts. Underline those residues that favor contacts with water.

W E A N I K Q R L S T E Y K Q

(e) Which of the following mutations in the original sequence is likely to most strongly destabilize the helix? Explain your reasoning.

The answer is 3, see below.

1. W, I, L, Y \rightarrow A will remove the hydrophobic patch on the helix surface (hence less favorable contact with the β -sheet) but will not destroy the helix

2. E, K, D, R \rightarrow A will remove surface charges and make the helix less hydrophilic, but is not likely to perturb the helix significantly

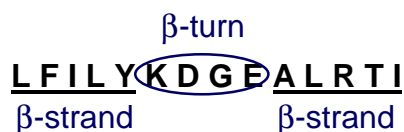
3. **E, S \rightarrow R** will place charges of same sign close in space to each other – their repulsion will destabilize the helix

4. W, A \rightarrow P Pro in the first turn of the α -helix should not significantly perturb it.

Problem 4. (20 points, 10 points each)

A 14-residue peptide, **LFILYKDGEALRTI**, in the context of a protein forms a β -hairpin, consisting of two antiparallel β -strands connected by a β -turn. You have found that the amide nitrogen of phenylalanine forms a hydrogen bond with the carbonyl oxygen of threonine. Based on your finding, characterize the structure of the β -hairpin:

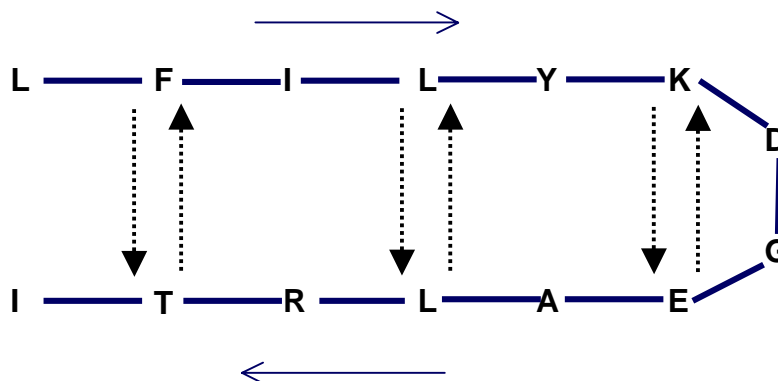
(a) In the peptide sequence below clearly indicate, which residues belong to the two β -strands and which to the β -turn. Explain your reasoning.



Since F is H-bonded to T, this leads to the H-bonding pattern shown below. This identifies the four residues in the middle as belonging to the β -turn: the CO of the first residue in the turn (K) is H-bonded to NH of the fourth residue (E).

(b) Draw *schematically* the hydrogen bonding pattern for this β -hairpin. Do not draw the side chains – just indicate what residues are connected by H-bonds. How many hydrogen bonds between the backbone atoms are present in the structure? Explain your reasoning.

Here is a simplified scheme of H-binding in the β -hairpin. Arrows are used to indicate the direction from donor (NH group) to acceptor (CO group). More detailed as well as less detailed drawing were considered OK, if they correctly addressed the question.



Altogether there are 6 H-bonds in this hairpin: two per each residue that has its NH and CO groups oriented toward the interacting strand. Due to the nature of the β -conformation, the NH and CO groups of alternate residues (here L1, I3, Y5, A10, R12, and I14) are oriented outside the area between the strands and, therefore, cannot participate in the H-bonding between these strands.

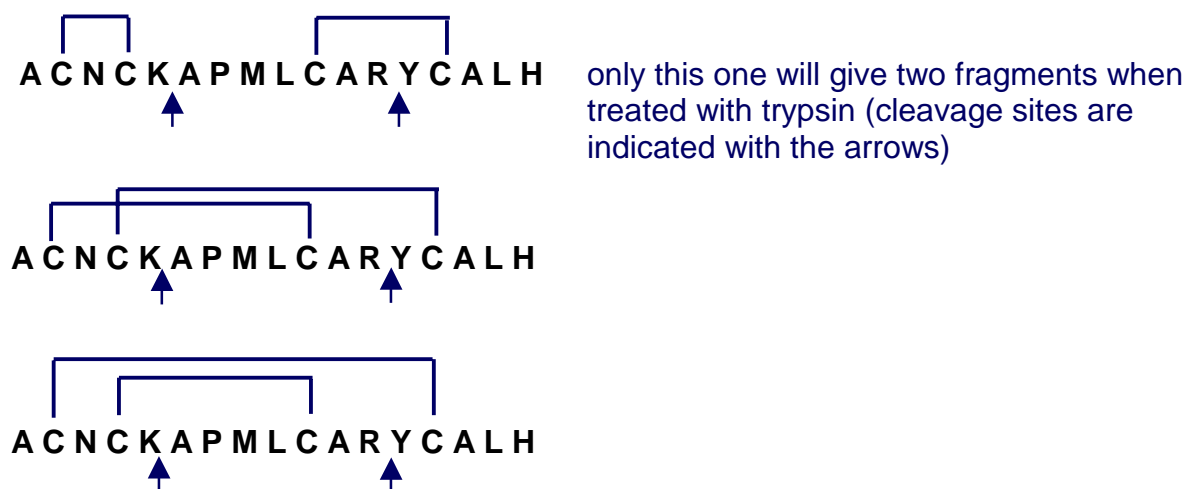
Bonus Problem. (20 points)

A small protein has a sequence **ACNCKAPMLCARYCALH**. It is known that the structure of this protein is stabilized by two disulfide bonds, but their exact location is unknown. In order to determine their location, you treated the protein with trypsin. Trypsin cleavage gave two peptides.

(1) Based on this result, determine the location of the –S-S- bonds. Show them schematically on the sequence by connecting the corresponding residues. Explain your reasoning.

ACNCKAPMLCARYCALH

There are three possible –S-S- bonding patterns for this peptide:



(2) What other cleavage reagent(s) (from the Table on the cover page) could be used to *unambiguously* verify your guess? Explain your reasoning.

Chymotrypsin will not help – the result will be one fragment in all three cases.

Asp-N-protease will be of no use here.

Cyanogen bromide (cleaves at M) treatment will result in two fragments in the first case and in two fragments in the other two cases, hence this will help.