Answers to PS 2 Chem 5.07

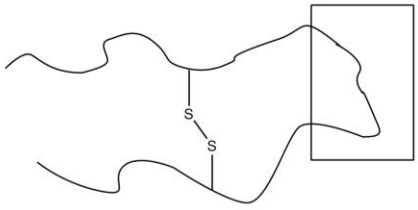
Primary sequence problem

1. Proteins are often associated with membranes by attachment of fatty acids such as palmitate ($CH_3(CH_2)_{14}CO_2^-$) to the amino terminus of the protein, in a post-translational modification step. A group of investigators was interested in a peripheral membrane protein (a protein weakly associated with the membrane) that plays an important role in a signaling pathway inside the cell. The molecular weight of the protein was determined to be 35 KDa by ESI MS and in an effort to obtain some sequence information **the protein was initially cleaved with the protease trypsin**. [Recall trypsin is a protease that cleaves adjacent to positively charged amino acids). The resulting peptides were separated and isolated by HPLC (high pressure liquid chromatography) and one of these peptides was further investigated. You are assigned the task of determining the primary sequence of this peptide given the following pieces of information. [On the last page of the exam is some additional information that might be useful to solve this problem].

a. Total acid hydrolysis of this peptide in 6N HCl, subsequent to reduction of the peptide with dithiothreitol, resulted in the following amino acid composition:
   2 F, M, A, V, 2 K, S, 2 C and D

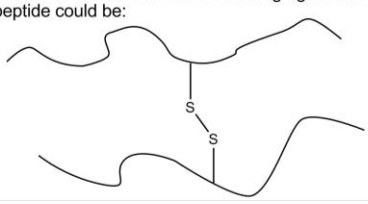
Recall that some amino acids are destroyed during this process and may not be detectable.

ANSWER: Acid hydrolysis may destroy S, T, Y (depending on how long the hydrolysis is performed) and will completely destroy W, so these residues, if present, may not be observable by HPLC analysis. The following amino acids are observed: 2F, M, A, V, 2K, S, 2C, D. The fact that two C residues are observed indicates that there may be a disulfide bond in the structure. Possible structures are shown below:



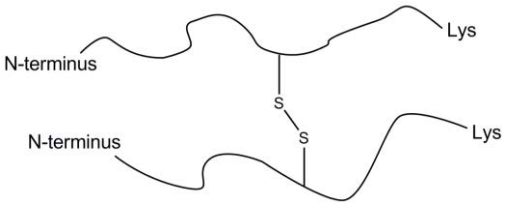
We don't know if this is one continuous peptide or two peptides that are simply covalently linked through a disulfide bond.

Since the protein was not treated with a reducing agent such as DTT prior to treatment with trypsin, the peptide could be:



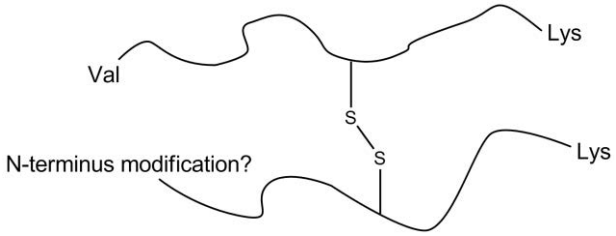
,

One clue to help resolve the structure of the peptide is that there are 2 Ks from the acid hydrolysis experiment. Because the original protein was cleaved with trypsin, each K must be at the C-terminus of a peptide. Recall or look up that trypsin cleaves peptides adjacent to positively charged amino acids such as K or R. The only way for this to work is in the following peptide:



b. One round of Edman sequencing gives V.

ANSWER: One round of Edman sequencing yielded V. From the acid hydrolysis experiment we know that there is only one V, so that means that there is only one unmodified N-terminus, otherwise Edman sequencing would have given two different amino acids. In the introduction to the problem we are reminded that palmitate ($CH_3(CH_2)_{14}CO_2^-$) can modify the N-terminus of proteins. This result suggests that the one the N-termini of our peptide could be modified. A possible structure at this point is:



c. The intact peptide was reduced with dithiothreitol and then treated with the protease called chymotrypsin. Chymotrypsin cleaves peptide bonds after aromatic amino acids or of **very** large hydrophobic amino acids. Five peptides were isolated subsequent to this cleavage and the amino acid composition of each was determined:

(W, V); (C, F); (D, K); (M, A, K, C) and a fifth peptide which upon total acid hydrolysis gave S and F and a third product that is not an amino acid. This third product is hydrophobic and upon examination by MALDI TOF mass spectrometric analysis was shown to have a molecular weight of 256 Da. [Hint think about how a protein might be peripherally associated with a membrane]

ANSWER: First the peptide is treated with DTT, and then treated with chymotrypsin. DTT reduces the disulfide bond to two cysteines. Chymotrypsin cleaves after aromatic amino acids (F, W, Y) or amino acids with very large hydrophobic side chains (L, I, V). If you look at the binding pocket of chymotrypsin (in your text book), you will notice that there are no charged residues as there is for example with trypsin, which has Asp189. Because like attracts like, hydrophobic residues that are large (the binding pocket is deep) and will preferably bind to chymotrypsin and undergo peptide hydrolysis. Trypsin has a negatively charged Asp residue which will lead to the preferential binding, followed by cleavage, of positively charged residues (K, R).

Five peptide fragments were isolated from this experiment:

W, V

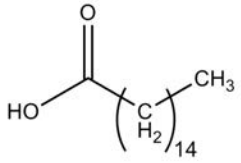
C, F

D, K
M, A, K, C
S, F, X (where X is not an amino acid)

Because chymotrypsin was used to cleave this peptide we can assign the sequences for some of the peptides:

W, V is VW: The W is at the C-terminus
C, F is CF: F is at the C-terminus
D, K is DK (remember that K is at the C-terminus of each peptide fragment) therefore we also know the order of this dipeptide.
M, A, K, C is (M, A, C)K (we cannot determine the order of M, A, and C at this point)
S, F, X is XSF (the rational for this sequence is described below), but F again is at the C-terminus.

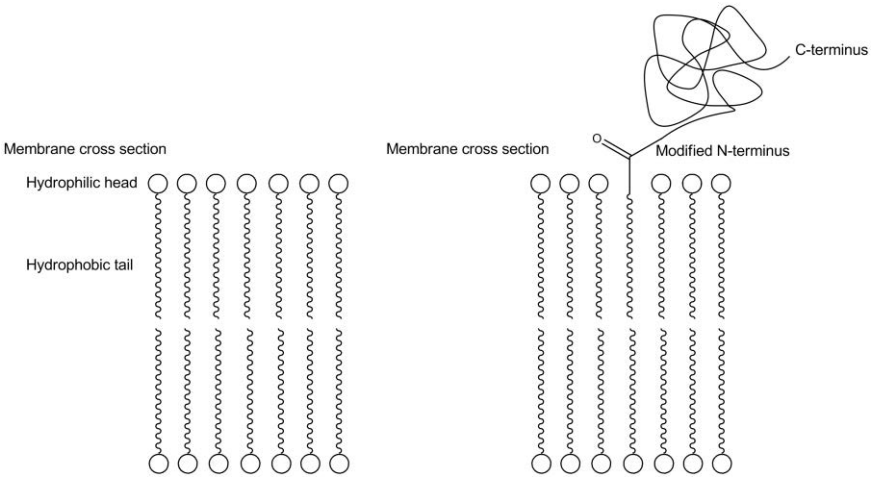
The mass of the unknown product is 256 Da, which corresponds to the species liberated on hydrolysis of the N-terminus:



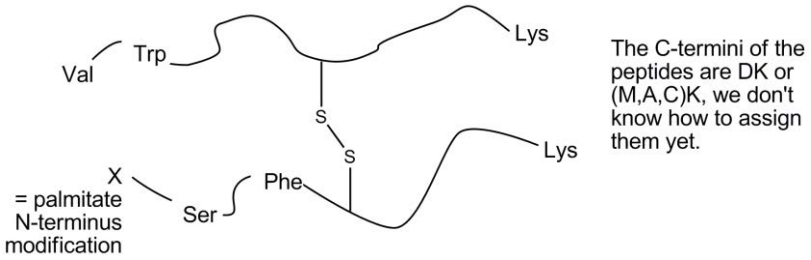
Palmitate obtained after determination of amino acid composition by either acid or base hydrolysis.

Thus we can identify X as palmitate, which is our N-terminal modification. The sequence for the (S,F,X) peptide fragment must be XSF.

From the lecture on water (Lecture 2) you learned membranes are composed of lipid bilayers, each layer has a hydrophobic and hydrophilic section. The hydrophobic tails organize on the inside of a membrane with the hydrophilic head oriented outward. A protein with an N-terminally modified tail that has a long hydrophobic tail and a hydrophilic head can be inserted within this membrane as shown on the right of the figure below.



From this step the following structure can be proposed:



d. The intact peptide was also cleaved with a chemical reagent rather than an enzyme. The reagent used was cyanogen bromide (CNBr). This reagent, discussed in your text book but not in class, cleaves adjacent to methionines and leaves an unusual amino acid (a homoserine lactone). Reduction of the intact peptide with dithiothreitol and cleavage with CNBr gave three peptides which could be separated and the amino acid composition of each determined:

(A, K); (W, F, V, D, K, C); and (F, S, C, and an unnatural amino acid)

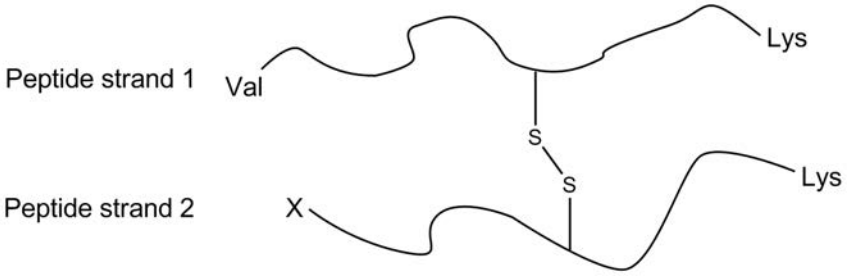
ANSWER: In this step, the disulfide bond is reduced with DTT and then the peptides are treated with cyanogen bromide, which cleaves after Met. This reaction results in a modified Met residue; it is now a homoserine lactone, which would be detected by HPLC as a nonstandard amino acid. The following three peptide fragments were isolated: To the right of the arrow are the known sequences of these peptides based on what we know from this step.

Fragment 1: A,K → AK
Fragment 2: W, F, V, D, K, C → V(W, F, D, C)K
Fragment 3: F, S, C, and an unnatural amino acid (cleavage product of M) → (F, S, C)M

The unnatural amino acid is the modified Met.

Using the information in a-d and additional information in the introduction to the problem, draw the structure of the primary sequence (using the one letter code, a chemical structure is only necessary for parts of the molecule that are not normal amino acids) of this isolated peptide. **Show how each piece of information gave you insight that led to the final structure.** This analysis (telling us what you learned from a through d etc) and the role of trypsin is critical to receive partial credit in case you get the incorrect structure.

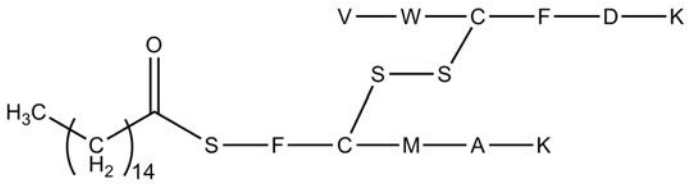
FINAL ANSWER: To determine the sequences of the two peptide strands we need to compare our results from c and d.



For peptide strand 1, we know that it starts with V and ends with K, so it must only be composed of amino acids found in fragment 2 of part d. Based on the observation of the three fragments in part c (VW, CF, and DK) (all of these amino acids are present in fragment 2 of part d) we can assemble fragment 2 into VWCFDK. This is peptide strand 1.

For peptide strand 2, we know that we have an XSF from part c, where X is the N-terminal modification. This result means that fragment 3 above must have the sequence: XSFCM. Furthermore, from part c, we know that we have a sequence that looks like (M,A,C)K, combining this with XSFCM gives: XSFCMAK.

The peptide is:



2. Problem on weak non-covalent interactions: OMP decarboxylase is a protein catalyst, enzyme, that catalyzes the conversion of OMP to UMP shown below. (See the first lecture on catalysis to observe the tremendous rate acceleration for the enzyme catalyzed over the non-catalyzed reaction. To obtain a structure of this enzyme it was co-crystallized with 6-azaUMP.

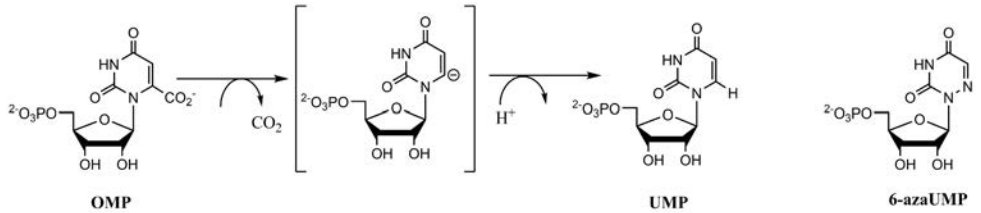
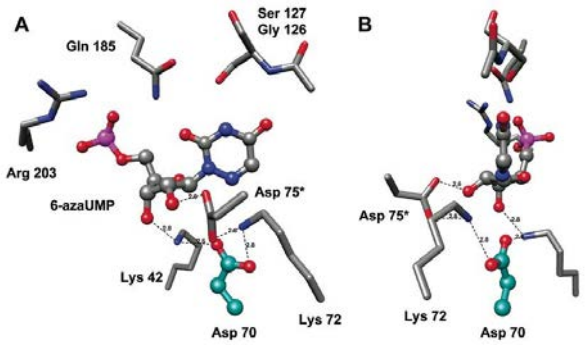


Figure 1. Structure of OMP decarboxylase with 6-azaUMP bound in the active site. A and B are the same structure examined from a different perspective.



Questions:

i. Redraw the Figure in A putting in all the appropriate charges. What are the pKas of the side chains of the amino acids that you have drawn in your figure?

Note: this problem was taken from Biochemistry 2009 48, 5518-31. The enzyme is of great interest because of its tremendous rate acceleration relative to the non-enzymatic reaction. If you look at the actual substrate that has a carboxylate in the C-6 position of the pyrimidine, it seems unusual that Asp75 would be sitting adjacent to this carboxylate. This electrostatic repulsion of the two negative charges (assuming that the structure is similar with the actual substrate bound), would destabilize substrate binding and thus be unfavorable for substrate binding, but perhaps enhance catalysis by destabilization of the ground state of the reaction. What is noteworthy in the above structure is that there are multiple lysines (positively charged) and carboxylates (negatively charged) in close vicinity. Do they form favorable electrostatic or salt bridge interactions (NH_3^+ with CO_2^-) or do they form repulsive electrostatic interactions (CO_2^- with CO_2^- or NH_3^+ with NH_3^+). The issue of appropriate interactions requires a careful look at the 3-dimensional structure. There have been 5 groups arguing over the interpretation of these interactions with respect to binding and chemistry.

ii. Weak non-covalent interactions within macromolecules and between small molecules and macromolecules play an essential role in structure and function. Show with structure two types of weak non-covalent interactions in the structure.

The two types of interactions are electrostatic salt bridges and H bonding and really depends on how you protonate the groups in the active site.

iii. From a chemical perspective, using your drawing in part ii, what is unusual about the amino acid side chains that are adjacent to the site where the chemistry occurs (Eq. 1)?

What is most unusual is the negatively charged carboxylate of Asp75 next to the carboxylate of the substrate OMP or the putative carbanion intermediate in Eq. 1.

3. 2, 3Bisphosphoglycerate (BPG, Figure 1) is present in red blood cells at 4 mM concentrations, about the same concentration as hemoglobin (Hb, 2 mM). Without BPG, Hb is an extremely inefficient O_2 transporter, releasing only about 8% of its cargo to the tissues (Figure 2). A structure of deoxyHb bound to BPG is shown in Figure 3. BPG does not bind to oxyHb.

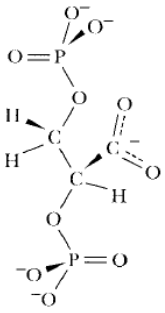


Fig. 2 BPG

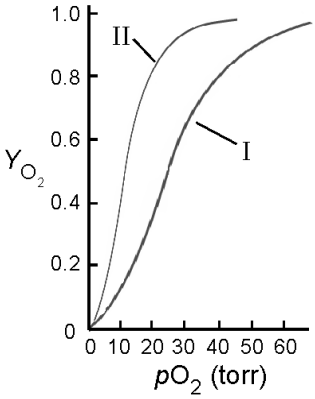


Fig. 3 Purified Hb and Hb in whole blood

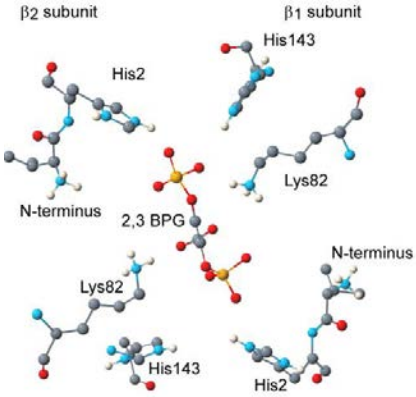


Fig. 4 BPG bound to deoxyHb

The globin gene in fetuses differs from adult Hb in that it has two α and two γ chains. Recall that adult Hb has two α and two β chains. The fetal γ chain is thought to have arisen from the β chain via gene duplication and mutation. Noteworthy is the substitution of His143 in the β chain with a serine in the γ chain (see Figure 4).

Questions:

1. Given the information above, label purified Hb (purified implies the small molecules have been removed from the protein) and Hb in whole blood in Figure 3. Describe in one clear sentence the basis for this labeling.

Purified Hb is curve II and whole cell Hb is curve I. If you examine the x axis, at 20-30 torr the pO_2 in the tissues, most of the O_2 is still bound in II, but 20 to 50% is bound in I. BPG binds to the deoxy form of Hb (intersubunit domain as shown in Fig 4 above. Look in your book for additional pictures). Binding specifically to deoxyHb shifts the equilibrium to the T state and thus dumps more O_2, which is in accord with the experimental observation. deoxyHb has lower affinity for O_2 (T state, weak binding state), while R state has high affinity.

$$Hb(O_2)_n + BPG \leftrightarrows deoxyHb{\bullet}BPG + O_2$$

2. From Figure 4 what is (are) the major type of weak non-covalent forces that allow tight binding of BPG to deoxyHb? Clearly show all the charges in Figure 4.

You see in Fig 4., 2His (143 and 2) and 1 Lys on each β subunit. All of these residues can be charged at physiological pH. From Fig 2 you can see that BPG has 3 to 5 charges (recall that the pKa of the second phosphate is about 6.8 and the extent of ionization depends on the pH of the solution). Thus the binding of BPG is electrostatic in nature (and includes H bonding), probably with the negative and positive charges balanced.

3. The binding curves for O_2 to adult and fetal Hb are shown in Figure 5.

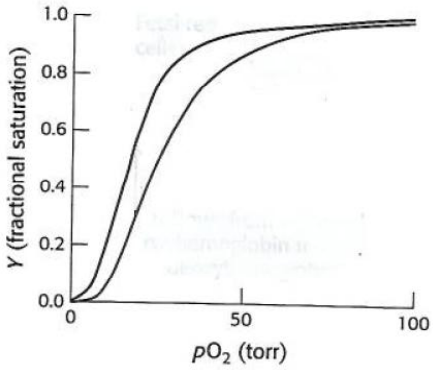


Figure 5

Label the curve associated with HbF (fetal) and HbA (adult). Explain in one sentence the basis for your choice of labeling. Explain why the mutation of His143 to a Ser, could have the observed effect.

HbF is the curve to the left and Maternal Hb is the curve to the right. O_2 flows from the maternal Hb to the fetal deoxyHb, that is, fetal red blood cells have a higher O_2 affinity than the maternal Hb. The basis for this distinction is that fetal Hb does not bind BPG as well as the maternal Hb. This is because replacement of His143 with Ser, potentially removes up to two positive charges and hence there is less electrostatic and H bonding interactions with the fetal protein.

5.07SC Biological Chemistry I
Fall 2013