1). 8 points total

T or F (2 points each; if false, briefly state why it is false)

____T____ The pH at which a peptide has no net charge is its isoelectric point.

____T____ Hydrophilic side chains of amino acid residues normally locate themselves toward the exterior of globular proteins.

____F____ Proteins are made up of D-amino acids.  **Proteins are made up of L-amino acids**

____F____ Mutations always affect biological activity of a protein. **Some mutations are silent, affecting neither structure or function. These are usually amino acids substitutions that are similar in size, shape or charge.**

2) 12 points total

Fill in the blank:

The α-helix is an example of a **secondary** structural element. It is a helical structure formed by **hydrogen bonding** between groups in peptide bonds. In particular, the **carbonyl carbon** of the *i*th residues of an α-helix is bonded to the **amide nitrogen** of the *(i+4)*th residue. The result is α-helical structure that repeats every 3.6 residues. The **pitch** of the helix, or distance along the axis per turn, is 5.4 nm. The arrangement orients the amide planes **parallel** to the helix axis and the side chains **perpendicular** to the helix axis.
Myoglobin was the first globular protein whose structure was solved by x-ray crystallography. The sequence of helix E is given below. Using the helix wheel, map out the amino acid side chains.

E Helix Sequence: SEGDLKHLGATVLTLGGIL
What is the nature of the $\alpha$-helix? Where would it most likely be found?

By plotting out the residues, what you would find is that every 3-4 residues is hydrophobic. Hence, we end up with an ampiphilic helix, which would be found on the surface of Mb.
4). 3 points

Consider what we know about amino acid structure. Briefly describe three factors about amino acids that do not foster $\alpha$-helical formation. (I will give a bonus point for describing four factors).

I was looking for you to address three of the four items listed below. If you address all four, you will get an extra point 😊

1). Stretches of big bulky amino acids like Phe, Trp, Tyr are not conducive to helix formation due to steric hinderance.
2). Proline, because of its cyclic structure, can not assume the correct values of $y$ and $f$ for helix formation.
3). Stretches of charged amino acids (i.e., all glu or asp or all lys or arg) is not conducive to helix formation due to charge repulsions.
4). Glycine can assume too many degrees of freedom, hence it is not compatible with the structured nature of the $\alpha$-helix.
5). 7 points total

For the sequence drawn below, write the one letter code and determine the pI. What is the charge on this molecule at pH 12?

The sequence spells out ELVIS

<table>
<thead>
<tr>
<th>Group</th>
<th>pKa</th>
<th>Charge at pH 0</th>
<th>Charge at pH 3</th>
<th>Charge at pH 4.5</th>
<th>Charge at pH 12</th>
</tr>
</thead>
<tbody>
<tr>
<td>amino terminus</td>
<td>~10</td>
<td>+1</td>
<td>+1</td>
<td>+1</td>
<td>0</td>
</tr>
<tr>
<td>Glu side chain</td>
<td>4.3</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>Carboxy terminus</td>
<td>~2</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>Sum of charges</td>
<td></td>
<td>+1</td>
<td>0</td>
<td>-1</td>
<td>-2</td>
</tr>
</tbody>
</table>

\[ pI = \frac{(2+4.3)}{2} \sim 3.15 \]

Charge at pH 12 = -2
6). 6 points
Determine the amino acid sequence of the following oligopeptide from the experimental data below.

1. The amino acid composition is found to be [ala, lys, phe, met, cys, plus some decomposition products].
   *We have at least 5 amino acids. Because acid hydrolysis is so harsh, we can't unambiguously detect every amino acid.*
2. The peptide has a molecular weight around 700 Da and absorbs at 280 nm.
   *700 Da ÷ 100 dalton/residue ~ 6 residues*
3. Treatment with carboxypeptidase results in tryptophan and a peptide.
   *Carboxypeptidase cleaves successively from the carboxy-terminus.*
   Therefore we can start to assign a sequence as
   
   __ __ __ __ __ W

4. CNBr treatment yields a tetrapeptide and a dipeptide.
   *CNBr cleaves after Met*
   So we have:
   
   __ W and __ __ __ M
   OR
   __ M and __ __ __ W

5. Trypsin digestion produces an amino acid and a pentapeptide with met on the amino end.
   *Trypsin cleaves after basic residues.*
   If we look at the results from point 4, the only way that we can have a pentapeptide with Met on the **amino** end is
   
   K and M __ __ __ W

6. Chymotrypsin digestion yields a dipeptide and a tetrapeptide.
   *Chymotrypsin cleaves after aromatics.*
   So, we have
   
   K M __ F __ W
   There is not enough information to assign Ala and Cys so the final product could be either
   
   K M A F C W or K M C F A W
Determine the subunit composition of a protein based on the following information. Be explicit in describing the subunit composition.

Molecular mass by gel chromatography: 200 kD
Molecular mass by SDS-PAGE: 100 kD
Molecular mass by SDS-PAGE with 2-ME present: 40 kD and 60 kD

Gel chromatography provides the mass of the native protein. So we know that the native protein is 200 kD.

SDS disrupts non-covalent interactions. It tells us that there is a polypeptide chain of 100 kD. This suggests to us that our protein consists of 2 chains of 100 kD, held together via non-covalent interactions.

β-mercaptoethanol reduces disulfides to free cysteine residues.

So, our 100 kD chains are composed of 2 chains, one of 40 kDa and one of 60 kD, held together by a disulfide bond.

Hence our protein is composed of four chains total. A 40 kD and 60 kD chain are held together covalently to give a 100 kD chain. The 100 kD chain associates via non-covalent interactions with itself to give the 200 kD protein.
Homework questions based PDB file and Swiss-PDB Viewer Tutorial (attached at end of homework).

8). 17 points total

From the PDB file, answer the following questions.

a). What is the chemical classification for carbonic anhydrase II and what does this general type of protein class do (from class notes)? Lyase (oxo-acid). Lyases add chemical groups to double bonds (in this case hydration of $CO_2$).

b). What experimental method was used to determine the structure? x-ray crystallography.

c). What is the source of CA2? human

d). How many amino acid residues comprise this protein? 259

e). What is (are) the main hetero-atoms? Zn and H$_2$O

f). Were all amino acids detectable in the structure? No - the remarks line tells us no.

REMARK 4 N-TERMINAL RESIDUES SER 2 AND HIS 3 AND C-TERMINAL RESIDUE 1CA2 79
REMARK 4 LYS 261 WERE NOT LOCATED IN THE DENSITY MAPS AND, 1CA2 80
REMARK 4 THEREFORE, NO COORDINATES ARE INCLUDED FOR THESE RESIDUES. 1CA2 81


g). How many W, Y, F respectively are in the structure? Would you expect this protein to have an absorbance at 280 nm? W=7; Y=8; F=12; yep - all are aromatics and would absorb in the UV.

h). How many $\alpha$-helices, $\beta$-sheets and turns are defined in the structure? (Try to find them in the structure using Swiss-PDB Viewer).

7 $\alpha$-helix; 10 b-sheets; 6 turns.

i). How many water molecules are found in the structure? 166

j). As a thought question, not related to the PDB file, would you think that many of the waters are found on the inside of the structure? Why or why not? No, the interior of a protein is hydrophobic and there are few chances to satisfy hydrogen bonds.
Question related to the Swiss-PDB tutorial.

9). 2 points

a). The Zinc hetero-atom in the structure of CA2 plays a role in catalysis. Described its location in the structure. The Zn is located in a cleft in the molecule.

b). What statements about the location of an enzyme’s active site can be made based on this information? Active sites of enzymes are not directly on the surface of the protein, because we need a protected environment in which to carry out chemistry so that unwanted reactions and by-products do not occur. In the case of hemoglobin and myoglobin, which are not enzymes, but bind O₂, the heme pocket prevents the iron from oxidation (FeII → FeIII) which would abolish O₂ binding, so this type of arrangement can also protect the functional integrity of sensitive functional groups at active sites. As such, active sites are often times found in crevices, grooves and pockets that are accessible to the surface (we need to get substrate in and products out).
Protein Data Bank and Swiss-PDB Viewer Homework

The ultimate goal behind this tutorial is to get familiar with working with Swiss-PDB Viewer. An additional tutorial that is recommended by one of my colleagues who does x-ray crystallography (but which I haven’t personally used) is located at http://www.usm.maine.edu/~rhodes/SPVTut/index.html.

Why Swiss-PDB Viewer? It appears to be the most used, and best supported structural imaging software available. It is available free at http://us.expasy.org/spdbv/mainpage.html. From that site you can download versions for Mac, PC, Linux and SGI. To date, I have used a different program to view molecules (Rasmol), but it isn’t as powerful as Swiss-PDB Viewer and it is no longer supported. This is a learning experience for me also.

The best way to learn about any software is to "play" with it. Try out the different menus and see what happens. Don’t be afraid to explore the program. This tutorial is based on what I did to learn basic familiarization with the program. Answer the homework questions as you go through this exercise.

A). We will start with the Protein Data Bank located at http://www.rcsb.org/pdb/.

Download the structure for carbonic anhydrase II. The pdf id is 1CA2.

In the homework, there are a series of questions related to the pdb file.

B). Open the structure with Swiss-PDB viewer.

One homework question related to this section.

Rotate the structure: You have a stick structure of the atoms in front of you.

Color scheme: White - carbon
      Blue - nitrogen
      Red - oxygen
      Yellow - sulfur

Can you identify amino acids based on their chemical structure?

Look for histidine, tyrosine

What sulfur-containing amino acids does the structure contain?

C). Manipulate the image:

Under the "Display" menu:

Show backbone oxygens: Turn off and on; see how the structure is altered.
Play with other options to check them out; some options alter the image, some are not active right now.

D). Control panel window:

The control panel allows you to highlight and manipulate selected amino acid residues. We are going to manipulate the image through the control panel. You have the option of affecting individual amino acids by clicking in the appropriate column next to the amino acid, selecting all amino acids by holding down the shift key and clicking at the top of the column or selecting specific groups of amino acids by using options on the menu and then manipulating the image through the control panel.

**Column 1:** show - hides or shows selected residues. Try hiding the first histidine - you can see it disappear from the structure. Select all the residues by holding down the shift key & clicking at the top of the column - your entire structure should disappear! Bring your structure back!

**Column 2:** side - hides or shows selected side chains. Deselect all the side chains. You can start to see the "roadmap" of the protein backbone (It becomes even clearer if you deselect backbone oxygens under the Display menu).

At this point, can you identify the Zn atom?

You can also now start to play with selected types of atoms:

Under the Select menu:

Select → Group property → Basic

Notice that the basic amino acids are highlighted in the command window.

Click on the word "side" in the command window.

The basic amino acids are highlighted in the structure.

Let's color them green.

Click on the words "ribn col" in column 4. A box of crayons comes up; select green; click OK.

Rotate your structure; check out their location (surface or interior?).

Do the same thing for the other Group Properties.

Do the same thing for the individual amino acids: Select → Group Kind → Cys (for example).

**Column 3:** labl - puts a label on the selected amino acids.

Under the labl column: select His4 and Phe 260 (the amino-terminus and carboxy-terminus of the protein, respectively).
Column 4: (has a bunch of dots for symbol). This puts different types of surfaces on the selected atoms.

Under the Select menu:

Select → Group property → Basic
Play around with the 3 different surface representations (van der Waals, accessible and molecular).

Column 5: ribn & color box

Under the word ribn, you can shift-click to select all. You will obtain a ribbon trace of your protein. Try to follow the α-helices and β-sheets through the protein. You can make the ribbon display look "nicer".

Under the following menus:

Display - render in 3D
Prefs - ribbons

Set the following commands to have the same color for α-helices (and a 2nd color for β-sheets).

Use this top
Use this side
Use this bottom

Look at the hetero-atom:

In the Control window:

Select Zn262 - make it green. Answer the homework question on Zn.

Stereo-viewing:

Display - stereo.

If you can de-focus your eyes to see in 3D, you have a great advantage over the rest of the world. Here is how I was taught. Tilt your head up so that you are looking down the line of your nose. You should start to see three images. Ignore the two outside images and keep focussing on the center image. Eventually, it should snap into focus and you should see the image in 3D. Once it snaps into focus, you will be able to move your eyes around the structure and you will be able to rotate the structure without losing the 3-dimensionality. It took me about 2 days of practicing to be able to do this - you might want to start with a 3D image on paper.