1.) Please answer each question briefly; unless explicitly stated, full sentences are not required. (3 points each)

a) Many integral membrane proteins have one or more "membrane-spanning" alpha helices which start on one face of the plasma membrane and extend straight through the membrane to the other side. In some cases, the antigen receptor on T lymphocytes, the membrane is spanned by a single helix, and in others, like the B-adrenergic receptor on neural cells, a protein may have several helices running back and forth across the membrane.
Envision the different environments that a membrane-spanning helix runs through. Consider the impact this will have on amino acid composition in the helix. Please **list three amino acids that you would expect to find in the mid-section of such a helix**, the part that is in the middle of the lipid bilayer.

Any hydrophobic amino acid should work: I, L, V, M, F, Y (G and A are also acceptable – though sub-optional).

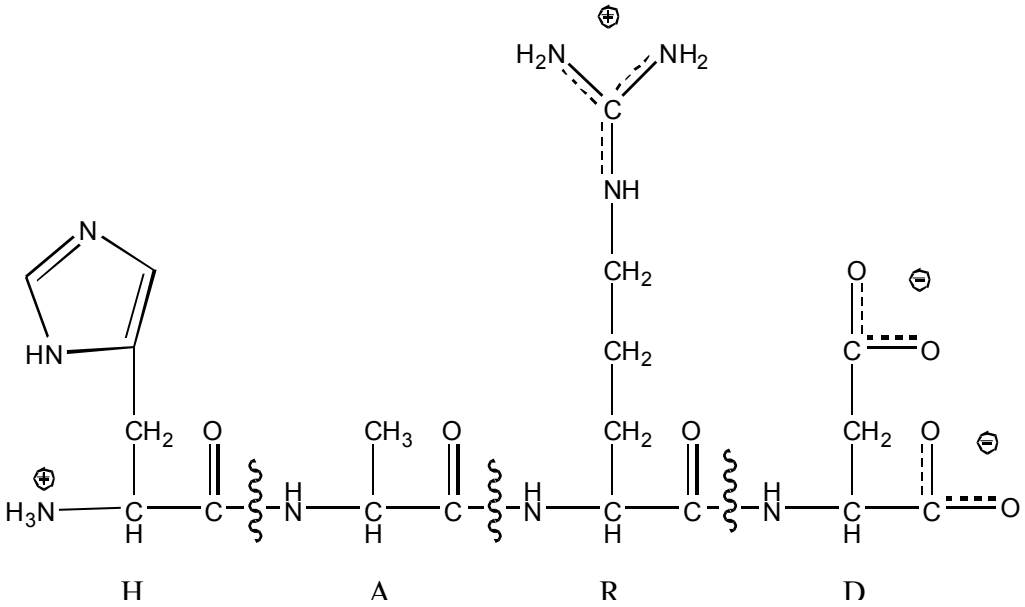
b) Please **name and draw** one amino acid that you would expect to find **at one end of the membrane-spanning segment as it passes by the phospholipid head groups**.

This must be a hydrophilic amino acid and is apt to be positively charged such as R or K.

c) A fraction of a cell extract is passed over a gel-filtration column. The second and seventh protein-containing fractions to come off the column look interesting and so are analyzed by SDS-PAGE. Each fraction is shown to contain only a single polypeptide when the gel is stained with Coomassie blue. **Which fraction** -- #2 or #7 -- should give a **band closer to the bottom of the gel**?

By gel filtration, fraction #2 must be larger than fraction #7. Thus, on SDS-PAGE, if both are single polypeptides, #7 should migrate faster and be closer to the bottom.

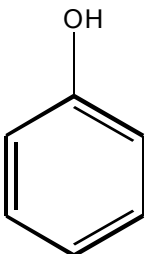
Question #1, continued...
e) Please draw the structure of the following peptide: HARD. Indicate approximate charges (including partial charges, if any) appropriate to pH 6.5.



f) Please **sketch** a graphical representation of a polypeptide whose **tertiary structure includes three globular domains**. Indicate *all* amino- and carboxy-termini in your sketch as appropriate.

Critically: domains distinct, only one polypeptide

g) Persons interested in purifying nucleic acids from whole cell extracts often remove contaminating proteins with organic solvents. Typically, proteins are first extracted from the aqueous phase into phenol and the phenol discarded, followed by extractions of the aqueous phase with chloroform to remove the traces of phenol. If phenol is so effective in purifying proteins out from other cellular macromolecules, why don't we use it as an initial purification step for proteins?



Organic, hydrophobic

Proteins are normally hydrophilic on the OUTside and hydrophobic on the INside. Putting a protein in hydrophobic solvent, then, tends to turn it inside out. This is denaturing…bummer

i) Consider two proteins mixed together in a buffer at pH 6.0. If protein A has a pI of 4.5 and protein B has a pI of 7.7, **which one** (or both or neither) **will be retained by an anion-exchange matrix?**

At pH 6, A is negative and B is positive. Only A will be retained by this resin.

j) As the result of our shared ancestry, we have many proteins in common with all other known life forms. However it is perhaps not surprising that due to evolutionary divergence, human proteins have many differences in primary structure relative to proteins performing the same functions in, say, a corn plant. Indeed it is often more surprising to find a single amino acid position that is absolutely identical between comparable proteins from all species. **What is one specific factor that might cause a glutamic acid to be found at the same position in the primary structure of comparable proteins from all species?**

If this amino acid is conserved widely, it probably plays a key role in protein structure or function. For example, it may form a salt bridge with a lysine or arginine.

k) The protease thermolysin contains a Zn^{+2} metal ion that plays a critical role in its enzymatic activity, much like the enzyme carboxypeptidase presented by one of our groups last week. The zinc ion is held in the proper position in the protein by close interaction with three amino acid side chains: H69 (i.e., a histidine at position 69 in the primary structure), E72, and H196. In a sentence or two, please **explain how three amino acids that are not immediately adjacent to each other in the sequence** of the thermolysin protein **can still occupy points in space that are within a few angstroms of each other**.

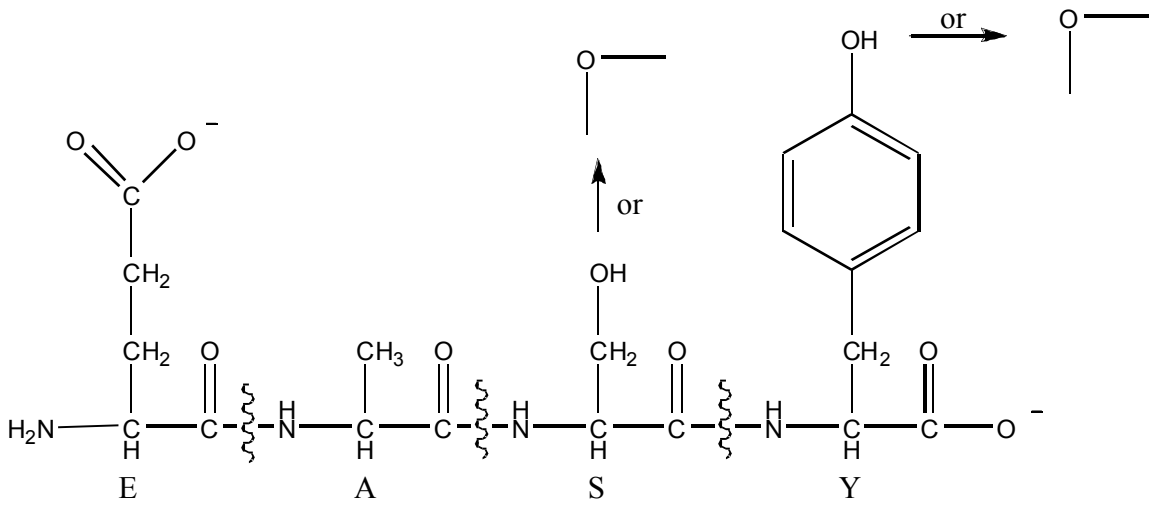
Proteins do not usually occur as elongated chains; rather they fold into compact balls. Thus, amino acids that are separated in 1° structure may be next to each other in 3° structure.

2.) For each of the following, please answer the question in one or a few words, or indicate that there is INSUFFICIENT INFORMATION for you to be sure of the answer. Although you need not do so for full credit, if you would like to ensure that I understand the reasoning behind your answer, you may justify your choice in one short sentence.

a) A particular dye molecule is green in hexane, yellow in methanol, and orange in water. If this dye molecule, like a heme, can be bound to a protein for some biological function, in what part of a protein would you expect to find it if it were yellow?

In an area that is polar but not fully exposed to water, such as near the surface or in an extensively H-bonded area such as a β sheet.

b) Please draw the structure of the following peptide: EASY. Indicate charges appropriate to pH 12.0.



d) Consider two proteins mixed together in a buffer at pH 6.0. If protein A has a pI of 4.5 and protein B has a pI of 7.7, which one (or both or neither) will be retained by an anion-exchange matrix?

At pH > pI, lower [H^+], net deprotonated, charge –
Protein A will be – and B will be + at this pH. Only protein A will be retained by an anion-exchange column.

e) As the result of our shared ancestry, we have many proteins in common with all other known life forms. However it is perhaps not surprising that due to evolutionary divergence, human proteins have many differences in primary structure relative to proteins performing the same functions in, say, a corn plant. Indeed it is often more surprising to find a single amino acid position that is absolutely identical between comparable proteins from all species. What is one specific factor that might cause a glutamic acid to be found at the same position in the primary structure of comparable proteins from all species?

Several possibilities exist, including

- **it is engaged in an ionic interaction with a positively-charged amino acid**
- **it plays a catalytic role that requires a negative charge**

f) Suggest a possible amino acid substitution (a novel amino acid in place of the naturally-occurring one) that would have the same impact on protein structure as one of the four treatments in question #1 part (a). Be sure to indicate 1) which Tube # you are matching the effect of, 2) what amino acid might have been there naturally, and 3) what amino acid you would switch it to.

Easiest to imagine is a mutation that replaces a cysteine involved in the α/β disulfide bond with, say, a serine (like tube #3).

g) Consider the two peptides WEAK and YEAR. If you want to **separate these two peptides by ion-exchange chromatography**, what would be an appropriate pH to choose?

WEAK and YEAR will both be net neutral around pH 7. Best pH for separation by ion-exchange chromatography would be a pH between the pK_A of lysine and that of arginine (e.g. pH 12).

h) The protease thermolysin contains a Zn^{+2} metal ion that plays a critical role in its enzymatic activity, much like the enzyme carboxypeptidase presented by one of our groups last week. The zinc ion is held in the proper position in the protein by close interaction with three amino acid side chains: H69 (i.e., a histidine at position 69 in the primary structure), E72, and H196. In a sentence or two, please explain how three amino acids that are not immediately adjacent to each other in the sequence of the thermolysin protein can still occupy points in space that are within a few angstroms of each other.

Proteins in their native state are not linear beasts. Rather, they are folded into complex 3D forms, with the result that amino acids that are far apart in primary structure may be very close in tertiary structure.

i) True or False? The structure of alpha helices is produced by the packing of hydrophobic amino acids down one "face" of the helical "cylinder".

False. While some alpha helices may be amphipathic (with one hydrophobic and one hydrophilic face), what really produces the alpha helical structure is the array of hydrogen bonds between backbone functional groups aligned along the helical axis.

j) Suppose that I pour a mixture of three proteins over an ion-exchange column and find that one adheres to the column while the other two flow through the column. I set the two aside for now and focus on the one on the column. I increase the salt concentration (just as you did in lab) to elute my protein, and when I check my eluate fractions, I am amazed to find that I have two proteins instead of just one. What could be going on here?

It seems likely that the protein that stuck to the column has (at least) two subunits that are held together by salt-labile bonds (e.g. H-bond, ion pairs).

k) Consider a membrane protein whose amino- and carboxy-termini have both been experimentally found to be on the cytosolic side of the plasma membrane. This protein also contains a heme-based prosthetic group, which is found on the extracellular side of the membrane. It is suggested that the transmembrane portions of this protein are alpha-helical. Given what you know about the sizes of alpha helices and lipid bilayers, what would you predict would be the absolute minimum length of this protein?

If a lipid bilayer is about 40 Å thick and one turn of a helix is about 5.5 Å (3.6 residues), then each transmembrane helix is about 7 turns long, or about 25 residues long. Two of these helices plus a few residues on the N- and C-termini plus a dozen on the heme side gives a minimum of about 65 AA.

l) Please sketch a graphical representation of a polypeptide whose tertiary structure includes three globular domains. Indicate *all* amino- and carboxy-termini in your sketch as appropriate.

There should be only one N- and one C- terminus.

Figure1:

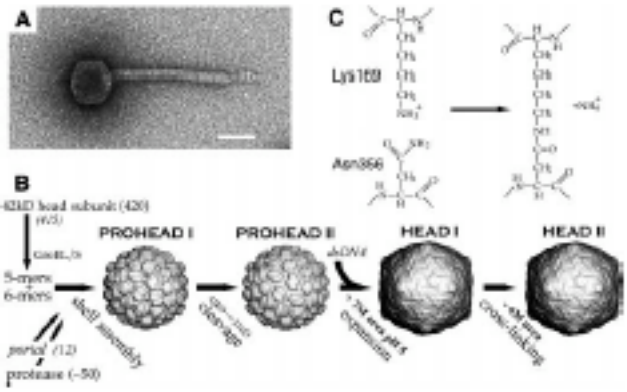


Fig. 1. HK97 assembly and maturation. (A) Negatively stained electron micrograph of a mature dsDNA-filled capsid, with noncontractile tail and accessory proteins. (B) Steps in capsid assembly and maturation (see text; in vitro conditions are in bold, and in vivo conditions or components that differ from the in vitro conditions are in italic) (9). (C) Chemistry of the cross-linking reaction.

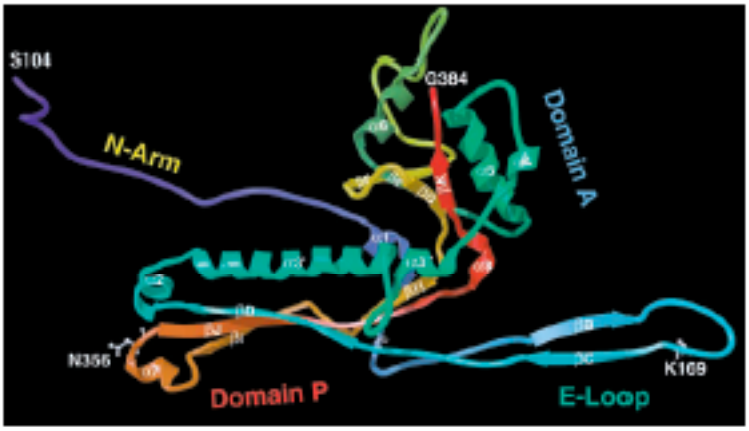


Fig. 2. Structure of one gp5 subunit, color ramped from the NH_2-terminus (violet) to the COOH-terminus (red) (label colors correspond to Fig. 3A domain colors). The Head II NH_2-terminus becomes Ser^{104} by maturational proteolysis in the Prohead I to II transition. The subunit is organized into A and P domains, plus the extended N-arm (violet) and E-loop (cyan). Lys^{169}, on the E-loop, forms an isopeptide bond with Asn^{356} on a neighboring subunit.

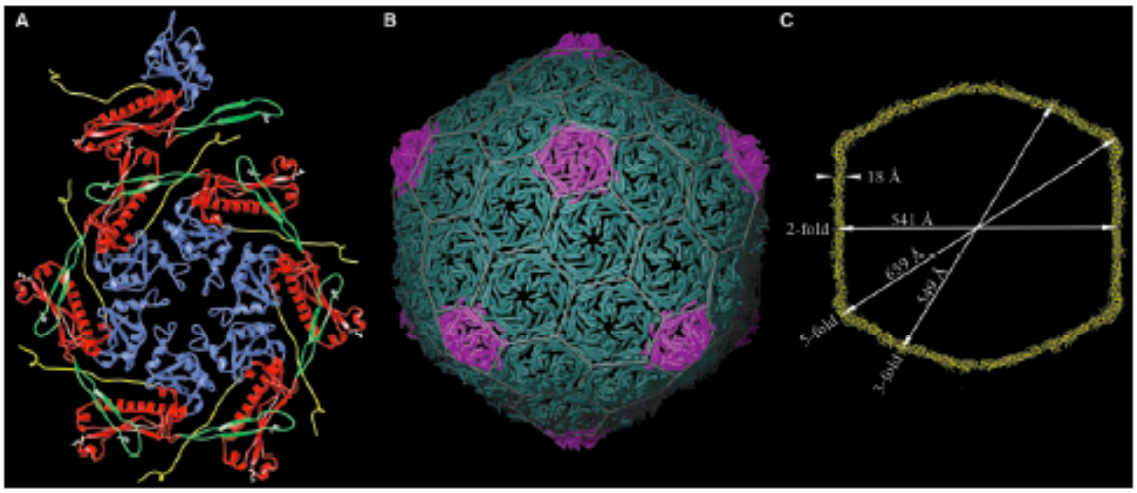


Fig. 3. Capsid organization. (A) The capsid asymmetric unit (A domain: blue; P domain: red; N-arm: yellow; E-loop: green). The capsid is a $T = 7$ arrangement of 420 subunits, organized into hexamers (one shown) and pentamers (one pentamer subunit is shown). The subunits wrap around each other in an intricate arrangement. Cross-links cannot form between subunits within the asymmetric unit, because the cross-linking residues (Lys^{169} and Asn^{356} in white) are not in close proximity. (B) The complete capsid from the particle exterior (each subunit backbone is a smoothed tube). The hexamers (green) are flat, with most of the particle curvature at the concave pentamer (magenta), producing the distinctive icosahedral capsid shape. A $T = 7$ cage (gray) indicates the quasi symmetry axes. The pentagon and hexagon vertices are icosahedral or quasi threefold axes, with icosahedral or quasi twofold axes equidistant between them. (C) Cross section through the unusually thin empty capsid, which despite its large size (659 Å along fivefold), is only 18 Å thick. Icosahedral symmetry axes are indicated.

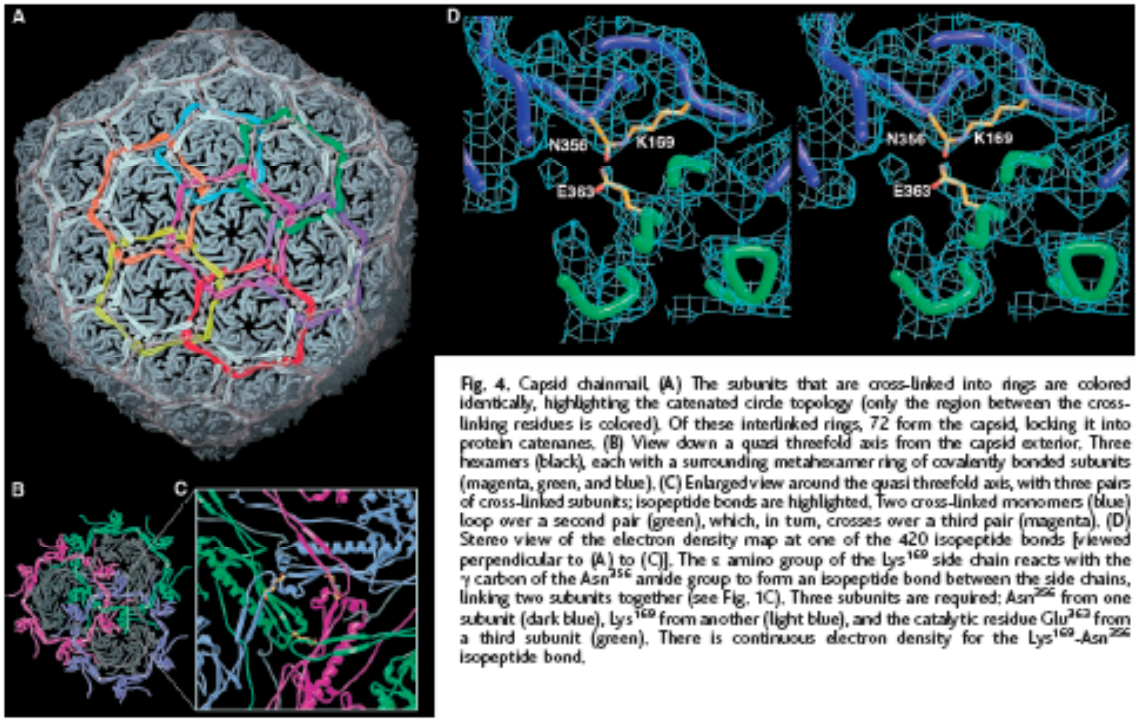


Fig. 4. Capsid chainmail. (A) The subunits that are cross-linked into rings are colored identically, highlighting the catenated circle topology (only the region between the cross-linking residues is colored). Of these interlinked rings, 72 form the capsid, locking it into protein catenanes. (B) View down a quasi threefold axis from the capsid exterior. Three hexamers (black), each with a surrounding metahexamer ring of covalently bonded subunits (magenta, green, and blue). (C) Enlarged view around the quasi threefold axis, with three pairs of cross-linked subunits; isopeptide bonds are highlighted. Two cross-linked monomers (blue) loop over a second pair (green), which, in turn, crosses over a third pair (magenta). (D) Stereo view of the electron density map at one of the 420 isopeptide bonds [viewed perpendicular to (A) to (C)]. The ε amino group of the Lys^{169} side chain reacts with the γ carbon of the Asn^{356} amide group to form an isopeptide bond between the side chains, linking two subunits together (see Fig. 1C). Three subunits are required: Asn^{356} from one subunit (dark blue), Lys^{169} from another (light blue), and the catalytic residue Glu^{363} from a third subunit (green). There is continuous electron density for the Lys^{169}-Asn^{356} isopeptide bond.

3.) Please look carefully at the structural model of the head subunit shown in figure 2.

a) The authors suggest that the unusual makeup of the HK97 capsid is due in part to the need for a strong structure to balance the fact that this capsid is "very thin". Although no scale bar is given in figure 2, you should be able to estimate the size of this protein quite accurately from details of its structure. Please tell me **how long** (left to right) and **how wide** (top to bottom) **in**

Angstroms you expect this polypeptide to be. Please be sure to rigorously justify your answer to this (and every other!) question. (7 points)

Using α helix #3 as a yard stick of (~ 9 helical turns × 5.4 Å/turn ≈ 49 Å), this protein appears to be about 130 Å long and about 70 Å wide.

b) Please draw arrows pointing to specific locations in the structural model that satisfy each of these conditions. (6 points)

- identify a place in the head subunit where you might expect to find a glycine residue
  any tight turn
- identify a place in the head subunit where you might expect to find a phenylalanine residue
  buried in domain A or P
- identify a place in the head subunit where you are *guaranteed* to find intra-chain hydrogen bonding
  within an alpha helix or beta sheet

c) The authors report that the loop between helices α3' and α3" is stabilized by an electrostatic amino acid interaction. (Let me know if you can't find this on your copy of the figure!!) Please explain the sense in which this interaction "stabilizes" the loop, **and** tell me which of the following amino acid pairs would be reasonable candidates for such an interaction: H-E, K-N, H-K. **Please show fully-detailed amino acid structures to support your answer.** You may assume pH = 7. (10 points)

i. **an ion-pair might 'stabilize' this loop by providing favorable enthalpy for the folded form vs. the unfolded. This OH is apt to be larger than the OH gained by forming H-bonds with solvent.**

ii. **At pH=7, H is slightly (+), E is (-), K is (+), and N is neutral. Thus, the only possible pairing here, even though it won't be great at pH=7, is H-E.**

4.) In discussing the novelty of the "protein chainmail" structure, the authors state: "The HK97 capsid protein has one clear homolog, the capsid protein of the *Pseudomonas* phage D3, which also forms crosslinks. They have 41% sequence identity and the cross-linking Lys^{169} and Asn^{356} residues…are conserved."

What are the authors inferring from the presence of overall sequence similarity and the perfect matching of these two residues between the two bacterial virus particles?

If you were able to mutate the position in D3 that corresponds to the HK97 K169, would you expect there to be any impact on the stability of the D3 virus?

Please explain! (12 points)

a.) From overall similarity and especially from conservation of these two key AAs, the authors are inferring a commonality of evolutionary history and structure between these two viral capsids.
b.) It is a logical prediction then that tweaking the D3 equivalent of HK97's K169 would weaken the D3 capsid, just as tweaking the K169 would destroy the key K-N bond and weaken the HK97 capsid.

3.) Consider how the inter-subunit peptide bond shown in figure 1c differs from other interactions between protein subunits that we have discussed. Some of the reagents that are commonly used to induce disassociation of polypeptides in multi-subunit proteins are β-Mercaptoethanol, moderately acidic buffers, and salts such as NaCl.

Please address the following three points, either individually or in a unified paragraph. (25 points)

a) Why would **each** of these reagents **work or not work** to separate the HK97 subunits? That is, what does each of these reagents do?

βME reduces disulphide bonds. The HK97 inter-subunit bond is an amide bond, and so will not be affected by βME. Moderately acidic buffers may affect the protonation status of sidechain and mainchain ionizable groups on proteins. Thus, such reagents could disturb electrostatic (and hydrogen bonding) interactions, but are unlikely to have much impact on peptide bonds within or between peptides. Salts will impact electrostatic and hydrogen-bonding interactions, but again are not going to separate the HK97 subunits.

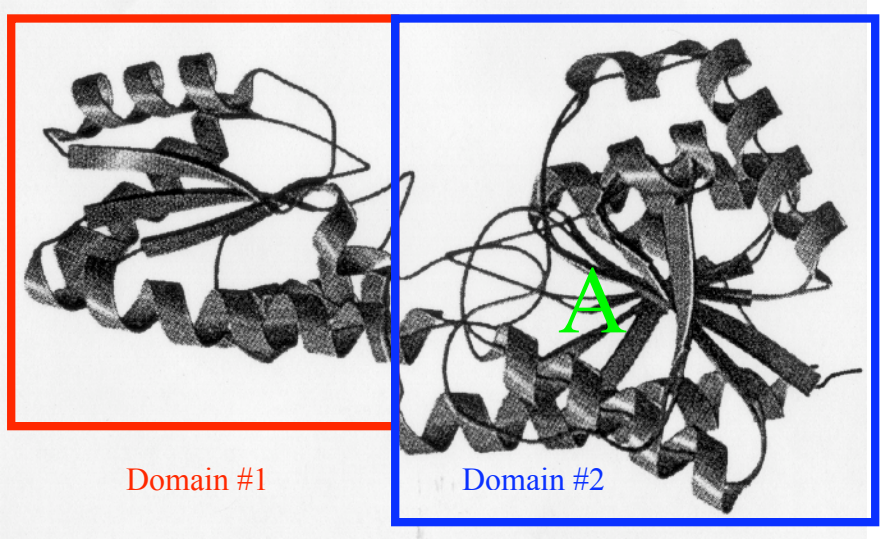
b) What other reagent *might* work here? What would it need to do in a chemical sense in order to separate the subunits?

To work here, a reagent must cause the hydrolysis of peptide bonds. Strong acid should therefore work, but of course that would hydrolyze all of the backbone peptide bonds as well. An enzyme – a protease – that would specifically hydrolyze this bond alone would be the best reagent, but it may not exist naturally.

c) For each of the reagents above, please identify the level(s) of protein structure (1^o, 2^o, 3^o, 4^o) on which you would expect it to have a significant impact.

βME will affect tertiary and quaternary structures that are held together by disulfides. Moderately acidic buffers and salts will primarily affect secondary and tertiary structures and may also disrupt quaternary associations that depend on charge-charge interactions and hydrogen bonding.

5.) Below is a diagram of the structure of a protein that I'll call simply "protein X".



a) Please indicate on the diagram the dominant secondary structural elements, the discernible domains, and any subunit boundaries that you can identify. (5 points)

2°: helices (presumably α-helix), β sheets, there are also some coiled regions
No subunit boundaries can be discerned, at least not from the available information

b) What types of forces would you expect to be responsible for maintaining each of the tertiary structural elements that you have identified? Please illustrate each of these types of interactions with an example using relevant amino acids. (12 points)

Many examples possible for all but disulfide: hydrogen bonds, hydrophobic interactions, covalent bonds (disulfides), ionic interactions

c) What types of amino acids would you expect to find in the region marked "A"? Please give two examples of this type of amino acid (name only is sufficient). (5 points)

"A" is at the core of a globular domain. Thus, it is reasonable to assume that it is shielded from solvent and composed primarily of hydrophobic amino acids. Examples are F, Y, W, L, I, V, M.

Situation One: You have a mixture of three peptides:
A: ELVISISKING (0)
B: ALIENSAREHERE (-2) at pH ~7
C: ISTHISGYFRREAL (+1)

6.) If you wanted to separate these peptides via ion-exchange chromatography, what type of resin(s) (anion-exchange or cation-exchange or both) would you use, and what pH would you make your buffer?

At pH 7, peptide A should not stick to either cation- or anion-exchange resin.

7.) In order to make 1.0 liter of a 0.050 M Phosphate buffer at the pH you chose in question #1, how many grams of what phosphate species would you need? The pKs for the phosphate system are 12.7, 7.2, and 2.1. Formula masses are: Phosphoric acid, 97 g/mol; Potassium dihydrogen phosphate, 136 g/mol; Potassium hydrogen phosphate, 174 g/mol; Potassium phosphate, 212 g/mol. (If you didn't answer #1, pick any pH you'd like; if you gave a range of pH values in #1, such as "between 4 and 10" then pick something, like "6")

At pH 7: $[H_2PO_4^-] + [HPO_4^{-2}] = 0.05\ M$

moles $H_2PO_4^-$ × [(1 mol KH_2PO_4)/(1 mol $H_2PO_4^-$)] × [(136 g KH_2PO_4)/(1 mol)] = 4.2 g KH_2PO_4

moles HPO_4^{-2} × [(1 mol K_2HPO_4)/(1 mol HPO_4^{-2})] × [(174 g K_2HPO_4)/(1 mol)] = 3.3 g K_2HPO_4

8.) If your well-meaning but criminally clumsy Professor K. accidentally pours 10 mL of 12 M HCl into your liter of buffer, what will happen to the pH?

$$0.01\ \text{L} \times \frac{12\ \text{mol HCl}}{\text{L}} = 0.12\ \text{mol HCl} = 0.12\ \text{mol strong acid}$$

This will 'push' the equilibria to the left. The exact effect will vary with starting pH, but the moles of acid are >2× the moles of buffer, so the effect will be extreme.

Situation Two: The figure below is a form of data presentation known as a Hydropathy Plot or (after the two guys who came up with it) a Kyte-Doolittle Plot. The plot shows the level of hydrophobicity or hydrophilicity along a polypeptide, from the N-terminus to the C-terminus. Sub-regions of the polypeptide primary structure that have an overall hydrophobic nature thus show up as curves ABOVE the center-line, and those that have an overall hydrophillic nature show up as curves BELOW the center-line. Some of these sub-regions have been marked with letters.

10.) The protein for which the Hydropathy Plot is shown above is a membrane-associated protein with globular domains. In general terms, where would you expect regions A, B, and C to be in the protein structure?

These are very hydroPHOBIC regions. Therefore, they should be away from solvent, either buried in the globular protein core or buried in the fatty-acid portion of the membrane.

11.) Region D can be shown to form an alpha helix consisting of amino acids ALSEIVRNLGTQASTVQT. One of your classmates suggests that this helix crosses through the plasma membrane, a distance of ~30 Angstroms, connecting the intra- and extra-cellular domains of this protein. Based on (a) what you know about these amino acids and (b) what you know about membranes, what comments would you make about your classmate's suggestion?

This looks like an amphipathic helix, with one face likely to be composed of charged residues. As such, it seems a poor candidate for a membrane-spanning helix, which we would expect to be more like region A – hydrophic but spanned by hydrophilic regions. Size-wise, it is about right, but that won't cut it here…

12.) Comparison of sequences of this protein from a broad evolutionary range of organisms shows that the following residues are universally present: a C at position 26, an S at position 126, an H at 141, a D at 145, a P at 188, a C at 234, and a K at 299. Please indicate whether you think that the conservation of each of these is apt to indicate a role in the structure of the protein or the function of the protein. Please justify your answers. Note that "I can't be sure" is a perfectly acceptable answer if your justification is good.

It seems reasonable to suggest that:

A. **C26 and C234 are involved in a structurally-important disulfide bond**
B. **P188 is required to form a structural "kink"**
C. **S126 could be involved in a structurally important H-bond, but more likely it is involved somehow in protein function**
D. **Similarly, H141, D145, and K299 could be involved in functionally-important H-bonds or salt-bridges (ion-pairs) or they could play a role in protein function**

13.) Region B in the plot above is normally found folded into one of the domains of the protein. What net enthalpic change do you think would accompany the unfolding of this peptide out into the solvent? Would this have any meaningful impact on the net free energy of folding? [EXTRA CREDIT: Indicate another region for which your answer would be different.]

Typically, unfolding of hydrophobic regions like B is ENTHALPICALLY favorable (-ΔH) or neutral. However, this effect is small compared to the large negative impact on ENTROPY brought about by unfolding a hydrophobic region into aqueous solvent. Therefore, the ΔH effect is NOT likely to have much impact on overall free energy of folding.

14.) The peptides from Situation One and the protein from Situation Two are mixed into one solution with cytochrome C and an uncharacterized mixture of sugars by your evil lab instructor. Using protein purification techniques that we've discussed, suggest a sequence of steps that should recover the protein from Situation Two in sufficient purity for further analysis. Please be as detailed as you feel you can afford to be.

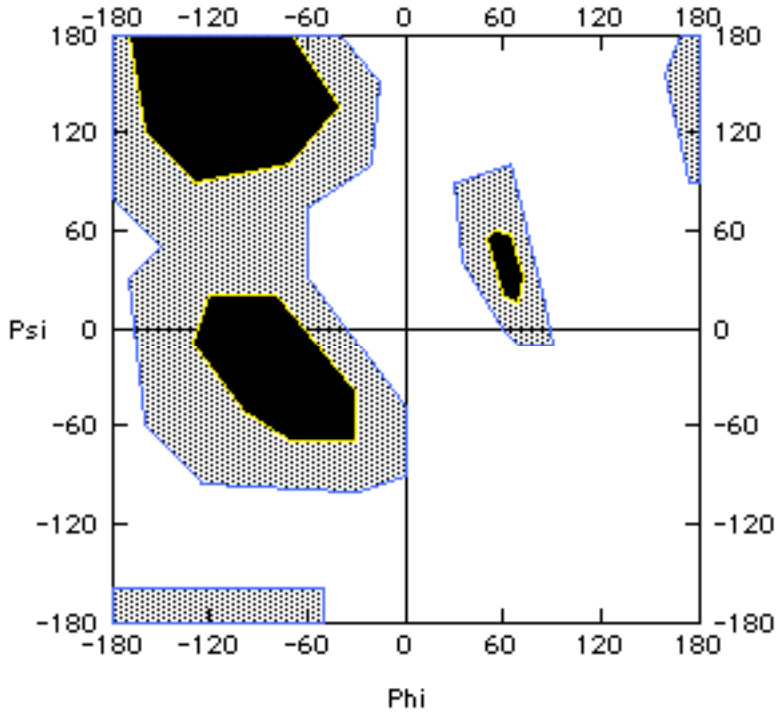
The easiest approach here (though creative answers are accepted) is to simply run a gel-filtration column like the one we ran in lab. The Sit.2 protein is ~3 times the size of cytochrome C and MUCH bigger than peptides or sugars, so it should speed through the column relative to the other components and emerge virtually pure.

15.)
a) Please draw the complete structure of the peptide WALTR. For all ionizable groups, indicate average charges that you would expect to find in solution at pH 7.5. In cases of partial (less than

90%) protonation or deprotonation, please estimate the fractional charge (an estimate accurate to within +/- 20% is fine). (9 points)

Three charged groups, term amine ~0.75+, R 1.0+, term carboxyl 1.0-

16.) (3 points each)
The figure below shows a Ramachandran plot for yeast cytochrome *c*. The shading indicates the frequency with which each Phi/Psi combination is observed in this protein -- white: not at all, stippled: low frequency, black: high frequency. Please examine the figure before answering questions (a) and (b).



a) Why is so much of the figure white?

Simply put, Steric Hindrance. As your text notes, steric interference between amino acids rules out 75% of all possible phi-psi combinations for the average polypeptide.

b) The two main darkly-shaded regions on the left half of the Ramachandran plot represent α-helices and β-sheets. What causes these two regions of Phi-Psi combinations to be disproportionately represented in a protein like cytochrome *c*?

These regions are more thermodynamically favorable because they are stabilized by regular pattens of backbone hydrogen bonding (regardless of R group identity).

c) Please briefly explain the difference between a protein ***domain*** and a protein ***subunit***.

A domain is a folded unit of tertiary structure, composed of amino acids from one polypeptide; a subunit is one polypeptide component of a protein whose parts are encoded by several genes.

d) Hydrogen bonds are plentiful in proteins, and are stronger than hydrophobic interactions. So why are hydrophobic interactions more important than hydrogen bonding in driving the process of protein folding?

Because h-bonds can form as well with water as with other parts of the protein, there is no net enthalpic advantage to the folded state relative to the unfolded. Only hydrophobic interactions -- through reduction of the unfavorable ordering of the solvent in the unfolded state -- are net favorable for folding.

e) Many integral membrane proteins have one or more "membrane-spanning" alpha helices which start on one face of the plasma membrane and extend straight through the membrane to the other side. In some cases, like the antigen receptor on T lymphocytes, the membrane is spanned by a single helix, and in others, like the β-adrenergic receptor on neurons, a protein may have several helices running back and forth across the membrane.
Envision the different environments that a membrane-spanning helix runs through. Consider the impact this will have on amino acid composition in the helix. Please **name three amino acids that you would expect to find in the mid-section of such a helix**, the part that is in the middle of the lipid bilayer.

Anything hydrophobic -- L, I, V, M, Y, W, F

f) Please **name** one amino acid that you would expect to find **at one end of the membrane-spanning segment as it passes by the phospholipid head groups**.

**Something charged -- either K or R to interact with the phosphate,
or E or D to interact with the head group.**

Folding and Stabilization of Protein Structure

1.) Please pick any five of the following eight types of conditions to which a protein might be subjected in the course of laboratory manipulations, and fill in the table to indicate their effects on the protein's structure. (4 points for each agent)

Conditions:

- incubation at 37°C
- incubation in a buffer containing 3.0 M βME
- incubation in a buffer containing SDS
- incubation at pH 3.2
- incubation at elevated temperature in *strong* acid (pH < 1)
- extraction into organic solvent, such as chloroform
- incubation in a buffer containing a high concentration (> 5 M) of a chaotropic salt
- vigorous shaking

Choice #	Condition	…will interfere with…	…by…	…and will therefore *primarily* affect this/these level(s) of structure
For example	*3.5 M NaCl*	*ionic interactions (and H-bonds)*	*competing with interactions between charged amino acids (or H-bond donors/ acceptors)*	*The most accessible amino acids will be most strongly affected, so effect on $4^o > 3^o$ (> 2^o)*
1	**3.0 M βME**	**Disulfide bonds**	**Reducing them to paired cysteines**	**3°, 4°**
2	**SDS**	**Hydrophilics, ionics, H-bonds**	**Competition with both charged and nonpolar**	**2°, 3°, 4°**
3	**Shaking**	**Hydrophobics especially**	**Energy to break weak bonds, plus exposure to air/water interface**	**3°, 4° (and some 2°)**
4	**Chaotropic Salt**	**Hydrophobics (some H-bond)**	**Unspecified mechanism**	**3°, 4° (and some 2°)**
5	**Organic Solvent**	**Hydrophobics (some H-bond)**	**Non-polar solvent/inversion of normal structure**	**3°, 4°,**
6	**Hot, strong acid**	**Peptide bonds**	**Hydrolysis**	**All**
7	**37°C**	**Ionic, H, hydrophobic**	**Thermal energy**	**4° > 3° > 2°**

8	**3.2 pH**	**Ionic, H bond**	**pK_A effect, H bonds**	**2°, 3°, 4°**

2.) Imagine an organism living in a freshwater spring that is made slightly acidic (pH ~ 5) by its high level of dissolved carbon dioxide gas. Suppose further that the cells of this organism contain a protein known as perriase. Investigation of perriase reveals that:

- it is a water-soluble, globular protein
- it is encoded by three genes (*a*, *b*, and *c*)
- the gene products are arranged in a quaternary structure that could be described as a dimer of heterotrimers: $(\alpha\beta\gamma)_2$ (Please ask me for clarification if this is not clear!)
- the alpha subunit has a molar mass of 24 kDa and a secondary structure consisting entirely of beta-pleated sheets (and connecting turns)
- the beta subunit has a mass of 79 kDa and is associated with alpha through a series of hydrogen bonds and a single disulfide bond
- the gamma subunit has a molar mass of 47 kDa and associates with beta through a series of ion-pair interactions (often called "salt bridges"). The ionic interactions depend on histidines of the beta subunit interacting with glutamic acid residues on the gamma subunit.

a) Please fill out the following table to indicate what you predict would happen if equal aliquots of perriase were dissolved in four tubes at 4°C. In the "difference(s)" column, please specify differences **both** in terms of the **interactions** affected and the **level of protein structure** affected. (4 points per "tube")

Tube Number	**Contains a solution of…**	**Difference(s) in protein structure relative to Reference Tube**	**What polypeptides will be associated?**
Reference	100 mM acetate buffer, pH 5.0	-	300 kDa
1	butanol	**Will disrupt hydrophobic interactions and denature all subunits; 3° and 4°**	**24 47 79 (disulfide bonded = 103 kDa)**
2 2 cont.	100 mM sodium phosphate, pH 7.0	**Will deprotonate histidines and possibly affect H-bonds; potentially 2°, 3°, and 4°**	**103 αβ 47 γ Dimers? unlikely**
3	3.0 M βMercaptoethanol in 100 mM acetate buffer, pH 5.0	**Will reduce disulfides; will hurt 4° and maybe 3°**	**24, 79 (unlikely to stick sans disulfide) 79, 47 (still stick = 126) Dimers? maybe**
4	0.1 M phosphate buffer at pH 2.0, with 1.0 M dissolved potassium chloride	**H^+ will protonate glutamates, possibly other key AA as well, and salt will mess up H-bonds and ionics; 2°, 3°, and 4°**	**24+79=103 47**

b) Please sketch cartoon images of the structure of the perriase ***alpha subunits*** as they would appear in the reference tube above, and in a sample which had been dissolved in a buffer containing a strong detergent and 2-mercaptoethanol and then *briefly* boiled.
(3 points per sketch)

Reference tube (acetate buffer, pH 5.0)	**Boiled in detergent and βME**

4.) Please pick any three of the following eight types of conditions to which a protein might be subjected in the course of laboratory manipulations, and fill in the table to indicate their effects on the protein's structure. (4 points for each agent)

Conditions:

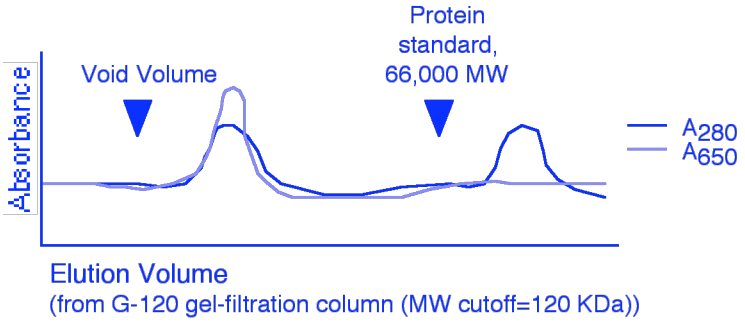
- incubation at 37°C
- incubation in a buffer containing 3.0 M βME
- incubation in a buffer containing SDS
- incubation at pH 3.2
- incubation at elevated temperature in *strong* acid (pH < 1)
- extraction into organic solvent, such as chloroform
- incubation in a buffer containing a high concentration (> 5 M) of a chaotropic salt
- vigorous shaking

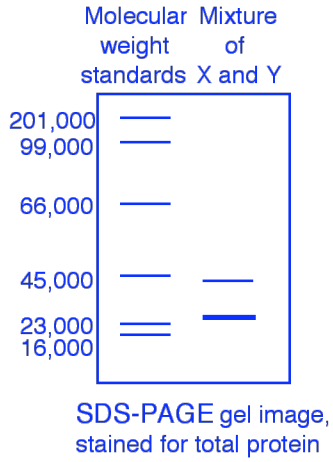
Choice #	**Condition**	**…will interfere with…**	**…by…**	**…and will therefore *primarily* affect this/these level(s) of structure**
For example	*3.5 M NaCl*	*ionic interactions (and H-bonds)*	*competing with interactions between charged amino acids (or H-bond donors/ acceptors)*	*The most accessible amino acids will be most strongly affected, so effect on $4^o > 3^o$ (> 2^o)*
	37 C	**H-bonds, hydrophobic, dispersion forces**	**jiggling them apart**	**4^o, 3^o, 2^o**
	chaotropic salts	**hydrophobics H-bonds**	**poorly understood**	**4^o, 3^o (2^o)**
	shaking	**hydrophobics**	**creation of non-aqueous air pockets**	**4^o, 3^o**
	BME	**disulfides**	**reducing -S-S- to -SH HS-**	**4^o, 3^o**
	SDS	**(ionics), H-bonds, hydrophobics**	**competition**	**4^o, 3^o, 2^o**
	low pH	**ionics, H-bonds**	**altering protonation state, competition**	**4^o, 3^o (2^o)**
	strong acid	**peptide bonds (and others via pH and heat)**	**hydrolysis**	**all**
	chloroform	**hydrophobics**	**competition**	**4^o, 3^o (2^o)**

Application: Protein Purification Techniques

1.) Bonnie the Biochemist is trying to purify two proteins from a bacterial cell culture. After several steps, she has succeeded in isolating a mixture of two proteins: Protein X and protein Y, a bacterial photoreceptor protein. Protein Y is covalently linked to an organic ring structure that strongly absorbs light in the orange region of the spectrum (thus, the protein is green). The pH at which each protein has a net charge of zero is pH 7.8 for protein X and 6.7 for protein Y. Both of these proteins are unusually stable over a wide range of pHs, from pH 2 to pH 12.

a) If Bonnie presented you with the following data from gel filtration and SDS-PAGE (under denaturing conditions), what conclusions could you reach concerning the molecular weight and subunit composition of each protein? (I'm sure she'd want justification for your hypothesis!)





Gel Filtration:

- **Gel filtration separates by size, with large things eluting first. Therefore, Peak #1 is larger than 66 kDa, and Peak #2 is smaller than 66 kDa**
- **Both peaks contain protein, as indicated by the A_{280}, and Peak #1 contains something that absorbs orange light (@ 650 nm).**
- **Thus, Peak #1 must be Y and Peak #2, X**

SDS-PAGE:

- **There is no protein > 66 kDa, so Y must have multiple subunits**
- **The band @ ~ 40 kDa is probably protein X (a single subunit)**
- **The strong band @ ~ 25 kDa represents multiple subunits of about the same size**

There are several ways to interpret the overall data, but the simplest is that...

- **X is a single-subunit protein of ~ 40 kDa**
- **Y is a homotetramer of overall mass ~100 kDa. Hence, alone the subunits run at ~25 kDa, but together they give a size about right for the gel filtration data**

b) If Bonnie was displeased with the gel filtration results and proposed instead to separate these two proteins by ion-exchange chromatography on CM-agarose at pH 10, what advice could you

give her about her choice of pH for this resin? If you feel another pH would be more effective, please indicate both what pH that would be (if you have a range in mind, pick a point in the middle of that range), and why. If you feel that her choice is sound, please indicate why. (6 points)

CM-agarose is for cation exchange. It is negatively charged at pH above ~ pH 4 (where we'd expect it to be ~ ½ deprotonated). Thus, @ pH 10, it should be entirely deprotonated and useful for separating positively-charged species. At pH 10, however, we are above the pI for both X and Y, and so they will each have net negative charge. Thus, pH 10 with CM-agarose is unlikely to work. Better would be to choose a pH at which one protein is positively charged and one is neutral or negative. That is the case at pH 6.7, pH 7, or anything in that immediate pH neighborhood.

3.) After listening to your advice in problem #2, Bonnie the Biochem Goddess is ready to carry out the ion-exchange chromatography experiment, but she needs to mix up a liter of buffer first, at a concentration of 100 mM. Being the geek/goddess that she is, she pulls out her handy pocket buffer chart, reproduced below.

Compound	**pK**
Oxalic acid	1.27
H_3PO_4	2.15
Citric acid	3.13
Formic acid	3.75
Succinic acid	4.21
$Oxalate^{-1}$	4.27
Acetic acid	4.76
$Succinate^{-1}$	5.64
MES	6.09
Carbonic acid	6.35
PIPES	6.76
$H_2PO_4^-$	6.82
MOPS	7.15
HEPES	7.47
Tricine	8.05
Tris	8.08
Glycylglycine	8.25
Bicine	8.26
Boric acid	9.24
$Bicarbonate^{-1}$	10.33
Piperidine	11.12
HPO_4^{-2}	12.38

(data from Dawson et al (1986), *Data for Biochemical Research* (3rd ed), pp. 424-425)

Acceptable choices lie within ± 1 pH unit of target:

PIPES and $H_2PO_4^-$ best for pH 6.7

$H_2PO_4^-$ and MOPS best for pH 7.0

$Bicarbonate^{-1}$ best for pH 10

[If you have not answered #2b, then use the default pH of 10 for this question; if you suggested a different pH, then use that pH.]

a) Which of these compounds would be reasonable choices for her to use for her buffer? Why? Assume that Bonnie has access to all of these compounds in her stockroom. (7 points)

For a buffer to be effective, it must consist of components with a pK_A near the target pH of the buffer. See above.

b) Being the wonderful soul that you are, you offer to make Bonnie's buffer for her. Please describe how you would do this, using only the compounds listed in the chart plus NaOH (5 M solution) and HCl (6M solution). Do not worry about the molar masses; simply give me mole quantities. (15 points)

As indicated at the time of the test, adding any of these components involves adding the protonated form – i.e., the conjugate acid. Thus, to use any of these for buffer construction, some NaOH will need to be added to convert some of the acid to the conjugate base. For example, using PIPES at pH 6.7: (and calling, for convenience, the protonated form "H-PIPES")

[PIPES] + [HPIPES] = 0.1 *M* → 1.871 [HPIPES] = 0.1
[HPIPES] = 0.0534 *M* and [PIPES] = 0.0466 *M* → STOP! Is this right? Base < Acid, so pH should be < pK_A – check!

Therefore, to make one liter of this buffer, I should start with 0.100 moles of protonated PIPES, which I will dissolve in 800 mL of water. I will then add 0.0466 moles of NaOH (9.32 mL at 5 *M*) to forcibly deprotonate 0.0466 moles of HPIPES, leaving 0.0534 moles protonated and 0.0466 moles deprotonated. I will then complete the job by topping off (up to 1.0 L) with water.

4.) The jellyfish *Aequorea victoria* glows green as the result of a pretty remarkable chain of biochemical events. In the first, the protein aequorin converts chemical energy into blue light with a maximal intensity at 470 nm. This light is then absorbed by a second protein known simply as GFP, ***g***reen ***f***luorescent ***p***rotein. Aequorin is a 21 kDa protein with a pI of 4.5; it is dependent on an attached non-protein chromophore for blue light generation. GFP is a 27 kDa protein with a pI of 6.2. Its ability to fluoresce comes from a remarkable reaction between three amino acids that occurs spontaneously when they are folded into the right conformation. GFP strongly absorbs light at 395 nm and 475 nm. The latter absorbance is probably the biologically important one, allowing the jellyfish to convert the blue fluorescence of aequorin into the green glow of GFP (green seems to be evolutionarily preferred, though it's not clear why). Aequorin loses the ability to chemiluminesce if treated with acid or ether; GFP denatures at pH above 11.

a) Please describe to me what would happen if you ran a solution containing a mixture of aequorin and GFP over the gel filtration column that you poured in lab. Accompany your narrative description with a diagram(s) of the data you would expect to obtain from this experiment, including data that confirm the successful separation of the two proteins. Be sure to label your diagram(s) so that interpretation is unambiguous (e.g., what are your axes? if there are multiple curves, what do they represent?). (10 points)

GFP is 27 kDa; aequorin is 21 kDa. Therefore they should be resolved by your gel filtration column. Monitoring absorbance at 280 nm should show two peaks; the first will be GFP and the second will be aequorin. To confirm these identifications, the first peak should absorb light at 395 and 475 nm as well as 280, but the second should only absorb at 280 (it also emits light at 470 nm).

b) Please describe an alternative method of separating these two proteins in a way that preserves their function and would be suitable for work at the preparative scale. You may assume that you have access to

- reagents for making buffers over a wide pH range
- a broad selection of gel filtration resins
- carboxymethyl-cellulose (pKa of 3.8)
- diethylaminoethyl-cellulose (pKa of 9.5)
- reagents and apparatus for 1- and 2-dimensional PAGE
- reagents and apparatus for isoelectric focusing
- reagents and apparatus for low- and high-speed centrifugation
- reagents and apparatus for the Bio-Rad protein assay
- spectrophotometers, pH meters, and other common biochemistry lab equipment

As above, provide sketches of key data, and tell me how you would confirm that the proteins have been separated. (10 points)

The best way (though there are many possibilities) would probably be to use ion-exchange chromatography. At pH 5, using DEAE, aequorin will be negatively charged and will stick to the column, but GFP will be positively charged and will flow through. These identifications could be confirmed by spectrophotometry (as in part a), or by SDS-PAGE, or...

c) Suppose that a fraction of a cell extract is passed over a gel-filtration column. The second and seventh protein-containing fractions to come off the column look interesting and so are analyzed by SDS-PAGE. Each fraction is shown to contain only a single polypeptide when the gel is stained with Coomassie blue. **Which fraction** -- #2 or #7 -- should give a **band closer to the bottom of the gel**? (3 points)

#2 elutes off the gel-filtration column before #7, and is therefore larger. That means that #7 should be closer to the bottom of the gel in SDS-PAGE because it is smaller.