Answer each of the following questions in the space provided; circle the correct answer or answers for each multiple choice question and circle either ‘True’ or ‘False’ to indicate whether each true/false statement is correct or not. (70 points)

Draw a di-peptide (any two amino acids) and label the peptide bond that links the two amino acids. (1 point)

Example:

![Di-peptide Example]

The following amino acids are found within a globular cytoplasmic protein. Identify the amino acids whose side chains would be expected to be on the outside of the globular protein by labeling them with an “O”. Identify the amino acids whose side chains would be expected to be clustered on the inside of the globular protein by labeling them with an “I”. Finally, label the ‘amino terminus’ listed as item h below with either an “O” or an “I” to indicate whether you believe this part of the polypeptide chain would be located on the exterior or interior (respectively) of the globular protein. (4 points)

a.) Tyrosine—O
b.) Alanine—I
c.) Valine—I
d.) Proline—I
e.) Threonine—O
f.) Tryptophan—I
g.) Aspartic Acid—O
h.) Amino terminus—O

True / False. The two polypeptides below will have the same shape. (1 point)

1. N-Ala-Val-Phe-Gly-Lys-Lys-Trp-Ile-Met-Ser-Cys-Gly-Asp-Gln-Glu-Phe-His-Ile-Leu-C

Although they contain the same combination of amino acids (although in a reversed order), they will NOT fold similarly and will be shaped differently.
One type of covalent bond that is common in maintaining protein structure is a **disulfide** bond. These bonds contribute to a protein’s **3°/4°** (1°/2°/3°/4°) structure. (2 points)

Answer each of the following True / False questions with respect to the following amino acid sequence: (4 points total; 1 point each)

\[
\text{N-Ala-Glu-Val-Tyr-Cys-Lys-Leu-Tyr-Ile-Ser-Met-Asn-Cys-Asp-Gly-Phe-Gln-Leu-His-C}
\]

Use the helix wheel to the left if necessary.

**True / False.** The amino acid sequence could form an \(\alpha\) helix.

**True / False.** The amino acid sequence below could form an amphipathic \(\alpha\) helix.

**True / False.** The amino acid sequence could be a strand in an amphipathic \(\beta\) sheet.

I messed up near the end of the amino acid sequence such that two hydrophobic amino acids (Gly and Phe) are adjacent to one another. Aside from this, the answer is true, but either answer is acceptable due to my error.

**True / False.** The amino acid sequence could contain an intrachain disulfide bond.

Answer each of the following True / False questions with respect to the following amino acid sequence: (4 points total; 1 point each) Use the helix wheel above if necessary.

\[
\text{N-Ala-Gly-Val-Met-Cys-Leu-Leu-Trp-Ile-Pro-Met-Val-Cys-Ala-Gly-Phe-Leu-Val-C}
\]

**True / False.** The amino acid sequence could form an \(\alpha\) helix.

**True / False.** The amino acid sequence below could form a hydrophobic \(\alpha\) helix.

**True / False.** The amino acid sequence could be a strand in a hydrophobic \(\beta\) sheet.

**True / False.** The amino acid sequence could contain an interchain disulfide bond.
You wish to produce a human enzyme, protein A, by introducing its gene into bacteria. The genetically engineered bacteria make large amounts of protein A, but it is in the form of an insoluble aggregate with no enzymatic activity. Which of the following procedures might help you to obtain soluble, enzymatically active protein? (1 point)

a.) Make the bacteria synthesize protein A at a slower rate and in smaller amounts.

b.) Dissolve the protein aggregate in urea, then dilute the solution and gradually remove the urea.

c.) Treat the insoluble aggregate with a protease.

d.) Heat the protein aggregate to denature all proteins, then cool the mixture.

I meant to add ‘not’ to the question, but inadvertently left the word out, so all answers EXCEPT c could produce a soluble protein that is functional.

The activity of many proteins is dependent upon molecular binding. How is it possible that binding to a small molecule, for example cAMP, can alter the ability of a protein to function? (2 points)

Shape change. Binding of the small molecule alters a protein’s shape and since shape is related to function, binding alters function.

True / False. Binding to cAMP can result in protein inactivation. (1 point)

Protein Y has been activated by binding to cAMP. How can (or when will) protein Y be inactivated? (2 points)

When cAMP concentrations are low enough such that it is unlikely to be bound to protein Y. This occurs when less cAMP is produced or cAMP is destroyed at a more rapid rate.
Which of the following protein pairs binds the most tightly to each other? (1 point)

a.) Protein A & Protein B: kD = 55nM
b.) Protein B and Protein C: kD = 4nM
c.) Protein C and Protein D: kD = 0.2\mu M
d.) Protein C and Protein A: kD = 100\mu M
e.) Protein D and protein A: kD = 0.1mM

In the question above, briefly describe the meaning of the term, “kD”? (2 points)
kD is a number that represents relative affinity between two molecules. It actually is the concentration at which $\frac{1}{2}$ of the available protein is bound to the other; the other half of the available protein is free in solution at the kD concentration.

A friend tells you that she has just discovered that the protein responsible for causing dogs to chase cars is a member of the MAP kinase family of protein kinases. In response to your blank stare, she adds that the yeast protein Ste7p, which is involved in response to a yeast hormone, is also a MAP kinase family member. Although you still have no idea of what either a MAP kinase or Ste7p is (but you’ll learn soon enough), which of the following statements can you safely predict to be true? (1 point)

a.) The dog protein and Ste7p have mostly similar amino acid sequence throughout their amino acid sequences.
b.) The dog protein and Ste7p catalyze the transfer of a phosphate group to another molecule.
c.) The dog protein phosphorylates the dog homologue of the yeast substrate of Ste7p.
d.) The dog protein and Ste7p are both involved in response to hormones.

What three amino acid side chains can be modified by the addition of a phosphate group? What family of enzymes is responsible for reversing this addition? Two of the three amino acids can be substrates of the same enzyme while the third requires a separate enzyme to act upon it. Hypothesize which of the three side chains requires its own enzyme. (5 points)

Tyrosine, Threonine and Serine are the amino acid side chains to which a phosphate can be added. The enzymes that reverse this are known as phosphatases. Because the shape of the Tyrosine side chain is very different than that of both serine and threonine (a ring versus a short chain, respectively), it’s likely that the kinase responsible for phosphorylating tyrosine is different from the kinase that phosphorylates both serine and threonine.
Some of the enzymes that oxidize sugars to yield useable cellular energy (for example, ATP) are regulated by phosphorylation. For these enzymes, would you expect the inactive form to be the phosphorylated form or the dephosphorylated form? Explain your answer.
(3 points)
It is likely that the enzyme is inactive when it is phosphorylated and active when de-phosphorylated. This is because when ATP is sparse, it would be necessary that the enzyme still be active such that more ATP could be made. Since ATP is the source of the phosphate that is transferred to proteins, low concentrations of ATP would result in a decreased ability to phosphorylate things. Thus, we would hope that an enzyme necessary for the production of ATP would be turned off (via transfer of a phosphate) when ATP is plentiful and turned on (in a dephosphorylated form) when ATP concentrations are low.

G proteins: (1 point)
 a.) form a covalent bond with guanine nucleotides.
 b.) are generally activated by factors that increase their rate of GTP hydrolysis.
 c.) immediately release the GDP produced by GTP hydrolysis.
 d.) “reset” themselves by phosphorylating bound GDP.
 e.) do not rapidly exchange bound GDP for GTP unless stimulated to do so by exchange factors.

Complete the following statement (2 points):
Proteins are typically categorized into families based upon:
Either similar function or similar structure/shape.

You want to amplify the DNA between the two stretches of sequence shown below. Of the listed primers, choose the pair that will allow you to amplify the DNA by PCR. (2 points)

```
5’-GACC TGT GGAAGC-------------------------------CATACGGGATTGA-3’
3’-C TGGACAC CT T CG-------------------------------GTATGCCCTAACT-5’
```

 a.) 5’-GACCTGTGGAAGC-3’
 b.) 5-CTGGACACCTTCG-3’
 d.) 5’-GTATGCCCTAACT-3’
 e.) 5’-TGTTAGGGCATAAC-3’
 f.) 5’-TCAATCCCGGTATG-3’
 g.) 5’-CGAAGGTGTCCAG-3’
 h.) 5’-GCTTCCACAGGTC-3’
 i.) 5’-CATACGGGATTGA-3’

~ 5 ~
The automated system, which utilizes color-labeled nucleotides for DNA sequencing, is malfunctioning; a call from the sequencing service company informed you that it would be two weeks before it was operational. Rather than wait for your results, they offered to perform the sequencing reaction the ‘old fashioned way’, using radioactively-labeled nucleotides instead—with the caveat that YOU must read the sequencing gel. A diagram of the gel is below. Each of the ‘rows’ represents a single nucleotide, so if two rows are stacked adjacent to one another, read them as two separate nucleotides. With this information, what are the first 51 nucleotides of the sequence this reaction obtained? What are the first 12 nucleotides of the template that was used to produce this product? (4 points)

The first 51 nucleotides of the sequence obtained are read from bottom → top as follows: 5’-ATGGCAGCACTCGAGTTAACTCATACAATAAAGGCTA GTAAGATCAATGTC-3’

The first 12 nucleotides of the template” can be interpreted any number of ways dependent upon how one interpreted “first”. Thus, the following are some possible options:

- 5’-GACATTTGATCTT-3’
- 3’-TTCTAGTTACAG-5’
- 5’-GACTCGTGCCAT-3’
- 3’-TACCGTGCTCAG-5’
- 5’-CTCGACCTCGAT-3’
- 3’-GAGCTGGAGCTA-5’
True / False. Regarding the sequence obtained in the previous question, this carries the same sequence as the coding strand that will be produced during transcription. (1 point)

In addition to deoxyribonucleoside triphosphates (dNTPs), what components are necessary to include in the tube when performing a DNA sequencing reaction? (4 points)
- ddNTPs (di-deoxynucleotides)
- DNA polymerase (usually Taq)
- Two primers
- A template strand of DNA

Design a DNA ‘probe’ that could be used to detect the DNA produced in the sequencing gel when performing a Southern blot. Assume that, to be effective, your probe would have to contain at least 60 total hydrogen bonds. (2 points)
Any sequence of DNA that is represented in the image shown on the previous page, as well as any sequence that is reverse-complementary to it, suffices, as long as 60 hydrogen bonds are contained within it. Most students were able to produce a sequence with 60 total hydrogen bonds (3 for each G/C and 2 for each A/T), but it was unclear whether the sequences provided as answers were contained within (or were reverse-complementary to) the sequence depicted on the previous page. Although I attempted to locate these sequences, in many cases I was unable to do so; in this case, the answer was marked as incorrect. However, if you feel that the sequence you provided as an answer was, in fact, contained within the sequence depicted on the previous page and/or was complementary to it, please submit your test for regarding and indicate specifically to which sequence on the image your probe would bind.

Which of the following is a limitation on the use of PCR to detect and isolate genes? (1 point)
- a.) The sequence at the beginning and end of the DNA to be amplified must be known.
- b.) It also produces large numbers of copies of sequences beyond the 5’ or 3’ end of the desired sequence.
- c.) It cannot be used to amplify a particular sequence from a mixture of mRNAs.
- d.) It cannot be used to amplify cDNAs.
- e.) It will amplify only sequences present in multiple copies in the DNA sample.
For each of the 5 situations described below, identify which of the molecular techniques discussed in class would be most appropriately used. Explain how you would use the technique you name to complete the task you are given in each case. (15 total points; 3 points each scenario)

1. A gene that is known to be expressed in normal liver tissue has been found to be over-expressed in certain liver cancers. You are tasked with determining which of the six liver tissue samples you are given is potentially cancerous.

   RT-PCR, qRT-PCR or Northern Blotting. It is necessary to convert the mRNA of each of the six liver tissue samples to cDNA via reverse transcription using the reverse transcriptase with a poly-T primer. Following this, specific primers for this cancer-causing gene must be used with PCR to amplify it. When these PCR products are run on a gel, it’s possible to distinguish relative amounts of mRNA (comparing the six liver tissue samples) by the intensity/darkness of the band acquired. The darker the band, the more mRNA was present in the original tissue sample. Similarly, Northern Blotting could also be used by isolating the RNA, separating it by electrophoresis, transferring it to nitrocellulose, incubating the nitrocellulose with a labeled probe that is complementary to the mRNA sequence for this liver cancer-causing gene. The greater the expression, the more mRNA—the greater the mRNA the greater the amount of probe that would be bound. Again, a darker band would be detected when more mRNA (i.e., expression) was present.

2. You are tasked with determining which other genes (besides the one in scenario 1 above) might be up-regulated (i.e., more highly-expressed) in liver cancer.

   DNA Microarray. This is the only technique through which relative expression levels of multiple genes is able to be simultaneously assessed. To do this, RNA would have to be purified from normal liver tissue as well as liver cancer tissue. Each liver RNA sample would have to be color-labeled (with a different color) and then the RNA samples would be mixed. The mixed RNA samples would then be incubated with a microarray slide (on which DNA samples from many different genes had been ‘dotted’). It would be washed and analyzed. When analyzing the color intensities of each of the microarray dots, the color that was associated with the normal tissue RNA would be more prominent when normal tissue had a higher expression of the gene that was encoded by the ‘dot’. Conversely, the color that was associated with the cancer tissue RNA would be more prominent when the cancer tissue had higher expression of the gene and the color would be yellow, indicating equal expression.

3. You are tasked with determining where else in the body this putative “liver cancer” gene is expressed.

   RT-PCR, Northern Blotting, in situ hybridization, or Western Blotting. See above for RT-PCR and/or Northern Blotting, though the samples being compared would be those isolated from various body tissues. For in situ hybridization, tissue sample sections (for microscopy) would be incubated with labeled DNA ‘probes’ complementary to the gene. Wherever the gene was being expressed, the probe would be identified in the cytoplasm (it would normally be in the nucleus because all of the cells’ DNA would contain the gene). If more of the label (color, radioactivity) were identified in one tissue relative to another, that tissue would be expressing the gene at a higher level than that of the other samples. For Western Blotting, rather than assessing RNA, the amount of protein in each of several tissue samples would be compared by subjecting each tissue protein to gel electrophoresis (SDS-PAGE), transferring the protein to nitrocellulose, incubating the nitrocellulose with an antibody toward the liver cancer-causing protein (after ‘blocking’
background binding with milk proteins), possibly incubating the nitrocellulose with a tagged secondary antibody and ‘developing’ the blot using some sort of color or chemical reaction.

4. Although expression levels between individuals may vary, you are tasked with determining whether all individuals have the same number of gene copies in their genomes. Southern Blotting, Q-PCR or FISH on chromosomal DNA. DNA or chromosomes from multiple individuals would have to be acquired. For Southern Blotting, the DNA would be subjected to electrophoresis and the DNA would be transferred to nitrocellulose. The nitrocellulose would be incubated with a labeled DNA probe complementary to the liver cancer-causing gene; the presence of the probe would be detected by some chemical reaction or development of radioactivity to determine how much of the probe had bound. If more DNA were present in the one or more of the original samples (i.e., more copies were present), more probe would be detected. Alternatively, for q-PCR, the DNA would be subjected to PCR using primers that were specific to the gene. Following amplification, the PCR products would be subjected to electrophoresis and the amount of DNA would be evaluated based upon the darkness/intensity of the band. If more copies of the DNA were present in the initial sample, the resulting PCR band would be more intense, indicating more copies of the gene. Finally, for FISH, chromosomes would be incubated with a labeled DNA probe and the label would be detected by some chemical reaction or development of radioactivity to determine how much of the probe had bound. If more DNA were present in the one or more of the original samples (i.e., more copies were present), more probe would be detected.

5. The protein encoded by this “liver cancer” gene is secreted from the cell and is known to interact with several other extracellular proteins. You are tasked with identifying these other proteins. Far Western Blotting. The protein encoded by the liver cancer-causing gene would be purified and labeled. Subsequently, extracellular proteins from liver samples would be acquired and subjected to SDS-PAGE. The separated proteins would be transferred to nitrocellulose and the nitrocellulose would then be incubated with the labeled protein. The nitrocellulose would be washed and the labeled protein would be detected using an antibody or some chemical reaction or development of radioactivity. The labeled protein would be bound to a region on the nitrocellulose at which a potential binding partner would be located. This would help identify IF such a binding partner existed and, if so, the approximate molecular weight of the binding partner.

You have just read about a transgenic mouse in which a mutant version of the gene for elastin (one in which a glycine amino acid was replaced with an asparagine) is cloned into a random site in the mouse genome. For what scientific question(s) might the phenotype of this mouse provide an answer? (3 points)
Possibilities include determining what effect this mutation might have on the assembly and function of elastic matrices. Recall from class that elastin protein assembles in the extracellular space to produce a matrix that, when stretched, exposes hydrophobic side chains to a polar aqueous environment. Thus, if one adds a version of elastin that is not completely hydrophobic, it might assemble with the normal (still present) version of elastin to disrupt the function of the elastic matrix. But, this disruption has to be to the whole matrix because elastin proteins do not function independently but rather in a complex with other elastin proteins.

The transgenic mouse described in the question above is an example of: gene deletion / gene replacement / gene addition (circle one; 1 point).
For three bonus points: Hypothesize what you might observe in the transgenic mouse described above and provide an explanation to support your hypothesis.

The mutated elastin gene will exist in this mouse along with two copies of the normal elastin gene, but we don’t know if all of the elastin protein that is produced will be functional. We also don’t know if the mutant version of the elastin protein will interfere with the function of the normal elastin protein, since the two versions of the protein will be mixed together in all elastic matrices. IF the mutant elastin does, in fact, interfere with the normal function of elastin, one might expect that elastic matrices will not function as well—thus, tissues that are supposed to be stretchy and recoil (lungs, skin, blood vessels) may not work (e.g., may not be as stretchy or may not recoil as well). IF the mutant elastin protein functions just like the normal elastin, having 3 copies of the gene may result in production of MORE elastin—this might result in elastin-containing tissues being MORE stretchy than normal.