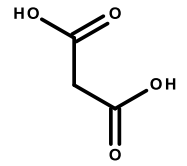


This exam has **14 pages** and is out of **150 points**. You should allot 1 min/point. On questions with choices all of your attempts will be graded and you will receive the best grade. Use the space provided, or the back of the preceding page.

1. (5 pts) Maleic acid contains two carboxylic acid groups, one with a pK_a of 2.0 and a second with a pK_a of 4.0. The fully protonated form of maleic acid is shown on the right.

i) Why do the two pK_a values differ? (1 pt).

ii) Briefly describe how you would prepare 1 L of a 0.1 M buffer at pH=2.0, assuming that you are starting with the disodium salt of the acid (4 pts).



2. (3 pts) All weak acids have “buffer” regions near any of their pK_a values. Briefly explain why the pH of the solution is resistant to change in these regions.

3. (7 pts) Most proteins generally consist of secondary structural elements.

i) Name the **two** common secondary structures. (2 pt)

ii) Describe the overall structure of **one** of these, including the position of sidechains (3 pts).

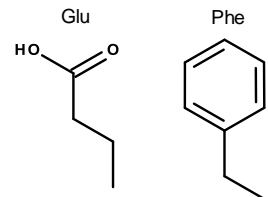
iii) What is the principle interaction that stabilizes both of these structures? Briefly describe how the interaction is related to the overall geometry of the secondary structure that you have chosen (2 pts).

4. (14 pts) Draw the chemical structure of a tri-peptide, i.e. three amino acids linked together. The sidechain of the first amino acid is a methyl group ($-\text{CH}_3$), the second is just a hydrogen atom, and the third is an isopropyl group ($\text{CH}_3\text{-CH-CH}_3$) (4 pts).
- Label a peptide bond in your drawing and indicate whether it is drawn in the cis or trans form. (2 pts)
 - Which form of the peptide bond is more stable, cis or trans, and why? (1 pt)
 - Label the amino and carboxy terminus of the protein (1 pt).
 - Give the names of the amino acids that you have drawn and write out the sequence of the protein (1 pt)
 - For the middle residue, indicate which bonds are freely rotatable and which are not. If any bonds are not rotatable, please explain why. (Hint: The peptide bond should have been one of the bonds you selected). (3 pts)
 - Sketch the energy contours that you would expect to find on the Ramachandran plot for both the first and second residue. Do you expect them to be different or the same, why? (2 pts).

5. (4 pts) The following amino acid sequence is found in a soluble globular protein:

Glu-Phe-Glu-Phe-Glu-Phe

- What is the likely secondary structure of this sequence? Why? (2 pts)
- Where do you expect the Phe residues to face: the interior of the protein, or to the solvent? Why? (1 pt)
- If this was an integral membrane protein, where would you expect to find the Phe residues? Why? (1 pt)



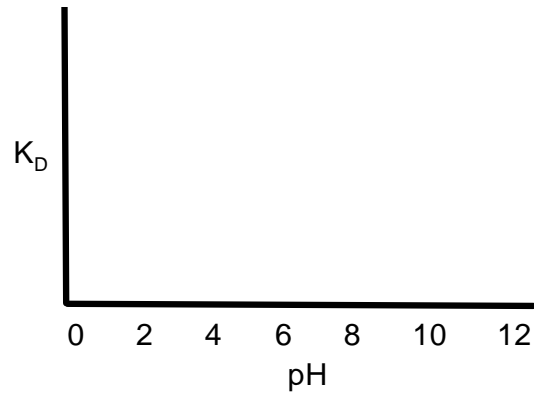
6. (8 pts) A DNA binding protein binds to DNA in a non-specific manner. The protein contains lysine (-NH₂) and glutamic acid (-COOH) residues, **both** of which are close to the DNA.

i) Which residues on the protein are most likely involved in stabilizing the DNA-protein complex, the lysine or glutamic acid residues? (1 pt)

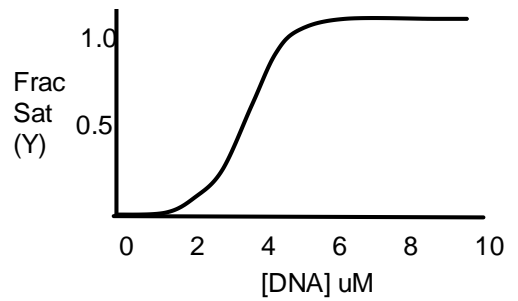
ii) What group on the DNA would this residue most likely interact with? (1 pt)

iii) Assume that you measured the dissociation constant (K_D) as a function of pH. Plot the K_D as a function of pH over the range of 0 to 12. Briefly justify your answer and indicate the pKa values you assumed for both Lys and Glu. Be sure to consider the role of both the Lys and Glu residues in your answer (4 pts).

iv) In general, does the K_D increase or decrease as the interactions between proteins and their ligands become more favorable? Why? (2 pts)



7. (7 pts) The binding curve for the same protein in the previous question is shown on the right. The protein has a molecular weight of 100 kDa protein binds a short DNA oligonucleotide (say 10 bases, 6 kDa). The binding curve is shown on the right. Please answer the following questions:



- i) What is the K_D for binding? (1 pt)
- ii) Based on the shape of the binding curve, how many DNA molecules are likely binding to the protein, one or more? Justify your answer. (2 pts)
- iii) What additional plot might you do to confirm your hypothesis to part ii? How would this plot be used to confirm your hypothesis? (2 pts)
- iv) How would you expect this curve to shift if the salt concentration was increased? Why? (2 pts)

Bonus 1: How could you determine the binding constant using gel filtration (size exclusion) chromatography? (1 pt)

8. (6 pts) Please do **one** of the following choices. Be sure to discuss **all** the important enthalpic **and** entropic considerations in your answers.

Choice A: A mutation in a protein converts a buried phenylalanine residue to a threonine residue (Threonine sidechain is $-C-C-OH$). How will this affect the stability of the protein?

Choice B: A mutation in a DNA sequence converts a G-C base pair to an AT base pair. How will this affect the stability of the double stranded (ds) form of DNA? Will it make it more stable or less stable? Why?

9. (4 pts) Please do **one** of the following choices. Do all parts within a choice.

Choice A:

- i) Describe or draw any phospholipid (1 pt)
- ii) Explain why corn oil is a liquid while corn oil margarine is a solid. (3 pts)

Choice B:

- i) What is the general structure of cholesterol? (1 pt)
- ii) How does cholesterol affect the phase transition (melting) of lipid bilayers? (1 pt)
- iii) Why is this effect important for the normal function of biological membranes? (2 pts)

10. (8 pts) Please answer the following questions on enzymatic activity. You can use an example from class to illustrate your answer, but it is not absolutely necessary to give specific details about any particular enzyme.

- i) Why are enzymes specific for particular substrates and what is the relationship between k_{off} , k_{CAT} for good and bad substrates? (1 pt).
- ii) How do enzymes increase the rate of catalysis? Provide a general principle that holds for all enzymes. (6 pts).
- iii) Why are transport proteins (e.g. K^+ channel) considered (at least by me) to be enzymes, even though they do not change the chemical structure of the substrate? (1 pt)

11. (6 pts) Please do **both** parts of this question:

- i) Briefly describe the most important characteristics of allosteric systems, including activators and inhibitors (4 pts).
- ii) Illustrate your answer with **one** of the following topics from the course: a) Oxygen delivery, b) altitude adjustment, c) enzyme inhibitors (one specific type), d) metabolic regulation (glycogen or glycolysis), or e) regulation of DNA transcription (2 pt).

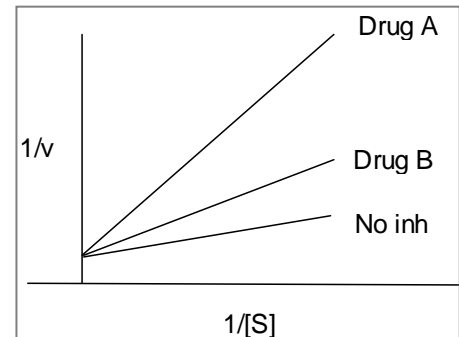
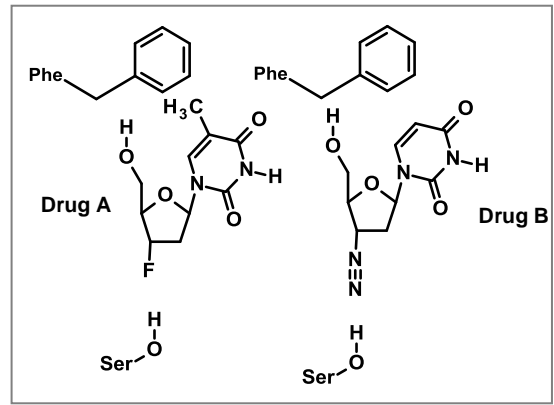
12. (3 pts) You are trying to purify a protein from bacteria. You have inserted the DNA that codes for the amino acid sequence into an expression vector and the yields of the protein are reasonably high. However, no matter what type of ion exchange or gel filtration chromatography you do, you always get contaminating proteins.

- i) How could you modify the expression vector, and therefore the resultant protein, to decrease the likelihood that you will have a contaminated protein? (2 pts)
- ii) How would this modification affect the **specific activity** of your final purified protein? (1 pt)

13. (1 pt) You are given a sample of either protein or a nucleic acid. What simple methods could you use to determine whether it is protein or nucleic acid [Hint: You could also measure the concentration with this method]?

14. (8 pts) The two drugs on the right are used to inhibit the growth of viruses, such as Herpes virus (which causes cold sores). The Ser and Phe are amino acids from the polymerase that interact with the drugs.

- Are these drugs based on purine or pyrimidine bases? (1 pt).
- Do you anticipate that these drugs would inhibit DNA replication or RNA synthesis? Why? [Hint: are they more similar to dNTPs or NTPs?] (1 pt)
- The ability of the drugs to inhibit polymerization was measured using steady-state enzyme kinetics with a constant amount of inhibitor. The resultant double reciprocal plots are shown on the right. Which inhibitor is more effective? Drug A or Drug B? Justify your answer with reference to **both** the double reciprocal data **and** the interaction between the drug and the polymerase (4 pts)
- Are these drugs competitive inhibitors or mixed type? Justify your answer (2 pt).

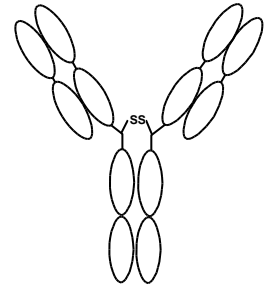


Bonus 2: These drugs are actually pro-drugs, in that they need to be converted to another compound by cellular enzymes before they are effective. What modifications to these drugs would likely occur to make them bind effectively to a polymerase? What type of enzyme would perform these modifications? (1 pt)

Bonus 3: Although these drugs bind to the DNA or RNA polymerase and reduce its activity, they also affect polymer growth, how? (1 pt)

15. (5 pts) Please do **one** of the following choices.

Choice A: The quaternary structure of the immunoglobulin was determined by simple techniques 20 years before the X-rays structure of an immunoglobulin was determined. What techniques would you use to determine the quaternary structure before X-ray diffraction could be used? Briefly describe the techniques, the data you would obtain, and how you would use this data to substantiate the structure shown on the right. (Note, the two heavy chains are linked by a disulfide bond).

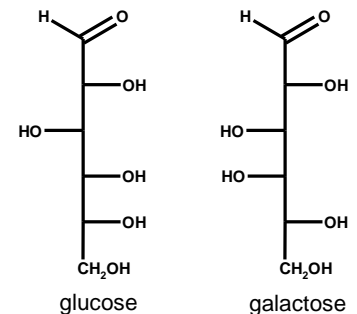


Choice B: The diagram on the right shows two simple diatomic molecules that differ only in their bond lengths. Explain why the scattered X-rays from these two molecules would be different such that their structures could be determined by X-ray diffraction.



16. (6 pts) The structure of glucose and galactose are shown on the right.

- Are these sugars aldoses or ketoses? Why? (1 pt)
- Draw β -glucose in its pyranose form using the reduced Haworth representation. (1 pt)
- Indicate the location of the new chiral centers on the pyranose ring of glucose, what is this new center called? (1 pt)
- Sketch, or describe, the chemical structure of any **one** of the following carbohydrates (2 pts)
 - lactose
 - sucrose
 - glycogen
 - cellulose
- Which of the above (a-d) could be used as an energy source? Indicate all possibilities (1 pt).



17. (3 pts) Please do **one** of the following choices.

Choice A: Yeast can produce ethanol obtained from glucose. What biochemical pathways are involved in the production of ethanol? What growth conditions would maximize the production of ethanol?

Choice B: Discuss the role of the organic electron carriers NAD^+/NADH and Q/QH_2 in the production of ATP.

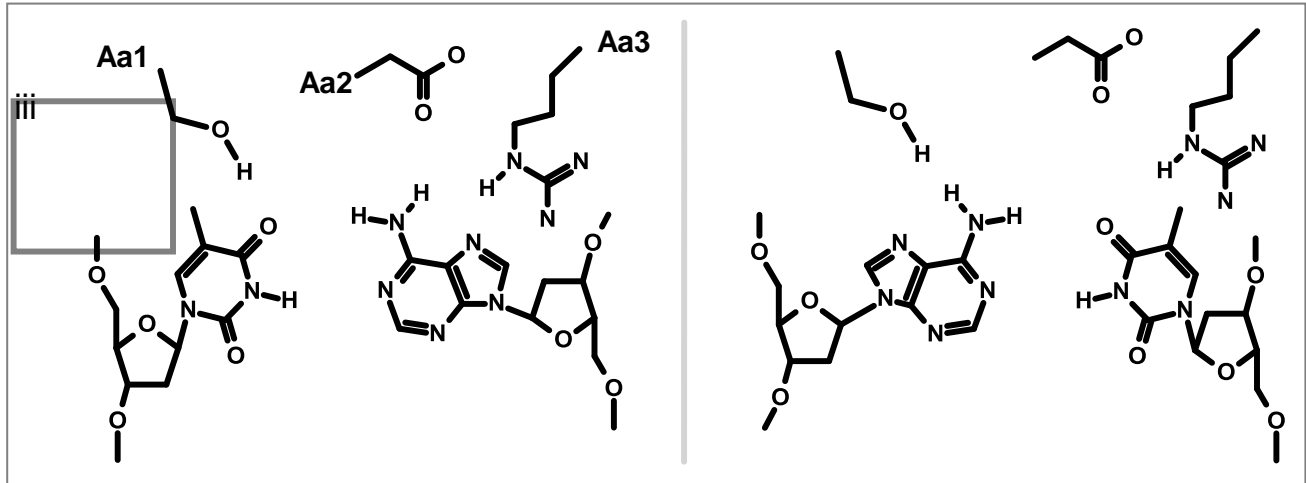
18. (6 pts) Please do **one** of the following questions.

Choice A: Pretend you just finished the Pittsburgh marathon. As a consequence, your glycogen levels and ATP levels in the liver are quite low. Discuss the process, with the major focus on regulation in your answer, by which your glycogen levels and ATP levels are restored as you eat lots of carbohydrates after the race.

Choice B: Pretend that you didn't run the Pittsburgh marathon, but lounged around all morning eating pancakes (with maple syrup of course). Consequently, the ATP levels in your liver cells are high. You are walking to campus and a ferocious dog, with rather large teeth, begins to chase you. What hormone is released and how does this hormone affect your ability to escape from the dog? You should discuss how this hormone will affect the regulation of metabolic pathways.

19. (10 pts) A protein binds to a specific sequence of double stranded nucleic acid. Part of the interaction between the protein and the nucleic acid is shown on the left side of the diagram. The amino acid sidechains from the protein are labeled Aa1, Aa2, and Aa3. The reversal of the two bases is shown on the right part of the diagram, along with a duplication of the protein shown in the left panel. The right panel will be useful for parts vii and viii.

i) Label the 5' and 3' carbons of left-most base (1 pt).



ii) Identify the glycosidic bond on the left-most base (1 pt).

iii) Place the appropriate missing atoms in the box labeled “iii” that would be required to connect this residue to the previous residue (1 pt).

iv) Indicate the “Watson-Crick” hydrogen bonds on the left-most base pair (1 pt).

v) Indicate H-bond donors and acceptors *on the protein* that could potentially interact with the DNA bases (1 pt)

vi) Is this protein binding in the major groove or the minor groove? How did you determine this? (1 pt).

vii) How would the binding affinity change if the protein bound to the reversed basepair (shown on the right)? You should assume that the structure of the protein does not change. (2 pts)

viii) If a protein (not necessarily the one shown in the diagram above) used a similar type of interaction in the other groove, i.e. if the amino acid sidechains entered from the lower part of the diagram, would your answer to part vii change? Why? (2 pts)

20. (4 pts) The DNA sequence on the right was treated with the restriction endonuclease EcoR1 (G^AAATTC).

GCGAATTCCCG
CGCTTAAGGGC

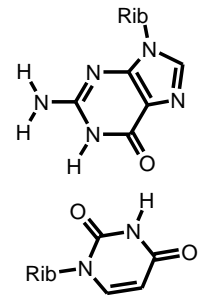
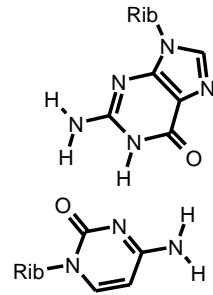
- Draw the expected products of the reaction (1 pts)
- Explain why the recognition sequence is the same on both the upper and lower strand (2 pt).
- Could the sticky ends generated by EcoR1 be ligated to those generated by the enzyme BamHI (G^AGATCC). Why or why not? (1 pt)

21. (5 pts) Please do **one** of the following choices.

Choice A: Describe the basic reaction mechanism for a typical DNA polymerase. Discuss why the Gibbs energy for the overall reaction is negative and also comment on the fidelity of the reaction, or why the polymerase is more likely to incorporate the correct base (**Note:** do not discuss removal of an incorrect base, see choice C).

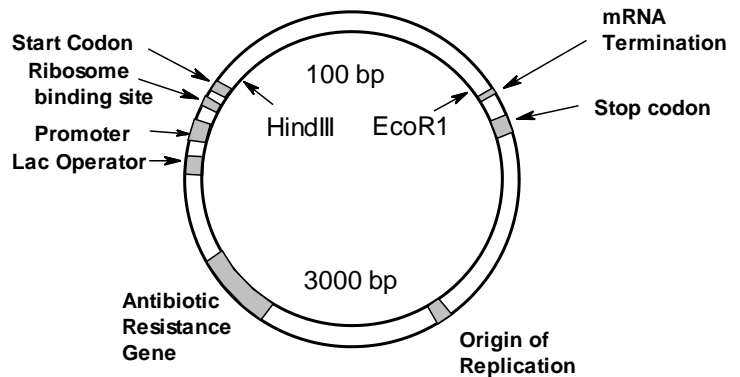
Choice B: Describe the mechanism by which tRNA molecules are charged and then briefly discuss how the same tRNA can be used to translate more than one codon. The diagram to the right may be helpful.

Choice C: Discuss the mechanism by which some DNA polymerases remove incorrectly incorporated bases. What is the consequence of lack of this function in a polymerase and how does it affect the treatment of HIV?



22. (5 pts) The diagram to the right shows an expression vector with a number of essential DNA features labeled.

- Two pairs of the labeled DNA sequences are in the wrong order. Identify one of them, and give the correct order. Briefly justify your answer. (4 pts).
- Why is there an “antibiotic resistance gene” in all plasmids? What is its role? (1 pt)



23. (5 pts) The following DNA contains a protein coding sequence of HIV reverse transcriptase. You would like to produce this protein in bacteria so that you could study drug resistant strains of the HIV virus. The sequence, along with the protein translation, is given in bold. The entire coding region is 500 nucleotide bases in length.

```

|-----500 bases-----|
AGCTGCTCATGCTCCCACATGCAATTCCTCCCA. . . . .GTGAGGGGAAATTAACCGCCGGCG
TCGACGAGTACGAGGGGTGTACCTTAAGGAGGGT. . . . .CACTCCCCCTTTAATTGGCGGCCGC
MetLeuProThrCysAsnSerSerPro ValArgGlyLysStop

```

The expression vector (complete diagram shown in the previous question) contains a HindIII site just after the start codon and a EcoR1 site adjacent to the stop codon that are contained in the vector, i.e. the vector sequence is:

```

HindIII                               EcoR1
--TTAGTAGGGCACCTCAATGAAGCTT---100 bases---GAATCTTAA---
```

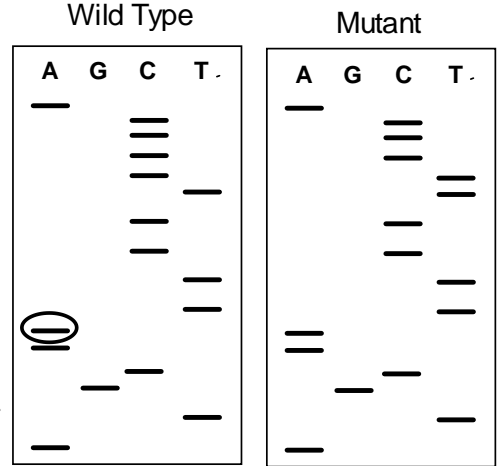
Because the start and stop codons are already in the vector, it is not necessary to amplify the start and stop codons from the DNA, just the sequence that codes for the remaining amino acids (LeuPro----GlyLys).

- Give the sequence of **both** the left and right primers that would generate the desired PCR product. Make the total length of your primers 12 bases (3 pts).
- Calculate the T_M for the **left** primer (1 pt). $T_M = 81.5 + 0.41*(\%GC) - 625/N$
- Based on this T_M what annealing temperature would you use for PCR? (1 pt)

24. (7 pts) The wild-type and mutant reverse transcriptases in previous question are sequenced using the primer: ATGCTCCCCAC.

Please answer the following questions.

- i) Briefly explain how the third DNA molecule in the “A” lane (circled) was generated (1 pt)
- ii) Identify the amino acid change associated with the mutation. The codon table is provided on the back of your formula page (6 pts)



The sequence of the reverse transcriptase gene is provided here for your convenience:

```

AGCTGCTCATGCTCCCCACATGCAATTCCTCCCCA . . . . . GTGAGGGGGAAATTAACCGCCGGCG
TCGACGAGTACGAGGGGTGTACCTTAAGGAGGGGT . . . . . CACTCCCCTTTAATTGGCGGCCGC
      MetLeuProThrCysAsnSerSerPro      ValArgGlyLysStop
    
```

25. (6 pts) Please do **one** of the following choices:

- Choice A:** Discuss the roles of the ribosome binding site, the start codon, and the stop codon on the overall process of protein synthesis.
- Choice B:** Briefly discuss the elongation step in protein synthesis. Be sure to address the role of the tRNA binding sites in the process.

26. (4 pts) You would like to make a human peptide hormone in yeast cell by using an expression plasmid that has the necessary signals for expression of proteins in eukaryotic cells, such as yeast. The peptide hormone will be used as a drug to treat individuals lacking this hormone. The sequence of the hormone is:

Met-Ala-Gly-Phe-Trp-Arg

The DNA sequence for this hormone is not available and thus you have to chemically synthesize the DNA instead of performing PCR. Assume that you are using the same expression vector as above, which contains the start and stop codons, separated by the EcoR1 and BamH1 sites.

- i) What DNA sequence would you have synthesized by the DNA synthesis company? Is this sequence unique? (3 pts)
- ii) Why would you use yeast or mammalian cells to produce this “protein drug” instead of E. coli (bacteria)? (1 pt)

Bonus 4. In what way is the ribosome like an apple?

Bonus 5. In what way is a peppermint patty like IPTG?

Bonus 6. Why is it necessary to produce T7 RNA polymerase when using the T7 expression system?

(Use back of page if necessary).

Page	Score
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
TOTAL	