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 pKa of 2.0 and a second with a pKa 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).
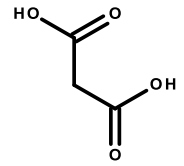
i) Ionization of the first group introduces a negative charge on the molecule that makes it more difficult for the second to ionize because that would produce two like charges in a small space, therefore the pKa increases (weaker acid).

ii) Since the pH is the same as the pKa, the fraction protonated = fraction deprotonated = 0.5. Since this is a diprotic acid, and you are beginning with the fully deprotonated species (disodium salt), you would have to add 1.5 equivalents of HCl. One equivalent to fully protonate the first group and then ½ equivalent to protonate the second –COO⁻ groupl.

Moles of HCl are 1.5 eq X 1 L X 0.1 moles/L = 0.15 moles (-1 pt if just ½ equivalent was added)

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

In this region, any added base (or acid) is used to deprotonated (or protonate) the weak acid, thus the concentration of hydrogen ions does not change much, so the pH does not change much.

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).

Alpha helix = hydrogen bonds parallel to helix axis, sidechains project out.
Beta sheet = hydrogen bonds perpendicular to stand directions, sidechains project above and below the sheet.

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 (-CH$_3$), the second is just a hydrogen atom, and the third is a isopropyl group (CH$_3$-CH-CH$_3$) (4 pts).

    i)   Label a peptide bond in your drawing and indicated whether it is drawn in the cis or trans form. (2 pts)

       Trans form

    ii) Which form of the peptide bond is more stable, cis or trans, and why? (1 pt)

       Trans – reduces unfavorable van der Waals between mainchain atoms.

    iii) Label the amino and carboxy terminus of the protein (1 pt).

       See sketch

    iv) Give the names of the amino acids that you have drawn and write out the sequence of the protein (1 pt)

       See sketch

    v) 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)

       Peptide bond cannot rotate because it has partial double bond character, p$_z$ orbital of N overlaps with p$_z$ orbital of carbon.

    vi) 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).

       The first residue has a sidechain, so its Ramachandran plot would consist of three low energy regions. The second residue does not have a sidechain, so most of the area of the Ramachandran plot would be low energy.

       The presence of the sidechain can cause van der Waals repulsion between atoms for certain configurations
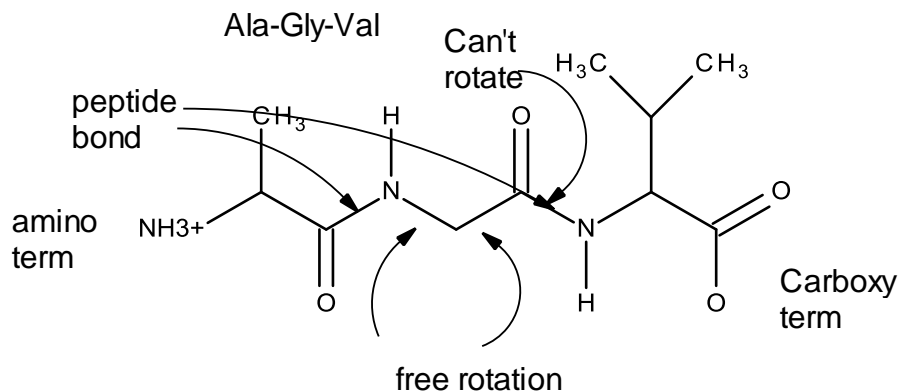
5. (4 pts) The following amino acid sequence is found in a soluble globular protein:

                                    Glu-Phe-Glu-Phe-Glu-Phe

    i)   What is the likely secondary structure of this sequence? Why? (2 pts)

    ii) Where do you expect the Phe residues to face: the interior of the protein, or to the solvent? Why? (1 pt)

    iii) If this was an integral membrane protein, where would you expect to find the Phe residues? Why? (1 pts)
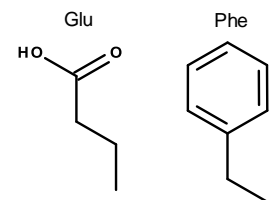
i) B-sheet, because of the alternating polar – non-polar nature of the sidechains,

ii) The non-polar Phe would face the interior, due to the hydrophobic effect.

iii) Facing the non-polar lipids.

6. (8 pts) A DNA binding protein binds to DNA in a non-specific manner. The protein contains lysine ($-NH_2$) 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)

K_D vs pH plot: Strong binding at low pH; "Neg charge repulsion" in middle; "poor binding, weak interaction between Lys and Phos (Hbonds)" at high pH. x-axis: pH 0 2 4 6 8 10 12

i) lysine residues, these are positive when protonated and would interact with the negative charge on the phosphate.

ii) phosphate groups, due to their negative charge.

iii) At low pH, the glutamic acid residues are protonated, so there will be a favorable interaction with the DNA due to the positive charge on the Lys residues. Giving the smallest Kd

When the glutamic acid residues become deprotonated, the charge repulsion will increase the Kd (poorer binding).

When the lysine residues become deprotonated the binding will be further weakened because of the loss of charge-charge interactions, only H-bonds between the lysine and DNA phosphates are possible. (-1 pt for neglecting neg charge on Glutamic acid)

iv) More favorable interactions = lower Kd
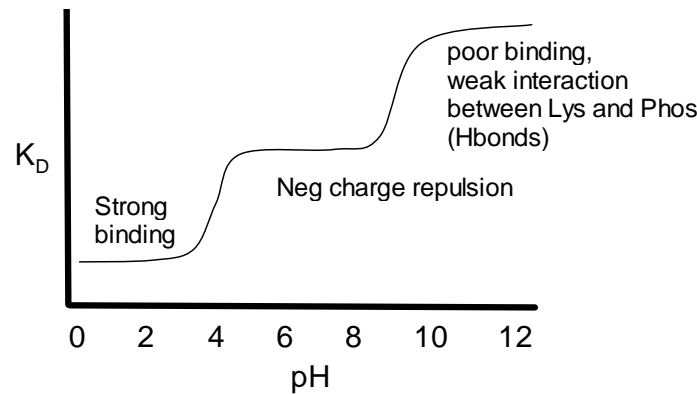
7. (7 pts) The binding curve for the same protein in the previous question is shown on the right. The protein has a molecular weight of 100 kDa protein binds a short DNA oligonucleotide (say 10 bases, 6 kDa). The binding curve is shown on the right. Please answer the following questions:
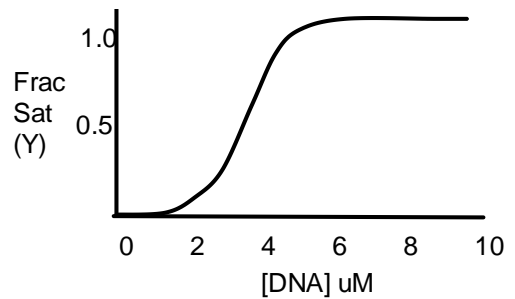
i)   What is the K_D for binding? (1 pt)
ii) Based on the shape of the binding curve, how many DNA molecules are likely binding to the protein, one or more? Justify your answer. (2 pts)
iii) What additional plot might you do to confirm your hypothesis to part ii? How would this plot be used to confirm 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)

i) 3 uM, where Y = 0.5
ii) It is cooperative, so more than one DNA must be binding.
iii) Do a Hill plot, the slope of the line when it crosses the axis should be something other than one.
iv) Salt will probably reduce the binding due to interfering with any electrostatic interactions between the phosphate and the lysine. Therefore the Kd will increase and the curve will shift to the right.

Bonus 1: Gel fitration can be used to separate the protein from the DNA protein complex since they differ in size, giving free protein and bound protein: Y=bound/total.

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?

Choice A:
Large non-polar residue is replaced by a smaller polar residue.
Less stable because of loss of van der Waals and the hydrophobic effect.

Choice B:
Less stable:
two Watson crick H-bonds instead of three
weaker van der Waals between the stacked bases.

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)

Choice A:
i) Glycerol with phosphate on C1, two fatty acids esterified to C2 and C3. A "cartoon" diagram was acceptable, provided it was clear that there were two non-polar tails and a polar headgroup.
ii) Corn oil contains cis double bonds in the fatty acids, kink in acyl chain reduction in van der Waals lowers melting point, therefore it is a liquid. Margarine is made by adding hydrogen to the double bonds, removing the kink, increasing van der Waals interactions, raising the melting point, therefore it is a solid.

Choice B:
i) OH group attached to four rigid rings, flexible non-polar tail.
ii) It broadens the phase transition.
iii) It keeps membranes fluid over a wide temperature range, allowing for the diffusion of compounds in the membrane.

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)

i) Good substrates have many interactions with the active site, leading to **small off-rates**. Therefore the substrate is bound long enough for the chemical transformation to occur, i.e. $k_{off} < k_{cat}$. Full credit for relating many interactions to a slow off-rate.

ii) They reduce the energy of the transition state, which increases its concentration and therefore the rate (3 pts).
This occurs with all enzymes because functional groups are pre-ordered, so there is no decrease in energy as the substrate-enzyme complex goes to the transition state. (3 pts)

iii) They catalyze the reaction of moving the ion across the membrane.

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).

i)
Enzyme is in two forms – relaxed (active) or tense (inactive).
Activators and inhibitors binding away from the active site, changing the shape of the enzyme.
Inhibitors shift the equilibrium to the tense form/decreasing the activity
Activators shift the equilibrium to the relaxed form/increasing the activity.

a) oxygen transport – oxygen is an allosteric activator, the binding of one oxygen increases the binding of subsequent oxygens – allowing full saturation in the lungs and effective release to the tissues.
b) altitude adjustment – bisphosphoglycerate is an allosteric inhibitor, it is increased at high altitudes. It shifts the binding curve to the right (lower affinity) but changes its shape so an equivalent amount of oxygen is delivered.
c) Mixed type inhibitors bind at a different location than the active site and cause an allosteric change in the active site that inhibits the enzymes.
d) ATP is an allosteric inhibitors of PFK. F26P, AMP, ADP activate PFK. Protein phosphorylation is a form of allosteric regulation – used to regulate glycogen synthesis and degradation.
e) Lac repressor is released from DNA by IPTG, IPTG causes a change for the relaxed DNA binding form to the tense non-DNA binding form.

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)

i) Add additional codons to either export the protein (leader sequence at the amino terminus), or codons to add a his tag (6 histidine residues) to allow affinity chromatography.
ii) The specific activity of the final purified protein should be higher, since you should have removed contaminating proteins.

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]?

Preferred answer: Protein and DNA have different absorption maxima, DNA = 260nm, protein =280nm. Simply take the absorption spectra.
Partial credit answer (1/2 pts): DNA is negatively charged, so use an anion exchange column (but since the protein may also be negatively charged, this may not work.)

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.
  i)   Are these drugs based on purine or pyrimidine bases? (1 pt).
  ii) Do anticipate that these drugs would inhibit DNA replication or RNA synthesis? Why? [Hint: are they more similar to dNTPs or NTPs?] (1 pts)
  iii) 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)
  iv) Are these drugs competitive inhibitors or mixed type? Justify your answer (2 pt).
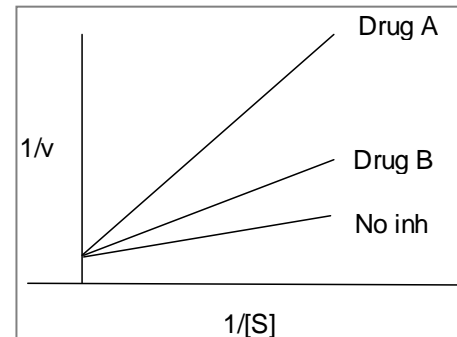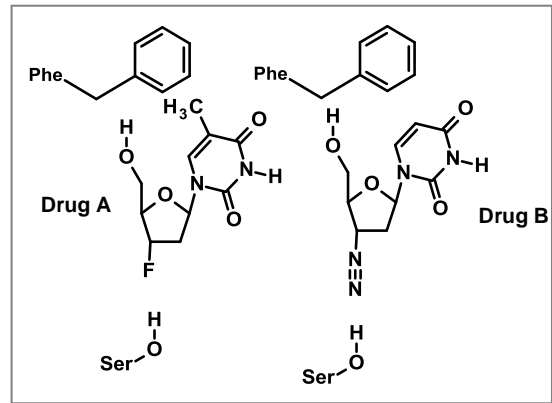
i) pyrimidine bases.

ii) They lack a 2'OH group, so they are likely to inhibit DNA polymerases.

iii) Drug A is the better inhibitor:
    slope of plot is larger, therefore α is larger, therefore K_I is SMALLER. (2 pts)
    Drug A has a more favorable H-bond between the Ser and the electronegative fluorine (1 pt)
    Drug A has a more favorable vdw & hydrophobic effect between Phe and $-CH_3$ group. (1 pt)

iv) Competitive – they look like the substrate (nucleotides) (1 pt) and the enzyme kinetic data shows no change in Vmax (same intercept on y-axis) (1 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)

Conversion to nucleoside triphosphates by kinases.

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

They lack a 3' OH group, and therefore once incorporated the growing DNA chain would be terminated, in the same fashion as would occur with ddNTPs.

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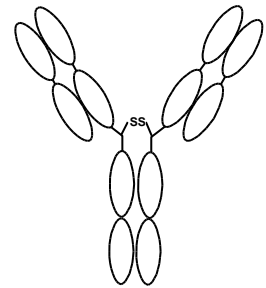
Choice B: The diagram on the right shows two simple diatomic molecules that differ only in their bond lengths. Explain why the scattered X-rays from these two molecules would be different such that their structures could be determined by X-ray diffraction.

Choice A:

- The overall molecular weight would be obtained by gel filtration chromatography.
- The subunit molecular weights would be obtained by SDS-PAGE.
- One experiment would be done without BME, which would give the molecular weights of any chains linked by disulfide bonds, in this case the two heavy chains would be linked together, but the two light chains would be separate (not the case for most immunoglobulins, the light is also disulfide bonded to the heavy).
- SDS PAGE with BME (beta mercaptoethanol) would give the isolated chains, in a ratio of 1:1 of light and heavy chains.
- In order to agree with the overall molecular weight from gel filtration, there would have to be two light and two heavy chains.

Choice B:

- The relative position of the atoms affects the interference between the X-rays scattered from each atom.
- The difference in interference leads to a difference in intensities.
- The difference in intensities can be used to determine the relative positions.

16. (6 pts) The structure of glucose and galactose are shown on the right.
  i) Are these sugars aldoses or a ketoses? Why? (1 pt)
  ii) Draw β-glucose in its pyranose form using the reduced Haworth representation. (1 pt)
  iii) Indicate the location of the new chiral centers on the pyranose ring of glucose, what is this new center called? (1 pt)
  iv) Sketch, or describe, the chemical structure of any **one** of the following carbohydrates (2 pts)
      a) lactose   b) sucrose   c) glycogen   d) cellulose
  v) Which of the above (a-d) could be used as an energy source? Indicate all possibilities (1 pt).
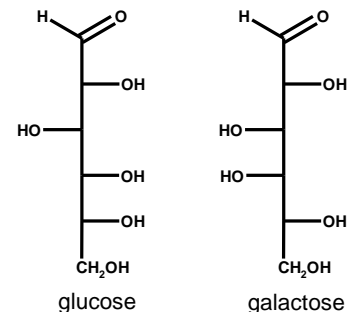
i) both are aldoses because they have an aldehyde group.

ii) Six member ring, OH on C1 up, then down, up, down.

iii) The C1 is now chiral and is called the anomeric carbon

iv) lactose = galactose+glucose, sucrose=glucose+fructose, glycogen = glucose a(1-4) + a(1-6) branches, cellulose = glucose β(1-4) linkages, not branched.

v) Preferred answer: All by cellulose can be used by humans for energy.
   Accepted answer: All are possible energy sources, cellulose can only be used by organisms that have cellulases.

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.

Choice A

- Glucose enters glycolysis and is converted to pyruvate.
- Under low oxygen conditions the pyruvate is converted to ethanol to regenerate NAD+ as an electron acceptor in glycolysis.

Choice B

- NAD+ is used to accept electrons produced in the oxidation of organic molecules in metabolic pathways.
- NADH gives its electrons to complex I which then passes them on to conzyme Q and then ultimately to oxygen.
- This process generates a hydrogen ion gradient across the inner mitochondrial membrane that is used to generate 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.

Choice A:

- Blood glucose levels would become high
- Causing release of insulin
- Insulin would cause dephosphorylation of enzymes in cell.
- Glycogen synthase is active and glucose would be stored in glycogen

- F26P levels will rise, activating PFK in glycolysis.
- Oxidation of glucose will restore ATP levels.

Choice B:

- Epinephrine would be released from your central nervous system
- Binding to its receptor would signal enzyme phosphorylation due to the activation of kinases.
- This would activate glycogen phosphorylase, causing glucose to be released from glycogen.
- High levels of ATP would allow synthesis of glucose from pyruvate by gluconeogenesis (F26P levels are low so bisphosphatase in gluconeogenesis can be active)

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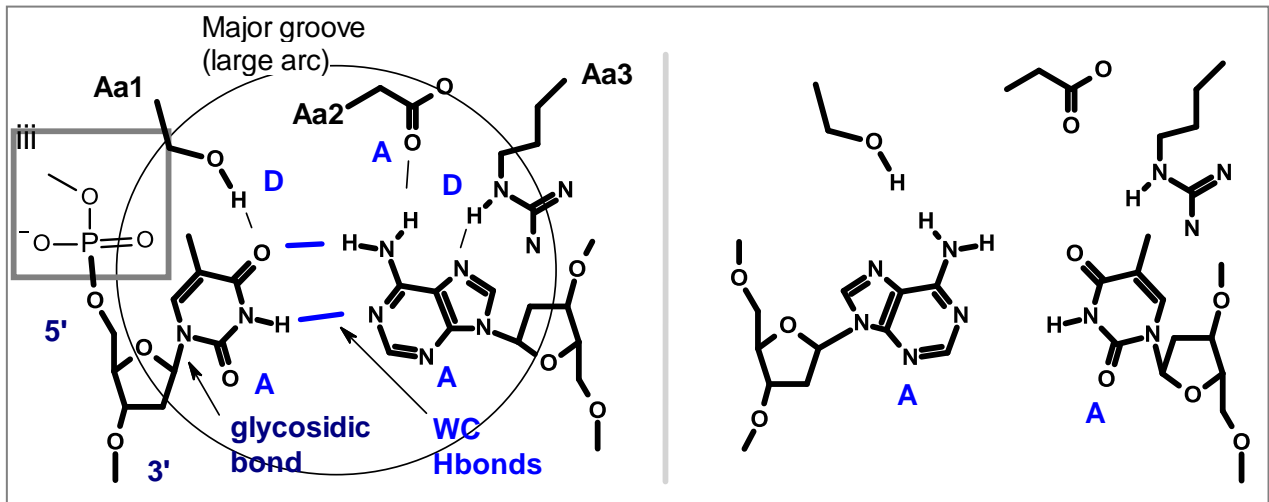
i) Label the 5' and 3' carbons of left-most base (1 pt). See diagram

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

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). See diagram

iv) Indicate the "Watson-Crick" hydrogen bonds on the left-most base pair (1 pt). There are two, labeled.

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

There are three potential H-bonds between the protein and the DNA. The order of these on the protein is: Donor   Acceptor   Donor

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

Major groove – large arc if you drew a circle through the riboses.

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)

The affinity would be weaker, because the hydrogen bond donors and acceptors no longer line up (e.g. donor paired with a donor).

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)

The affinity would be the same for either TA or AT because either basepair shows the same pattern of hydrogen bond acceptors in the minor groove, two acceptors on the DNA, in approximately the same location.

10

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

> GCGAATTCCCG
> CGCTTAAGGGC

i) Draw the expected products of the reaction (1 pts)

ii) Explain why the recognition sequence is the same on both the upper and lower strand (2 pt).

iii) Could the sticky ends generated by EcoR1 be ligated to those generated by the enzyme BamHI (G^GATCC). Why or why not? (1 pt)

i) GCG   AATTCCCG
    CGCTTAA   GGGC

ii) The enzyme is a homodimer with two fold rotational symmetry; one monomer recognizes the upper strand, one the lower strand.

iii) No, the products after digestion with BamHI are

    G   GATCC
    CCTAG   G

The sticky end from EcoR1 "AATT" is not complementary to the sticky end from BamH1

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?
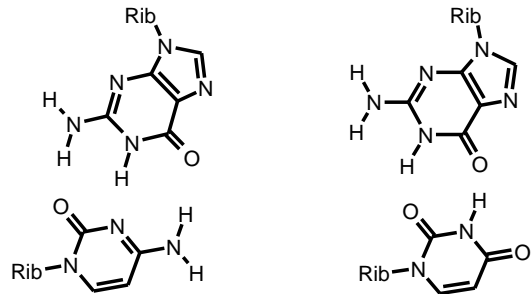
Choice A:
- Require a template with an annealed primer, with a 3'OH
- dNTPs are brought into the active site,
- matching based on WatsonCrick Hbonds, plus purine-pyrimidine size
- 3'OH attacks phosphate, forming new phosphodiester bond, releasing pyrophosphate
- Pyrophosphate hydrolyzed to inorganic phosphate – large neg free energy drives polymerization by indirect coupling.

Choice B:
- A specific aminoacyl tRNA synthetase first binds the correct amino acid and ATP, forms an ester to the amino acid.
- Correct tRNA is bound and the amino acid is transferred to the 3'OH of the tRNA
- A single anti-codon loop on a tRNA can recognize more than one codon by virtue of the fact that there are multiple possible base-pairings for the 3rd base of the codon. For example the normal GC pairing can be replaced by a GU basepair.

Choice C:
- The mismatched base is removed by hydrolysis, by 3' – 5' exonuclease activity.
- If this activity is not present in the polymerase, then mutations will be introduced into the DNA sequence
- These changes can cause changes in the aminoacid sequence of HIV enzymes, such that the drug that inhibits the enzyme can no longer bind.

11

22. (5 pts) The diagram to the right shows an expression vector with a number of essential DNA features labeled.
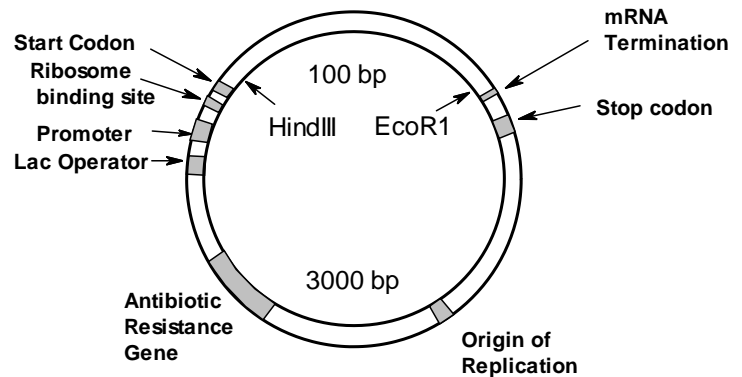
  i) **Two pairs** of the labeled DNA sequences are in the wrong order. Identify one of them, and give the correct order. Briefly justify your answer. (4 pts).

  ii) Why is there an "antibiotic resistance gene" in all plasmids? What is its role? (1 pt)

[Diagram: circular plasmid labeled with Start Codon, Ribosome binding site, Promoter, Lac Operator, mRNA Termination, Stop codon, Antibiotic Resistance Gene, Origin of Replication, HindIII, EcoR1, 100 bp, 3000 bp]

i)
The lac operator sequence and the promoter are switched. The promoter should come before the operator so that the binding of the lac repressor to the lac operator can regulate mRNA production.

OR

The stop codon and the mRNA termination signals are reversed, as it is, the mRNA would have no stop codon so the protein would not be terminated correctly.

ii) The antibiotic resistance gene makes bacterial that contain the plasmid resistant to antibiotics – so that only bacteria with the plasmid can grow in the presence of the antibiotic, thus preventing the growth of bacteria without the plasmid. The bacteria without the plasmid will not produce the protein of interest.

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------------------|
  AGCTGCTCATGCTCCCCACATGCAATTCCTCCCCA......GTGAGGGGGAAATTAACCGCCGGCG
  TCGACGAGTACGAGGGGTGTACCTTAAGGAGGGGT......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---GAATTCTTAA---
```

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).

  i) 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).

  ii) Calculate the T_M for the **left** primer (1 pt). $T_M = 81.5 + 0.41*(\%GC) - 625/N$

  iii) Based on this T_M what annealing temperature would you use for PCR? (1 pt)

i) The left primer sequence will begin with the required restriction site (HindIII), followed by the sequence of the top strand at the left boundary. The right primer will begin with the EcoR1 site, followed by the sequence of the bottom strand at the right boundary.

Left primer = AAGCTTCTCCCC [Note start codon is omitted]

Right primer = GAATTCTTTCCC [Note: stop codon is omitted]

ii) Tm = 81.5 +0.41(100 × (7/12)) – 625/12 = 53.3

iii) The annealing temperature should be 5 degrees below this, or 48.3.

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)

Wild Type   **Mutant**
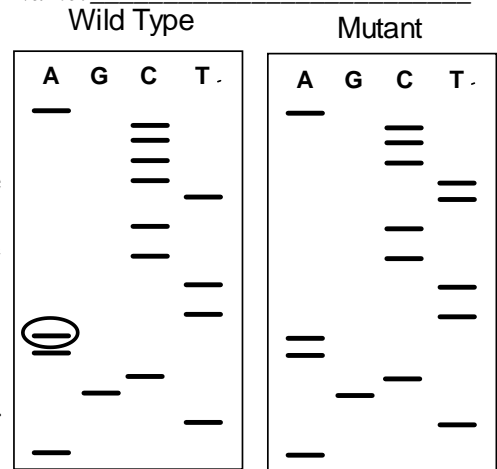
    The sequence of the reverse transcriptase gene is provided here for
    your convenience:
```
        Primer→ATGCTCCCCAC
        AGCTGCTCATGCTCCCCACATGCAATTCCTCCCCA......GTGAGGGGGAAATTAACCGCCGGCG
        TCGACGAGTACGAGGGGTGTACCTTAAGGAGGGGT......CACTCCCCCTTTAATTGGCGGCCGC
               MetLeuProThrCysAsnSerSerPro       ValArgGlyLysStop
```

i) This band is generated due to the presence of ddATP in the reaction, which is a chain
   terminator, this band represents the addition of ddATP to the growing strand.
ii)
```
                      Cys Asn Ser Ser Pro
    Wild-type sequence:  A TGC AAT TCC TCC CCA
    Mutant sequence:     A TGC AAT TCC TTC CCA
                      Cys Asn Ser Phe Pro
```
The primer (shown above) indicates that the sequencing should start with the underlined (and
   highlighted) "A". This "A" is the last base of a Thr codon, so the TGC is the first full codon
   and every three bases after that corresponds to a codon. **The change is from Ser to Phe.**

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.

Choice A:
Ribosome binding site: binds to the rRNA in the 30S small subunit and positions the start
   codon at the right position on the ribosome, in the P tRNA site.
Start codon: Binds the first tRNA, a special tRNA containing the modified amino acid fMet.
Stop codon: Recruits the protein termination factor to the A site, causing hydrolysis of the
   protein from the last tRNA.

Choice B:
The peptide, attached to the last tRNA used, is in the P site.
The charge tRNA for the next codon enters the A site.
Peptide bond formation occurs, the amino group of the amino acid in the A site attacks the
   carboxyl end of the growing peptide, transferring the protein to the A site.
The ribosome shifts, moving the peptide-tRNA to the P site.

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:

<div align="center">Met-Ala-Gly-Phe-Trp-Arg</div>

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)

i) You would need to have:
   a) EcoR1 site
   b) 6 codons that specified the amino acid sequence, missing the start codon, since it is provided by the vector.
   c) Bam site

Note: no stop codon is required, because it would be provided by the vector.

The sequence is not unique, because there is more than one codon for each amino acid. A possible sequence is:

```
GAATTC GCT GGT TTT TGG CGT GGATCC
EcoR1  Ala Gly Phe Trp Arg BamH1
```

Many students may have just given the codons required for the protein (-1/2 pt)

```
ATG GCT GGT TTT TGG CGT TAA
Met Ala Gly Phe Trp Arg stop
```

ii) Proteins produced in bacteria can have toxic impurities.

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

Similar in shape, stem represents the growing polypeptide coming out of the exit tunnel.

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

It induces the lac repressor to become unbound from the DNA by causing an allosteric change. In this case a person gripping the DNA is the bound repressor and they must let go to bind the patty.

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

The T7 expression system replaced the normal promoter with the T7 promoter which can only be recognized by T7 RNA polymerase.

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