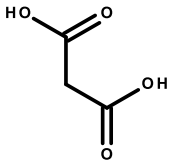
This exam has **13 pages** and is out of **150 points**. You should allot 1 min/point. On questions with choices all of your attempts will be graded and you will receive the best grade. Use the space provided, or the back of the preceding page. A codon table and codon usage information is given on the 14th page.

1. (6 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. Briefly describe how you would prepare 0.25 L of a 0.1 M buffer at pH=2.0, assuming that you are starting with the disodium salt of the acid. You should give the moles of sodium malate and HCl that you would need to make this buffer.



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.

3. (10 pts) Most proteins generally consist of secondary structural elements.

i) Name the **<u>two</u>** common secondary structures. (2 pt)

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

iii) What are the principle interactions that stabilizes both of these structures? (2 pts).

iv) How are the interactions that you discussed in part iii) depicted on a Ramachandran plot? (3 pts).

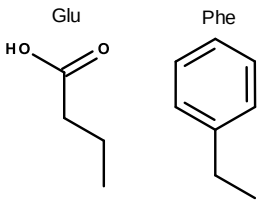
Points:_____

4. (10 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 an 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)
ii) Which form of the peptide bond is more stable, cis or trans, and why? (2 pt)
iii) Label the amino and carboxy terminus of the protein (1 pt).
iv) Give the names of the amino acids that you have drawn and write out the sequence of the protein (1 pt)

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

Glu-Phe-Glu-Phe-Glu-Phe



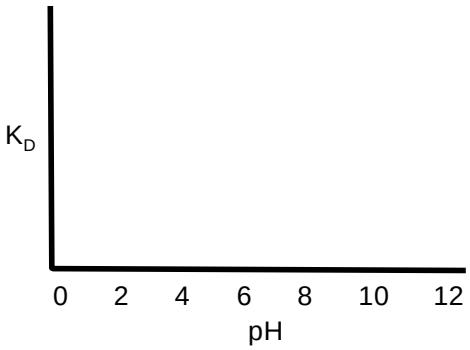
i) What is the likely secondary structure of this sequence? Why? (2 pts)
ii) Where do you expect the Phe residues to face, the interior of the protein, or to the solvent? Why? (1 pt)
iii) If this was an integral membrane protein, where would you most likely expect to find the Phe residues? Why? (1 pts)

Points:_____

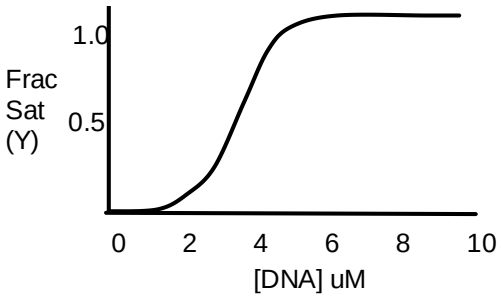
6. (7 pts) A DNA binding protein binds to DNA in a non-specific manner. The protein contains lysine (-NH_2) residues which are close to the DNA.

ii) What group on the DNA would the lysine 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. You can assume the pKa values of Lys is 9. Please justify your answer (6 pts).



7. (6 pts) The binding curve for the same protein in the previous question is shown on the right, for a single pH value. The protein has a molecular weight of 100 kDa protein binds one or more short DNA oligonucleotide (say 10 bases, 6 kDa). 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 could you determine the binding constant using size exclusion chromatography? (1 pt)

Points:_____

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 A-T 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. (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 necessary to give specific details about any particular enzyme.

i) Why are enzymes specific for particular substrates, 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).

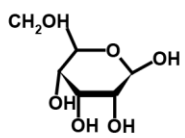
iii) Why are transport proteins (e.g. K^+ channel) considered (at least by me) to be enzymes, even though they do not change the chemical structure of the substrate? (1 pt)

Points:_____

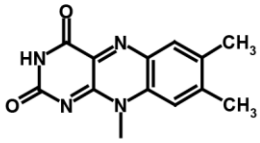
10. (8 pts, 1/2 pt each) On the right are a series of 15 biochemical structures (A-O), on the left is a list of names or descriptions. Indicate the correct match by writing the letter next to the description or name. Note that a structure *should* only be used once and one item should not have a structure associated with it.

Description	**Match (A-O)**
1. Product of glycolysis	
2. Product of anaerobic metabolism in humans.	
3. Product of anaerobic metabolism in Yeast.	
4. Fatty acid	
5. Triglyceride	
6. Phospholipid	
7. Cholesterol	
8. Six carbon aldose.	
9.Saccharide found in bacterial cell walls	
10. Disaccharide	
11. Electron carrier in the TCA cycle and fatty acid oxidation.	
12. Electron carrier in electron transport chain.	
13. Final electron acceptor in most species.	
14. Nucleotide normally found in DNA	
15. Nucleotide normally found in RNA	
16. Nucleotide that is used in DNA sequencing.	

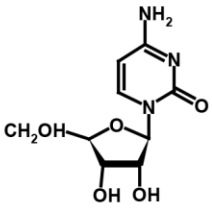
A



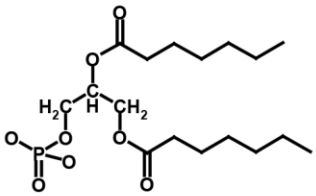
B



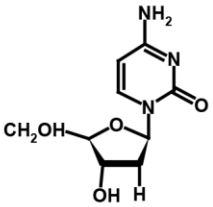
C



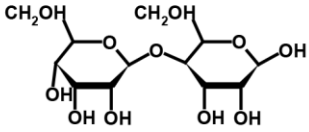
D



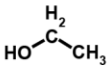
E



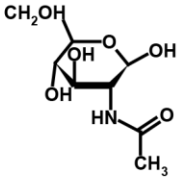
F



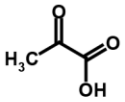
G



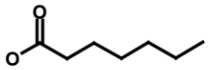
H



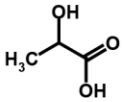
I



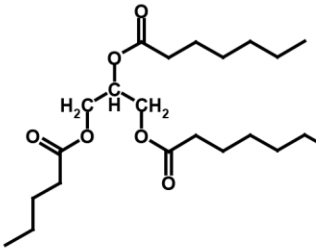
J



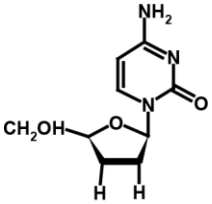
K



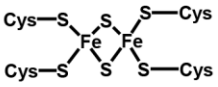
L



M



N



O

11. (1 pt) You are given a sample of either protein or a nucleic acid. What simple method could you use to determine whether it is protein or nucleic acid [Hint: You could also measure the concentration with this method]?

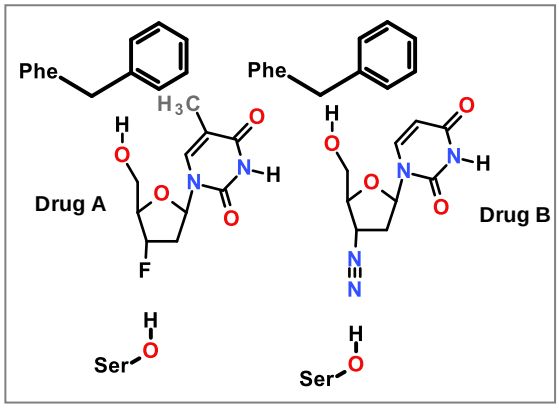
Points:_____

12. (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, d) metabolic regulation (glycogen or glycolysis), or e) regulation of DNA transcription by the lac repressor (2 pt).

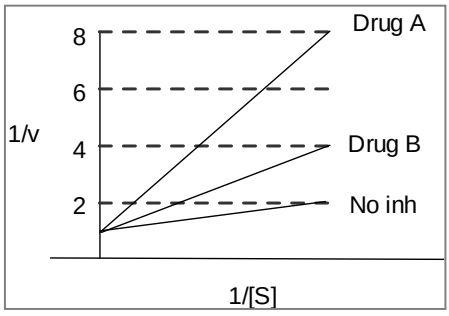
13. (8 pts) The two drugs on the right are used to inhibit the growth of viruses by inhibiting the DNA polymerase that the virus uses to replicate its genetic material. The Ser and Phe are amino acids from the polymerase that interact with the drugs.



i) Are these drugs based on purine or pyrimidine bases (circle correct answer)? (1 pt).

ii) Are these drugs competitive inhibitors or mixed type? Justify your answer (2 pt).

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)



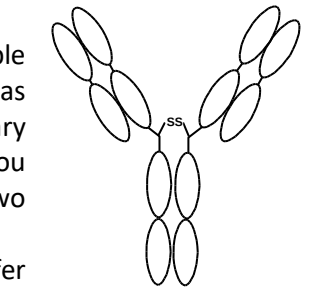
Points:_____

Bonus 1: 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 2: Although these drugs bind to the DNA polymerase and reduce its activity, they also affect polymer growth by acting as suicide inhibitors. How does this occur? (1 pt)

14. (5 pts) Please do **<u>one</u>** of the following choices.

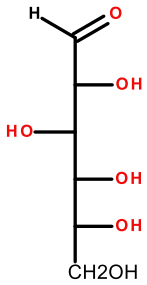
Choice A: The quaternary structure of the immunoglobulin was determined by simple techniques 20 years before the X-ray structure of an immunoglobulin was determined. What techniques were initially used to determine the quaternary structure? Briefly describe the techniques, the data you would obtain, 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.

Points:_____

15. (5 pts) The structure of glucose is shown on the right.



glucose

i) Draw β-glucose in its ***furanose*** form using the reduced Haworth representation. (1 pt)

ii) Indicate the location of the new chiral centers on the ring of glucose, what is this new center called? (1 pt)

iii) Sketch, or describe, the chemical structure of any **one** of the following carbohydrates (2 pts)

a) lactose b) sucrose

c) glycogen d) cellulose

iv) Which of the above (a-d) could be used as an energy source? Indicate **all** possibilities (1 pt).

16. (4 pts) 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?

Points:_____

17. (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 that produce glucose.

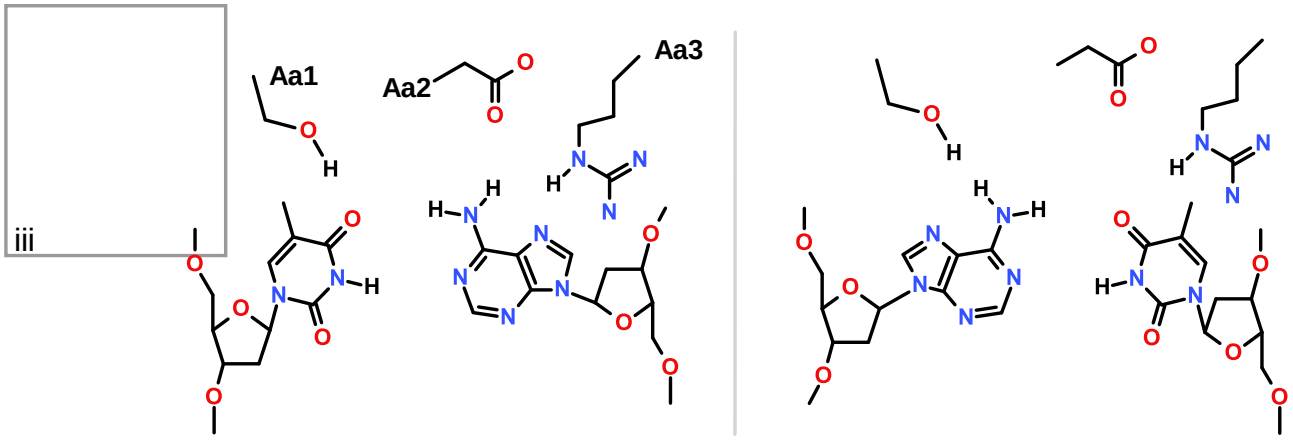
18. (5 pts) Please do **one** of the following choices.

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 B).

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

Points:_____

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 T residue to the previous residue. Include all atoms on the previous residue, and indicate that base with "B"(1 pt).

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

v) Indicate H-bond donors (D) and acceptors (A) *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).

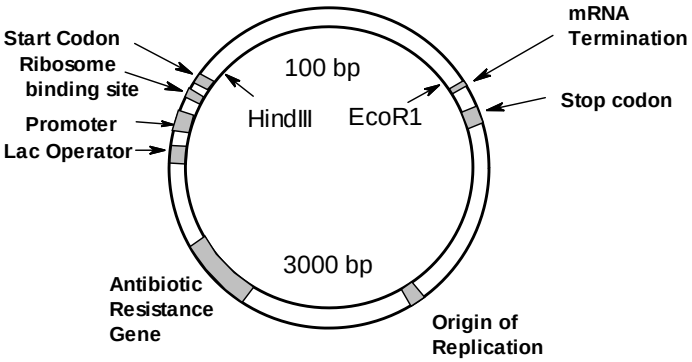
vii) How would the binding affinity change if the protein bound to the reversed basepair (shown on the right)? You should assume that the structure of the protein does not change. (2 pts)

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)

Points:_____

20. (5 pts) The diagram to the right shows an expression vector with several essential DNA features labeled.

Two of the labeled DNA sequences are in the wrong order. Identify **both**, and give their correct position. Briefly justify your answer with a description of the role of that DNA sequence in the production of mRNA and/or protein and why its location is incorrect.



21. (2 pts) Please do **<u>one</u>** of the following:

Choice A: Why is there an "antibiotic resistance gene" in all plasmids? What is its role?

Choice B: Why is there an "origin of replication" in all plasmids? What is its role?

22. (6 pts) Please do **<u>one</u>** 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.

Points:_____

23. (8 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 of the enzyme is 500 nucleotide bases in length, this is contained within the 10,000 bases of the entire HIV genome.

```
            |-------------------500 bases------------------|
---AGCTGCTC ATGCTCCCCACATGCAATTCCTCCCCA......GTGAGGGGGAAATTA ACCGCCGGCG---
---TCGACGAG TACGAGGGGTGTACCTTAAGGAGGGGT......CACTCCCCCTTTAAT TGGCGGCCGC---
            MetLeuProThrCysAsnSerSerPro       ValArgGlyLysStop
```

The expression vector (complete diagram shown question 20) 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---
```

(i.e. the start and stop codons are already in the expression vector/plasmid).

i) Give the sequence of **both** the left and right primers that would generate the desired PCR product so that you could insert the PCR product into the vector. The total length of your primers 12 bases (3 pts).

ii) Briefly explain *why* you wanted to add HindIII and EcoR1 sites to your PCR product (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? Why? (1 pt)

Left primer: 5'_____ _____ _____ _____ _____ _____ _____ _____ _____ _____ _____ _____

Right primer: 5'_____ _____ _____ _____ _____ _____ _____ _____ _____ _____ _____ _____

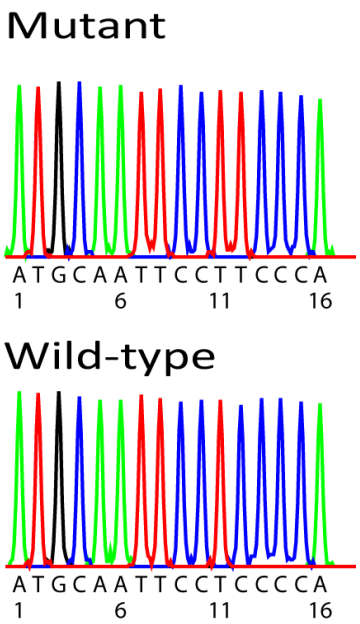
24. (2 pts) Please do **<u>one</u>** of the following choices. Regardless of your choice, briefly describe how the process works to purify the protein (i) or cause its export out of the cell (ii).

Choice A: You are trying to purify the mutant reverse trancriptase and cannot obtain pure protein using size exclusion or ion exchange chromatography. How could you modify the expression vector to allow affinity chromatography?

Choice B: High levels of reverse transcriptase that are produced from the plasmid are toxic to the cell. How would you modify the expression vector to cause the reverse transcriptase to be exported out of the cell?

Points:_____

25. (5 pts) The wild-type and mutant reverse transcriptases are sequenced using the primer: ATGCTCCCCAC**.** Identify the amino acid change associated with the mutation. The codon table is provided on the last page.



The sequence of the reverse transcriptase gene is provided here for your convenience (primer sequence underlined):

AGCTGCTC**ATGCTCCCCAC**ATGCAATTCCTCCCCA**.....GTGAGGGGGAAATTA**ACCGCCGGCG
TCGACGAG**TACGAGGGGGTGTACCTTAAGGAGGGGT.....CACTCCCCCTTTAAT**TGGCGGCCGC
MetLeuProThrCysAsnSerSerPro ValArgGlyLysStop

Bonus 3: Briefly explain how the third "peak", corresponding to a G, was generated

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-Ala

The DNA that codes 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 in Q20, which contains the start and stop codons, separated by the EcoR1 and BamH1 sites. Give the DNA sequence that you would request from the DNA synthesis company. Codon usage frequencies for yeast are given on the last page. Write your answer below:

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

Bonus 5. Why is it necessary to produce T7 RNA polymerase when using the T7 expression system?

(Use back of previous page if necessary).

Points:_____

5' Base	Middle Base				3'
	T	**C**	**A**	**G**	
T	Phe	Ser	Tyr	Cys	**T**
	Phe	Ser	Tyr	Cys	**C**
	Leu	Ser	**Term**	**Term**	**A**
	Leu	Ser	**Term**	Trp	**G**
C	Leu	Pro	His	Arg	**T**
	Leu	Pro	His	Arg	**C**
	Leu	Pro	Gln	Arg	**A**
	Leu	Pro	Gln	Arg	**G**
A	Ile	Thr	Asn	Ser	**T**
	Ile	Thr	Asn	Ser	**C**
	Ile	Thr	Lys	Arg	**A**
	Met	Thr	Lys	Arg	**G**
G	Val	Ala	Asp	Gly	**T**
	Val	Ala	Asp	Gly	**C**
	Val	Ala	Glu	Gly	**A**
	Val	Ala	Glu	Gly	**G**

Yeast Codon Usage - Middle column gives the percentage in highly expressed genes.

Ala	GCT	14.52	27.54	22.86
	GCC	27.62	16.14	23.67
	GCA	19.63	24.01	31.27
	GCG	38.23	32.30	22.19

Gly	GGT	32.91	50.84	31.79
	GGC	43.17	42.83	24.51
	GGA	9.19	1.97	24.75
	GGG	14.74	4.36	18.95

Phe	TTT	55.09	29.08	67.14
	TTC	44.91	70.92	32.86

Points:_____