

Biochemistry I - Spring 2000 - Final Exam & Solutions**Section A: Multiple Choice. Please circle the BEST answer (24 questions, 48 points total, 2pts each).**

1. A hydrogen bond is:
 - a) only found in water.
 - b) only found in DNA .
 - c) only found in proteins.
 - d) found in both DNA and proteins.**
2. For all practical purposes, a buffer will control pH over which range:
 - a) at pH values = pKa \pm 1.**
 - b) at pH values = pKa \pm 2.
 - c) at pH values = pKa \pm 0.1.
 - d) at any pH value.
3. The peptide bond in proteins is
 - a) planar, but rotates to three preferred dihedral angles.
 - b) cleavable by restriction endonucleases.
 - d) planar, and usually found in a trans conformation.**
 - e) cleavable by lysozyme.
4. Which of the following elements of secondary or super-secondary structure are most likely to be found in an integral membrane protein?
 - a) single β -strands.
 - b) isolated β -hairpin.
 - c) α -helices.**
 - d) β - α - β structure.
5. An "oil drop with a polar coat" is a metaphor referring to the three dimensional structure of:
 - a) an integral membrane protein.
 - b) collagen.
 - c) a globular protein.**
 - d) glycogen.
6. Immunoglobulins are proteins that
 - a) are involved in the determination of blood groups**
 - b) only bind small molecules with little specificity
 - c) only bind large proteins with little specificity
 - d) commonly perform catalytic functions.
7. The active site of an enzyme
 - a) contains amino acids that confer substrate specificity.
 - b) is flexible.
 - c) contains residues important for catalysis.
 - d) all of the above are true.**
8. The transition state of a reaction is
 - a) highly populated.
 - b) lowered in energy in enzyme catalyzed reactions.**
 - c) raised in energy in enzyme catalyzed reactions.
 - d) bypassed in enzyme catalyzed reactions.
9. Allosteric enzymes
 - a) are specific for the allo stereoisomer of amino acids.
 - b) must be tetrameric.
 - c) always display positive cooperativity
 - d) can be monomeric.**

10. The high rate of the formation of HIV viruses that are resistant to drugs is due to:
- Induction of mutations in the viral genome (DNA) by the drugs.
 - Interference of drugs with proofreading ability of PolI.
 - Interference of drugs with proofreading ability of HIV reverse transcriptase.
 - d) Lack of proofreading by HIV reverse transcriptase.**
11. Molecular weight determination of either proteins or DNA by gel electrophoresis relies on which of the following:
- A constant charge to mass ratio of the particles.
 - migration of charged molecules in an electric field.
 - slower migration of larger particles due to the gel.
 - d) all of the above.**
12. The subunit molecular weight as well as the number of subunits in the quaternary structure can be determined by:
- SDS-PAGE electrophoresis.
 - Affinity chromatography.
 - Gel filtration chromatography.
 - d) combining information from a) and c).**
13. Cholesterol is essential for normal membrane functions because it
- spans the thickness of the bilayer.
 - b) keeps membranes fluid.**
 - catalyzes lipid flip-flop in the bilayer.
 - plugs up the cardiac arteries of older men, including Dr. Rule.
14. Which of the following membrane structures function in active transport?
- peripheral proteins.
 - cytochrome C.
 - c) integral proteins.**
 - Coenzyme Q.
15. The glycosidic bond
- is found in oligosaccharides.
 - is found in DNA and RNA.
 - is found in proteins.
 - d) both a) and b) are correct.**
16. Which of the following are forms of energy storage in biochemical systems?
- NADH & FADH₂
 - Concentration gradients across membranes.
 - ATP.
 - d) all of the above.**
17. The hormones, glucagon and epinephrine, stimulate glycogen breakdown to G-1-P
- directly, by binding to glycogen phosphorylase.
 - b) indirectly, by first stimulating adenylate cyclase to make cAMP.**
 - only in the liver.
 - only in muscle cells.
18. The sequence of DNA and RNA molecules is always written:
- a) beginning at the 5' end and going to the 3' end.**
 - beginning at the 3' end and going to the 5' end.
 - both of the above are acceptable.
 - neither of the above.

19. The major reason for A pairing with U is:
a) complementary hydrogen bonds.
 b) a purine-pyrimidine pair fits well in the double helix.
 c) efficient stacking of this arrangement of bases in the helix.
 d) recognition of non-'Watson-Crick' hydrogen bonds by DNA polymerases
20. An expression vector or expression plasmid
 a) always contains an origin of replication.
 b) usually contains a gene that confers antibiotic resistance to the bacterial host.
 c) always contains DNA segments for the regulation of mRNA production.
d) all of the above.
21. Restriction fragments cut with *Sau3A* (X/GATCX; X is any base) and *BamHI* (G/GATCC)
 a) can be efficiently joined because they both cut the DNA at the same location.
b) can be efficiently joined because they both give the same cohesive ends.
 c) can not be joined because they both give different cohesive ends.
 d) can be efficiently joined because they both begin with G and end with C.
22. Replication in *E. coli* is initiated by the generation of short RNA primers because
 a) RNA polymerase is a more efficient enzyme at starting replication.
b) DNA polymerases require a priming strand with a 3'-OH group.
 c) The initial error rate is lower with RNA than with DNA.
 d) DNA polymerases don't become activated until phosphorylated by DnaA.
23. During replication, overwinding or overtightening of DNA is removed by:
 a) Ribosomes.
 b) DNA polymerase.
 c) DnaB.
d) Gyrase.
24. A promoter is a
 a) a specific sequence of DNA to which core RNA polymerase binds.
b) a specific sequence of DNA to which holo RNA polymerase binds.
 c) a specific sequence of DNA to which a repressor binds.
 d) a specific DNA sequence to which DnaA binds.

Part B: Please do All of the following Problems.

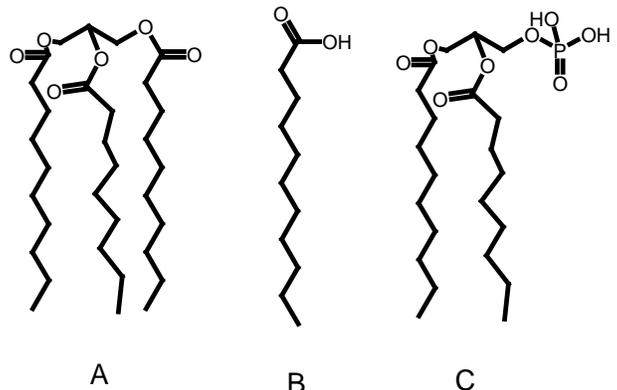
B1: (5 pts) Three molecules are shown to the right, labeled A, B, and C.

i) Below are the names of two lipid molecules. Write the letter that corresponds to the structure of the molecule.

Triglyceride: A

Phospholipid: C

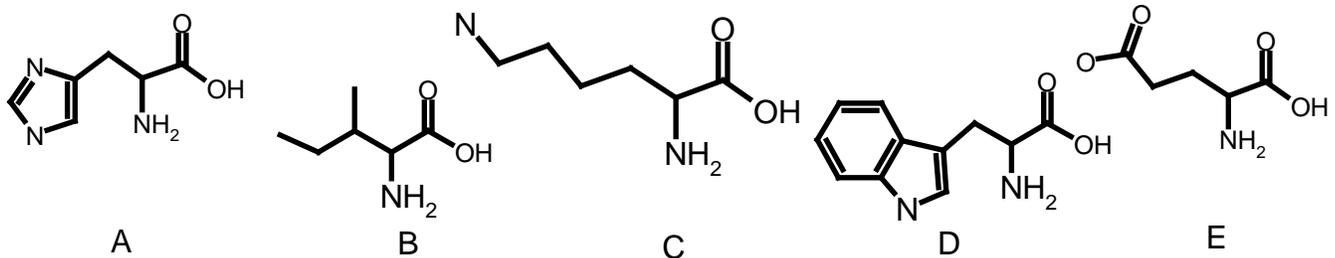
ii) Which compound(s) is(are) charged at neutral pH? B & C



B2: (4 pts) The structural hierarchy of proteins and nucleic acids (i.e. DNA and RNA) ranges from primary to quaternary structure. The following table lists this hierarchy for proteins and Nucleic acids. Two of the entries have been provided, supply **four of the remaining 6**. You can give either a short description or provide an example.

	Primary	Secondary	Tertiary	Quaternary
Protein	Seq of amino acids	α -helix/ β -strand	<i>Folded form of proteins, ie. myoglobin</i>	Hemoglobin
Nucleic Acids	Seq of bases	duplex DNA	tRNA	<i>Protein-nucleic acid complexes</i>

B3: (10 pts) The following are the structure of five amino acids.:



i) Write, underneath **two(2)** of the structures, the name of the amino acid (three letter code is quite acceptable)

A=His, B=Ile, C=Lys, D=Trp, E=Glu

ii) For the following statements, write the letter of the amino acid to which the statement best applies. Note, you may want to use the same amino acid more than once. If two (or more) are equally acceptable, write both of them. In some cases none of these may apply. If so, just write 'none'.

The best answer is given first, the remaining answers were acceptable

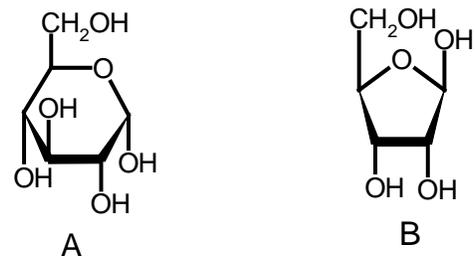
- a) Sidechain is hydrophobic and found buried in the core of protein: B,D
- b) Sidechain absorbs UV light strongly: D
- c) Sidechain found at active sites of some enzymes: A,E
- d) Sidechain can form electrostatic interactions with DNA: C,A
- e) Sidechain can form hydrogen bonds with nucleotide bases: E,A,C,D
- f) Sidechain that has a pK_a of around 4: E
- g) Sidechain that has a pK_a of around 7: A
- h) Can measure its sidechain pK_a with NMR: A,C,E,D

B4. (5 pts)

i) Which of the following two sugars is ribose (A or B)? **B is Ribose**

ii) Draw a box around the 2-OH of ribose. **Start counting at the oxygen, clockwise. The second carbon bears the 2-OH**

iii) Circle the carbons on ribose that participate in the formation of phosphate esters in the backbone of DNA and RNA. **The 3 and 5 positions, using the numbering discussed in part ii.**



B5. (14 pts) A list of energetic terms or 'forces' that play a central role in Biochemistry is given below:

1. Hydrogen bond
2. _____
3. Electrostatics
4. Van der Waals
5. Configurational Entropy

i) Line 2 is blank, which *very important* energetic term has been omitted?(2pts) **Hydrophobic effect.**

ii) Pick **any two of the above five** and give a brief description of its molecular nature in relationship to biochemical structures. (4 pts)

Choice #1

Hydrogen bond: electrostatic interaction between partial positive charge on hydrogen on partial negative charge on Oxygen (ie. C=O in proteins) or nitrogen (ring nitrogen in bases)

Choice #2:

Hydrophobic effect - Entropy gained by exclusion of water from surface of non-polar groups

Choice #3:

Electrostatics. Attraction between unlike charges, repulsion between like charges. Strength depends on dielectric constant

Choice #4:

Van der Waals - induced dipole in uncharged atoms, leading to attractive force

Choice #5

Configurational entropy. $S=R \ln W$. Denatured state favoured due to + entropy as a result of a large number of configurations of a chain.

iii) Compare and contrast the role of **two(2) of the above five** energetic terms in their contribution to protein and DNA Stability. Use the following table to guide your answer. Indicate in the left column the energetic terms you have selected to discuss (Each box is worth two points, *be brief*).

Force or Energetic Term:	Protein Stability	DNA Stability
Hydrogen bonds	Slight stabilization of folded form. The high energetic cost if they are not made is responsible for specificity of hydrogen bonding in secondary structure	Ditto
Hydrophobic effect	Very important in stabilization of folded form due to the large number of non-polar groups buried when proteins fold.	Less important for stability
Electrostatics	Play a very weak role in the stabilization of proteins.	Very destabilizing due to the charge on phosphate
Van der Waals	Moderately important for folding due to close packing of side-chains in the interior of the protein	Very important, in the form of pi-pi stacking.
Configurational Entropy	Large number of possible conformations in the unfolded state decrease the stability of the folded state	ditto

B6. (14 pts) Due to incompatible blood types there is always a shortage of human blood for medical treatment. Your responses to the following questions will lead you through the steps involved in the production of human hemoglobin in bacteria (This has actually been done by Dr. Chien Ho at CMU).

i) Briefly describe the steps involved in obtaining the required DNA fragments for the construction of the library. Discuss how you would obtain the hybridization probe to identify the hemoglobin genes and the requirement for an antibiotic resistance gene on the plasmid (4 pts).

i) *Cut up DNA with restriction enzyme*

ii) *Ligate into plasmid with an origin of replication and a gene that encodes an antibiotic.*

a) *origin so that it can replicate inside bacterial cells*

b) *antibiotic so that cells containing the expression plasmid can be identified*

iii) *Use known protein sequence to determine the sequence of the DNA probe.*

iv) *Probe library with DNA probe.*

ii) The final expression plasmid has to have several elements, or sequences, that control both the production of mRNA as well as protein. The line drawn below represents the DNA sequence of the expression plasmid that contains the hemoglobin gene as well as the required regulatory features for mRNA and protein synthesis. The start of mRNA synthesis is indicated by the →. Place the remaining regulatory features on this diagram in the correct order and with approximately the correct spacing. Draw a **neat** square or rectangle around those that are involved in mRNA production (or its regulation) and a circle around those that are involved in protein synthesis (or its regulation) (8 pts).

a) Termination codon

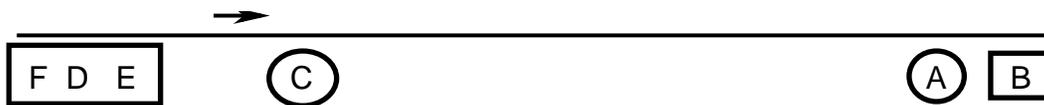
e) operator site

b) mRNA termination site (ρ-factor binding)

f) -35 region of the promoter

c) Start codon

d) -10 region of the promoter



iii) High levels of recombinant hemoglobin would be lethal for the cell. Briefly describe how you would control the expression of the hemoglobin in the bacteria.(2 pts)

The operator site would be the lac operator. It would bind the lac repressor protein in the absence of lactose (or IPTG) preventing the formation of mRNA. Addition of IPTG would induce expression.

iv) *Briefly describe how you would purify this recombinant protein from bacteria *in a single step*. You can assume that both the α and β chains are synthesized in the same bacterial cell. In addition, you can assume that the bacteria cannot provide sufficient heme, thus most of the recombinant hemoglobin in the lysate will be without heme.(2pts).

Affinity chromatography with heme as the ligand on the column

B7. (9 pts).

i) Sketch the tertiary structure of a tRNA molecule, indicating the location of the anti-codon loop and the acceptor stem.(2 pts).

Looking for an 'upside down' L, with the anticodon loop at the bottom and the acceptor stem on the top:

ii) The assembly step in protein synthesis results in the formation of the complex between the 70S ribosome, the mRNA, and fMet-tRNA in the peptidyl(P) site. A representation of this complex is given in the left-most figure at the bottom of this page. The diagram, proceeding from left to right, show the complete sequence of steps in elongation of the peptide by one amino acid. The first and last diagrams have been completed for you. Do the following:

A) Complete the drawing of the two middle steps.(2 pts) **See the lecture notes on protein synthesis.**

B) In one of the diagrams include the following labels, in their proper place:(3 pts) **See diagram.**

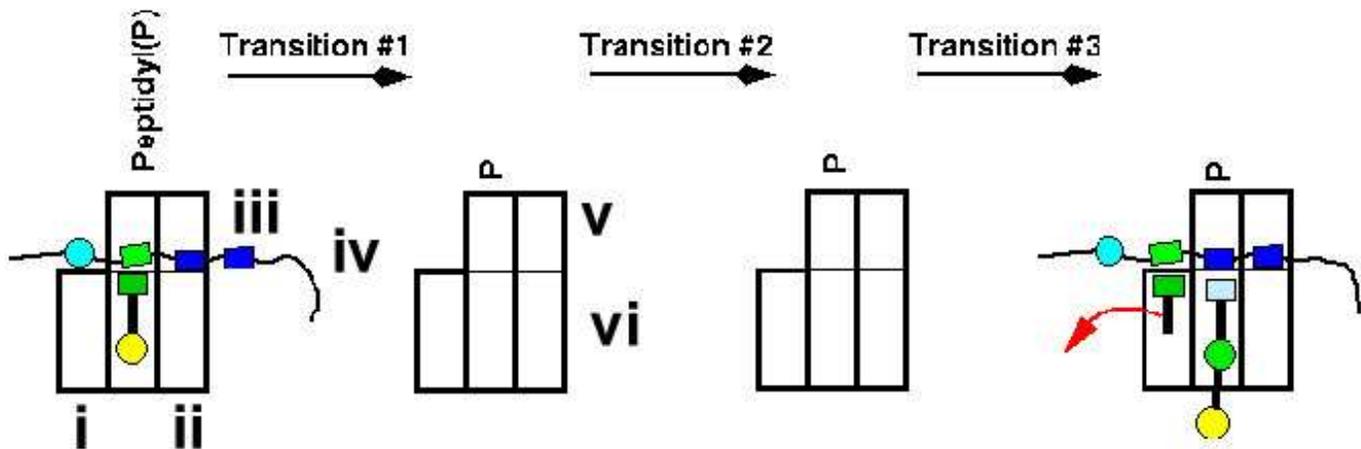
- i) Exit Site (E)
- ii) Aminoacyl Site(A)
- iii) Codon (Start or otherwise)
- iv) mRNA
- v) 30S subunit of the ribosome
- vi) 50S subunit of the ribosome.

C) There are three transitions between these four complexes. For each of these transitions, state one of the important events that occurs during the transition. The important events associated with the 1st transition are provided as an example, you need to do only transition #2 and transition #3.(2 pts)

Transition #1: *Charged tRNA binds to the aminoacyl-site, basepairing with the 2nd codon.*

Transition #2: *Formation of peptide bond*

Transition #3: *Translocation of ribosome*



B8 (8 pts): Enzyme Inhibition.

i) How do the enzyme binding sites for a competitive inhibitor and a non-competitive inhibitor differ?

*Competitive inhibitors bind to the active site
Non-competitive inhibitors bind elsewhere*

ii) Explain why a competitive inhibitor can only affect K_M and not V_{MAX} . A cartoon diagram of an enzyme might be helpful.

V_{max} is a property of the enzyme-substrate complex. Since the inhibitor can't be bound when substrate is bound, it cannot affect V_{max} . However, a higher substrate concentration is required to reach the same concentration of $[ES]$, therefore K_m is affected.

iii) Which type of inhibitor is more likely to serve as a feedback inhibitor in metabolic pathways? Why?

Since feedback inhibitors are those which are near the end of pathways, yet regulatory steps are near the beginning of pathways, the feedback inhibitor will NOT look like the substrate. Thus a non-competitive inhibitor is more likely.

B9:(8 pts) Enzyme Kinetics

The following is a plot of the initial velocity of isocitrate dehydrogenase (a TCA cycle enzyme) as a function of substrate concentration. No ATP or ADP was present in this reaction.

i) Estimate K_M from this graph. Explain your approach.

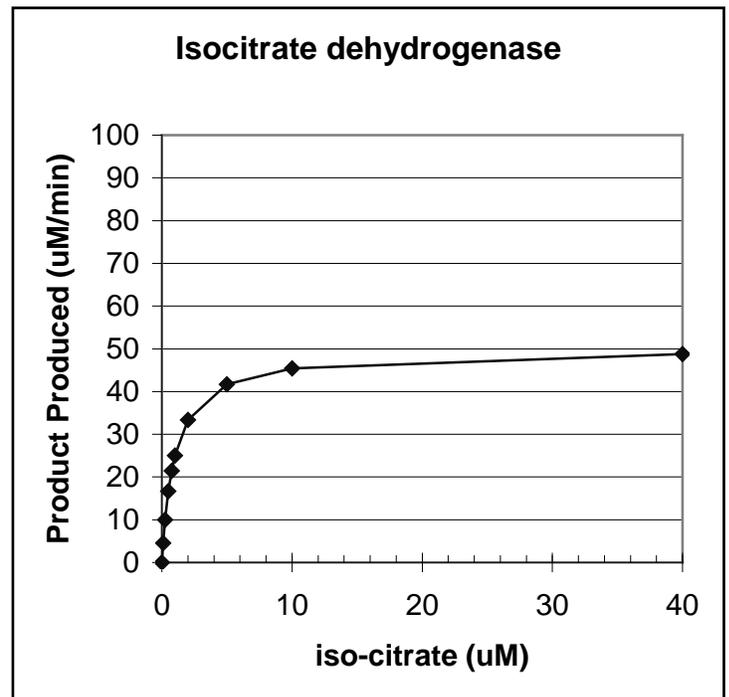
V_{max} is about 50 $\mu\text{M}/\text{min}$. K_m is the substrate concentration at $V_{max}/2$, or about 1 μM in this case.

ii) Sketch, using a **dashed line**, the data you would expect to obtain in the presence of ATP. Briefly justify your answer.

Since ATP is an inhibitor of isocitrate dehydrogenase, the dashed line would be below the curve (not shown).

iii) Briefly describe how you would obtain the K_M for this enzyme using a Lineweaver-Burk plot (a simple sketch would be fine).

A plot of $1/V$ versus $1/S$ is the Lineweaver-Burk plot. The x-intercept of this plot is $1/K_M$



Part C: Choice Questions:**Do either C1 or C2:(8 pts)**

C1. The metabolism of glycogen synthesis and degradation is tightly coupled to glucose synthesis and degradation in both liver and muscle tissues. However, glycogen metabolism is controlled by phosphorylation/dephosphorylation of enzymes while glucose metabolism is controlled by the effects of levels of fructose-2,6-phosphate (F2,6P,FBisP) on the activities of phosphofructokinase (PFK). Briefly describe how these two pathways are coordinately regulated. The following details may be useful in your discussion. F2,6P is synthesized from F-6-P by the enzyme phosphofructose kinase 2 (PFK-2) and degraded to F-6-P by the enzyme fructose-bis-phosphatase-2 (FBPase-2).

C1: PFK2 and FBPase2, like the enzymes of glycogen synthesis are regulated by phosphorylation. During glycogen degradation protein kinases are active and, in the liver, gluconeogenesis is also enhanced.

Since FBPase-2 is activated by phosphorylation the levels of F2,6 P will drop. this will inhibit PFK in glycolysis and activate PFKase in glyconeogenesis.

C2. Draw a simple diagram that illustrates the oxidative fate of the principle components of a bagel with cream cheese (i.e. glucose from the bagel, fatty acids and amino acids from the cream cheese). Your diagram should resemble a flow chart, showing only the *names* of the major metabolic pathways and how they are connected. The top of your 'flow chart' should begin with the **three** nutrients. The bottom of your 'flow chart' should end with H₂O.

Glucose ->glycolysis->pyruvate->acetyl-CoA->TCA

Fatty acids->beta-oxidation->acetyl-CoA->TCA

Amino acids->TCA

NADH+ FADH₂ from TCA, beta-oxidation & glycolysis ---->electron transport -->H₂O

Do either question C3 or C4 (8 pts)**C3.**

i) How frequently does the sequence encoding the start codon (AUG) occur in DNA? How does *E. coli* select the correct starting position to begin translation of mRNA into protein?

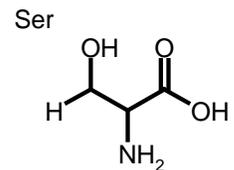
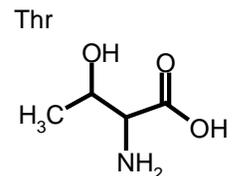
AUG occurs once every 64 basepairs. E. coli uses the SD sequence to position the mRNA on the ribosome.

ii) The charging of tRNAs with amino acids is a two step process. Draw, sketch, discuss, etc. these steps using the charging of tRNA^{Thr} with Thr as an example.

A) $Thr + ATP \rightarrow Thr-AMP \text{ (phosphate ester)} + PPi$

B) $Thr-AMP + tRNA^{thr} \rightarrow Thr-tRNA^{thr}$

iii) How does the cell avoid charging tRNA^{Thr} with Ser? Give a *plausible* description of the active site of the enzyme that charges tRNA^{Thr} (i.e. indicate the amino acids that might be present in the active site of the enzyme). Your description should discuss both activities present in the active site (hint: hydrolytic editing). The structures of Thr and Ser are given to the right.



The active site for step A above can accept either Thr or Ser. There might be a non-polar interaction with the methyl of Thr that would enhance the formation of Thr-AMP over that of Ser-AMP.

The editing site would be too small for Thr to fit into, thus preferentially hydrolyzing Ser-AMP.

C4. DNA synthesis on the leading strand is continuous, while DNA synthesis on the lagging strand is not. Discuss how *E. coli* accomplishes synthesis of the lagging strand. Briefly discuss the enzymes involved and highlight the important differences and similarities between the two DNA polymerases.

- i) DnaB opens up replication fork, exposing single stranded DNA on the lagging strand.*
- ii) Primase lays down an RNA primer*
- iii) PolIII synthesizes DNA in a processive manner until it reaches next RNA primer. PolIII cannot remove this primer because it has no 5'-3' exonuclease activity.*
- iv) PolIII exits*
- v) PolI, with its 5'-3' exonuclease activity, removes RNA primer and fills the resultant 'gap' with DNA*
- vi) PolI exits, being non-processive*
- vii) DNA ligase seals the gap.*

Do either question C5 or C6:

C5. (6 pts) Using either hemoglobin or phosphofructokinase (PFK) as an example, discuss the following aspects of allosteric effects in biological systems (Note, you need not use the same example when answering each part)

i) Necessary structural requirements of the protein for it to exhibit allosteric behavior (i.e. nature of binding sites, etc.).

Must have two binding sites, one catalytic and one regulatory. They can exist on the same peptide or on separate peptides. Examples, in hemoglobin, bisPG binds to α/β interface affecting oxygen binding at a distant site. PFK has a 2nd nucleotide binding site, so that AMP can regulate the enzyme activity.

ii) How the activity of a protein is affected by allosteric effects?

Conformational change due to binding of the allosteric compound binding. Hemoglobin is the clearest example, binding of bisPG stabilizes the low affinity conformation (T state) over the high affinity R state.

iii) Give an example of an allosteric effector and a brief description of its role in the regulation of the biological process performed by hemoglobin or PFK.

In hemoglobin, bisPG decreases oxygen affinity and its concentrations are used to fine tune oxygen affinity depending on the altitude.

In PFK ATP inhibits PFK to reduce glycolysis and energy production. AMP and F26 activate to enhance energy production.

C6 (6 pts): Discuss **two of the following three** features of enzyme catalyzed reactions. Indicate how these features are important for catalysis and provide a specific example of this feature in an existing enzyme. In your examples, you can use any enzyme you like (including ones not discussed in class) and you need not use the same enzyme for both of your answers.

i) Transition state stabilization.

The transition state is lowered in enzyme catalyzed reactions. Example, carbocation in lysozyme is stabilized by Asp. In the serine proteases the oxyanion hole stabilizes the tetrahedral intermediate.

ii) Substrate Specificity

Specific amino-acid side chains recognize specific groups on the substrate. Example, in trypsin, the Asp residue in the specificity pocket selects Arg/Lys over non-polar amino acids.

iii) Chemical mechanism.

Specific, usually polar, amino acids provide functional groups for catalysis. Ser in serine proteases, Glu in lysozyme, Asp in HIV protease

Do either question C7 or C8: (16 pts)

C7. The binding constant of single-stranded binding protein (SSB) to single-stranded DNA was measured in the following fashion. A number of different concentrations of DNA were added to a 1 μM solution of SSB and the absorbance of the solution was measured at each concentration of the DNA. The original absorbance of the protein solution was 1 absorbance unit. The **change** in absorbance of the protein, due to the addition of the DNA, is given in the 2nd column. A value of 100% would indicate protein that is fully complexed with DNA. The fractional saturation (Y) is also given in the 2nd column.

DNA Concentration (μM)	Change in Absorbance (%)	$\text{Log}(Y/(1-Y))$	$\text{log}[L]$
0	0 (Y=0.00)	-	-
0.1	2 (Y=0.02)	-1.70	-7
1.0	50 (Y=0.50)	0.00	-6
10.0	98 (Y=0.98)	1.70	-5
100.0	99 (Y=0.99)	2.95	-4

i) What wavelength of ultra-violet (UV) light was used in this experiment, 260nm or 280nm? (2pt)

Since the absorption of the protein is being monitored, 280 nm would be used.

ii) What is the extinction coefficient for this protein? (2 pts)

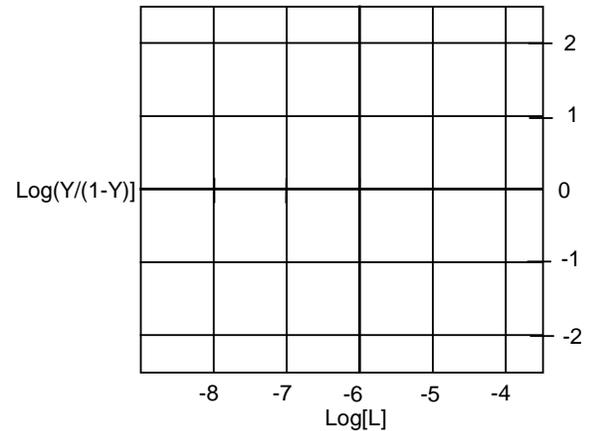
$$A = c\epsilon l, \quad \epsilon = 10^6 \text{ L/M-cm}$$

iii) Estimate the binding constant from these data, assuming that the binding is **not** cooperative. (Hint: You do not need to draw a graph of any sort here). Explain your approach. (4 pts)

For non-cooperative binding, the ligand concentration at Y=0.5 is the Kd. 1 μM in this case.

iv) Does the binding of DNA to this protein display either positive or negative cooperative behavior? Justify your answer. You should **not** have to do a Hill plot to determine this. However, if you must, the appropriate data ($\text{log}[y/(1-y)]$, $\text{log}[L]$) are given in the last two columns of the above table. Use the graph to the right. (4 pts)

A 10 fold increase in ligand binding (from 1 μM to 10 μM) results in an increase in Y from 0.5 to 0.98. This indicates positive cooperativity. In the case of non-cooperative binding the increase in Y for a 10 fold change in ligand concentration would be from 0.5 to 0.9.



v) Sketch, in the box to the right, the Hill plot you would expect to obtain from these data. You **do not** have to plot the data, just indicate the approximate value of the slope at $\text{log}(\theta)=0$. If you did this plot in part iv, go on. (2 pts)

The curve would intersect the $\text{log}(y/(1-y))=0$ line when $\text{log}[L]=-6$. The slope of this line would be greater than one (positive cooperativity), 1.7 to be exact.

vi) Given the role of SSB in DNA replication, does the particular DNA sequence used for this experiment matter? Why? (1 pt)

No, since SSB binds to DNA without any sequence specificity.

viii)**Provide a plausible model for the cooperative behavior observed here.(1 pt)

- a) Multiple binding sites (i.e. like hemoglobin)
- b) Binding of one SSB distorts the structure of the DNA to facilitate the binding of another
- c) Favourable protein-protein interactions between SSB proteins bound to DNA

C8 (16 pts): The restriction enzyme ZuluI binds to and cleaves the following DNA sequence: GATC. Below is a diagram of the interaction between an Asn residue of this restriction endonuclease and a TA base pair in its recognition sequence. The position of this base pair in the recognition sequence is indicated in bold and underlined (GAT**C**). Drawn below the TA base pair is a CG base pair to help you visualize what happens when the TA base pair is replaced by a CG basepair.

i) Label the major groove of the TA base pair.(1 pt) **Top part of bases.**

ii) Draw in the 'Watson-Crick' hydrogen bonds with **dotted** lines (1pt). **See diagram.**

iii) Draw in any hydrogen bonds that might occur between the protein and the DNA with **solid** lines.(1 pt) **See diagram.**

iv) The binding constant, K_{EQ} , of ZuluI to the sequence GATC is $10^{10} M^{-1}$. Calculate ΔG° for this binding reaction (assume $RT=2.5$ kJ/mol; $\ln 10^{10} = 23$)(2 pts).

$$\Delta G^{\circ} = -RT \ln K_{eq}$$

$$= -57.5 \text{ kJ/mol}$$

v)*Replacement of the TA base-pair with a CG base pair will cause the loss of a hydrogen bond between ZuluI and the DNA. What is the energetic cost of this hydrogen bond in kJ/mol? (Be careful when you answer this - don't forget to consider the water molecules)(3 pts)

Loss would be -20 kJ/mol since the original hydrogen bond to the water (in the unbound protein) is not replaced by an hydrogen bond in the protein-DNA complex.

vi) Replacement of the TA base pair in the sequence GATC with a CG base pair will give the sequence GACC. Calculate the ΔG° **and** K_{EQ} of binding of ZuluI to GACC. Assume that the only difference in the interaction of ZuluI with the GACC versus its interaction with GATC is the loss of the hydrogen bond to the adenosine residue that was discussed in part v. (5 pts)

$$\Delta G = -57.5 \text{ kJ/mol} + 20 \text{ kJ/mol} = -37 \text{ kJ/mol}$$

$$K_{eq} = e^{+37.5/2.5} = 3.2 \times 10^6$$

vii) Using the above affinity constants, calculate the fraction of ZuluI bound to its normal sequence (GATC) and compare this to the fraction of ZuluI that would be bound to the sequence (GACC) (3 pts).

$$f_B = K_{eq} / (1 + K_{eq}).$$

$$\text{For GATC: } f_B = 10^{10} / (1 + 10^{10}) = 1.0000$$

$$\text{For GACC: } f_B = 3 \times 10^6 / (1 + 3 \times 10^6) = 0.99999$$

Although the loss of free energy is enormous, the binding is still quite tight due to the large initial free energy.

