Biochemistry 465 **Your Name:** Key

Biological Information Processing **Prof. Jason Kahn**

Exam II (100 points total) **April 20, 2006**

You have 80 minutes for this exam.

Exams written in pencil or erasable ink will not be re-graded under any circumstances.

Explanations should be concise and clear. I have given you more space than you should need.

You do not need a calculator for this exam, and no other study aids or materials are permitted.

Generous partial credit will be given, *i.e.*, if you don't know, guess.

Honor Pledge: At the end of the exam time, please write out the following sentence and sign it, or talk to me about it:

"I pledge on my honor that I have not given or received any unauthorized assistance on this examination."

1. Protein-DNA Interaction, Bioinformatics (30 pts):

(a; 12 pts) Here is the binding equilibrium for a protein that dimerizes when it binds the DNA, which is the case for leucine zipper proteins. For simplicity, we assume there is no single binding (i.e. no stable PD) and that total $[P] = [P]_{free}$.

$$P + P + D \underset{}{\overset{K_d}{\rightleftharpoons}} P_2D$$

$$K_d = \frac{[P]^2[D]}{[P_2D]}$$

Recalling that we know the total DNA concentration D_T and that all of the DNA must be either free or bound, substitute for $[D]$ to obtain an equation including only K_d, $[P]$, $[P_2D]$, and D_T. Solve for the fraction of DNA bound by protein. What concentration $[P]$ would give 50 % binding?

The conservation equation for DNA is $D_T = [D] + [P_2D]$. (+2)

Substituting, we have: $K_d = \frac{[P]^2 \left(D_T - [P_2D]\right)}{[P_2D]}$ (+3)

The fraction of DNA bound by protein is $\Theta = [P_2D] / D_T$ (+1)

So $K_d = \frac{[P]^2 \left(D_T - [P_2D]\right)}{[P_2D]} = \frac{[P]^2 D_T}{[P_2D]} - [P]^2 = \frac{[P]^2}{\Theta} - [P]^2$

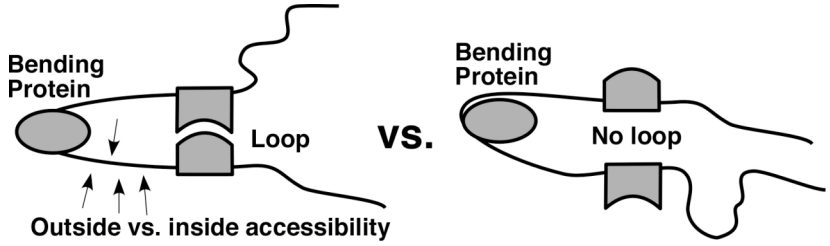
Rearranging, we have $\Theta = \frac{[P]^2}{[P]^2 + K_d}$ (+3), which should look familiar.

At $\Theta = 0.5$ we must have $[P]^2 = K_d$, so $[P] = \sqrt{K_d}$ (+3). K_d has units of M^2.

Score for the page_____

(b; 8 pts) DNA looping is important in transcriptional regulation. How can a bending protein exert "action at a distance" by binding within a loop? Looping can be detected by DNAse I footprinting. One can sometimes distinguish looping from independent binding to two sites by observing periodic changes in the footprinting of the DNA between the two sites. What do you think might cause these changes?

Bending can increase or decrease the free energy cost of loop formation, depending on the helical phasing of the bend direction with respect to the bending needed to form the loop. (+3) This does not require protein-protein contact (+2). We discussed this in the context of IHF modulation of NtrC looping. Periodic changes in the footprint are probably due to successive compression and expansion of the minor groove through the bent DNA in the loop (+3), by analogy with periodic changes in the footprints of positioned nucleosomes.



(c; 10 pts) The results below are from a BLAST search. What is the meaning of the "Expect" value? What do the "+" signs mean in the output? What about the dashes in the second query line?

```
Score =  201 bits (511),  Expect = 2e-50

Query  1    MRIILLGAPGAGKGTQAQFIMEKYGIPQISTGDMLRAAVKAGSELGLKAKEIMDAGKLVT  60
            MR++LLG P AGKGTQA  I+ KY IP ISTGDMLR+ +K G+ LG KAKE MD G LV
Sbjct  1    MRLLLLGPPSAGKGTQASGIVNKYHIPHISTGDMLRSNIKQGTALGNKAKEYMDQGLLVP  60

Query  61   DELVIALLKERITQEDCRDGFLLDGFPRTIPQA----DAMKEAGIKVDYVLEFDVPDELI  116
            DELV+A++++R+ Q+DC++GFLLDGFPRT+ QA    D + + G+ +D V+  +VP   +
Sbjct  61   DELVVAIVEDRLQQDDCQEGFLLDGFPRTVVQAKALDDVLDKMGVTLDKVVSIEVPKGTL  120
```

Too many points for this question, given to reward doing the homework. Note that the actual alignment is longer.

The "Expect" value is the probability of the observed match occurring by chance: the observed match here is clearly not random! (+4)

The + signs indicate conservative substitutions (+3) of amino acids in the protein.

The dashes in the query line indicate a gap in the alignment (+3). What might a 4-amino acid gap mean in terms of protein 3D structure?

Score for the page_____

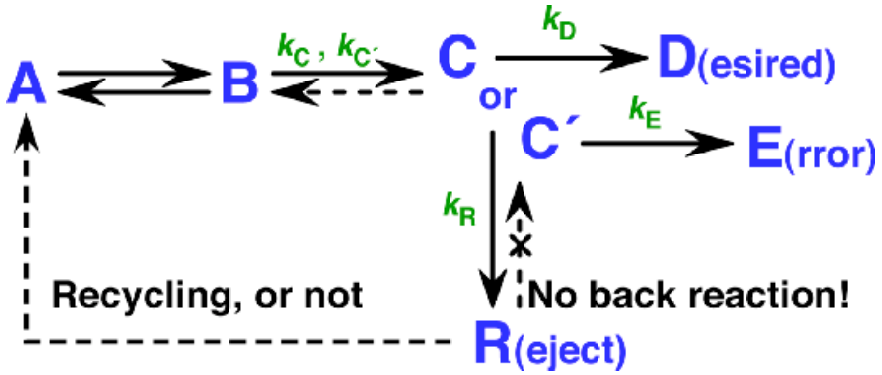
2. DNA Replication (39 pts):

(a; 4 pts) What aspect of the replication fork is captured by the phrase "trombone model?"

The dimeric Pol III complex (+1) anchors a lagging strand loop that changes in size (+2) during Okazaki fragment synthesis (+1).

(b; 11 pts) Why does kinetic proofreading require an irreversible step? What is the irreversible step in DNA polymerase proofreading? How can DNA polymerase proofreading be described in terms of a molecular clock?

From the class web page:



Irreversibility in the error rejection branch of the pathway is required to prevent "backflow" through the error rejection pathway that would otherwise introduce errors. (+3)

The irreversible step in DNA polymerase fidelity is the hydrolysis of a phosphodiester bond by the 3´→5´ exonuclease activity of the polymerase (C or C´→R). (+3)

The idea of a molecular clock is that an excited state of a macromolecule decays over time. If the next step in the pathway is faster than the decay, it is likely to occur (+2). Otherwise, the excited state decays and the system starts again. In this case, the exonuclease reaction is the timer and the next step is extension of the primer terminus. If extension is faster than hydrolysis, the polymerase moves forward. If extension is slow (i.e. at a mismatched terminus), then the clocks ticks down before is occurs and the last nucleotide is removed.(+3)

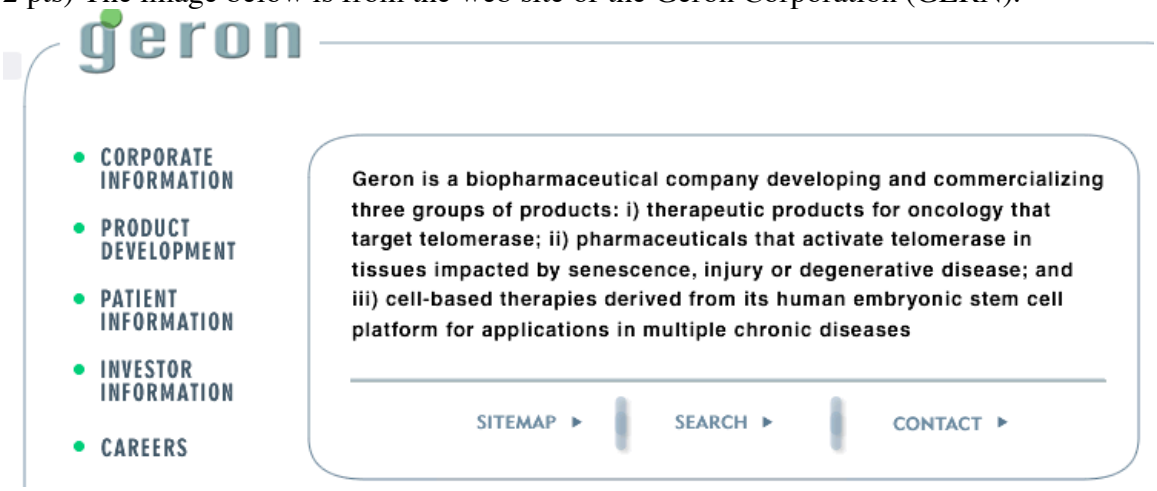
Score for the page_____

(c; 6 pts) Why did primers for lagging strand replication evolve to be made of RNA?

The early steps in polymerization are likely to be low-fidelity, as there is no helix to build on. (+3)
By making the primers RNA, they can be recognized and removed to avoid creating mismatches. (+3)

The "frozen accident" explanation for RNA primers was given +2.

(d; 12 pts) The image below is from the web site of the Geron Corporation (GERN).



Why would one want to treat cancer by inhibiting telomerase? What might be some of the risks, especially for young people? Why might one want to activate telomerase? What are the risks?

1. Inhibition of telomerase might prevent cancer cells from proliferating, as eventually they would hit the end-replication wall. Cancer cells are often observed to make telomerase. (+3)
2. But inhibition of telomerase would prevent its appropriate activation, especially in the germ line. So one's children would be born old. (+3)

3. Activation of telomerase would allow additional proliferation of tissue that other wise might senesce or die. So one could live forever. (+3)
4. But see (1) above. Activation of telomerase could potentiate tumor growth. So one might live forever as a cancer patient (+3)

Score for the page_____

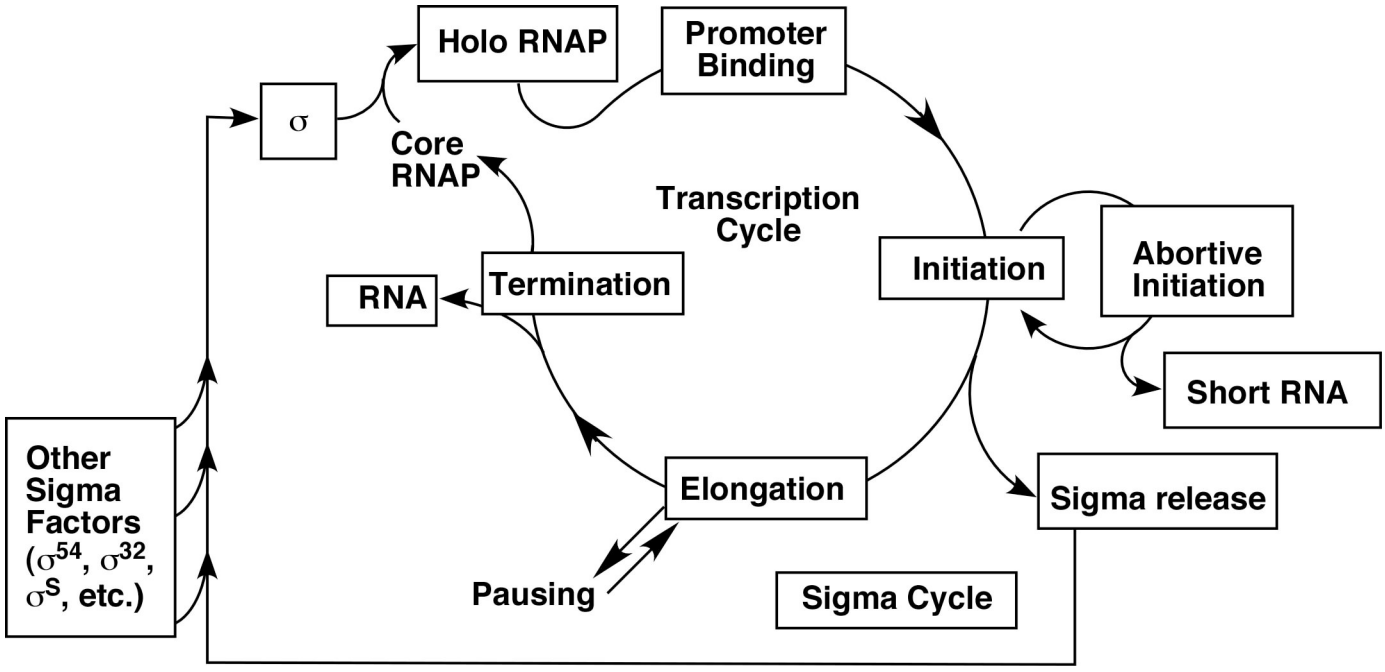
(e; 6 pts) Explain why a eukaryotic cell that has lost checkpoint control does not die, but may become hypersensitive to DNA damage agents.

The fundamental cyclin/CDK relay system for moving the cell cycle forward can work even when checkpoints are defective. If nothing is wrong with the DNA and the cell is healthy, this may well lead to successful cell division. (+3)

However, if the DNA is damaged and the checkpoints are absent, the cell is will attempt to divide even though the DNA is mutated or there are crosslinks, DSB's or whatever that have not been repaired. This will lead to chromosome mis-segregation or a high mutation rate, and eventually death (+3).

3. Transcription and Regulation (31 pts):

(a; 12 pts) Fill in the boxes in the schematic below. The same answer may appear more than once.



+1 per box except +2 each for Abortive Initiation and for Other Sigma Factors. Promoter Binding and Initiation were filled in for you.

Score for the page_____

(b; 6 pts) We speculated that abortive initiation may be mechanistically inevitable. What is abortive initiation? Why is promoter escape difficult?

Abortive initiation is the reiterative synthesis and release of short (<10 nt) RNA chains by a polymerase that's still at the promoter. (+3)

In order to escape from the promoter, many contacts that used to stabilize the open complex need to be broken (+3): sigma factor is loosened/released, the specific contacts to -35 and -10 and UP elements need to be broken. There should be some significant activation energy for breaking all these contacts.

(c; 13 pts) What is an AAA+ protein? One example we have looked at is the transcriptional activator NtrC, which is part of a two-component signalling system. What is the name for the downstream partner (i.e. NtrC) in these systems? Describe some evidence that the mechanistic basis of NtrC activation of the Eσ^{54} holoenzyme is that it alters the conformation of σ^{54}, as opposed to NtrC simply acting as a helicase to open up the DNA. Finally, how is the NtrC activation signal turned off?

An AAA+ protein uses the free energy of ATP hydrolysis to drive conformational change in a second macromolecule. (+3)

The downstream partner is the response regulator (upstream is the sensor kinase). (+2)

There are mutants of σ^{54} that are constitutively active, so it appears that the protein is autoinhibited rather than inherently incapable of open complex formation. There is mutagenesis, cryo_EM, and X-ray crystal structure data that show a direct interaction between σ^{54} and NtrC or related proteins. (+3 for any of these)

NtrC signalling shuts off when the activating aspartyl-phosphate residue hydrolyzes (+2), and then when glutamine levels recover NtrB ceases re-phosphorylating NtrC (+3).

Page	Score
1	
2	
3	
4	
5	
6	
Total	

Score for the page_____