Biochemistry 661 **Your Name:** Key

Nucleic Acids, Module I **Prof. Jason Kahn**

Final Exam (200 points total) **October 14, 2010**

You have 90 minutes for this exam.

Explanations should be concise and clear. I have given you more space than you should need. There is a extra space on the last page if you need it.

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 examination 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. Methods (40 pts):

(a; 24 pts) Explain why the electrophoretic mobility shift assay is better than footprinting for looking for very weak nucleic acid binding by proteins. Explain why footprinting requires much hotter DNA or RNA than EMSA (think about what the gels look like). On the other hand, list one of the several major advantages of footprinting over EMSA in terms of the methods' results or applicability.

(+8) - EMSA can detect just a few % bound or less, because one is looking for a shifted band vs very low background, whereas with footprinting we are looking at a decrease in a cleavage signal. It's like the difference between fluorescence and absorbance.

(+8) - In footprinting, label is spread out among dozens to hundreds of bands, ~~whereas~~ whereas with EMSA it is in two to several. Therefore to get ~~reasonable~~ reasonable signal over background requires hotter label for footprinting. (Higher specific activity or multiple labels near one end.)

(+8 for any)
- Footprinting locates where a protein binds on a fragment.
- " can be performed under equilibrium conditions – does not require a gel-stable complex.
- Footprinting can be done in vivo.

Score for the page

(b; 16 pts) Briefly describe the essential steps of the Roche 454 method of next-gen sequencing.

(+4) – DNA is sheared and known adapters are ligated to the ends

– The DNA is captured on a bead by virtue of binding a complementary oligonucleotide.

(+6) – Critical step: the beads are dispersed into an emulsion such that PCR can be done in the emulsion, separately on each bead.

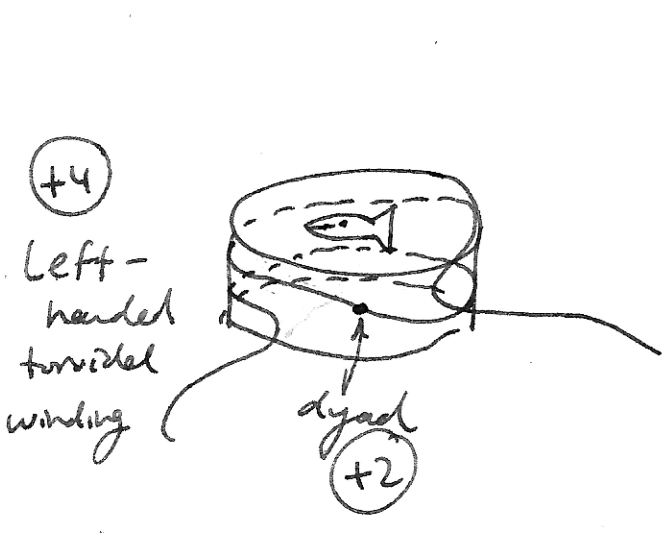
(+2) – Beads are distributed into individual wells of a plate

(+4) – Pyrosequencing is performed, such that there is a flash of light resulting from PP_i release

– Sequence is done on each well simultaneously by tracking the light from each well with a video camera.

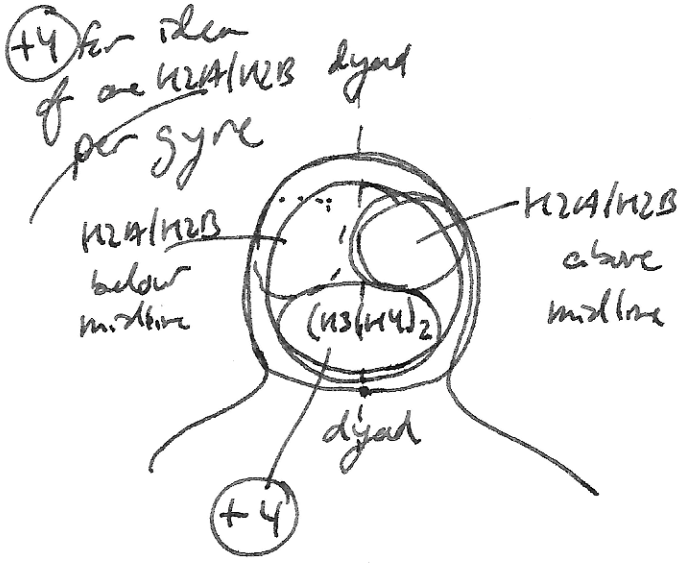
2. Nucleosome Structure and Stability (48 pts):

(a; 16 pts) Draw two sketches of the structure of the Nucleosome, one showing DNA gyres on a tuna can and one showing the disposition of the histones from above. Draw the $(H3/H4)_2$ tetramer as one large oval blob and each H2A/H2B dimer as a smaller circular blob.



(+4)

(+2)



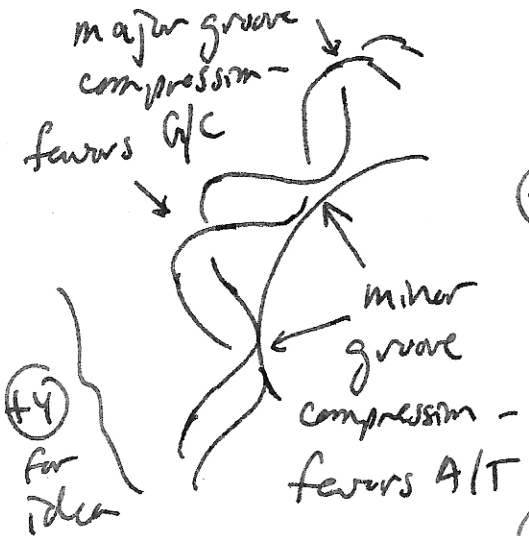
(+4) for idea of one H2A/H2B dyad per gyre

(+4)

(+2) for overall neatness

Score for the page

(b; 16 pts) Explain the origin of the observed periodicity of AT-rich vs. GC-rich sequences in sequences that form rotationally positioned nucleosomes. The periodicity in and of itself provides evidence for overtwisting in the nucleosome, because the average separation between e.g. AA dinucleotides is smaller than in free DNA. Physically, why does it make sense that nucleosome-bound DNA can have a smaller helical repeat? How does overtwisting resolve at least most of the "nucleosome paradox?"



- If A/T rich and G/C-rich regions alternate, the intrinsic bendability preferences make the best match to the shape needed to bind the nucleosome, and the sequence will have a preferred inward and outward orientation. (+4)

- Charge neutralization by the Arg/Lys residues ~~help~~ allow phosphates to be closer together. (+4)

(+4) $-(\pm)\Delta Tw + (\ominus) Wr = \ominus Lk$ but $|\Delta Lk| < |Wr|$

(c; 16 pts) Originally it was thought that nucleosome acetylation might have its effect via dramatic conformational changes in the nucleosome. This turned out not to be the case: what is the actual role (in general) of histone modification? What pair of enzymatic activities is generally associated with activation and repression of transcription? What method is used to track changes in histone modification state across a gene?

(+6) - Histone modifications allow for asynchronous communication among modifying enzymes and transcription factors and the transcriptional machinery - a "histone code".

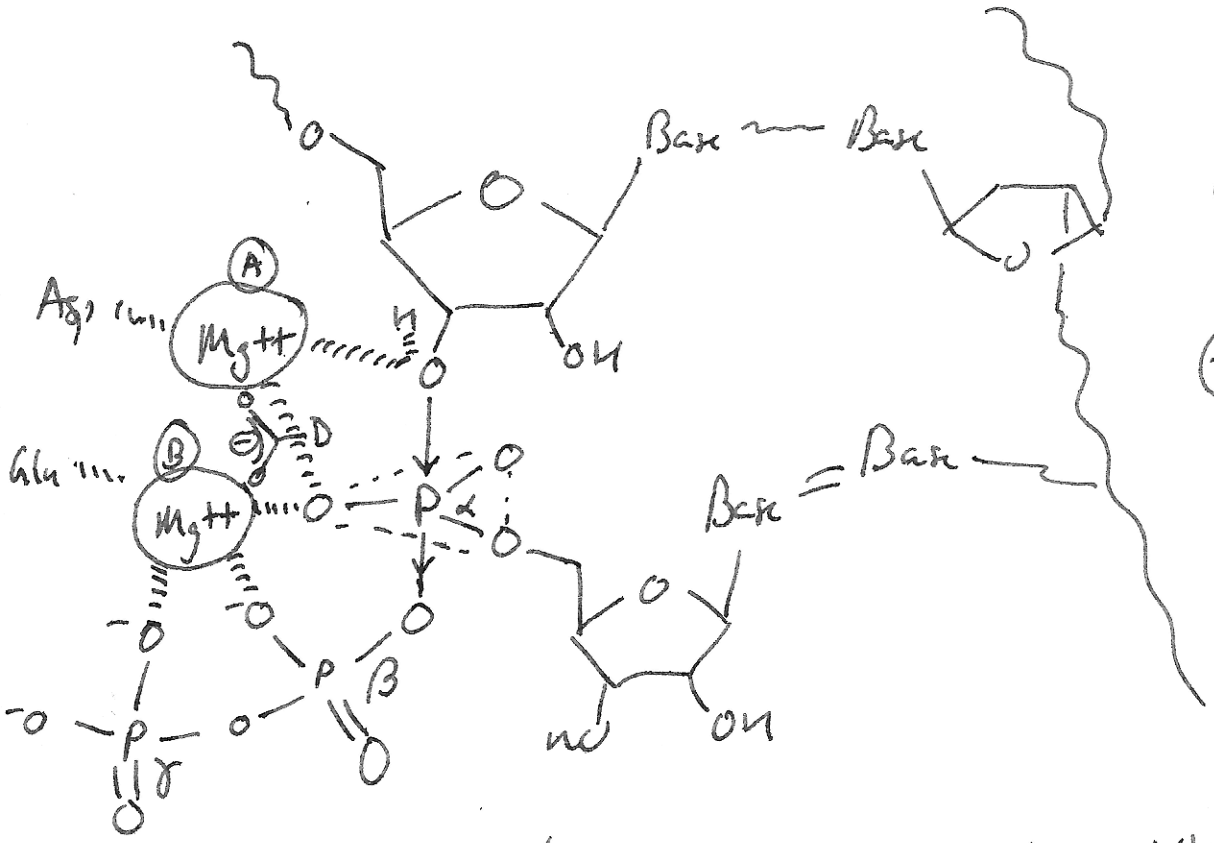
(+6)
- Histone acetyltransferase = HAT → activation
- Histone deacetylase = HDAC → repression

(+4) - ChIP = chromatin immunoprecipitation, especially ChIP on chip or ChIP-seq

Score for the page

3. DNA Polymerases and DNA Replication (56 pts):

(a; 20 pts) Sketch the two metal ion mechanism for RNA polymerization. Don't worry about the exact arrangement of protein side chains. State the function of each metal ion. One reason this mechanism has been so attractive, besides its universality, is that it may offer some supporting evidence for an RNA world. How so, for example in contrast to an enzymatic reaction mechanism with a covalent E-I intermediate such as the serine protease mechanism?



(+4) for correct chemistry

(+3) for reasonable sketch - details get points below

(+3) A: Activate 3'-OH as a nucleophile by stabilizing ⊖ charge on oxygen

B: Coordinate β/γ phosphates. (+1)

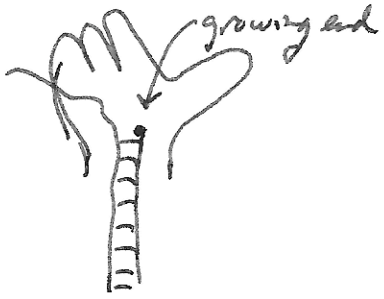
(+3) Activated PPi as a ~~leaving~~ leaving group by stabilizing developing ⊖ charge

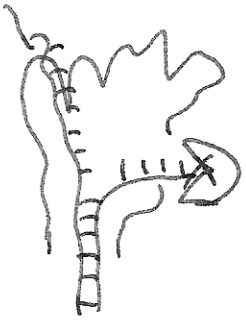
(+3) Both: chelate and orient oxygen of α-phosphate

— All the enzyme has to do is position the metal ions — no particular functional groups or enzyme-based nucleophile is required. (+3)

Score for the page

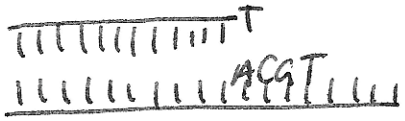
(b; 18 pts) We discussed the idea that slow extension of a mismatched primer terminus is fundamental to DNA polymerase fidelity. What rationale did we offer for the relative lack of specificity of the 3′→5′ exonuclease (hint: it's a separate active site)? We did not discuss the experimental evidence for slow extension of a mismatch, but it is actually reasonably straightforward to study if you have an oligonucleotide synthesizer (or IDT's web site) in hand. Describe how you would do such an experiment. It is also helpful to have an exonuclease-deficient mutant available – how does the use of an exo⁻ mutant polymerase simplify the results?

Synthesis mode: 

Exo mode: 

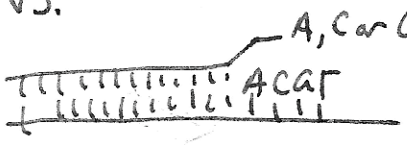
– The exonuclease acts on unpaired 3′ terminus – once the end of the growing helix frays, there is ~~not~~ no way to distinguish a ~~Sa~~ mismatch from a correct match. (+4)

Experiment: (+5) for basic idea of using synthetic primer/template



+dGTP or
~~dATP~~
+dATP, +dTTP, +dCTP

VS.



+ dGTP or
+ others

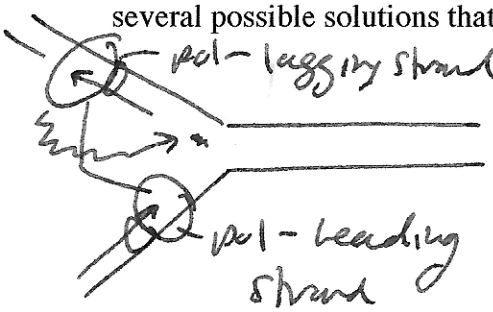
(+5) compare rates of extension w/ different combinations of correct/mismatched terminii; correct/m.m. next nucleotides

Experiment would be to look for extension by 1 or 2 bp

– With an exo⁻ mutant, don't have to worry about removal of primer terminus to get rid of mismatch, or also ~~a futile~~ a futile cycle of addition + removal of a mismatched dNTP (+4)

Score for the page

(c; 18 pts) The coordination of leading and lagging strand DNA synthesis at the replication fork is a subject of enduring interest. Describe the fundamental puzzle about the timing and list two of the several possible solutions that have been proposed.



(+4) - Lagging strand synthesis is discontinuous

- How does the lagging strand get back (+6) to the RNA primer fast enough for the lagging strand to be replicated at the same speed as the leading strand?

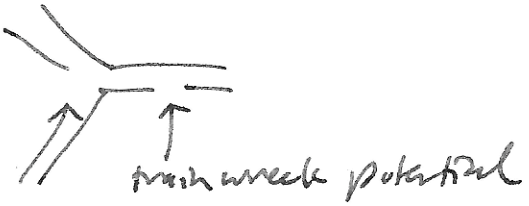
(+4) each for any two

1. Leading strand polymerase pauses upon lagging strand polymerase release at the end of an Okazaki fragment
2. There are gaps left in the lagging strand b/c primer synthesis triggers loop release
3. The rate of the leading strand is limited by the helicase, therefore lagging strand replication is faster intrinsically
4. There are actually two lagging strand polymerases - a trimeric pol complex

4. DNA Repair (36 pts):

(a; 8 pts) What common type of disaster is thought to be the reason for the low viability of RecBCD mutants? What is the general source of information used to fix lesions that affect both strands, like crosslinks or double-strand breaks?

(+4) - Replication fork collapse



- Somehow the information ~~must~~ must come from a clean (+4) copy of the sequence, i.e. a sister chromatid or the other replicated duplex, therefore typically requires recombination to fix.

Score for the page

(b; 28 pts) Mismatch repair (MMR) is essential for mopping up most of the few mistakes made by DNA replicative polymerases.

(+4) Name the protein in *E. coli* responsible for mismatch recognition: MutS

How does the MMR machinery know which strand to fix (what marks the daughter strand)?

(+4) - Hemimethylation - the unmethylated strand is the daughter strand to be fixed. or - MutH.

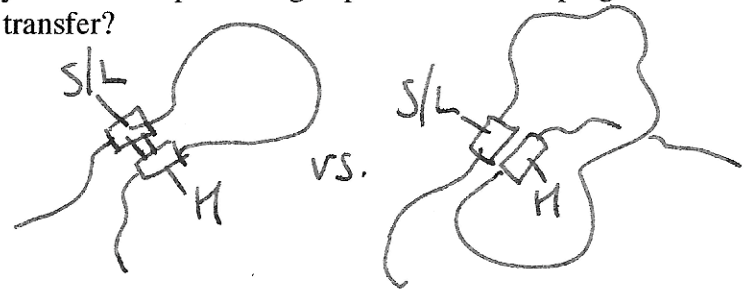
There have been many mechanisms proposed for the needed information transfer between the marker and the mismatch. Give two simple reasons that this information transfer is a challenge.

(+4) - The hemimethylated dam site (GATC) may well be far away from the mismatch

(+4) - The dam site could be to either side (5' or 3' of the mismatch).

Why doesn't simple through-space random looping make sense as a mechanism for information transfer?

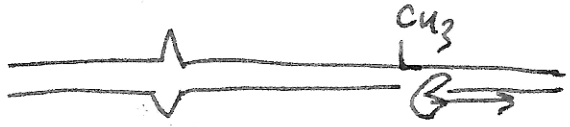
(+4)



Doesn't tell you which direction to chew in!

What would the consequences be of making the wrong "decision"? → (should have specified about what choice)

(+4)



the exonuclease needed to remove the "long patch" could set off into uncharted "good DNA" waters.

How does the MMR system illustrate the idea that cost is nearly no object when it comes to the integrity of the genome?

- The long patch could be ~1000 bp away - we are burning ≥1000 ATP-equivalents to fix one mistake.

Score for the page

5. Protein-DNA Recognition (20 pts):

(a; 20 pts) Some proteins can recognize specific nucleic acid sequences without hydrogen bonding to the DNA/RNA bases. Provide the name of a protein that recognizes DNA in this manner. Briefly describe where this protein interacts with the DNA, the type of primary interactions that are made, and a key features of the DNA sequence that facilitates binding.

(+5) – TBP or TATA Box Binding Protein

(+5) – Binds the minor groove of the DNA ~~where~~ (which is lacking in discriminatory H-bonding contacts).

(+5 for either) – Binds TATAAAA and related sequences A/T rich – ~~These~~ there are not minor groove $G-NH_2$'s to block binding. Also readily deformable

(+5 for either) – Primary interactions are largely van der Waals, including the partial intercalation to make two large kinks.

Page	Score
1	/24
2	/32
3	/32
4	/20
5	/18
6	/26
7	/28
8	/20
Total	/200

Score for the page 20