MCB 110. Midterm 1. 2006
ANSWER KEY

1. (20 pts.) Find the one letter below that best matches the following statements. You may use each letter only once.

A. A+G = C+T
B. fDNA
C. First replication fork protein loaded
D. Measure of sequence relatedness
E. binds pol III and helicase
F. Direction of hydrolysis for proofreading
G. xDNA
H. A+T=G+C
I. autosomal dominant
J. 0.01
K. Topoisomerase II
L. recA
M. propeller twist
N. 33 degrees
O. substrate for DNA polymerases
P. single-stranded DNA
Q. Substrate for DnaB helicase
R. X-linked recessive

i) **D** E-value
ii) **C** DnaB helicase
iii) **A** follows from Chargaff's rules
iv) **P** floppiest form of DNA
v) **M** angle between bases within a base pair
vi) **K** makes and repairs staggered, double-strand breaks in DNA
vii) **N** average helical twist for a RNA stem
viii) **E** clamp loader
ix) **O** deoxyriboATP
x) **I** phenotype does not skip generations

2. The restriction endonuclease NciI recognizes and cuts the five-base-pair sequence 5'-CC(G/C)GG-3' (G/C means either G or C will work in that position).

2A. (6 pts) How often, on average, would this sequence occur in random DNA? Assume the DNA is 25% A, G, T & C.

Once every 512 bases:

$$(1/4)^4 \times 1/2 = 1/512 \text{ bases}$$

2B. (4 pts.) Nci1 leaves a one=base 5' overhang. Schematically draw the products of a Nci1 digestion, including the location of the restriction-site bases.

```
5'   C C | G / C G G   3'
     G G C / G | C C
```

2C. (6 pts.) Assuming that the genomes of different individuals are 99.9% identical, approximately how many RFLPs would you expect for the Nci1 fragments of the human genome (3×10^9 bp).

$$3 \times 10^9 \text{ bp} \times 10^{-3} \text{ differences} \times \frac{1 \text{ site}}{512 \text{ bp}} =$$

$$5.86 \times 10^{-3} \times 10^9 \times 10^{-3} = 5.86 \times 10^3 \text{ RFLPs}$$

$$= 5860 \text{ RFLPs}$$

2D. (10 pts.) Your friend, Gene Watson, is starting a company to test for the presence of known human disease genes in patient DNA. Sequencing each sample is too expensive, so he needs to figure out which probes to use in Southern blots. He asks you for advice. In this case, you know the sequence of each disease gene of interest and the sequence of the human genome. Briefly outline an experimental procedure you would use to determine which, if any, of the Nci1 RFLPs could provide a diagnostic marker for a specific human disease gene. What results would the disease marker give in your experiment?

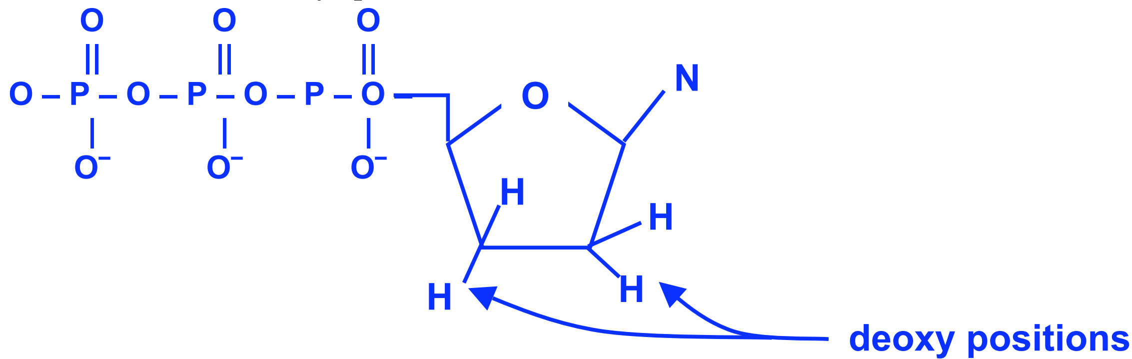
1. **Pick probes at increasing distances from gene X.**
2. **Southern blot Nci1 digested DNA from normal and affected individuals using each probe.**
3. **A probe covering a linked RFLP will produce a different pattern of bands in normal and affected DNA samples.**

3. The Global Ocean Survey is a project to sample the DNA sequence diversity of the oceans. Craig Venter and his crew sailed around the world, collected water samples from various locations in the ocean, and sent them to their lab in Maryland for sequencing.

3A. (8 pts.) Suppose you're running the lab in Maryland. Briefly list the steps of the procedure you would use to process the sea-water samples to create templates suitable for DNA sequencing.

1. **Extract DNA.**
2. **Cut with a restriction enzyme.**
3. **Clone into a plasmid that contains sites to bind sequencing primers (make a library)**
4. **Isolate individual clones from the library.**
5. **Purify plasmids from the individual clones.**

3B. (6 pts.) DNA sequencing utilizes dideoxynucleoside triphosphates. Draw the chemical structure of a dideoxynucleoside triphosphate, including the sugar and phosphate groups. You can represent the base with the letter N. Indicate the two "deoxy" positions.



3C. (6 pts.) Why do the dideoxy-NTPs terminate DNA strand synthesis?

They lack a 3' –OH nucleophile.

3D. (6 pts.) Why do sequencing reactions require a DNA polymerase that lacks 5' → 3' exonuclease activity?

The 5' → 3' exonuclease would chew away the primer to different extents.
This heterogeneity would eliminate the common start point of all labeled products.

3E. (8 pts.) You run a sequencing reaction using the following dideoxynucleotides, each with its own fluorescent marker: ddGTP-red, ddaATP-green, ddTTP-yellow, ddCTP-blue. You run the products of the reaction on a gel and the fluorescence reader scans the colors as the DNA fragments run off the bottom. The order of colors detected at the bottom of the gel (the positive pole) is:

5'

1. Red **G**
2. Yellow **T**
3. Yellow **T**
4. Green **A**
5. Blue **C**
6. Blue **C**
7. Red **G**
8. Blue **C**

3'

What is the sequence of the TEMPLATE strand for these products?

5' GCGGTAAC 3'

3F. (10 pts.) You discover using BLAST that one of the sequences from a sample taken at a depth of 26,000 feet in the Marianas Trench off the Philippines matches the sequence of the human protein kinase, PKA. The E-value of the match is 0.05. List two different ways (experimental and/or computational) you could further test to see if the protein encoded in the deep-sea organism is homologous to PKA.

1. **BLAST against a larger sequence library and find sequences homologous to both proteins.**
2. **Make a multiple sequence alignment and see if functional sites are conserved.**
3. **Express and purify the protein and see if it has kinase activity.**
4. **Solve the structure and see if it has a kinase fold.**

4. (10 pts.) Suppose you wanted to make a strain of *E. coli* with a high mutation rate. List two genes you would mutate in order to reduce the fidelity of DNA replication. Explain why you would pick those genes and what kinds of mutations (substitution, insertion or deletion) you would make in them. Be careful not to make lethal mutations!

- **Dna pol I or pol III – point mutation that eliminates 3' → 5' proof reading exonuclease.**
- **Delete or inactivate any mismatch repair protein – MutS, MutH…**
- **Delete pol I**

5. (8 pts.) Please mark the following statements true or false.

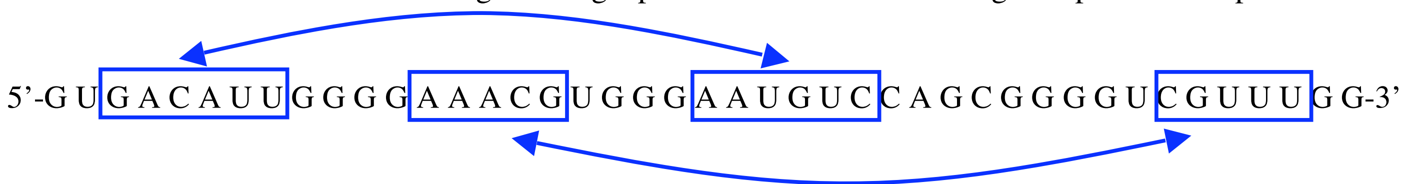
T In the absence of dNTPs, the Klenow fragment of DNA pol I will degrade the primer.

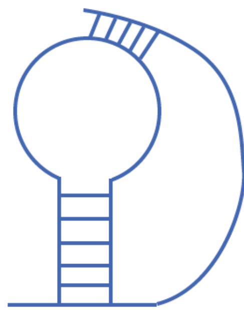
F DNA pol I is a great antibiotic target, because it is required for *E. coli* growth.

F The helicase loader, DnaC, uses deoxyATP in the loading reaction.

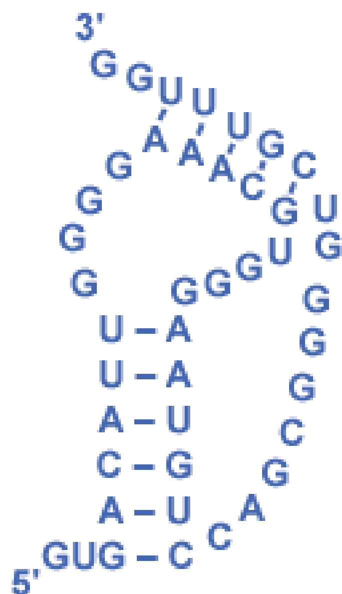
T The clamp loader links together the two pol III molecules at the replication fork.

6. (10 pts.) Experiments suggest that the 42-nucleotide RNA sequence below forms a pseudoknot. Draw schematically a possible pseudoknot structure of this RNA molecule. Indicate which bases form base pairs. Partial credit will be awarded for diagramming a pseudoknot without showing the specific base pairs.

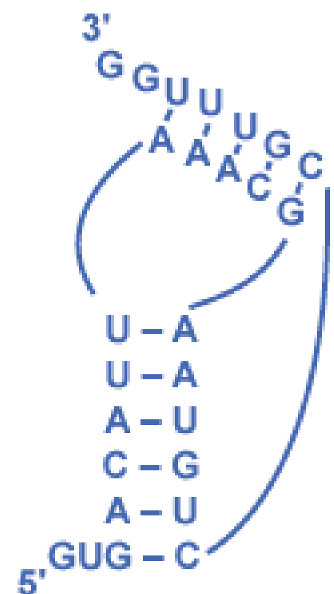




(5 pts)



or



7. Briefly contrast (explain the differences between)

7A. (8 pts.) the roles of metal cations in stabilizing DNA and RNA structures.

DNA: cations stabilize by screening electrostatic repulsion of PO_4s

RNA: metals bind specific sites to stabilize tertiary structures as well as screening backbone PO_4s

7B. (8 pts.) the roles of ATP in the reactions of DNA ligase and type II topoisomerases.

ligase: ATP donates AMP to a reaction intermediate that activates the DNA phosphate for attack.

topoisomerase: ATP binding and hydrolysis are coupled to conformational changes that order the reaction. ATP is not bonded to the DNA.

7C. (8 pts.) the roles of the sliding clamps on the leading strand and the lagging strand of a DNA replication fork.

Similar roles
Leading strand: maintains processivity
Lagging strand: maintains processivity, attracts pol III to primer, binds pol I to promote maturation of Okazaki fragments.

7D. (8 pts.) the roles of the folded parts of the core histones (H2A, H2B, H3 and H4) and their floppy, N-terminal tails.

Folded parts: bind and package DNA

Histone tails: provide sites of modification (Ac, PO_4, Ub, etc.) that form the basis of the histone code that signals the function of the DNA.