

**CHEM 331 Problem Set #7- Lehninger 5e, Chapter 8**  
**Due Friday, November 30<sup>th</sup>, 2012**  
**Answer Key**

1. The composition (mole fraction) of *one* of the strands of a double-helical DNA is  $[A] = 0.3$ , and  $[G] = 0.24$ . Calculate the following, if possible. If impossible, write "I." (9 pts.)
- a.  $[T]$  on the same strand                    I
  - b.  $[C]$  on the same strand                    I
  - c.  $[T] + [C]$  on the same strand            0.46  
(  $[A] + [G] = 0.3 + 0.24 = 0.54$ ; the rest must be T or C, so  $[T] + [C] = 1 - 0.54 = 0.46$  )
  - d.  $[A]$  on the other strand                    I
  - e.  $[T]$  on the other strand                    0.3  
(Chargaff's rules imply that  $[A]$  on one strand must equal  $[T]$  on the other.)
  - f.  $[A] + [T]$  on the other strand            I
  - g.  $[G]$  on the other strand                    I
  - h.  $[C]$  on the other strand                    0.24  
(Chargaff's rules imply that  $[G]$  on one strand must equal  $[C]$  on the other.)
  - i.  $[G] + [C]$  on the other strand            I
2. Write a double-stranded DNA sequence containing a six-nucleotide palindrome. (3 pts.)

**Ans:** Any double-stranded sequence that has the form:

1-2-3-4-5-6-6'-5'-4'-3'-2'-1'

1'-2'-3'-4'-5'-6'-6-5-4-3-2-1

where each number and its prime represent correctly paired bases (A with T, C with G) all along the double-stranded molecule.

Example: G-G-A-T-C-C

C-C-T-A-G-G

3. Why does lowering the ionic strength of a solution of double-stranded DNA permit the DNA to denature more readily (for example, to denature at a lower temperature than at a higher ionic strength)? (2 pts.)

**Ans:** Lower ionic strength reduces the screening of the negative charges on the phosphate groups by positive ions in the medium. The result is stronger charge-charge repulsion between the phosphate, which favors strand separation.

4. One strand of a double-helical DNA has the sequence  
(5')GCGCAATATTTCTCAAATATTGCGC(3')  
Write the base sequence of the complementary strand.  
What special type of sequence is contained in this DNA segment?  
Does the double-stranded DNA have the potential to form any alternative structures?  
(3 pts.)

**Ans:** The complementary strand is

(5')GCGCAATATTTGAGAAATATTGCGC(3')

(Note that the sequence of a single strand is always written in the 5'→3' direction.)

This sequence has a palindrome, an inverted repeat with twofold symmetry:

(5')GCGCAATATTTCTCAAATATTGCGC(3')

(3')CGCGTTATAAAGAGTTTATAACGCG(5')

Because this sequence is self-complementary, the individual strands have the potential to form hairpin structures. The two strands together may also form a cruciform.

5. Hairpins may form at palindromic sequences in single strands of either RNA or DNA. How is the helical structure of a long and fully base-paired (except at the end) hairpin in RNA different from that of a similar hairpin in DNA?  
(2 pts.)

**Ans:** The RNA helix assumes the A conformation; the DNA helix generally assumes the B conformation. (The presence of the 2'-OH group on ribose makes it sterically impossible for double-helical RNA to assume the B-form helix.)

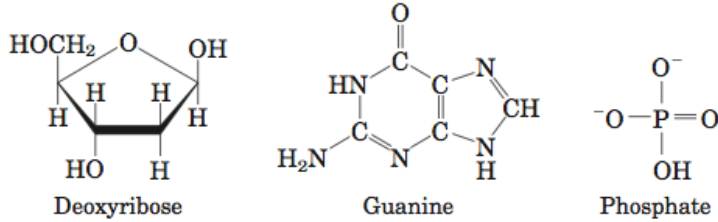
6. Hydrolysis of the N-glycosyl bond between deoxyribose and a purine in DNA creates an AP site. An AP site generates a thermodynamic destabilization greater than that created by any DNA mismatched base pair. This effect is not completely understood. Examine the structure of an AP site (see Fig. 8-33b) and describe some chemical consequences of base loss.  
(2 pts.)

**Ans:** Without the base, the ribose ring can be opened to generate the noncyclic aldehyde form. This, and the loss of base-stacking interactions, could contribute significant flexibility to the DNA backbone.

7. Draw the following structures and rate their relative solubilities in water (most soluble to least soluble): deoxyribose, guanine, phosphate. How are these solubilities consistent with the three-dimensional structure of double-stranded DNA?

(6 pts.)

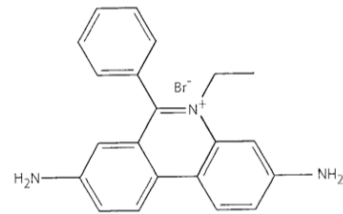
**Ans:**



Solubilities: phosphate > deoxyribose > guanine. The negatively charged phosphate is the most water-soluble; the deoxyribose, with several hydroxyl groups, is quite water-soluble; and guanine, a hydrophobic base, is relatively insoluble in water. The polar phosphate groups and sugars are on the outside of the DNA double helix, exposed to water. The hydrophobic bases are located in the interior of the double helix, away from water.

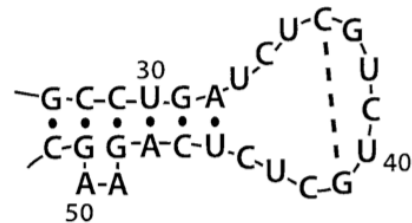
8. Ethidium bromide is a fluorescent molecule that interacts with DNA and is widely used in electrophoresis. Based on the structure of ethidium bromide, how does it interact with the DNA molecule and what effect would this interaction have with the melting temperature of the DNA? Justify your answer.

(2 pts.)



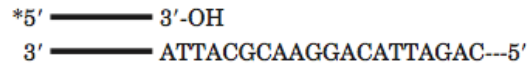
**Ans:** The ethidium bromide intercalates between bases and disrupts base stacking interactions. One would expect this disruption to decrease the melting temperature

9. A ribosomal protein, L5, binds to a 5S-ribosomal RNA molecule. The specific secondary structure of ribosomal RNA that interacts with three tyrosine residues in the protein is shown. Would you expect the tyrosine residues in the protein to interact with the helix or the loop portion of this hairpin structure? Explain your answer. (3 pts.)



**Ans:** Tyrosine has both an aromatic ring and -OH group. It could interact with the loop portion by H-bonding or stacking interactions. It would interact with the helix portion by polar interactions or H-bonding. More interactions are possible with the loop portion of the molecule.

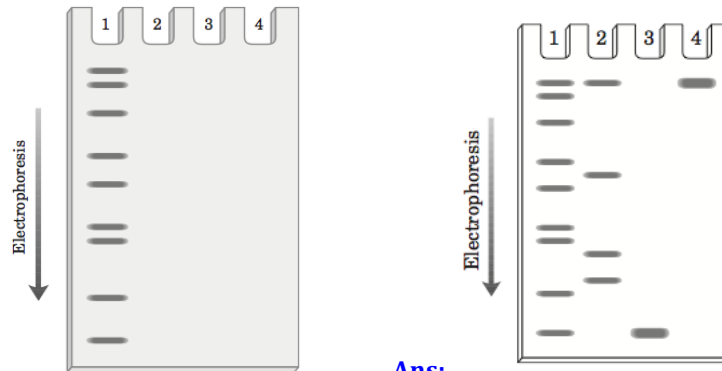
10. The following DNA fragment was sequenced by the Sanger method. The asterisk indicates a fluorescent label.



A sample of the DNA was reacted with DNA polymerase and each of the nucleotide mixtures (in an appropriate buffer) listed below. Dideoxynucleotides (ddNTPs) were added in relatively small amounts.

1. dATP, dTTP, dCTP, dGTP, ddTTP
2. dATP, dTTP, dCTP, dGTP, ddGTP
3. dATP, dCTP, dGTP, ddTTP
4. dATP, dTTP, dCTP, dGTP

The resulting DNA was separated by electrophoresis on an agarose gel, and the fluorescent bands on the gel were located. The band pattern resulting from nucleotide mixture 1 is shown below. Assuming that all mixtures were run on the same gel, what did the remaining lanes of the gel look like? (6 pts.)



**Ans:**

**Lane 1:** The reaction mixture that generated these bands included all the deoxynucleotides, plus dideoxythymidine. The fragments are of various lengths, all terminating where a ddTTP was substituted for a dTTP. For a small portion of the strands synthesized in the experiment, ddTTP would not be inserted and the strand would thus extend to the final G. Thus, the nine products are (from top to bottom of the gel):

- 5'-primer-TAATGCGTTCCTGTAATCTG
- 5'-primer-TAATGCGTTCCTGTAATCT
- 5'-primer-TAATGCGTTCCTGTAAT
- 5'-primer-TAATGCGTTCCTGT
- 5'-primer-TAATGCGTTCCT
- 5'-primer-TAATGCGTT
- 5'-primer-TAATGCGT
- 5'-primer-TAAT
- 5'-primer-T

**Lane 2:** Similarly, this lane will have four bands (top to bottom), for the following fragments, each terminating where ddGTP was inserted in place of dGTP:

- 5'-primer-TAATGCGTTCCTGTAATCTG
- 5'-primer-TAATGCGTTCCTG
- 5'-primer-TAATGCG
- 5'-primer-TAATG

**Lane 3:** Because mixture 3 lacked dTTP, every fragment was terminated immediately after the primer as ddTTP was inserted, to form 5'-primer-T. The result will be a single thick band near the bottom of the gel.

**Lane 4:** When all the deoxynucleotides were provided, but no dideoxynucleotide, a single labeled product formed: 5'-primer-TAATGCGTTCCTGTAATCTG. This will appear as a single thick band at the top of the gel.

11. Bacterial endospores form when the environment is no longer conducive to active cell metabolism. The soil bacterium *Bacillus subtilis*, for example, begins the process of sporulation when one or more nutrients are depleted. The end product is a small, metabolically dormant structure that can survive almost indefinitely with no detectable metabolism. Spores have mechanisms to prevent accumulation of potentially lethal mutations in their DNA over periods of dormancy that can exceed 1,000 years. *B. subtilis* spores are much more resistant than are the organism's growing cells to heat, UV radiation, and oxidizing agents, all of which promote mutations.

a. One factor that prevents potential DNA damage in spores is their greatly decreased water content. How would this affect some types of mutations? (2 pts.)

**Ans:** Water is a participant in most biological reactions, including those that cause mutations. The low water content in endospores reduces the activity of mutation-causing enzymes and slows the rate of nonenzymatic depurination reactions, which are hydrolysis reactions.

b. Endospores have a category of proteins called small acid-soluble proteins (SASPs) that bind to their DNA, preventing formation of cyclobutane-type dimers. What causes cyclobutane dimers, and why do bacterial endospores need mechanisms to prevent their formation? (2 pts.)

**Ans:** UV light induces the condensation of adjacent pyrimidine bases to form cyclobutane pyrimidine dimers. The spores of *B. subtilis*, a soil organism, are at constant risk of being lofted to the top of the soil or into the air, where they are subject to UV exposure, possibly for prolonged periods. Protection from UV-induced mutation is critical to spore DNA integrity.

12. Single stranded RNA often forms hairpin turns that allow it to base pair with itself.

a. Consider the following sequence:

5'-CUAGAUGGUAGGUACGGUUAUGGGUAACUCUG-3'

Suggest how this strand might form a hairpin. That is, which bases would pair, and which would be in the turn? More than one answer is possible.

(Hint: Try drawing a figure for your answer.)

(2 pts.)

This question can have a variety of answers. The basic principle is to follow the base-pairing rules. Any hairpin will have a loop connecting the paired strands. Note that enough bases should be paired (enough enthalpy or solvation entropy gained) to compensate for the loss of configurational entropy when you go from a random polymer strand to one in which certain parts of the molecule are "locked" into a particular configuration.

- b. The RNAFold server (<http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>) will make a prediction of the folding of any segment of RNA or single-stranded DNA. Submit the sequence above to this server. Compare your prediction to that of the server and comment on any differences. Here are a few definitions:

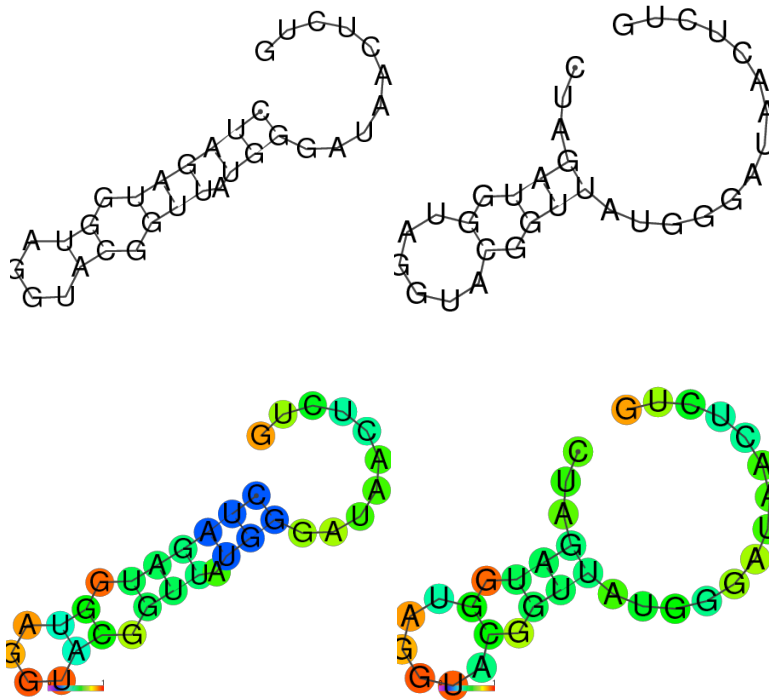
Minimum free energy structure= The MFE structure of an RNA sequence is the secondary structure that contributes a minimum of free energy. This structure is predicted using a loop-based energy model.

Centroid structure= The centroid structure of an RNA sequence is the secondary structure with minimal base pair distance to all other secondary structures in the Boltzmann ensemble.

Ans: (4 pts.)

**MFE**

**Centroid**



Note that the server's algorithm includes the non-Watson-Crick G=U base pair as a possibility. Also, the server predictions include multiple loops and regions with a single base mismatch.

13. Restriction endonucleases are enzymes that cleave both strands of DNA duplexes at specific sites. The recognition sites for these enzymes often are palindromic. Here are two palindromic DNA sequences:

5'-G-A-A-T-T-C-3'  
3'-C-T-T-A-A-G-5'

5'-G-A-T-A-T-C-3'  
3'-C-T-A-T-A-G-5'

The palindromes can be found by carefully placing a two-fold symmetry axis between the strands, perpendicular to the page.

- a. Locate the palindromes and the symmetry axes.

(1 pt.)

Palindromes are shaded in green, symmetry axis represented by \*.

5'-G-A-A-T-T-C-3'

3'-C-T-T-A-A-G-5'

5'-G-A-T-A-T-C-3'

3'-C-T-A-T-A-G-5'

- b. The EcoRI enzyme cleaves the first sequence between G and A. Identify the cleavage points. The resulting fragments are said to have "sticky ends". What does this mean? (2 pts.)

5'-G↓A-A-T-T-C-3'  
3'-C-T-T-A-A↑G-5'

The resulting pieces of DNA are as follows:

5'-G    A-A-T-T-C-3'  
3'-C-T-T-A-A                                  G-5'

Both of these fragments have a small region of unpaired, single-stranded DNA that are now available for a binding partner. This is what is referred to as "sticky ends."

- c. EcoRV cleaves the second sequence between T and A. Identify the cleavage points. The resulting fragments have "blunt ends". What does this mean? (2 pts.)

5'-G-A-T↓A-T-C-3'  
3'-C-T-A↑T-A-G-5'

The resulting pieces of DNA are as follows:

5'-G-A-T    A-T-C-3'  
3'-C-T-A    T-A-G-5'

Both of these fragments are fully base-paired, with no unpaired overhang. Hence, the ends are termed "blunt."

14. To examine the roles of hydrogen bonds and hydrophobic interactions between transcription factors and DNA, go to FirstGlance: <http://firstglance.jmol.org> . Enter the PDB id [1tgh](#) in the query box. This file contains the crystal structure of a human TATA-binding protein and a segment of double stranded DNA.

Once the structure loads, click the "Spin" button to stop the molecule from rotating. Then click the "Contacts" button. Select the radio button for "Chains". Then click on

any part of the protein (displayed in blue) to select it as the target.

Click "Show Atoms Contacting Target" to produce a list of contacts to display. Check only "Show putatively hydrogen-bonded non-water" to display hydrogen bonds between the protein and the TATA box.

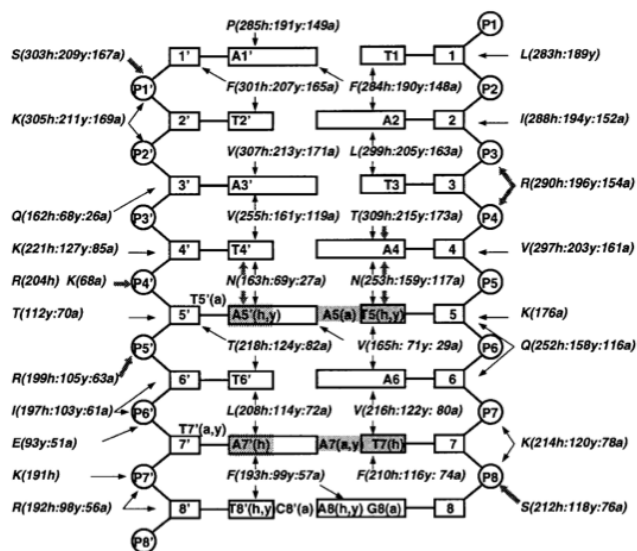
Then click the image viewing option: "Maximum Detail: Target and Contacts Balls and Sticks, Colored by Element"

Use the zoom and rotate functions to answer the following questions: (5 pts.)

- a. Which of the base pairs in the DNA form hydrogen bonds with the protein? Which contribute to the specific recognition of the TATA box? (Hydrogen bonds should be no longer than 3.3 Å.)  
The TATA box sequence is rich in Thymine and Adenine residues (hence the name "TATA"). Asparagine 163 makes both van der Waals contact and hydrogen bonds with Adenine and Thymine bases in the TATA sequence. There are many others; for a full listing, see figure below.
  
- b. Which amino acids in the protein interact with the TATA box base pairs? Residues that interact with the TATA box include Ser 212 and Lys 214, which hydrogen bond with the phosphate backbone, and Valines 216 and 165, which make hydrophobic contacts with the base pairs in the groove.
  
- c. Click "Return to Contacts", and check "Show hydrophobic (apolar van der Waals) interactions". Identify the hydrophobic interactions. Provide appropriate text and pictures to describe clearly how the molecules interact.  
The TATA binding motif is part of a transcription factor complex whose function is to promote gene transcription. It does this by binding to specific sequences and bending the DNA in that region, which promotes strand separation, making it easier for RNA polymerase to gain access to the region downstream of the TATA box. In addition to hydrogen bonds with the phosphate backbone, the TATA binding protein has four Phenylalanine residues (Phe single letter code=F)- 193, 210, 284, and 301- that project into the major groove and make hydrophobic interactions with the bases. The rather bulky aromatic side chain of the Phe is responsible for much of the helix distortion. Because the sequence is rich in A=T base pairs (rather than G≡C, which have one more hydrogen bond interaction), the strands are easier to separate.



A complete summary of interactions between the human TATA box binding protein and the TATA box DNA sequence comes from Juo et al., *J Mol Biol* 261:239-254 (1996) (on the next page)



**Figure 4.** Interactions of hTBP (human), aTBP (*Arabidopsis*) and yTBP (yeast) with their respective TATA boxes. The DNA sequence drawn is the symmetrical T-A-T-A-T-A-T-A used with hTBP, and variants in the other two structures are shaded. Thin arrows indicate van der Waals' contacts, and thick arrows hydrogen bonds. Numbers in parentheses are residue numbers of equivalent positions in the three TBP sequences.

15. Using Figure 27-7 to translate the genetic code, design a DNA probe that would allow you to identify the BRCA1 gene, which codes for a protein with the following amino-terminal amino acid sequence. (4 pts.)

H<sub>3</sub>N<sup>+</sup>- Met-Asp-Leu-Ser-Ala-Leu-Arg-Val-Glu-Glu-Val-Gln-Asn-Val-Ile-Asn-Ala-Met-Gln-Lys-

The probe should be 18 to 20 nucleotides long, a size that provides adequate specificity if there is sufficient homology between the probe and the gene.

**Ans:** Most amino acids are encoded by two or more codons (see Fig. 27-7). To minimize the ambiguity in codon assignment for a given peptide sequence, we must select a region of the peptide that contains amino acids specified by the smallest number of codons. Focus on the amino acids with the fewest codons: Met and Trp (see Fig. 27-7 and Table 27-3). The best possibility for a probe is a span of DNA from the codon for Ile to the codon for the C-terminal Lys. The sequence of the probe would be

5'-AU(U/C/A) AA(U/C) GC(U/C/A/G) AUG CA(A/G) AA(A/G)

At the positions where there are multiple options in parentheses, the synthesis would be designed to incorporate any of the bases listed at that position, producing a mixture of different 18-nucleotide probes.