7.014 Problem Set 2

1. Lipids

(a) Carbohydrates and lipids are both made up of long hydrocarbon chains, but they have very different properties. What is the major distinction between carbohydrates and lipids?

Lipids are not soluble in water - they are composed largely of long chains of hydrocarbons that “prefer” to interact with other hydrophobic molecules, while carbohydrates have polar properties. Carbohydrates are made up of monosaccharides that are linked together to form larger polymeric structures, while lipids do not form polymers in the same manner. Carbohydrates are generally made up of ringed hydrocarbons attached in a chain while lipids are made up of linear hydrocarbons.

(b) Membranes are made up of many components and, in some membranes, phospholipids are the major component.

(i) Below is an example of a saturated phospholipid (Where the intersection of two lines represents a carbon atom with hydrogen atoms to fill available covalent bonds). Draw a square around the portion of the molecule that is hydrophobic and a circle around the portion that is hydrophilic.

Phospholipids are also sometimes represented like this:
(ii) Using the simplified structure above, draw what a lipid bi-layer looks like in an aqueous environment. Explain what causes the lipids to create this formation.

Two lipid layers, with hydrophobic tails facing each other and polar head groups facing out. The polar groups interact with polar water molecules, while the hydrophobic tails pack together as a result of hydrophobic forces- it takes energy to disrupt interactions between water molecules and replace them with new interactions between the hydrocarbons and water, which would not be as favorable (entropy).

(iii) In your depiction of a lipid bi-layer drawn above, what type of force is acting…

- …between the phosphoglycerol (hydrophilic) component and the fatty acid (hydrophobic) component of the phospholipids?

  covalent bond- ester linkage formed via a dehydration reaction that links the fatty acid and the glycerol groups.

- …between the phosphogroup and the surrounding aqueous environment?

  hydrogen bonds- negatively charged phosphate groups can interact with the partial positive charges of hydrogen atoms in water.

- …between the fatty acid groups of different phospholipids molecules?

  hydrophobic forces- nonpolar substances interact with each other in an aqueous environment (see explanation in part ii)

(iv) In reality, membranes are composed of several different types of lipids, as well as proteins. One reason why there are multiple types of lipids is to ensure that the membrane remains fluid so that proteins, lipids and small molecules can move through and within the membrane. In particular, there is always a mixture of saturated and unsaturated phospholipids. Give a short explanation of why a membrane containing unsaturated phospholipids would be more fluid than a membrane made exclusively of saturated phospholipids.

Saturated fats have hydrocarbon chains that are saturated with hydrogen (each C atom bonded to a maximum number of H atoms), and there are no double bonds. Unsaturated fats have double bonds, which produce kinks in the molecule. These double bonds make parts of the molecule more rigid and the molecules cannot pack as closely/tightly. As a result of the looser packing between lipids, the bi-layer is more fluid (fats like butter and lard are more saturated than fats like olive oil).

2. There are thousands of different kinds of enzymes and each enzyme recognizes a specific substrate or substrates (substrates are the molecules upon which an enzyme acts). You are working in a lab that studies the activity of an enzyme called a nuclease that was purified from the gram-negative bacteria Serratia marcescens. A nuclease is an enzyme that cleaves or breaks apart DNA or RNA by hydrolyzing the phosphate-sugar backbone.

The work in your lab has shown that there are four residues important for binding or catalysis of the substrate.
Question 2 continued

(a) Below is the amino acid sequence of the *Serratia* nuclease. Those amino acids important for binding or catalysis are marked by being enlarged and bolded. These amino acids are Arg78, His110, Asn131 and Glu148.

\[
\begin{align*}
10 & \quad 20 & \quad 30 & \quad 40 & \quad 50 & \quad 60 \\
\text{MRFNNKLAL AALLFAAQAS ADTLESIDNC AVGCPGGSS NVSIVRÄHT LNNNSTTKFA}
\end{align*}
\]

\[
\begin{align*}
70 & \quad 80 & \quad 90 & \quad 100 & \quad 110 & \quad 120 \\
\text{NWVAYHITKDT TPASGKR NW KTDPALPAD TLPADYTGA NAALKVDRGH QAPLÅSÅG}
\end{align*}
\]

\[
\begin{align*}
130 & \quad 140 & \quad 150 & \quad 160 & \quad 170 & \quad 180 \\
\text{SDWESLNYLS NITPQKSDLN QGÅWARÅDQ ERKLIDRADI SSVYTVTGPL YERDÅMGKLPG}
\end{align*}
\]

\[
\begin{align*}
190 & \quad 200 & \quad 210 & \quad 220 & \quad 230 & \quad 240 \\
\text{TQKAHTIPSÅ YWKVIÅFINNS PAaNHYAAFL FDQNTPKGÅD FÇQÅFRTVDE IÅEKRTÅGLÅIÅW}
\end{align*}
\]

\[
\begin{align*}
250 & \quad 260 \\
\text{AGLPDDVQÅS LÅSKPÅGVLÅE LÅMGKÅN}
\end{align*}
\]

If the role of this enzyme is to cleave DNA and RNA, why does it make sense that Arginine (R) and Histidine (H) are two of the amino acids important for binding the substrate?

*R and H both have positively charged sidechains. It makes sense that they would be able to form interactions with the negatively charged phosphate groups in the backbone of DNA/ RNA molecules and thus help the enzyme bind to its substrate.*

(b) As stated above, it is known that these residues are important for binding or catalysis. You want to test for which of these functions (binding or catalysis) the amino acids Arg78 and His 110 is important. To perform this test you change Arg78 and His110 to different amino acids and then monitor if the nuclease can still cleave DNA. Below is the outline for the assay:

1. Incubate either or wild-type (wt) or mutant (mt) enzyme with DNA.

2. After several minutes, you isolate the DNA from the reaction.

3. Run the DNA pieces through an agarose gel matrix (see the Research Method box on pg. 319 in *Purves et. al.* using an electric current. At this time it is not important to understand the technique, but it is important to understand that the electric current causes DNA fragments of different sizes to separate from each other and that this separation can be visualized.

The chart below shows which enzymes you will use in the assay:

<table>
<thead>
<tr>
<th>Name of Enzyme</th>
<th>Amino Acid</th>
<th>Position of Amino Acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>wt*</td>
<td>Arg</td>
<td>78</td>
</tr>
<tr>
<td>mt1</td>
<td>Ala</td>
<td>78</td>
</tr>
<tr>
<td>mt2</td>
<td>Lys</td>
<td>78</td>
</tr>
<tr>
<td>mt3</td>
<td>Trp</td>
<td>78</td>
</tr>
<tr>
<td>wt*</td>
<td>His</td>
<td>110</td>
</tr>
<tr>
<td>mt4</td>
<td>Ala</td>
<td>110</td>
</tr>
<tr>
<td>mt5</td>
<td>Lys</td>
<td>110</td>
</tr>
<tr>
<td>mt6</td>
<td>Trp</td>
<td>110</td>
</tr>
</tbody>
</table>
Question 2 continued

* note that there is only one “wt” enzyme. “wt” is listed one time for each of the different amino acids you are studying.

Below is a representation of your agarose gel after you have separated the DNA. There are multiple lanes on the gel. Each lane contains the DNA from one of your different incubations. The version of the enzyme (wt or mt) that was incubated with the DNA in a specific lane is noted above the lane (no = no enzyme added to that DNA).

(i) Based on your results above do you think that Arg78 is more likely to be important for binding the DNA or for cleaving the DNA? Why?

Arg78 is more likely to be important for binding the DNA. When Arg is replaced with an amino acid with similar properties (Lys, which also has a positively charged sidegroup), cleavage is seen. This makes sense if the interaction between the negative DNA and positive amino acids is needed for binding of the DNA and if binding is necessary in order for cleavage to take place. The other amino acid substitutions do not conserve the positive charge so even if cleavage activity is still possible, cleavage is not seen because binding is not possible (due to the loss of the positive-negative interaction between enzyme and substrate).

(ii) Based on your results above do you think that His110 is more likely to be important for binding the DNA or for cleaving the DNA? Why?

two possible answers:
1) His110 is likely to be more important for cleaving the DNA. Assuming that the positive/negative interactions discussed previously are necessary for binding to occur, it would be expected that the Lys mutant would still cleave DNA if His was important for binding. The fact that all mutants are unable to cleave DNA suggests that the defect is in cleaving activity, not the binding activity.
2) The assay doesn’t differentiate- the fact that all the His mutants are the same means that it is difficult to interpret the results of the experiment. It could be that it is not the charge, but some other property of the amino acids that is important for binding/cleavage of DNA. If no cleavage occurs, it is difficult to tell if the problem is in binding (which is needed for cleavage to occur) or the ability of the nuclease to cleave DNA.
(d) Your lab also is interested in understanding how much energy is required to cleave the phosphodiester bond. To do this, you perform an assay similar to the one described above, but this time you allow the reaction to reach equilibrium. You then measure the amount of starting full-length DNA and the amount of cleaved DNA. The reaction can be described using the following equation:

\[
\text{1 full length DNA molecule} \quad \rightarrow \quad \text{2 shorter DNA molecules}
\]

or

\[
A \quad \rightarrow \quad 2B
\]

The final concentrations of your reactants and products are:

\[
A = 3 \text{ nM} \\
B = 4.5 \text{ mM}
\]

(i) Based on these measurements, determine the amount of standard free energy (\(\Delta G^\circ\)) for \(A \rightarrow 2B\). Show your work.

\[
\Delta G^\circ = -RT \ln (\text{Keq}) \\
\text{Keq}= \frac{[B]^2}{[A]} \\
RT= 0.6 \text{ kcal/mol (approx, from lecture)}
\]

\[
\Delta G^\circ = -(0.6 \text{ kcal/mol}) \times \ln \left( \frac{(4.5 \times 10^{-3} \text{ mol})^2}{(3 \times 10^{-9} \text{ mol})} \right) = -5.3 \text{ kcal/mol}
\]

(ii) What does your answer tell you about how the reaction proceeds under standard biological conditions?

The \(\Delta G^\circ\) is negative, which means that the reaction will proceed spontaneously towards the formation of products. The reaction is exergonic (releases energy) and the products are at a lower energy level than the reactants.

(iii) Why is the concentration of enzyme not important in determining \(\Delta G^\circ\)?

Enzymes do not change the starting/ending \(G^\circ\) values of the reactants/products, although the path that the reaction takes is affected. Enzymes lower the activation energy of a reaction, which is the level of energy corresponding to the high-energy transition state intermediate in the reaction. Enzymes will change the rate at which the reaction proceeds but will not affect things like the equilibrium constant of the reaction. The enzyme is neither a reactant nor a product, but is a biological catalyst.

(iv) Draw an energy diagram for the hydrolysis of the phosphodiester bond with and without the presence of the nuclease enzyme. Make sure to label the reactants, the products, the value of \(\Delta G^\circ\) and the activation energy.

\[
\Delta G^\circ = -5.3 \text{ kcal/mol}
\]

\[
\text{activation energy}
\]

\[
\text{activation energy}
\]
3. To answer the problem below, you need to use the StarBiochem, a java viewer for macromolecules. To get access to and directions for this viewer, go to the expanded problem statement online at http://web.mit.edu/vis/. The questions are reproduced below. To obtain credit for the problems, you must write your answers down in the spaces provided on the problem set that you will turn in to be graded.

We examine the primary, secondary, and tertiary structure of a specific protein, nitrogenase Fe (1XD8).

(a) Chain A consists of 289 amino acids.
   (i) List in order the 12 amino acids numbered 16 through 28 in chain A.
   ser, thr, thr, thr, gln, asn, leu, val, ala, ala, leu, ala, glu
   (ii) What level of protein structure does this represent?
   Primary structure, which is the linear sequence of amino acids in a polypeptide

(b) These 12 amino acids also make up an α-helix in nitrogenase.
   (i) What level of protein structure do α-helices and β-sheets represent?
   Secondary structure- local folding patterns including alpha- helices and beta pleated- sheets
   (ii) Do the side chains of the amino acids in a helix point into or out of the helix?
   Side chains point out of the helix
   (iii) What type of bond is primarily responsible for maintaining secondary structure?
   H- bonding
   (iv) What part of the amino acid participates in this bond (side chain or backbone)?
   Backbone interactions between partial positive charges of H atoms and partial negative charges of
   either the =O (carbonyl oxygen) or N of the backbone participate in the bond.

(c) The tertiary structure of a protein is formed by bending and folding of the amino acid chain, with the
    interactions between the amino acid side chains determining this structure. Below, we list four kinds
    of tertiary interactions between side chains that are possible and four sets of residues.
   (i) Match up which set of residues belongs to which type of bond
   A. Hydrogen Bond - II Residues 8 and 10 (Tyr, Lys)
   B. Ionic bond - I Residues 129 and 10 (Lys, Asp oppositely charged AAs)
   C. Disulfide bridge - III Residues 132 and 97 on Chain B (Cys, Cys covalent bond))
   D. Hydrophobic clusters - IV Residues 8, 126 and 135 (Tyr, Val, Phe all nonpolar)
   (ii) Which represents the strongest interaction: A, B, C or D?
   C- disulfide bridge- covalent bond between 2 cysteines
Question 3 continued

(iv) Which represents the weakest interaction: A, B, C or D?

A- hydrogen bond

(d) The residue found in chain B at position 129 in this protein is known to be involved in catalytic activity of this nitrogenase. Residue 129 helps with the binding of ADP.

(i) What type of residue is at position 129 (polar, nonpolar, acidic or basic)?

Aspartic acid is the residue at 129- acidic (carboxylic acid group). Can lose H+ to form negatively charged group at pH7.

(ii) What is likely the strongest bond found between residue 129 and ADP?

Hydrogen bond – Asp has a negative charge and so do the phosphate groups of ADP. Therefore, the strongest interaction that might occur is between the negative charge on Asp and the nitrogenous base of ADP.

(iii) A mutant version of this nitrogenase that is unable to bind ADP has a glutamic acid in this position instead of the wild-type residue. How would this substitution affect ADP binding?

Both amino acids are acids and could form similar interactions with ADP. The sidechains differ by a single methylene group, and it is possible that the change interferes sterically with the ability of ADP to fit into a binding site of the enzyme.

4. The cell harvests energy by metabolizing the carbohydrate glucose. The overall reaction for the breakdown of glucose (under aerobic conditions) is:

\[ \text{C}_6\text{H}_{12}\text{O}_6 + 6 \text{O}_2 \rightarrow 6 \text{H}_2\text{O} + 6 \text{CO}_2 \]

(a) When, during the process of glucose metabolism is \( \text{O}_2 \) used?

In the final step in electron transport chain \( \text{O}_2 \) is used as a terminal electron acceptor, when it is reduced to form water

(b) When, during the process of glucose metabolism is \( \text{CO}_2 \) released?

During the breakdown of pyruvate (when it is oxidized to acetate) and the citric acid cycle, as the carbon containing compound is further oxidized (one molecule released during the breakdown of pyruvate and two during the citric acid cycle)

(c) When, during the process of glucose metabolism is \( \text{H}_2\text{O} \) released?

During the final step in electron transport chain, when electrons and hydrogen ions are transferred onto oxygen. Also during glycolysis.

(d) What is/are the oxidizing agents during glycolysis?

\( \text{NAD}^+ \) is reduced (gains e-). Remember that the oxidizing agent is reduced. During glycolysis \( \text{NAD}^+ \) picks up e- and protons from glucose (and various metabolic intermediates) to form \( \text{NADH} + \text{H}^+ \). Glucose is oxidized (it is a reducing agent).
(e) Which process of the metabolism of glucose is blocked under anaerobic conditions?
The electron transport chain is blocked without a terminal electron acceptor. Instead, pyruvate is fermented.

(f) Do you think that the cell could use galactose \( \text{C}_6\text{H}_{12}\text{O}_6 \) as a substrate for glycolysis? Why or why not?
No. Although the molecules have the same chemical formula, they differ in structure, and the first enzyme in the process of glycolysis is specific for glucose. Enzymes are highly specific with regards to the substrates they recognize and use.

5. Back at your summer job at the Venter Institute, you are still working on characterizing Species M, Species I and Species T. Currently, you are working on determining the composition of the nucleotides in the DNA for two of the species, M and I.

(a) Before you can even begin your experiment for the day, your annoying bench-mate suggests that you should determine if DNA or RNA is the stable, heritable genetic material for these organisms. Give two reasons (based on the structure and/or chemistry of DNA or RNA) why DNA is a better choice to be the stable, heritable genetic material than RNA.

The double stranded nature of DNA lends itself well to a mechanism of replication, and the double stranded nature means that DNA always has a copy of itself (important for repair). In addition, the ribose of RNA has a free –OH group on the 2’ carbon. This –OH group can participate in hydrolyzing the backbone of RNA, which would be disastrous for an organism.

(b) After explaining to your bench-mate why RNA is a poor macromolecule to encode the precious genetic information of a cell, you begin your experiment. You first isolate DNA from each of the two species and aliquot some of the DNA into four tubes per strain. You label these tubes:

\[
\text{M-A} \quad \text{M-G} \quad \text{M-C} \quad \text{M-T} \quad \text{I-A} \quad \text{I-G} \quad \text{I-C} \quad \text{I-T}
\]

The DNA in each tube will be used to identify the percentage of one of the four nucleotides (A, G, C or T). For example, the DNA in tube M-A will be used to identify the percentage of adenine in the DNA of species M.

You go to another room to gather some reagents you will need to perform the different reactions. When you come back, you find that your bench-mate is not only annoying but clumsy too! While carrying some ethanol, he knocked over your tubes and spilled the ethanol on top of them. This has caused all the writing on your tubes to come off, and you no longer know which tube contains the DNA from the two different species. You pick up the tubes and re-label them #1- #8. You decide to perform the experiment anyway by choosing two tubes at random on which to perform each nucleotide-determining reaction. Below are your results:

<table>
<thead>
<tr>
<th>Tube # nucleotide determined</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>% A/G/C/T</td>
<td>30</td>
<td>30</td>
<td>20</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>33</td>
<td>30</td>
</tr>
</tbody>
</table>

What is the nucleotide composition for each of the two different species?
Values for A/T and G/C will be the same for an organism because of base-pairing (Chargaff’s Rule). In each case, all percentages should add up to 100%

<table>
<thead>
<tr>
<th>Species #1</th>
<th>Species #2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A ____<strong>30</strong></td>
<td>A __<em><strong>33</strong></em></td>
</tr>
<tr>
<td>G __<em><strong>20</strong></em></td>
<td>G __<em><strong>17</strong></em></td>
</tr>
<tr>
<td>C __<em><strong>20</strong></em></td>
<td>C __<em><strong>17</strong></em></td>
</tr>
<tr>
<td>T ____<strong>30</strong></td>
<td>T __<em><strong>33</strong></em></td>
</tr>
</tbody>
</table>

(c) Now, you need to identify which nucleotide composition belongs to which species. To do this, you use a small amount of DNA that is still labeled as being from Species M or Species I. To determine which species has which nucleotide composition, you do a DNA re-naturing experiment. Below is the outline of the experiment:

1. Combine DNA from Species M or Species I with DNA from at least two of your tubes labeled #1-#8.
2. Break the DNA into smaller pieces using a technique called sonication. The smaller pieces will have, on average, 500 base pairs.
3. Heat the DNA in the tube so that the hydrogen bonds between the two strands of the helix break and the strands come apart.
4. Incubate the tubes at room temperature, which will allow hydrogen bonds to reform.
5. Monitor the percentage of single-stranded DNA (ssDNA) that is present in the samples at different times during the room temperature incubation.
6. Graph the percentage of ssDNA over time.

(i) Before you begin, your bench-mate suggests that you use tubes #1 and #2. Why is this idea bad? *Tubes #1 and #2 contain DNA from the same organism- both tubes would yield the same results and would not give useful information.*

(ii) You choose to use tubes #3 and #4 instead. Below are the results from your experiment:

For tubes containing Species M DNA and DNA from tube #3 or #4:

<table>
<thead>
<tr>
<th>% ssDNA</th>
<th>time</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>#3</td>
<td></td>
</tr>
<tr>
<td>#4</td>
<td></td>
</tr>
</tbody>
</table>

For tubes containing Species I DNA and DNA from tube #3 or #4:

<table>
<thead>
<tr>
<th>% ssDNA</th>
<th>time</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>#3</td>
<td></td>
</tr>
<tr>
<td>#4</td>
<td></td>
</tr>
</tbody>
</table>
Which tube, #3 or #4 contains which species’ DNA (M or I)? Explain your reasoning based on your data.

Tube #3 contains DNA from Species_M___  Tube #4 contains DNA from Species_I___

*Complementary DNA strands (= from the same organism) will anneal to each other faster than will strands with different sequences.* Faster annealing means that single stranded DNA will disappear at a faster rate.

Bonus: Can you provide a reason for why the ssDNA went away faster for the incubation of DNA from Species I with DNA from tube #4 than the ssDNA for the incubation of the DNA from Species M with DNA from tube #3?

The time it takes for the DNA to reanneal will be affected by factors such as the size of the genome (more DNA- more time for complementary sequences to “find” each other) and the composition of the genome. If, for example, there are many repetitive elements in a genome, then reannealing will occur more quickly. If there are more unique sequences in a genome (more complex) then it will take longer for complementary pieces to find each other and anneal.
### Structures of Amino Acids

#### Generic Amino Acid:

Individual amino acids are linked through these groups to form the backbone of the protein.

- **Alanine** (ala)
- **Arginine** (arg)
- **Asparagine** (asN)
- **Aspartic Acid** (asp)
- **Cysteine** (cys)
- **Glutamic Acid** (glu)
- **Glutamine** (glN)
- **Glycine** (gly)
- **Histidine** (his)
- **Isoleucine** (ile)
- **Leucine** (leu)
- **Lysine** (lys)
- **Methionine** (met)
- **Phenylalanine** (phe)
- **Proline** (pro)
- **Serine** (ser)
- **Threonine** (thr)
- **Tryptophan** (trp)
- **Tyrosine** (tyr)
- **Valine** (val)

**Protein Synthesis**

Side chain, unique to each different amino acid

Individual amino acids are linked through these groups to form the backbone of the protein.