Fall Semester, 2002

Introductory Biology 111:
Cell and Molecular Biology

www.bio.davidson.edu/Biology/Courses/Bio111.html

Study Guide
(Revised: June 2001)

by

Dr. A. M. Campbell and Dr. K. Bernd
Davidson College

and

Dr. J. Serie
Macalester College

Table of Contents

Prologue 1: Overview of the Biological Sciences	1
Prologue 2: The Cell	5
Unit I: Cellular Communication	**7**
System 1 - Liver cells deliver glucose	7
System 2 - Heart cells pump harder	15
System 3 - Nerves tell skeletal muscles to run	25
System 4 - How does an egg know when it has been fertilized?	34
Other Communication Systems	41
Some Experimental Approaches	44
Unit II: Genetics	**50**
Mitosis	51
Meiosis and gametogenesis	51
Mendelian Genetics	53
Molecular Genetics	60
A Point Mutation can Change Your Life (WWW)	62
Gene Regulation	66
Structure Function Relationships	68
Linkage	71
RFLPs	74
Chromosome mapping	80
Molecular Causes of Diseases	97
An Article looking at Genetic Link to Alzheimer's	113
Unit III: Bioenergetics	**114**
Why did the US government spray Mexican marijuana with paraquat?	116
Follow up story to paraquat	130
Why do vegetarians eat tofu?	131
Why is Cyanide the terrorist's poison of choice?	134
Follow up story on poisoned grapes	146
Why would authorities ask you to update vaccinations after a flood?	14

	7
Two Research Questions and Approaches in Bioenergetics	150
Unit IV: Other Interesting Topics	148
Cancer	148
AIDS	163
Genetic Engineering	183

Prologue 2: The Cell

Brief Overview Reading:	Chapter 1
Focused Reading:	p 3-5 "Evolutionary Milestones" stop at “Controlling internal environments"
	Chapter 4: The Organization of Cells
WWW Focused Reading:	Virtual Plant Cell
	Purves6e, Chapter 4, Eukaryotic Cell Tour
	Immunofluorescence Labeling of ER
	Relative Sizes: from glucose to cells and larger

To look at images on the world wide web (WWW) you must use a computer that is connected to the campus network, that has Netscape or similar browser, and go to the Biology 111 Home Page. To reach the Biology 111 Home Page, up at the top of the screen in the box labeled “Location:”, you must type in the web address (called a URL)

<www.bio.davidson.edu/Biology/Courses/Bio111.html>

You must type this exactly as it appears above, except do not include the < and > symbols. Once you have reached the Biology 111 Home Page, all you need to do is click once on the appropriate underlined words to see the focused reading material. If you get a message that says “Bad URL” or the file cannot be found, then you have probably made a typo. If you wait for quite a while and nothing ever appears or you get the message that says the server could not be contacted, then the computer that provides this service (the web server) is too busy with all your classmates and cannot talk with you right now. You should try again later, in one or two minutes.

Once you have found this “Virtual Cell” you will see a cartoon of a plant cell. Point and click the mouse and you can see inside the cell. We strongly recommend that you use the “Hot Spots” option and then click on the picture. This will allow you to see more and more detail. You should also use the search function to select certain organelles.

Your text has supplemental tutorials and activities online. Connections between the text and online material are indicated in the book by ‘hand on a mouse’ icon. Web reading will also include some of these supplemental materials and will be seen in the study guide as “purves6e, Chapter, Title of the material”. To access this material you must type in the URL <www.whfreeman.com/purves6e/> , click on the chapter and then the title that are indicated in the study guide assignment.

Study Questions:

1. The cell theory was one of the first fundamental theories of the biological sciences. What are the elements of the cell theory and why is this theory so important? In what ways does this theory drive the modern biological sciences?

2. A major tenet in the biological sciences is that form follows function. Give an example that illustrates this point on the cellular level. Be able to explain how this example illustrates the point.

3. What are the differences between prokaryotes and eukaryotes? Give an example of each type of cell.

4. Approximately how big are typical prokaryotic and eukaryotic cells? About how big is this? What else is this size? How much smaller is a cell than say a marble or a bowling ball or a typed period -- "."? What are the limiting factors in cell size (i.e. why can't cells be larger than they are? Why aren't they smaller?)

5. Eukaryotic cells are full of little compartments called organelles. Why? What is adaptive or useful about having all these little compartments?

6. For each organelle or cellular structure described in Chapter 4

   A. Be able to give a very brief (a few words) description of its basic function(s).

   B. Be able to draw and label each organelle or accurately describe its structure. Pay close attention to the distinguishing features of the organelle (e.g. the curved, stacked cisternae of the Golgi apparatus, the small and large subunit structure of the ribosome, the double membrane surrounding the nucleus, the microtubular core of the cilia and flagella, etc.)

   (Study note: Online supplement includes 'flashcards' for each chapter that are good for definitions.)

7. Today's reading includes two animations of the 'same' content. Compare Virtual Plant Cell and Purves6e, Chapter 4, Eukaryotic Cell Tour. What are the strong and weak points of each?

8. Be able to describe or make a sketch of the structures of a chloroplast, a mitochondrion, a nucleus.

Occasionally, there will be places in the study guide where you will find fact-oids called "News Items". The information in these news items is not for you memorize, nor will it be tested. However, you might find them very interesting since they contain recent research findings based on the material you are learning. For example:

News Item: Researchers recently identified a giant sulfur bacterium, which they called *Thiomargarita namibiensis*. This organism is remarkable because it is a prokaryote but the cells grow to have diameters of 750 μm (see "Relative Size' website for illustration of how big, or small, this is). While this is still very small when compared to a house, prokaryotes have no internal membrane systems and so the exchange of nutrients and waste products must occur by diffusion. According to previous thought a prokaryotic cell this large is not possible--but there it is off the coast of Chile. How can this organism survive? You can't change the laws of nature so does it use mechanisms we haven't seen before? Since this organism was only discovered recently much more work must be done to answer questions like these. For more information see H.N. Shulz et al. 1999 *Science*. Vol 284 p493-5

----------STOP----------

Unit I: Cellular Communication

Brief Overview Reading: Chapter 2,3, & 5

Note: Yes three chapters is a lot. This reading will be discussed throughout this unit. We will focus discussion on different sections on different days. Briefly reading through all of the material now means that when you get to it as 'focused' reading it is not completely foreign.

Living beings can be composed of a single cell (e.g. bacteria, cyanobacteria and protists such as *Paramecium* and *Chlamydomonas)* or many cells -- the latter are called **multicellular organisms**. The human being is an example of a multicellular organism about which much is known. Estimates place the number of cells in the human body at 70 trillion. (If you counted these cells at a rate of one cell per second, it would take you over 2 million years to count every cell in the human body.) With a few exceptions (e.g. the red blood cells) each of these individual cells is a living entity with its own complete set of genes and life maintenance equipment. Each of these cells maintains its own existence, and also makes a vital contribution to the life of the multicellular organism.

In order for multicellular organisms to function properly, their cells must communicate. For instance, your muscles must contract only when your brain sends a message to contract and not any other time. Your salivary glands must secrete a lot of saliva when there is food in your mouth and only a little saliva at other times. Your heart rate must increase when you exercise, but not when you sleep. Unit I focuses on how cells communicate with one another in order to coordinate their functions. While we will focus most closely on cellular communication in multicellular creatures, you should keep in mind that communication is very important to unicellular creatures as well. For instance, they must swim toward nutrition or sunlight if they are photosynthetic and must be able to sense when conditions are right to reproduce (we will see this in the lab).

In this unit, we will consider four examples of cellular communication. Each system uses a slightly different communication system, and taken together, these four systems represent most of the communication systems we understand so far.

System #1: The Production of Glucose (sugar) by the Liver During Stress

Glucose is the primary sugar used by biological creatures for fuel. Humans, like other creatures, burn (oxidize) glucose to carbon dioxide and water and use the energy released by this process to perform life's many functions. To ensure that cells have enough glucose to burn (and, therefore, enough energy) the body maintains a constant supply of glucose in the blood at a level of about 100 mg of glucose per 100 mL of blood.

However, when we are scared or under stress, our bodies respond by increasing the blood glucose level to ensure that we have enough fuel to fight or flee from what is scaring us. This extra glucose comes from the liver. During meals, glucose enters the body and the liver takes it out of the blood and stores it for later use. To put glucose in storage, the liver hooks many glucose molecules together (**polymerizes** them) to form a large storage molecule called **glycogen.** Then, when glucose is needed later (either because you haven't eaten for a while or because you are scared), these big glucose

polymers will be broken down into individual glucose molecules, and the glucose will be dumped into the blood to provide fuel for all of the cells of the body.

Focused Reading: p 34-36 "Macromolecules: Giant polymers" stop at "Proteins:.."
p 43-46 “Carbohydrates…” stop at end of page
Focused WWWebreading: Purves6e, Chapter 3, Tutorial 3.1 (Carbohydrate section)

Study Questions:

1. What is **glucose** used for in biological creatures? What is **glycogen** used for? What is the relationship between glucose and glycogen?

2. What is a **polymer**? What is a **monomer**? Is glucose a polymer or monomer? What is glycogen? Explain. What is a **monosaccharide**? A **disaccharide**? A **polysaccharide**?

3. Be able to recognize a monosaccharide and polysaccharide when you see one drawn. (You do **not** have to be able to draw these molecules yourself.)

4. Glucose molecules are joined to form glycogen by a process called **dehydration synthesis** (or **condensation synthesis**). Glycogen is broken down to form glucose by the process of **hydrolysis**. "Hydro-" means water. What does water have to do with these two processes? Be able to illustrate both of these reactions including the breaking or forming of bonds and the involvement of water in the process.

5. Starch (made by plants) and glycogen (made by animals) are formed by joining glucose monomers by **alpha glycosidic linkages** while cellulose (made by plants) is made by joining glucose monomers by **beta glycosidic linkages**. What are the differences chemically, and what practical significance does this have in your own life?

During a meal, glucose molecules are joined together to form the polymer glycogen in the liver. This process is called **glycogenesis**. (The "genesis" or creation of glycogen.) An enzyme catalyzes the formation of each alpha glycosidic bond between each glucose molecule. This enzyme is called **glycogen synthase**. The following reading assignment is about the action of enzymes. In addition, because enzymes are proteins, this reading assignment also includes an explanation of proteins and protein structure. Also, in order to understand how proteins fold, you will need to understand about hydrophobic and hydrophilic groups.

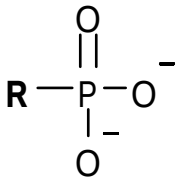
Focused Reading: p 20-25 "Chemical bonds..." stop at “Chemical reactions...”
p 36-42 “Proteins: Polymers of amino acids" stop at "Carbohydrates”
p 102-112 “Enzymes: ” stop at end of chapter

Glycogen synthase, then, lowers the activation energy barrier and allows glucose molecules to be joined together to form glycogen at a reasonable rate at normal body temperatures. **All chemical reactions in living things that involve the breaking or forming of a covalent bond are catalyzed by enzymes.** The rate at which enzymes perform their functions can be increased or decreased by

allosteric or **covalent modulators**. Thus, the rate at which glycogen is synthesized could be increased if the cell increased the rate at which glycogen synthase catalyzes the reaction.

The enzyme that breaks down glycogen to glucose is **glycogen phosphorylase**. The breakdown of glycogen to glucose is called **glycogenolysis**. (The "**lysis**" or degradation of glycogen.) The rate of this enzyme can also be controlled. Therefore, the cell can increase or decrease the rate at which glycogen is broken down simply by increasing or decreasing the catalytic rate of glycogen phosphorylase.

Glycogen synthase and phosphorylase are turned on and off by the process of **covalent modulation**. This process is like **allosteric modulation** or **regulation** (described on page 136 in your text), except that it depends on the process of **phosphorylation**. Phosphorylation is the covalent



addition of a phosphate group to an enzyme by dehydration synthesis. Phosphate groups look like → and "R" is used to indicate the rest of the molecule. “R” can be anything from a simple hydrogen to an enormous protein.

In proteins, phosphate groups are added onto the side chains (added onto an –OH group) of certain amino acid residues by standard dehydration synthesis (also called a ‘condensation reaction’ See p35 fig3.2a) . After a phosphate is added by a covalent bond, the protein is said to be **phosphorylated.** (By looking on page 37 in your text, can you figure out which three amino acids are the only ones that can be phosphorylated? Hint: Which three have an –OH in the side chain?)

Phosphorylation can either turn an enzyme on (increase its catalytic rate) or turn an enzyme off (decrease its rate.) Regardless of the direction of its action, phosphorylation is a kind of switch or signal that changes the rate of an enzyme's activity. This is not allosteric modulation which uses weak bonds to regulate the enzyme.

When you are scared, your liver slows the rate of glycogen synthesis and increases the rate of glycogen breakdown. Fear causes the phosphorylation of liver enzymes. In this example, phosphorylation inactivates glycogen synthase and activates glycogen phosphorylase. Therefore, when these two enzymes are phosphorylated by the cell, the rate of glycogen breakdown increases and the rate of synthesis decreases. When these two enzymes are **dephosphorylated** (phosphate is removed) by the cell (this happens when you get calm again), the rate of glycogen synthesis increases and the rate of degradation decreases.

Study Questions:

1. What two enzymes are responsible for synthesizing and breaking down glycogen in the liver? How is the rate of each enzyme controlled?

2. What is **glycogenesis**? What is **glycogenolysis**?

3. Draw a phosphate group and demonstrate how it is added to a protein during the process of phosphorylation.

4. Be able to recognize an amino acid and show how it is joined together by a **peptide bond** to form a **dipeptide** and finally a protein.

5. Proteins have many functions in living things. List as many of these functions as you can.

6. The many different functions of proteins are possible because these molecules can take many different shapes. Explain, <u>in chemical terms</u>, how proteins form their three-dimensional shapes.

7. Two proteins with different shapes will have different functions and different amino acid sequences. Explain how changing the amino acid sequence of a protein can change its function.

8. What is activation energy?

9. How do enzymes work? What do they do to cause reactions to proceed? What don't they do; that is, what are the limitations of enzymatic catalysis?

10. Explain <u>in chemical terms</u> how enzymes can be specific for their substrates. What are the biological consequences of enzyme specificity? What would the consequences be if enzymes were less specific or not specific at all?

11. Explain the catalytic cycle (E + S --> ES --> P + E). Using this explanation as background, <u>explain</u> how each of the following events would increase the rate of an enzyme catalyzed reaction:
    A. increasing the concentration of substrate
    B. increasing the affinity of the enzyme for its substrate
    C. increasing the temperature
    D. increasing the inherent catalytic rate of the enzyme

12. Using a scenario from the social sciences, humanities, fine arts or your everyday life, describe a situation that is <u>analogous</u> to the catalytic cycle. Your model is a good one if you can answer questions A-D above using this model.

13. How are enzymes turned on and off by **allosteric modulation**?

14. The first lab unit explores the effect of environmental conditions on the rate of an enzymatic reaction. Write out the reaction we'll be following using the 'E + S --> ES --> P + E' format. What is the enzyme in the reaction?

15. Give an example <u>not covered in class</u> of a system in which the control of the rate of an enzyme is important for the proper function of a biological system. (The enzyme system need not be explained in detail. Assume enzymes catalyze all chemical reactions that

break and form covalent bonds. Use your own experience as a guide and use your imagination.)

We now know how the liver will be able to increase the supply of glucose to the blood, but we are missing some very important elements of the system. How does the liver "know" that the body is under stress? It is sitting quietly in your abdomen -- it can't see or hear. It has to be "told" that stress is occurring. The systems which carry this kind of information to body cells are the integrating and controlling systems of the multicellular creature, the nervous system and the endocrine system (hormones).

In this case, the endocrine system plays a major role in "informing" the liver that the body is under stress and, therefore, needs more glucose. The endocrine system is a system of glands in the body that secrete **hormones**. A hormone is a messenger molecule that is made and secreted (released into the blood) by an endocrine gland and the hormone travels in the blood to a **target organ**. It binds to the cells of the target organ and causes some change to occur. Examples of hormones are insulin (lowers the blood sugar level among other things -- its absence causes diabetes mellitus), growth hormone (stimulates growth -- its absence causes dwarfism), thyroid hormone (increases metabolic rate -- low levels cause coldness, weight gain and lethargy.)

Brief Overview Reading: Chapter 41
Focused Reading: p 712-713 "Animal hormones" stop at "Some hormones act…"

In the case we are currently considering, the hormone mediating the response of the liver to fear is **epinephrine** (also called adrenaline). This hormone is made and secreted by the **adrenal gland** in response to stress. When something stressful happens (e.g. your boss yells at you, you are in a car accident, you have to give a speech), the information about this event enters your brain through your sense organs (you hear, see, touch, smell and/or feel the stressful event.) Your brain interprets this event as stressful, using memory and some genetic responses (such as aversion to pain), and your brain then sends a message, via a nerve, to the adrenal gland.

[Note: Interpretation of the event as stressful is an important step in this process. Some things are always stressful (pain, cold, dehydration, severe hunger, etc.) while other things have to be interpreted as stressful, (e.g. social situations, threatening words or gestures, pressure to succeed, etc.) One way to reduce the physiological response to stress (which may be related to such diseases as high blood pressure and cancer) is to stop interpreting things as stressful. Unfortunately, this is not very easy.]

The nerve impulse from the brain reaches the adrenal gland and causes cells to secrete epinephrine. (The interaction between the nerve and the cells of the adrenal gland is itself an example of intercellular communication. We will deal with signaling by the nervous system later.) Epinephrine enters the blood and goes everywhere -- to all the cells in the body.

Even though epinephrine goes everywhere, it does not affect every cell of the body. It only affects the cells that have **epinephrine receptors** on their surface. These receptors can bind specifically to epinephrine in the same way enzymes bind specifically to their substrates.

To summarize so far, epinephrine is secreted by the adrenal gland when the brain "decides" that something stressful has happened. This hormone goes everywhere in the body but only binds to cells, which bear epinephrine receptors on their surfaces, like liver cells. We also know that the liver will be able to deliver glucose to the blood in response to stress if two of its enzymes, glycogen synthase and glycogen phosphorylase can be phosphorylated. Somehow, then, the epinephrine bound to the epinephrine receptors on the liver cells' surface has to trigger the phosphorylation of these enzymes inside the cell; the process of getting an external signal communicated inside a cell is called **signal transduction**. Signal transduction is usually accomplished through a **second messenger system** which relays information from the hormone receptor to enzymes inside the cell.

Before we look at second messenger systems, however, we have to look more closely at the surface of the cell and how it is constructed.

Brief Overview Reading: Chapter 5
Focused Reading: p 77-82 Stop at "Membrane carbohydrates…"
p 49-51 "Lipids…" Stop at "Carotenoids and steroids"
p 36-42 "Proteins" (review)
p 282 "Receptors" Stop at end of page
p 284 "G protein linked receptors" Stop at "cytoplasmic..."
WWW Reading: Crystal Model of a lipid bilayer
Fluid Model of a lipid bilayer

Receptors are **integral membrane proteins** with their "active site" or **ligand-binding** site facing outward for binding with the extracellular ligand. [NOTE: A ligand is any small molecule that binds to a protein at a specific site. Hormones are a kind of ligand that binds to binding sites of hormone receptors.] Cells have in their membrane many copies of a receptor that binds a given hormone. In addition, each cell has many different kinds of receptors -- one kind of receptor for every different extracellular signal molecule with which the cell interacts. The liver, for instance, interacts with epinephrine, growth hormone, thyroid hormone, insulin, glucagon, and many other hormones. The cells of the liver, then, have in their membranes many copies of each of these different receptors, which can bind each of these hormones. This is part of the "mosaic" of the fluid mosaic model.

Study Questions:

1. What is a ligand? Give some examples.

2. Draw a diagram of a phospholipid which illustrates its distinguishing characteristics (Do not use the balloon with two tails model found in your textbook -- come up with a diagram of your own that conveys the important features of the phospholipid molecule.)

3. Explain, in chemical terms, why phospholipids are excellent molecular building blocks for membranes.

4. Describe the fluid mosaic model of membrane structure.

5. Describe, in chemical terms, how an integral membrane protein would differ in amino acid sequence from a soluble protein (one that floats freely in the cytoplasm.) How would an integral membrane protein have to be constructed (what types of amino acids would be in what places in the molecule) in order to be embedded and floating in the phospholipid bilayer?

6. Membrane receptors are one type of integral membrane protein. List other types of integral membrane proteins (see Fig 5.1, 5.6, 5.9 5.13.5.17). Be able to state the function and give a specific example for each type of protein.

   News Item: An international team has found that the same ligand (estrogen) can bind to two different estrogen receptors, called alpha and beta. When this common ligand binds to alpha, it activates gene activation, but when it binds to beta, it inhibits gene activation. Thus, the same ligand can give two different signals, depending on which receptor is present. (*Science* Vol. 277: 1508. September 1997)

7. Anti-estrogens are used for treating and preventing breast cancer. Hypothesize a molecular mechanism to explain how the same ligand could give two different signals, such as described above.

----------STOP----------

We are now ready to put the elements of this story together by introducing the **cAMP** (pronounced cyclic AMP) **second messenger system** which forms the link between the epinephrine receptor and the phosphorylation of liver cell enzymes. This system and others like it are called "second messenger systems" because they provide a second message to the cell. The hormone provides the first message by binding to its receptor on the cell surface. The information of this binding is relayed into the cytoplasm through the second messenger system. In addition to reading about the cAMP second messenger system, you will also read about nucleotides and nucleotide structure, since cAMP (and ATP and GTP) are nucleotides.

Focused Reading:	p 47-48 "Nucleic acids..." stop at “DNA is a guide..."
	p 728-729 “Mechanisms of hormone action" stop at "A cell can have..."
	Fig. 15.8 (p284), 15.17 (p291), 15.12 (p287)
	p291 “Enzyme activities…” Stop at “Different genes…”
WWW Reading:	ATP - Chime Image
	cAMP - Chime Image
	Purves6e, Chapter 15, Tutorial 15.1 Signal transduction pathway

If ever there was a "domino" effect, it is the release of glucose in response to fear. Here is a summary, with a few more details than are given in your book to help you understand what is happening at the molecular level. [Read this story sentence by sentence. Be sure you understand each event before you go on.]

1) A stressful thing happens and the brain sends a signal to the adrenal gland, which secretes epinephrine.

2) Epinephrine enters the blood and goes everywhere in the body.

3) When epinephrine reaches the liver, it binds to epinephrine receptors on the liver cell membrane.

4) Binding of ligand with receptor causes the receptor molecule to change its native conformation.

5) This change in native conformation reveals a binding site on the cytoplasmic side of the receptor, which binds to G protein (the stimulatory version of G protein called G_s.)

6) When G protein binds to the epinephrine receptor (allosteric modulation). The binding causes the G protein to change shape.

7) This change in shape causes a GTP binding site in the G protein to lose its affinity for GDP and gain affinity for GTP.

8) The GDP sitting in the site leaves and GTP binds to the site.

9) When GTP binds to the G protein, this causes the G protein to change shape again and become able to bind to **adenylyl cyclase**, an integral membrane protein that is an enzyme.

10) When the G protein binds to adenylyl cyclase, adenylyl cyclase changes shape and this activates an enzymatic site on adenylyl cyclase, which binds **ATP** and converts ATP to **cyclic AMP**.

11) cAMP floats away from adenylyl cyclase and binds to the allosteric modulating site of **cAMP-dependent protein kinases**. One particular cAMP-dependent protein kinase is **Protein Kinase A** also known as **PKA.**

12) The cAMP-dependent protein kinase becomes activated by this non-covalent **allosteric modulation**. **PKA** is activated by allosteric modulation caused by binding cAMP.

13) The activation PKA causes this protein kinase to phosphorylate an enzyme in the liver cell called **phosphorylase kinase**. [Valuable Hint at no extra charge: All **kinases** phosphorylate. The word before "kinase" in the enzyme's name usually tells you which molecule the enzyme phosphorylates. For instance, hexokinase phosphorylates a hexose (a 6-carbon sugar). Phosphofructokinase phosphorylates phosphofructose (another 6-carbon sugar).]

14) Phosphorylation activates phosphorylase kinase, which phosphorylates **glycogen phosphorylase** and **glycogen synthase**.

15) The phosphorylation of **glycogen phosphorylase** activates it, thus causing it to break down glycogen to glucose (**glycogenolysis**) at a faster rate. This releases more glucose into the blood and the blood levels of glucose rise. **Glycogen synthase** is inhibited by phosphorylation. Therefore, **glycogenesis** is inhibited in the presence of stress, thus helping to keep glucose in its monomeric form.

Study Questions:

1. What are the components of a **nucleotide** and how are the components of a nucleotide linked together? What is the difference between a **triphosphonucleotide**, a

diphosphonucleotide, and a **monophosphonucleotide**? Give examples of each. Chemically, how are these nucleotides converted into one another?

2. Proteins become activated and inactivated by ligand binding because they change their shape in response to the binding of ligands. Identify every protein in the cAMP second messenger system and describe how ligand binding affects that molecule. What is it able to do after ligand binding that it was not able to do before?

3. Describe how phosphorylation is used in the cAMP second messenger system. Which proteins are phosphorylated and how are they changed by this event?

4. The cAMP second messenger system represents an **enzyme cascade**. Why is it called a cascade? What is adaptive about such a cascade? Why didn't the system evolve in such a way that the activation of glycogen phosphorylase was directly linked to the epinephrine receptor? [NOTE: There is probably more than one plausible answer to this question. Don't stop till you've really thought about it.]

5. Be able to describe in chemical terms (as described above), the entire process of stress-induced plasma glucose elevation from the stressful event through elevation of blood glucose levels.

6. Choose an example from the social sciences, the humanities, the fine arts or your everyday experience that is analogous to the cAMP second messenger system. Your model is a good one if you can trace the entire pathway (outlined in #5) using this analogous system.

NEWS ITEM: A new protein has been identified as a critical player in the metabolism of glycogen. PTG (Protein Targeting to Glycogen) forms a large complex with the opposing enzymes phosphorylase kinase and glycogen synthase. Researchers hypothesize that PTG acts as a molecular scaffold bringing the key enzymes into a central location to facilitate their regulation. (See Printen *et al.*, *Science.* Vol. 275: 1475-1478. 7 March, 1997)

Now that the cAMP second messenger system has been activated, it must be deactivated! Otherwise, you could not go back to a "normal" state after your stressful encounter. You'd be permanently wired on a blood sugar high!

Epinephrine (and all hormones and signal molecules) binds to its receptor through non-covalent interactions, i.e. hydrophobic interactions, hydrogen bonds and ionic bonds. These bonds are fairly easily broken, and as the epinephrine molecule sits in the ligand binding site of the receptor, it eventually wiggles free simply because of its constant motion due to kinetic energy. If the epinephrine level is still high (your brain is still stimulating epinephrine release by the adrenal gland), then a new epinephrine molecule will take the old one's place and the receptor will remain activated. However, when the stress has passed, the brain will stop stimulating epinephrine release and the epinephrine levels will fall because the body quickly destroys the free molecule. Then, when a molecule of epinephrine wiggles free of the epinephrine receptor, there will be very few molecules to come in and take its place, and the hormone binding site on the receptor will remain unfilled or empty.

Another way of looking at this is through the "law of mass action" from chemistry. According to this law, if the concentration of reactants increases, the reaction rate will increase. We can look at the free hormone and its binding site in the following chemical notation:

[H] + [B] <----> [HB]

[H] = the concentration of free hormone in the blood
[B] = the concentration of free (empty) binding sites on the hormone receptors
[HB] = the concentration of binding sites containing bound hormone.

As the concentration of free hormone ([H]) increases due to adrenal gland release, the rate of the reaction increases and more hormone is bound to receptors. Likewise, as the hormone concentration decreases, the rate of the reverse reaction is increased and more hormone comes free of its receptor.

ALL HORMONES AND SIGNAL MOLECULES HAVE THIS RELATIONSHIP WITH THEIR RECEPTORS. THEREFORE, THE STRENGTH OR DEGREE OF SIGNALING DEPENDS ON THE HORMONE CONCENTRATION.

When the hormone concentration falls, the receptor has no hormone bound to the active site. This causes the receptor to go back to its original shape. In this original shape, the receptor cannot bind and activate G protein; no hormone bound, no G protein activated.

Also, the G protein very slowly (in about a minute) cleaves its GTP to GDP. (It removes the last or "terminal" phosphate from GTP. GTP thus keeps two phosphates and becomes GDP.) With GDP bound instead of GTP, G protein goes back to its original shape and loses its ability to bind to adenylyl cyclase. Therefore, no new adenylyl cyclases can be activated.

Also, cAMP is degraded to AMP (the cyclic connection between the phosphate and the third carbon of the ribose is broken) by an enzyme called **phosphodiesterase.** (The bond it breaks is a phosphodiester bond -- "diester" because it contains two oxygens (an ester linkage contains one oxygen) and "phospho" because it also contains a phosphate group.) This reaction is occurring constantly and breaks down cAMP quickly after it is formed. The mechanisms that stop the cAMP cascade allow signal transductions to be brief. In that way, if you want to continue to make extra glucose for the blood, your adrenal gland has to continue to release epinephrine in response to input from your brain. Your brain, therefore, has control over the whole process.

Study Questions:

1. Explain why non-covalent bonding between the ligand and the hormone receptor facilitates effective cellular communication. What problems would be caused if the hormone bound covalently to its receptor?

2. Explain in conceptual or chemical terms the relationship between hormone concentration and signaling strength. How does this system work, exactly?

3. How is the cAMP system stopped once it has started? Describe all the mechanisms involved. What is adaptive about this immediate inhibition of the system?

4. Describe how enzymes are named. How can you tell what an enzyme does, even though you haven't ever encountered it before? Here are some enzymes to practice on: Pyruvate dehydrogenase; ribulose bisphosphate carboxylase (hint -- what is a "carboxyl" group? See page 31); tyrosine kinase; DNA polymerase; peptidyl transferase; aminoacyl-tRNA synthase, phospholipase.

The cAMP second messenger system was the first one to be discovered. It was discovered by Earl Sutherland, who won the Nobel Prize for this discovery in 1971. Since then, we have learned that many cells use this system for signaling. Here are some other examples:

Secretion of thyroid hormone by the thyroid gland triggered by thyroid stimulating hormone
Secretion of cortisol by the adrenal gland triggered by adrenocorticotropic hormone
Secretion of progesterone by the ovary triggered by luteinizing hormone
Reabsorption of bone triggered by parathyroid hormone
Increased heart rate and force of heart contraction triggered by epinephrine
Increased water retention by the kidney triggered by antidiuretic hormone
Increased triglyceride (fat) breakdown triggered by epinephrine
Learning and memory
Mating in *Chlamydomonas* (and you'll see this one in lab)

----------STOP----------

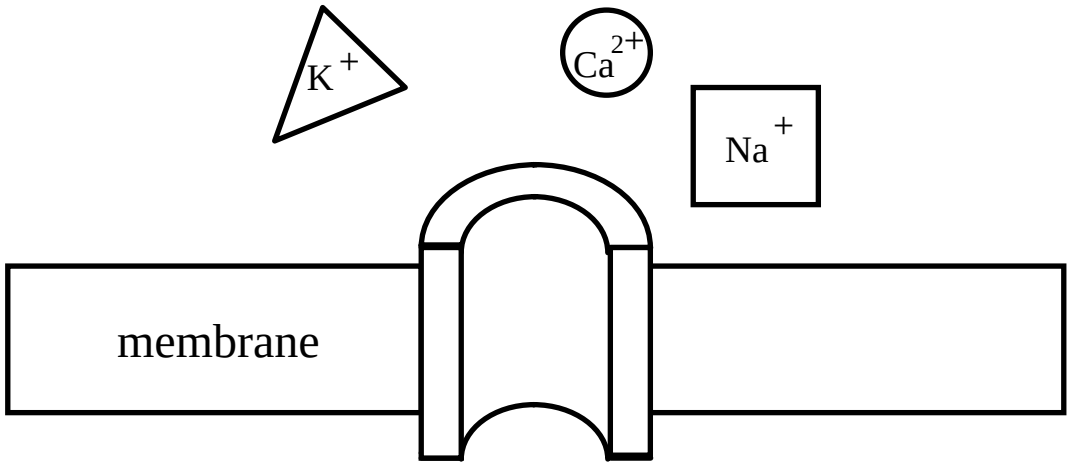
System #2: The Increase in the Force of Contraction of the Heart in Response to Fear

Everyone has experienced the "pounding" of the heart in the chest that occurs when one is afraid. Edgar Allen Poe frequently referred to this response in order to heighten the sense of fear in his readers. In nature, fear is an emotion that is produced in response to things that are physically harmful such as predators. The physiological response to fear, then, is to prepare you to flee or fight the thing that is making you afraid. In order to do this, you have to increase the supply of oxygen to your muscle cells, since they will be working especially hard in fleeing or fighting. (You also need more fuel for your muscles to burn for energy. This fuel, glucose, is supplied by the liver as you learned above.) Blood carries oxygen to the tissues and the only way to increase the oxygen supply is to increase the rate at which blood is delivered to the tissues. To do this, the heart beats faster and harder. It is the mechanism that produces the increase in contraction force that we will consider here. [NOTE: The explanation in this paragraph was teleological. The following explanation is causal.]

The force of heart contraction is also controlled by the hormone epinephrine, secreted by the adrenal gland in response to fear (a kind of stress.) Epinephrine goes everywhere in the blood including the heart. The heart muscle cells bear epinephrine receptors. These receptors are called **beta adrenergic receptors**. [There are also alpha receptors which do not work through a second messenger system. Receptors that bind epinephrine (adrenaline) are called "adrenergic."] This receptor triggers a cAMP second messenger system in exactly the same fashion as the liver epinephrine receptor does. The receptor-ligand complex activates a G protein, which activates adenylyl cyclase, which converts ATP to cAMP. The cAMP then allosterically activates cAMP-dependent protein kinase. However, this is where the similarities end. The cAMP-dependent protein kinase of cardiac muscle cells phosphorylates two proteins that we will look at here:

1) Ca^{2+} channels in the plasma membrane and
2) myosin heads.

We need to look carefully at each of these protein systems, beginning with the calcium ion channel in the plasma membrane of heart cells. **Ion channels** are protein molecules that span the membrane, are cylindrical in shape, and allow the passage of ions through the ion channel and thus the cell membrane. Each channel is **selective** for a given ion. ("Selective" means that they aren't quite as picky about what passes as they would be if they were "specific." So, they are pretty good at allowing only one type of ion to pass, but not as good as receptors are at binding only one ligand, or as specific as enzymes that bind only one substrate.) Below is a cut away view of a selective ion channel:



selective ion channel

All cells have Ca^{2+} channels, Na^{+} channels, K^{+} channels, Cl^{-} channels, etc., in their plasma membranes. Because ions are charged, they are extremely hydrophilic and are repelled by the fatty acid tails of the phospholipid molecules in the membrane. Therefore, THE ONLY WAY AN ION CAN CROSS A MEMBRANE IS WITH THE HELP OF A PROTEIN THAT SPANS THE MEMBRANE.
The following reading assignments are about ions and ion channels.

Brief Overview Reading: Chapter 2, 3, & 5
Focused reading: p 24 "Ions form bonds…" stop at "Polar and Nonpolar…"
p 85-86 "Passive processes of membrane transport..." stop at "Osmosis is passive water movement…"
p 795 fig 45.1
p88 fig5.9
WWW Reading: Simple Ion Channel - RasMol Image

Purves6e, Ch5, Tutorial 5.1 Membrane transport (passive transport section)

Ion channels are cylinders with water in the central channel. These ion channels allow ions to flow across the plasma membrane **down their concentration gradient**. Ions can only move down their gradient through channels; they can only move from an area of high ion concentration toward and area of low ion concentration. In the case of the Ca^{2+} channel we are considering, Ca^{2+} is in very high concentration on the outside of the cell (10^{-3} M) and very low concentration on the inside (10^{-7} M). Thus, Ca^{2+} moves from the outside of the cell toward the inside of the cell when the Ca^{2+} channel in the membrane is opened.

Note here that channel proteins cannot "pump" ions up their concentration gradient. Therefore, this channel cannot move Ca^{2+} from the inside of the cell toward the outside. It can, however, be open or closed. When closed, it does not allow the passage of Ca^{2+} toward the inside of the cell. When open, it does allow this passage. Such an ion channel is said to be **gated**. That is, it acts as if it has a gate that can be opened to allow ions to flow, or closed to stop the flow. The Ca^{2+} channel we are currently considering is gated and can be opened or closed.

The type of gated ion channel we are considering is opened and closed in response to a change in the **voltage** across the heart muscle cell's plasma membrane. These voltage changes occur rhythmically, around 80 times per minute, and are responsible for producing the normal heart beat. A bit later in this unit, we will consider how this voltage is created and how an ion channel might respond to changes in voltage. For now, all you need to know is that ion channels which open and close in response to changes in voltage are called **voltage-gated channels**. The Ca^{2+} channel we are considering is a voltage-gated channel. Other types of channels open and close in response to the binding of a ligand (**ligand-gated channels**) or stretch (**stretch-mediated channels**). We will consider ligand-gated channels later in this unit.

While the Ca^{2+} channel we are looking at is voltage-gated, it is also modified by being phosphorylated by cAMP-dependent protein kinase. When it is phosphorylated, it stays open longer than normal, thus allowing more Ca^{2+} than normal to enter the heart muscle cell. It is this higher concentration of Ca^{2+} inside the cell that produces a more forceful contraction.

Study Questions:

1. What is it about the atomic structure of an ion that makes it charged?

2. Ca^{2+} is 10^{-3} M on the outside of the cell and 10^{-7} M on the inside. How much of a difference is this? In other words, what is the magnitude of the Ca^{2+} gradient across the cell membrane? If you don't understand molarity, read p 29-30 in your text.

   [NOTE: The magnitude of concentration gradients is expressed in terms of the fold difference across the membrane, e.g. a 10 fold difference, a 30 fold difference -- that is 10 (or 30) times higher on one side than the other. Here is an example of why this makes sense:

	Concentration Inside	Concentration Outside	Arithmetic Difference	Fold Difference

Concentration Gradient A	1000 mM	900 mM	100 mM	1.11 times
Concentration Gradient B	200 mM	100 mM	100 mM	2.0 times

While both concentration gradients have an arithmetic difference of 100 mM, the gradient B is actually almost twice the size of A (2 fold versus 1.11 fold.) Substances will move almost twice as fast down gradient B as they will down gradient A.

3. Describe the chemical structure of an ion channel.

4. What do ion channels do? Why is this function necessary?

5. What does “gated” mean? What is a "gated channel?" What are the three types of ion channels 'gates'?

6. Choose something from your everyday life that could serve as a good model (analogy) for a gated channel. Explain why this item is a good model for a gated channel.

You know epinephrine activates the cAMP second messenger system in heart muscle cells (called **myocardial cells**) and this causes the Ca^{2+} channels in the plasma membrane to stay open longer than normal, allowing more Ca^{2+} to enter the cell down its concentration gradient. We need to look at a few more things before this makes any sense. First, why is Ca^{2+} in higher concentration outside the cell than inside? What creates this concentration gradient and why is it created? And secondly, how does more cytoplasmic Ca^{2+} help the muscle cell contract with greater force?

First, the Ca^{2+} gradient. As you will see throughout this unit, Ca^{2+} is widely used as an **intracellular signal** (a signal within the cell). Cells keep the intracellular, or cytoplasmic, concentration of Ca^{2+} very low when they are "at rest", that is, not receiving a signal. Then, if a communication molecule (e.g. a hormone) causes an increase in cytoplasmic Ca^{2+} concentration, this provides a signal to the cell. Low cytoplasmic Ca^{2+} levels means "don't secrete, or don't contract, or don't pump ions (whatever the cell does for a living -- don't do it!) High Ca^{2+} levels means, "Secrete! Contract! Pump ions! -- Do what you do!"

This Ca^{2+} signaling system must have two elements present in order for it to work correctly. First, the cell has to have a way to keep the cytoplasmic Ca^{2+} levels very low under resting conditions. Secondly, the cell has to have a way to quickly increase the cytoplasmic Ca^{2+} concentration when a signal arrives. Because of its signaling role in the cell, Ca^{2+} is frequently called a **second messenger** or a **third messenger** (though most people don’t distinguish between second and third).

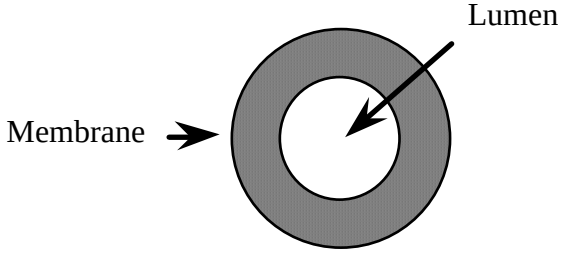
A rapid increase in cytoplasmic Ca^{2+} concentration is allowed by opening of Ca^{2+} ion channels. Something happens (ligand binding, cell stretching, or a voltage change) which causes the Ca^{2+} channels in the plasma membrane to open, thus allowing Ca^{2+} to flood into the cytoplasm. The longer the channel stays open, the more Ca^{2+} ions enter. [Note: We will talk more about exactly how ion channels are opened and closed later in the unit.] For now, however, let's look at how the cell maintains a low level of cytoplasmic Ca^{2+} at rest. This is maintained by an **active transport** system for Ca^{2+} in the cell membrane and in the membrane of the endoplasmic reticulum.

Focused Reading: p 88-91 "Active transport …" stop at "Endocytosis…"
p 126-128 "ATP couples..." stop at "Enzymes biological..."

WWW Reading: Immunofluorescence Labeling of the Sarcoplasmic Reticulum
Animation of Calcium Pump
Purves6e, Ch5, Tutorial 5.1 Membrane Transport (Active transport sections)

Active transport is the movement of substances up their concentration gradient. This violates the second law of thermodynamics (that everything tends toward maximum randomness or entropy – p97 of the text) and therefore REQUIRES THE INPUT OF ENERGY FROM THE CELL. When the cell burns glucose to carbon dioxide and water, energy is given off. The cell harvests this energy and stores it in the phosphate bonds of ATP. When ATP is converted to ADP, the stored energy is released and cellular work can be performed using this energy. Very frequently, though not always, the terminal phosphate released in this reaction is covalently bonded to another molecule (e.g. glucose, or a protein). The molecule is thereby **phosphorylated**. You have encountered phosphorylation previously in its ability to activate or inactive enzymes by covalent modulation. Now you are encountering it again. This time, phosphorylation will be used to provide the energy required to "pump" ions against their concentration gradient. It takes one ATP to move two Ca^{2+} ions into the ER.

In the myocardial cells, there are two sets of **Ca^{2+} pumps** or active transport systems that remove Ca^{2+} from the cytoplasm. One pump is in the plasma membrane and it moves Ca^{2+} from the cytoplasm toward the outside of the cell. The other pump is in the membrane of the endoplasmic reticulum and it moves Ca^{2+} from the cytoplasm into the lumen of the ER. A **lumen** is the inside of a tube or hollow ball. The lumen of a balloon is the space where the air is; the lumen of a garden hose is the space where the water is, etc. In muscle cells, the ER is called the **sarcoplasmic reticulum** or **SR** [sarco = muscle] Therefore, sometimes the ER and SR are referred to jointly as the **SER**.



Both pumps remove Ca^{2+} from the cytoplasm, either by pumping it outside the cell or into an organelle. The latter process is called **sequestering** Ca^{2+} because the concentration of Ca^{2+} becomes very high in the SER. Both pumps cycle by the following mechanism (see the diagram on the next page for an illustration of this mechanism) which is outlined in steps 1 - 7 below:

1. We begin our study of the Ca^{2+} pump cycle at a random point: the pump is dephosphorylated and its Ca^{2+} binding sites are facing the cytoplasm. The calcium binding sites have a very high affinity for calcium ions at this point.

2. Ca^{2+} floating in the cytoplasm binds to the binding sites on the pump, which causes a conformational change in the pump. Even though there is very little Ca^{2+} present in the

cytoplasm, the few molecules present that bump into the binding sites bind tightly to the pump and stay there. The conformational change causes ATP to bind to the ATP-binding site on the pump (it had been vacant).

3. When ATP binds, its terminal phosphate is transferred to the pump, **phosphorylating** it (thus ATP becomes ADP).

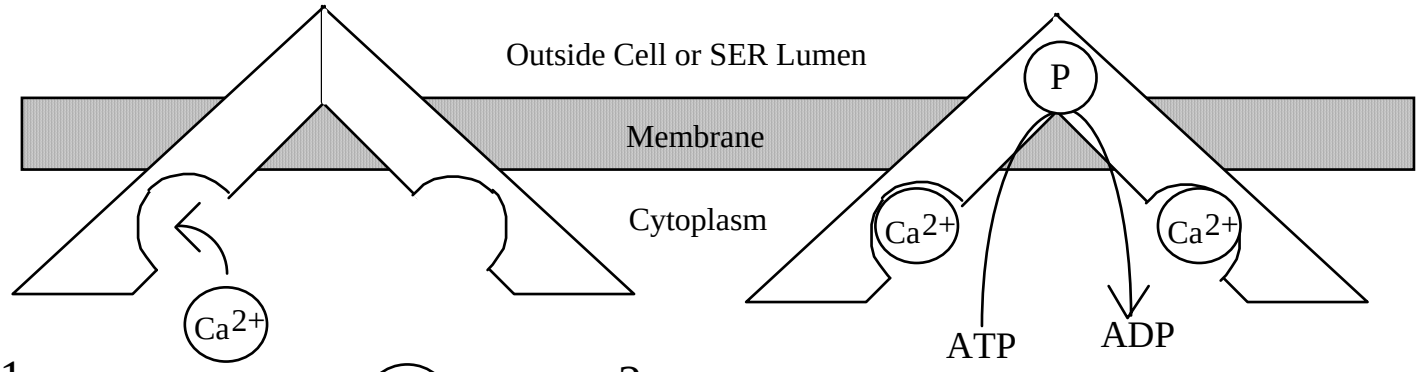
4. This phosphorylation causes the pump to change conformation and present the Ca^{2+} binding sites to the other side of the membrane. (For pumps in the plasma membrane, this flips the Ca^{2+} binding sites outside the cell while the SER pumps now have their Ca^{2+} facing the lumen of the SER. If this seems confusing, review the Purves6e online tutorial Ch5, tutorial 5.1.)

5. Flipping its Ca^{2+} binding sites to the other side of the membrane causes the Ca^{2+} binding sites to have a lower of the affinity for Ca^{2+}. This causes the Ca^{2+} to diffuse off of the Ca^{2+} binding sites and away from the pump.

6. The release of Ca^{2+} from the binding sites causes a conformational change in the pump. This conformational change causes the pump to become **dephosphorylated**.

7. When the pump becomes dephosphorylated, it changes its conformation which makes the Ca^{2+} binding sites flip to the other side of the membrane so they are facing the cytoplasm, which results in the binding sites having a high affinity for Ca^{2+} again. The cycle repeats from step #1.

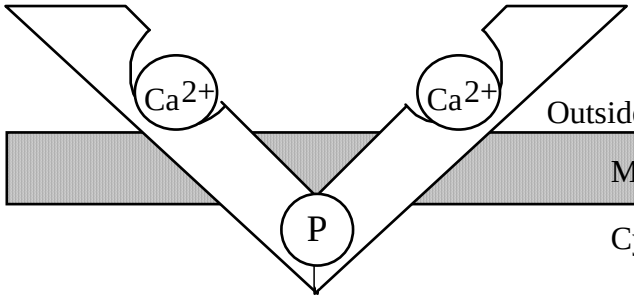
ATP-dependent Ca^{2+} pump moving 2 ions across a membrane



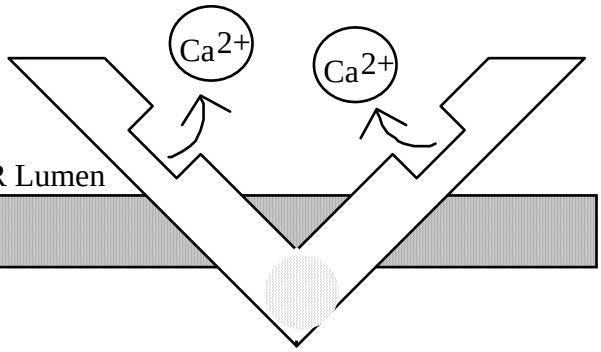
1 Ca^{2+} binding sites face the cytoplasm and have a high affinity for Ca.

Ca^{2+}

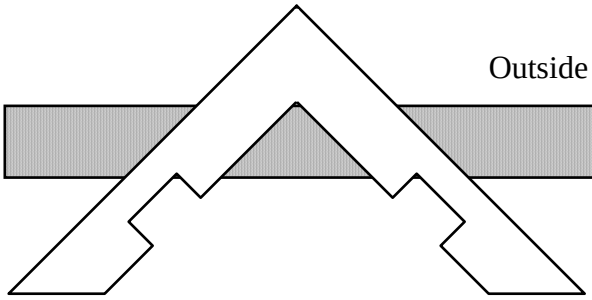
2 Calcium ion binds to the binding sites. This binding causes a conformational change in the pump which causes it to bind ATP and become phosphorylated.



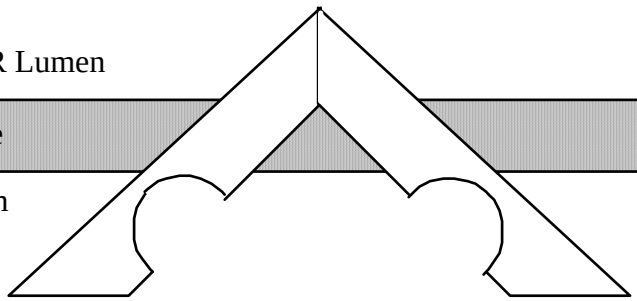
3 The phosphorylation flips the Ca^{2+} binding sites to the other side of the membrane.



4 Flipping the binding sites causes the binding sites to lower in affinity, the ions float free, and the pump becomes dephosphorylated.



5 The pump changes conformation again so that the ion binding sites face the cytoplasm again.



6 Finally, the ion binding sites have high affinty for calcium again.

This process is called **ATP-dependent Ca^{2+} transport** and the pump is called an **ATP-dependent Ca^{2+} transporter.** The role of ATP in the process of transport outlined here is to provide the energy required for the pump to "flip" -- that is, open to the opposite side of the membrane; the flipping event changes the affinity of the ion binding site. The loss or gain of the ion causes changes which allow phosphorylation or dephosphorylation. We spend a lot of energy pumping ions. It has been estimated that we spend 10-20% of all the calories we consume in the active transport of ions. While this active transport accomplishes several other things, one of its main functions is to facilitate signaling.

Study Questions:

1. Explain why heart muscle cells (and all cells, in fact) spend energy pumping Ca^{2+} across their membranes. Explain how Ca^{2+} is used as a signal in cells.

2. Explain the mechanism by which Ca^{2+} is pumped across the plasma membrane and the membrane of the SER. This process requires ATP for energy. How, specifically, is ATP involved in this process?

3. This question provides a slightly different way of looking at the answer you gave in #2. The ATP-dependent calcium transporter changes conformation three times during each pump cycle: 1) The transporter flips toward the inside and outside of the cell; 2) it changes the shape/affinity of its Ca^{2+} binding sites; and 3) it changes the shape/occupancy of its phosphorylation site. What causes each of these changes to occur? (e.g. what causes the pump to flip to the outside, what causes the affinity of the binding site for Ca^{2+} to decrease, etc.) Likewise, each of these changes in conformation causes something to happen. What does each of these changes cause? (e.g. what happens when the pump flips to the outside? What happens when the shape of the phosphorylation site changes?

4. Again, use an analogy to explain the ATP-dependent Ca^{2+} pump. Try to develop an analogy that models all the aspects of the pump.

5. Develop an analogy to explain how Ca^{2+} is used as a signal molecule in the cell. Make sure your analogy can be used to explain how Ca^{2+} is handled by the cell when it is "at rest", i.e. not being signaled.

----------STOP----------

Back to the myocardial cell. To summarize so far, the brain has interpreted something in the environment as frightening, and it has sent a nerve impulse to the adrenal gland to get it to secrete epinephrine. Epinephrine levels in the blood and tissue fluid have risen, and epinephrine has bound to the beta adrenergic receptors on the myocardial cells' plasma membranes. This has triggered the cAMP second messenger system, which has activated cAMP-dependent protein kinase that has phosphorylated (using ATP as the source of the phosphate) the voltage-gated Ca^{2+} channel in the cell

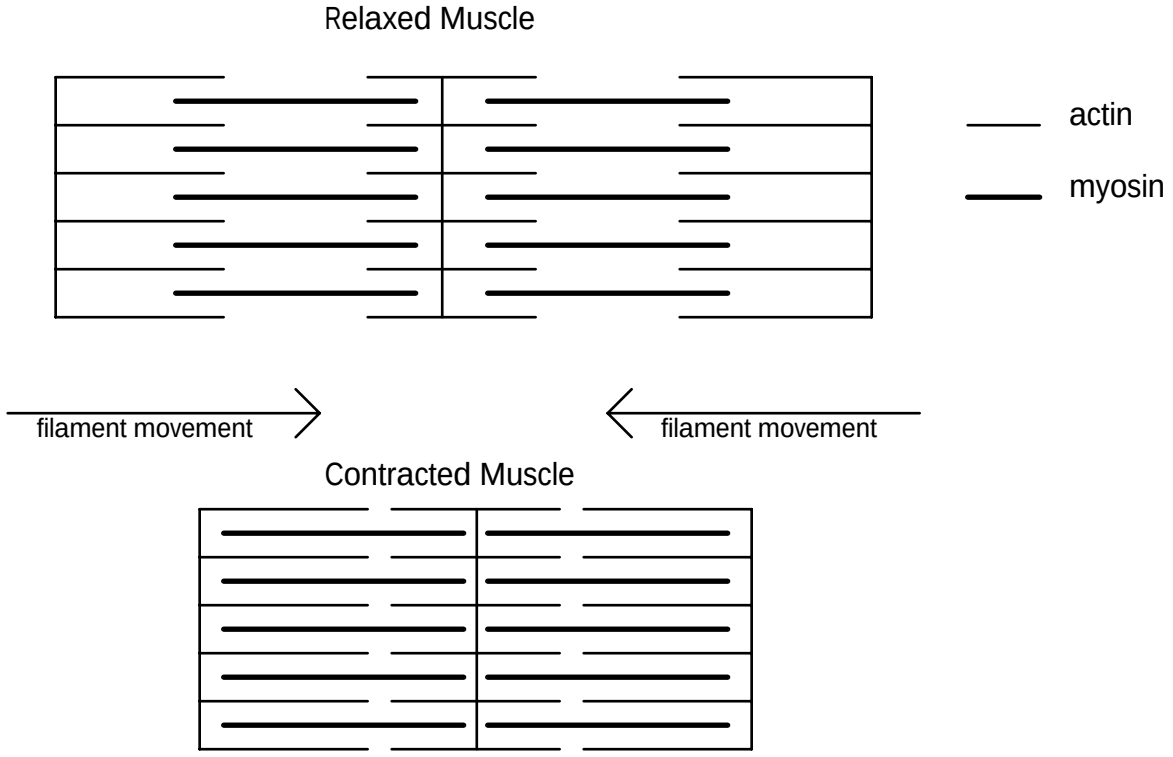
membrane. This phosphorylation has caused this channel to remain open longer than normal. Ca^{2+} has moved down its concentration gradient into the cell. The resting cell maintained this gradient when cellular proteins actively pumped Ca^{2+} into the extracellular space and inside the SER lumen. The phosphorylated Ca^{2+} channel remains open longer than normal allowing more Ca^{2+} than normal to enter the muscle cell.

So, how does this extra Ca^{2+} in the heart muscle cause an increase in contraction strength? In order to look at this question, we need to look at how muscle cells contract. All cells use their **cytoskeleton** to move. But cells which are specialized for contraction have very highly organized cytoskeletal components. These components are specialized **microfilaments** (described in general on p 72-73 and seen in fig. 47.7 p836) called **actin** and proteins that act as motors called **myosin**.

Focused Reading: p. 835-839 "Skeletal muscle" stop at "Single muscle twitch..."
Note Figures 47.8 through 47.10.

WWW Reading: Purves6e, Ch47,Tutorial 47.1
How the Calcium Pump Fills the SER with Calcium

The mechanism used by muscle cells to contract was described by Hodgson and Huxley in the **sliding filament theory** for which they won the Nobel Prize. According to this theory, muscles contract when actin and myosin filaments slide past one another like this:

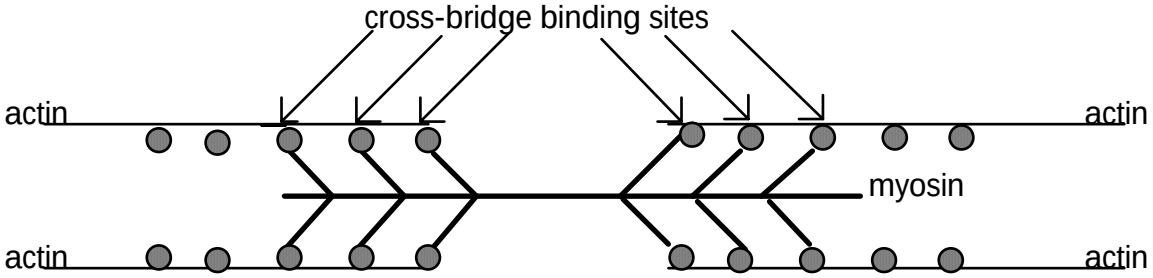


The actin and myosin protein fibers overlap one another. They slide past one another to produce contraction. This sliding movement is very energetic and requires significant amounts of ATP. The sliding is produced by the vigorous movement of parts of the myosin filament, called **myosin heads**, which form **cross-bridges**. These myosin heads bind to the actin and pull, then release and reset,

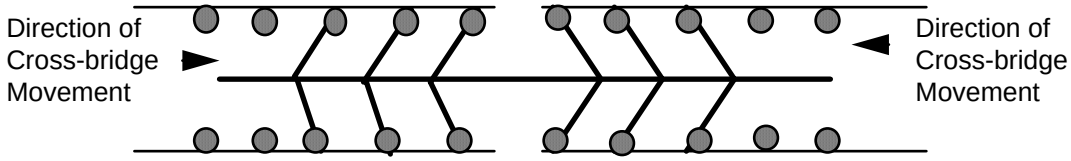
then bind and pull, then release and reset. This process is very much like rowing a boat. There is a diagram of this process on the next page.

So, what does this contraction process have to do with Ca^{2+}? Well, here's the story. When the myocardial cell is at rest (not contracting) the myosin head binding sites on the actin filaments are covered by a protein called **tropomyosin**. Unless tropomyosin is moved, the myosin head crossbridges cannot form and contraction cannot occur. Sitting on the tropomyosin is a second protein called **troponin**.

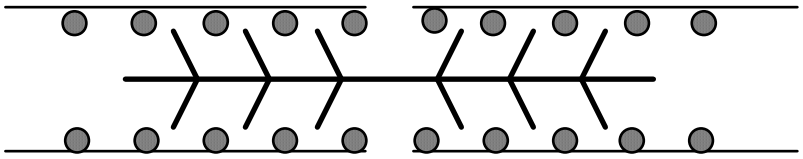
NEWS ITEMS: Mutations in troponin and tropomyosin have been shown to be two causes of Familial Hypertrophic Cardiomyopathy - an autosomal dominant disorder that causes an enlarged heart, disorganized sarcomeres, and ultimately death. (See Thierfelder et al., *Cell*. Vol. 77, 701-712. 3 June, 1994.) Mutations in the actin gene have been shown to cause idiopathic dilated cardiomyopathy. (*Science* 280: 750. 1998)



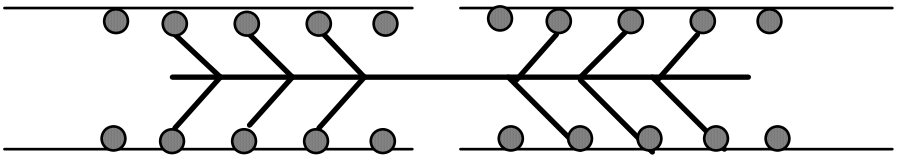
The myosin heads make cross-bridges when they bind to actin.



The myosin heads move (stroke) while bound to the actin. This causes the actin filaments to slide together causing the muscle to shorten.



The myosin heads then break their cross-bridges and swing back to their original position.



The cross-bridges form again and the cycle repeats.

When cytosolic Ca^{2+} levels are low, Ca^{2+} is <u>not</u> bound to troponin (for the same reason that ligands are not bound to proteins when their concentrations are low -- same concept, different example.) When Ca^{2+} is not bound to troponin, troponin has a particular shape that allows tropomyosin to cover the cross-bridge binding sites. The myosin heads can't bind and contraction cannot occur. This all

changes when Ca^{2+} levels are high. Under these conditions, Ca^{2+} binds to troponin. This causes troponin to change shape and this change in shape pulls tropomyosin away from the cross-bridge binding sites. The myosin heads (always ready to bind) can now bind and contraction continues to occur until Ca^{2+} levels fall causing the tropomyosin to cover over the cross-bridge binding sites again.

Ca^{2+} plays a regulatory role in the strength of cardiac muscle contraction. Myocardial cells will not contract at all unless cytoplasmic Ca^{2+} levels rise and tropomyosin is moved out of the way of cross-bridge formation. So increases in cytosolic Ca^{2+} concentration occur $\geq$70 times per minute (on average) in the heart when you are not frightened; this produces the regular heartbeat. However, epinephrine's effect on the plasma membrane Ca^{2+} channel allows it to remain open longer, allowing more Ca^{2+} than normal into the heart muscle cells. Higher Ca^{2+} levels allows more cross-bridge binding sites than normal to be exposed. This allows cross-bridges to be formed -- more oars pulling in the water means more strength -- thus, the force of the cardiac contraction is increased.

Finally, cAMP-dependent protein kinase has an additional action -- it phosphorylates the myosin heads. These phosphorylated heads are capable of "rowing" at a faster rate than when they are not phosphorylated. Therefore, they can produce more strokes per millisecond. Since the limiting factor in this system is the amount of time the cross-bridges are formed, increasing the stroke rate of the cross-bridges increases the amount of movement (i.e. force) that can be generated per unit of time.

Study Questions:

1. Using the sliding filament theory, explain how muscles contract.

2. What role does Ca^{2+} play in muscle contraction?

3. How does epinephrine increase the strength of cardiac contraction? Explain this in detail, as you would for a traditional exam question. Then explain it in simple terms as you would to a younger brother or sister.

4. You have now encountered many ways that ATP is used in the cell. List them and give a brief explanation of each.

5. You have now encountered several examples where an event in the cell is triggered by a change in conformation or shape of a protein. List all the examples you have encountered and briefly describe the effect of the conformational change in each system.

6. "Beta blockers" are drugs that block the beta adrenergic receptor so epinephrine cannot bind to the receptor. These drugs are used to lower blood pressure and to ease the strain on a weakened heart. Explain the mechanism by which beta blockers produce the latter effect.

7. In what ways can a cell increase its permeability to a particular ion? List all the mechanisms you can think of. (As always, answer this in chemical terms.)

NEWS ITEMS: When a muscle is stretched, it tends to contract back to it resting size. The molecule responsible for the ability to spring back is called titin, due to its giant size. Titin can be stretched to 4 times its resting length before it must spring back. This has been visualized for the first time using some fancy microscopy call optical tweezers,

which we will talk about later. (See Erickson. *Science*. Vol. 276: 1090-1092 16 May 1997, and related research articles.)

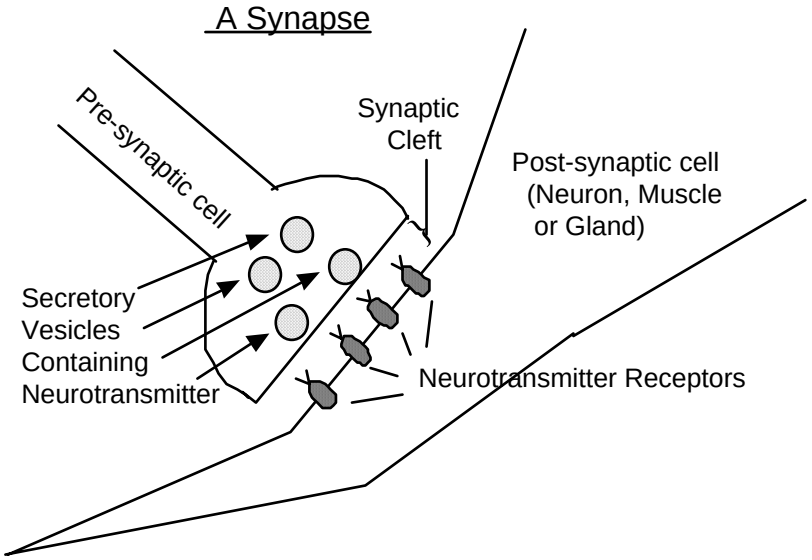
An interesting side note: Psychophysiologists at Univ. Southern California have observed a correlation between low heart rate at rest and aggressive and antisocial personality traits. A colleague commented that this was an interesting finding but what do you do with this information? Is this a good hypothesis?

----------STOP----------

System #3: The Contraction of a Skeletal Muscle in Response to an Impulse from the Nervous System

Brief Overview Reading: Chapter 5 and 44

The previous two systems we looked at involve a hormone as the communication molecule. Now we are going to focus on communication mediated by the other major integrating system, the nervous system. The nervous system and endocrine system both communicate through **chemical messengers.** The endocrine system uses hormones while the nervous system uses **neurotransmitters**. In both cases, the communication molecules are secreted by one cell, travel to the target cell, and bind to specific receptors in the plasma membrane of the target cell. This binding triggers changes in the target cell but unlike hormones, neurotransmitters travel only a very short distance to get from the cell that secreted them to their target cell. When two neurons are communicating (a nerve cell is called a **neuron**), one neuron secretes a neurotransmitter that travels about 0.1 nm to get to the next neuron. ("nm" means nanometer. A nanometer is a billionth of a meter, 10^{-9} meter.) This small gap between neurons, across which the neurotransmitter diffuses, is called the **synaptic cleft** and the area where one neuron interacts with another cell is called a **synapse**. The neuron that secretes the neurotransmitter is called the **pre-synaptic neuron** and the one that bears the neurotransmitter receptor (and binds the neurotransmitter -- the target cell) is called the **post-synaptic neuron**. Here's a picture and there is also one on page 789 in your text (Figure 44.14)



When a neuron is communicating with a muscle cell rather than another neuron, the synapse is called a **neuromuscular junction**. The process of communication from the pre-synaptic neuron to the post-synaptic muscle cell is essentially the same as when two neurons communicate.

Focused Reading: p775-776 "Neurons are…" stop at "Glial Cells…"
p776-777 "Neurons: generating…" stop at end of page 777
p785-786 "Neuron, Synapse and Communication" stop at "The arrival"
p 789 fig44.13

Just as there are many different hormones, there are also many different neurotransmitters. The one we will look at here is **acetylcholine**, which is used as the chemical messenger between the nervous system and skeletal muscle cells. A list of neurotransmitters can be found on page 790 in your text.

In looking at this system, we will begin at the beginning of the process and look at the cellular and chemical events that produce a nerve impulse. Nerve impulses are electrical events, that is, they are caused by the flow of charged particles (in this case ions move instead of electrons as is the case for electrical current in power lines.) The flow of charged particles is called **current**. The force that moves charged particles (causing them to flow (causing current)) is **voltage** or **electrical potential**. Voltage is a separation of charge. According to the second law of thermodynamics, charged particles will move in such a way that electrical neutrality is produced (an equal distribution of positively and negatively charged particles). Thus, if you separate positive particles from negative particles -- create concentration gradients of negative and positive charges, you have created a voltage -- a potential force that will compel charged particles to move to correct this imbalance -- to create electrical neutrality. In doing this, negatively charged particles will move toward the concentration of positive charge and *vice versa* until they are completely mixed and the solution is electrically neutral. When voltage exists, then, there is always a negative **pole** and a positive pole -- like a battery. The negative pole (the **cathode**) attracts positively charged ions (called **cations**) and the positive pole (the **anode**) attracts negatively charged ions (called **anions**.) The bigger the separation of charge, the bigger the voltage (the more current will flow between the poles).

At rest (that is, when no signal is being sent or received), the plasma membranes of all cells, including neurons, have a **voltage** across them. The outside of the cell is the positive pole and the inside of the cell is the negative pole. The separation of charge across the membrane is small with a voltage of only **-60 millivolts (mV)**. [By convention, the voltage is given the sign of the pole that is inside the cell. So a voltage of -60 mV means that the magnitude of charge separation is 60 mV with the inside of the cell negative with respect to the outside.] -60 mV then is said to be the **resting membrane potential**. This voltage exists in all cells at rest and in neurons when they are not propagating an impulse.

It is this voltage that allows the creation and propagation of a nervous impulse. This voltage is used in other cell types to transmit signals as well. But before we look at that, we have to look at how this voltage is created in the first place. Separation of charge can be thought of as a charge concentration gradient. Just as the Ca^{2+} gradient was created by an active transport system, so is the membrane voltage.

Focused Reading: p 88-91 "Active Transport" stop at "Endocytosis"
Pay special attention to Fig 5.12, p90

p 777-778 "Ion pumps..." stop at end of page 778.
p 777 Fig 44.5 (Research Method)

The **ATP-dependent Na^+/K^+ pump** operates exactly like the ATP-dependent Ca^{2+} pump we have already considered. It is a bit more complicated because it transports two ions in opposite directions across the membrane. [This pump is, therefore, called an **antiporter** (two substances, opposite directions.) An antiporter is an example of a **cotransporter** (two substances transported at once in any direction (both inward, both outward or one in and one out.)) The Ca^{2+} pump, on the other hand, is called a **uniporter** because is transports only one substance.]

Even though the Na^+/K^+ pump transports two ions, the same rules apply here as they did in the calcium ion pump:

1. Phosphorylation and dephosphorylation cause the pump to flip

2. The flip causes a change in the ion binding sites' affinities

3. The loss or gain of ions into the binding sites causes the pump to be phosphorylated or dephosphorylated

4. The cycle repeats.

The relationship of K^+ to phosphorylation and site affinity is exactly opposite that of Na^+ since they are being transported in opposite directions across the membrane.

It is the unequal transport of potassium ion versus sodium ion that CREATES THE RESTING MEMBRANE VOLTAGE. The Na^+/K^+ pump transports three sodium ions (3+) to the outside of the cell for every two potassium ions (2+) it transports to the inside of the cell. Therefore, the pump separates charge -- that is, it pumps more positive charge to the outside than it does to the inside. The pump is, therefore, said to be **electrogenic** (it generates voltage). This makes the outside of the cell membrane slightly positive and the inside slightly negative. The magnitude of this charge difference is 60 mV (technically, -60 mV since the inside is negative).

Note that the Na^+/K^+ pump has to keep pumping constantly because the membrane has Na^+ and K^+ channels in it that "leak." This situation is analogous to bailing a leaking boat. You have to keep bailing because the water keeps leaking back into the boat. But if you bail as fast as the boat leaks, you can stay afloat. Likewise, at "rest" (this means -- no signal is being sent. It does not mean the cell is inactive), the Na^+/K^+ pump bails as fast as the channels leak -- so a steady state is maintained. In this steady state, because of the action of the pump, three significant conditions exist:

1. There is a concentration gradient for Na^+ across the cell membrane. The concentration of Na^+ is very high on the outside of the cell and very low on the inside of the cell. This gradient is produced by the Na^+/K^+ pump moving Na^+ from the inside of the cell to the outside.

2. There is a concentration gradient for K^+ across the cell membrane. The concentration of K^+ is very high on the inside of the cell and very low on the outside of the cell. This gradient is produced by the Na^+/K^+ pump moving K^+ from the outside of the cell to the inside.

3. There is a voltage across the membrane. This voltage is produced by a separation of positive charge such that more positive charge is placed on the outside of the cell than on the inside. This voltage is produced by the unequal pumping of Na^+ and K^+ by the Na^+/K^+ pump.

Study Questions:

1. What is voltage? What is current? How are these two concepts related?

2. Explain the concepts of voltage and current using an analogy.

3. Describe the mechanism the ATP-dependent Na^+/K^+ pump uses to move ions across the membrane.

4. How is the resting membrane potential created? What causes the outside of the cell to be positive and the inside to be negative?

5. Explain why the Na^+/K^+ pump has to pump ions all the time. Use an analogy (other than the leaky boat analogy) to describe this phenomenon.

So here the neuron sits with its resting membrane potential at -60 mV and its concentration gradients for Na^+ and K^+ well established. The neuron that will ultimately synapse on the muscle cell and get it to move is called a **motor neuron**. It is called "motor" because it causes movement (as opposed to a sensory neuron that carries sensation). If you want to move your leg, you send an impulse from your brain down to the spinal cord. In this area sit the motor neurons that communicate with your leg muscles. The neurons coming down from the brain synapse on the motor neurons and secrete a neurotransmitter onto their membrane. (The type of neurotransmitters used at this synapse is not known, but probably several different neurotransmitters are involved.) This causes a change in the motor neurons which causes nerve impulses (called **action potentials**) to be transmitted across the motor neuron out to the muscles of the leg.

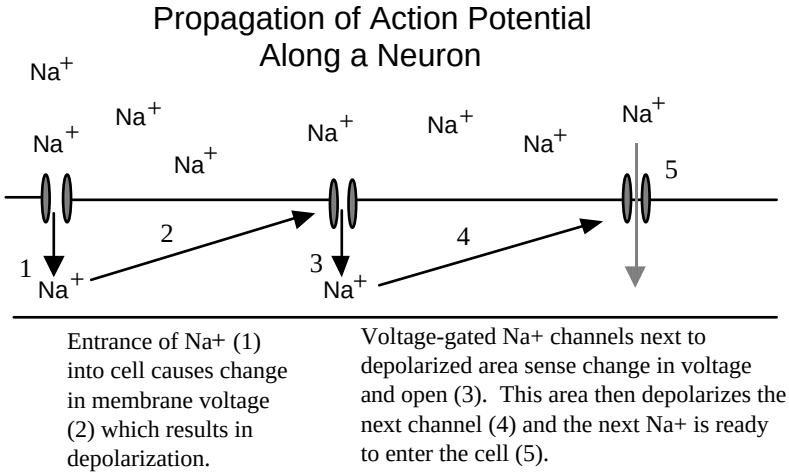
Focused Reading: p 777-781 "Simple electrical..." stop at "Action…"

When the motor neuron receives the message from the brain, the neurotransmitter binds to its receptor in the plasma membrane of the motor neuron. This neurotransmitter receptor is physically linked with a **Na^+ channel**. (See Fig 44.15 on page 786). When the neurotransmitter binds, it causes a change in shape (surprised?) in the receptor. This change in shape causes the Na^+ channel to open. This is an example of a **ligand-gated channel**. The binding of a ligand (the neurotransmitter) causes the opening of the channel.

When the channel opens, Na^+ is free to move quickly down its concentration gradient. Because it is in high concentration outside the cell and low concentration inside, it quickly moves into the cell. When Na^+ moves, it carries its positive charge with it. Therefore, the inside of the cell becomes more

and more positive as more and more Na^+ enters. This movement of positive charge causes a change in the membrane voltage. At rest, it is -60 mV but as sodium enters it becomes -59, -58, -57, -56, etc., until it reaches -40 mV. -40 mV is called the **threshold potential**. At this point, the voltage is sensed by **voltage-gated Na^+ channels** in the neuron's plasma membrane near the recently opened ligand-gated channel. These voltage-gated Na^+ channels change shape when the voltage reaches -55 mV. This change in shape causes them to open allowing <u>even more</u> Na^+ to flood into the cell. Now the voltage continues to become more positive (-40, -30, -20, etc.) until the inside actually becomes positive with respect to the outside at +50 mV. This is called **membrane depolarization** because the original poles (negative inside and positive outside) have been obliterated. (see left side of fig 44.8 pg780)

There are voltage-gated Na^+ channels all the way along motor neurons from their bodies in the spinal cord to their synaptic terminals in the arm muscles. As each area of membrane reaches threshold voltage (-40 mV) the depolarization is sensed by the neighboring voltage-gated Na^+ channels and triggers the channels to open one by one all along the way to the end of the neuron. This propagation of a wave of depolarization is called an **action potential** - a "nerve impulse." This will be the command signal to the muscle to contract. This process can be diagrammed as follows:



Entrance of Na+ (1) into cell causes change in membrane voltage (2) which results in depolarization.

Voltage-gated Na+ channels next to depolarized area sense change in voltage and open (3). This area then depolarizes the next channel (4) and the next Na+ is ready to enter the cell (5).

Do you see why nerve impulses are described as "waves of depolarization?"

Study Questions:

1. Explain how the resting membrane potential makes the action potential possible. Describe this in actual chemical terms, and then describe it using an analogy.

2. Ligand-gated and voltage-gated ion channels are <u>both</u> involved in the generation of an action potential. Describe the role played by each type of channels.

3. Explain how the opening of Na^+ channels in the plasma membrane produces a change in membrane voltage. Why is this change called depolarization?

[Note: The cell membrane has to be returned to its resting state before it can send another signal through an action potential. This membrane **repolarization** is produced by voltage-gated K^+

channels that open after the voltage-gated Na^+ channels. This causes K^+ to flow down its concentration gradient towards the outside of the cell. K^+ carries its positive charge with it and makes the inside of the membrane more negative as it leaves. This (along with the closing of the Na^+ channels) repolarizes the membrane.]

Focused Reading: p 781-83 "Action potentials are..." stop at "Ion channels and…"
p 782 Fig 44.11
p 781 Fig 44.10
(If you can explain these to your non-science friends, you understand action potentials.)

Study Questions:

1. What role does K^+ play in an action potential?

2. Make a list of the similarities between K^+ and Na^+ in an action potential. List the differences
between these two ions in an action potential.

> NEWS ITEM: A research team from Australia and Germany has cloned a gene that is expressed only in the brain and causes a certain type of epilepsy, called benign familial neonatal convulsions. The convulsions start about 3 days after birth, but usually disappear within a few months. The culprit is a mutation in a potassium channel gene that in one family had a 5 base pair insertion, which cause the deletion of 300 amino acids at the end of the protein. If the channel is not full length then it cannot work properly (structure – function relationship again). (*Science* Vol. 279: 403. January 1998)

3. What would happen to a neuron that contained the mutation described above?

----------STOP----------

Eventually the wave of depolarization, the action potential, reaches the synaptic terminal of the motor neuron. This neuron terminates at a skeletal muscle cell in a leg muscle; the terminal structure is called a **neuromuscular junction**. This neuromuscular junction looks very much like the synapse you have already encountered, where the neuron from the brain told the motor neuron to start its action potential. When the action potential reaches the synaptic terminal of the motor neuron, it causes the membrane of the pre-synaptic cell to depolarize (just like all the rest of the membrane all the way down from the spinal cord.) However, the synaptic terminal contains **voltage-gated Ca^{2+} channels** in its membrane. When the membrane depolarizes, these voltage-gated Ca^{2+} channels open and Ca^{2+} flows down its chemical concentration gradient into the synaptic terminal. (NOTE: When the cell was "at rest" the Ca^{2+} gradient was produced by the same plasma membrane Ca^{2+} pump that works in the heart muscle. Virtually all cells pump Ca^{2+} out of the cytoplasm using this pump.)

The synaptic terminal of the motor neuron contains secretory vesicles full of the neurotransmitter **acetylcholine**. When Ca^{2+} enters the terminal, these vesicles fuse with the plasma membrane and release their contents into the synaptic cleft. The process of secretion is an example of **exocytosis**.

Focused Reading: p91-92 stop at "Membranes are not simply…"
p92-94 Membranes are dynamic" to end of chapter
p 91 Fig 5.14b
p 787 fig 44.16

WWW Reading: Movie of Calcium Influx into a Neuron
Purves6e, Chapter42, Tutorial 42.2 Synaptic transmission

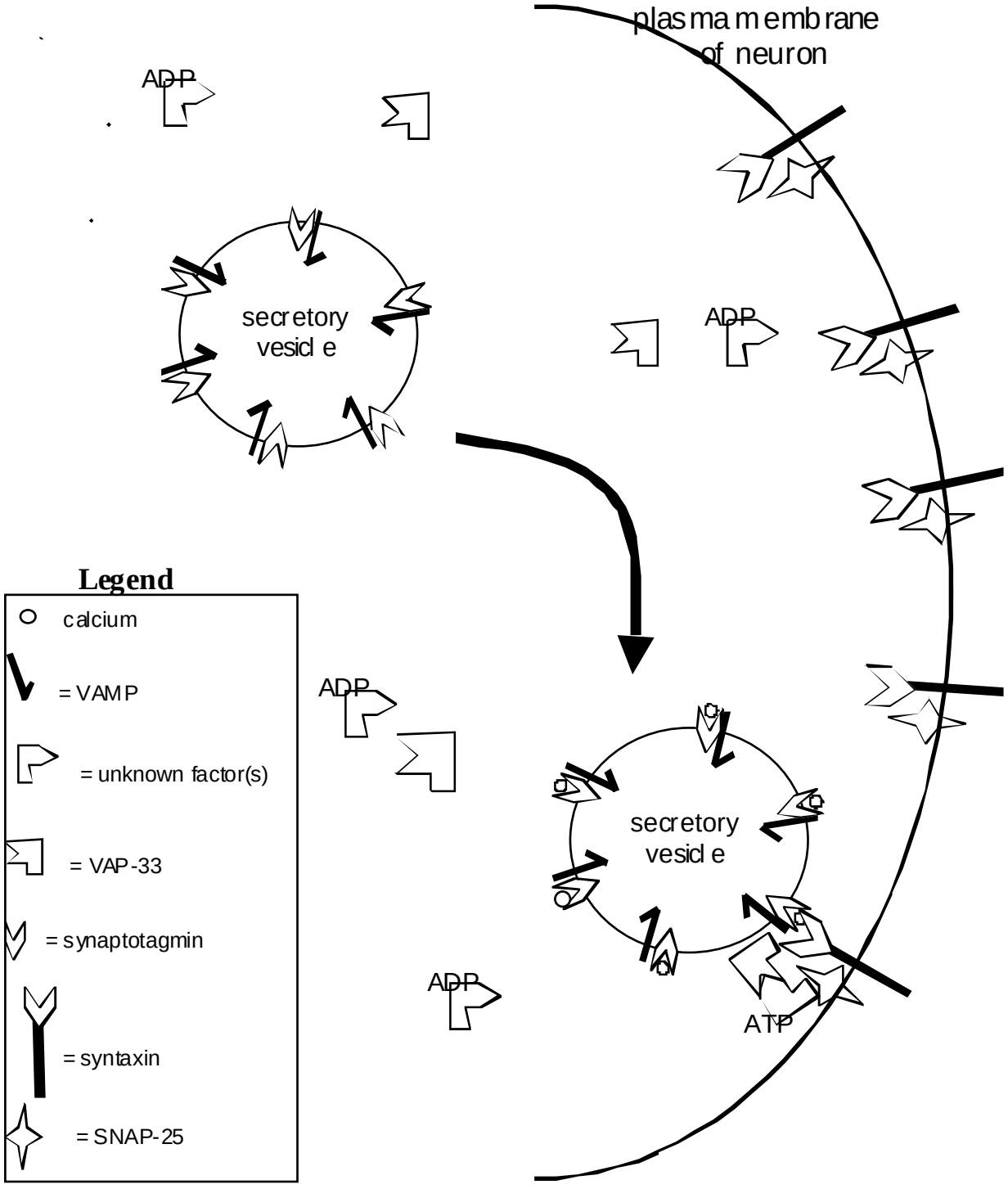
Exocytosis and endocytosis are mirror images of one another. Cells use exocytosis to secrete products (e.g. hormones, neurotransmitters, cell wall components, milk, digestive enzymes, sweat, tears, etc.). Cells use endocytosis to engulf cells and other substances, usually for utilization by the engulfing cell. Cells engulf bacteria, viruses, dead cells from one's own body, proteins, water, iron, etc. Cells also use endocytosis to retrieve membrane added during exocytosis and *vice versa*. This process is called **membrane recycling** or **membrane traffic**. Some cells sit and secrete constantly. Their secretion is said to be **constitutive**, that is, it occurs constantly and is not regulated. Other cells, such as neurons, store their secretory product and wait for a signal to secrete; this is **regulated secretion**. The signal to secrete is usually a rise in intracellular Ca^{2+}.

So, how does a rise in the level of cytoplasmic Ca^{2+} trigger secretion in most cells and, specifically, in motor neurons? We don't know the complete answer, but a story has emerged which is currently very popular. (Note: Results from this field of investigation are so new that investigators have not yet settled on a common set of terms for the molecules involved. The terms presented here are from one lab investigating neurons.) According to this theory, there are membrane proteins in secretory vesicles called **VAMP**s (vesicle associated membrane protein) that can bind membrane proteins in the plasma membrane called **syntaxin** (proteins discovered in synaptic vesicles that facilitate vesicle movement (taxis = movement.)). Both VAMP and syntaxin are **integral** membrane proteins (that means they go through a membrane, remember). Most of the mass of VAMP is on the cytoplasmic side of the vesicle membrane. Most of the mass of syntaxin is on the cytoplasmic side of the plasma membrane. The cytoplasmic portions of VAMP and syntaxin act as 'handles' that allow the vesicle to come in contact with the plasma membrane. When these two proteins (VAMP and syntaxin) bind to one another, the vesicle and plasma membranes fuse and the contents of the vesicle (in this case, acetylcholine) is secreted into the synaptic cleft. The binding of VAMP to syntaxin is facilitated by some additional cytoplasmic proteins (called NSF and SNAPs) and the entire process is ATP-dependent. The VAMP-syntaxin-cytoplasmic protein(s) complex falls apart when ATP is hydrolyzed to ADP by an enzymatic site on one of the proteins. (See figure on next page.)

NEWS ITEMS: Release of neurotransmitter has been visualized for the first time by using genetic engineering and the lightening bug enzyme luciferase. Every time these modified cells secrete neurotransmitters, they also produce a small spark of light, which can be seen through a microscope. This allows researchers to determine how many vesicles fuse with the plasma membrane for any given stimulus. (See Miesenbock and Rothman. *Proc. Natl. Acad. Sci. USA.* Vol 94: 3402. April, 1997.)

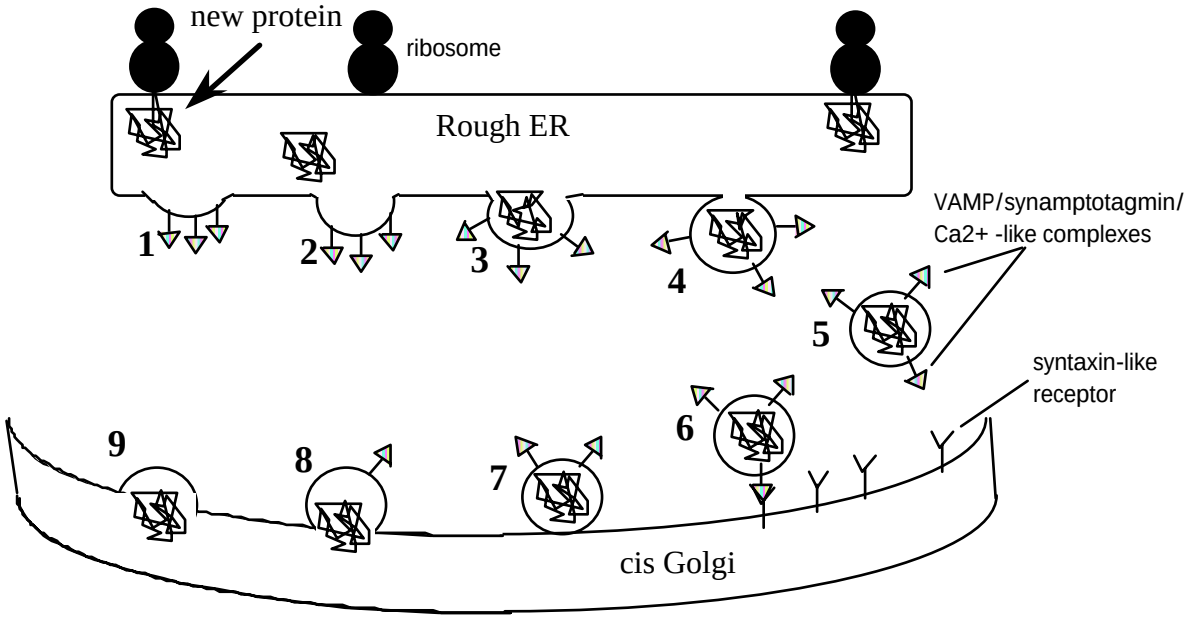
Nicotine that is found in cigarette smoke has been shown to bind to presynaptic acetylcholine receptors and cause a rise in intracellular calcium at the nerve terminus. This leads to an increase secretion of other neurotransmitters. (Science Vol 269: 1692. September 1995) However, a more recent study has shown that chronic exposure of nicotine can cause two of the three known versions of the acetylcholine receptor to become permanently inactivated. The third receptor is still functional which leads to increased neurotransmitter (dopamine) release and thus the craving for nicotine is sustained. (Olale *et al*., *J. Pharmacology and Experimental Therapeutics* 283: 675. 1997)

The protein complex is hypothesized to look something like this:



While high cytoplasmic Ca^{2+} levels clearly trigger this binding and subsequent membrane fusion, the mechanism of calcium triggering exocytosis is not completely understood. Control is likely to occur through a second protein bound onto the vesicle membrane called **synaptotagmin**, a Ca^{2+}-binding protein. Investigators hypothesize that, when synaptotagmin binds Ca^{2+}, it changes shape. This change in shape causes a change in shape in the VAMPs which allow them to bind to syntaxin on the inside of the cell membrane, thus allowing secretion to occur.

In addition to mediating secretion, this process of protein-mediated binding followed by membrane fusion is the way substances are transported and specifically sorted within the cell. For instance, some cells secrete protein (e.g. insulin and other protein hormones, digestive enzymes). This protein is made on ribosomes which are "docked" onto the surface of the ER. As the protein is made, it is translocated ('moved across') into the lumen of the ER. From there, the protein must travel to the Golgi apparatus and then to secretory vesicles for secretion. The protein is first concentrated into a specific region of the ER. Small vesicles containing the protein then bud off the ER and fuse with the **cis face of the Golgi.** The protein product then transported through the Golgi by budding off each Golgi **cisterna** and fusing with the next cisterna until the **trans face** is reached. Then, secretory vesicles bud off the trans face, ready to release their contents outside the cell upon receiving the appropriate signal. The process looks like this:



Evidence is building that all of the budding and coalescing processes (called **vesicular transport**) are mediated by protein complexes that function like the VAMP-syntaxin-cytoplasmic protein complex that mediates secretion. Thus, if one asked how the vesicle that buds off the ER "knows" to fuse with the Golgi and not with a mitochondrion (or the nucleus or the plasma membrane), the answer is probably that this ER vesicle contains a VAMP-like protein that is specific for (that is, complementary in structure to) a syntaxin-like molecule on the Golgi membrane. Thus, the transfer of proteins within the **endosomal system** (the system of organelles in the cell that includes the Golgi, ER, lysosomes, phagocytic vesicles, and secretory vesicles--reviewed in pages 64-66) is probably mediated by specificity of membrane-bound "docking" proteins.

NEWS ITEM: A protein called syntaxin 5 has been identified as a necessary molecule for the formation of Golgi and the delivery of vesicles to the Golgi. (*Science* Vol. 279: 696. January 1998)

The endomembrane is not just a one-way street for protein synthesis and secretion. Vesicles also travel in the other direction and are sometime used 'against us'. The toxin produced by *Shigella dysenteria* (Shiga toxin) enters the cell by endocytosis and causes hemorrhagic colitis. Shiga toxin and has recently been shown to travel all the way 'down' the

pathway to the ER before escaping into the cytoplasm and wreaking havoc on the cell. F. Mallard et al. 1998 *JCB* vol 143 973-990.

But we digress, back to the neuromuscular junction. After acetylcholine is secreted into the synaptic cleft, it diffuses the very short distance to the post-synaptic membrane of the skeletal muscle cell where it binds to an acetylcholine receptor. Acetylcholine receptors are in part **ligand-gated Na^+ channels**. (Again, see page 786; Figure 44.15 for an illustration.) Thus, the binding of acetylcholine to the muscle cell membrane triggers an action potential that is spread across the muscle cell membrane in exactly the same way that the action potential was spread down the neuron. The mechanism of generation and propagation are the same.

The action potential that spreads across the muscle cell membrane triggers a rise in (guess what?) intracellular Ca^{2+} levels in the muscle cell (this is a very hot area in research today). By the same mechanism as in heart muscle, this Ca^{2+} binds to troponin causing it to pull tropomyosin away from the cross-bridge binding sites on the actin filaments. Contraction is sustained for as long as cytoplasmic Ca^{2+} levels remain high. And cytoplasmic Ca^{2+} levels remain high as long as an action potential is being propagated along the muscle cell membrane. And an action potential is propagated as long as acetylcholine is bound to its receptor. And acetylcholine is bound to its receptor as long as it is secreted by the pre-synaptic neuron. An enzyme in the synaptic cleft called acetylcholinesterase destroys acetylcholine almost immediately. Therefore, the pre-synaptic cell must provide a continual supply of the neurotransmitter if the receptor is to remain activated.) And the pre-synaptic neuron secretes acetylcholine as long as action potentials continue to reach the synaptic terminal. And action potentials reach the synaptic terminal as long as they are generated at the cell body in the spinal cord, which continues as long as the brain is telling you to flee the bear chasing you. Now at this point things get a little vague. But neuroscientists practice their craft believing that everything, even the will to contract, is caused by a chemical process. Practitioners of neuroscience have a lot of work to do before they truly understand these higher neurological functions.

One more thing -- how does the action potential in the muscle cell membrane actually cause an increase in cytoplasmic Ca^{2+} levels in the muscle cell? We don't know. We do know that most of the Ca^{2+} in this process comes from inside the SR where it has been pumped by the ATP-dependent Ca^{2+} pump while the cell was at rest. Thus, the action potential in the muscle cell membrane must trigger the opening of a Ca^{2+} channel in the membrane of the SR. Ca^{2+} then is free to flow down its concentration gradient into the cell. We do not know the exact mechanism whereby the membrane action potential communicates with the Ca^{2+} channel in the SR. However, recent research has pointed to a protein called **triadin** which spans the gap between the plasma membrane and the Ca^{2+} channels in the SR (triadin spans the triad and we are not referring to 3 cities in NC!). Therefore, triadin is in the right place to bridge the gap but no one has any idea how this could be done. (See figure 44.9 pg. 981 for a picture of where the T tubule meets the SR.) Whoever figures it out first will become very famous since this is the last big mystery about muscle contraction.

Study Questions:

1. What events are triggered by the arrival of the action potential at the synaptic terminal?

2. Describe the process of exocytosis.

3. List some ways in which cells use exocytosis and endocytosis. How are these two processes used together to ensure that the cell's size does not change?

4. Describe the current theory that explains how increased Ca^{2+} concentrations trigger secretion.

5. Describe the process by which protein travels from the ER through the Golgi and into secretory vesicles. How is this process controlled so that the correct vesicles coalesce with the correct target organelle?

6. Muscle cells and neurons are physiologically more similar than one might think. In what ways are these cells similar in their chemical responses? What types of membrane receptors and channel proteins do both types of cells have? In what ways are these two cell types different in their chemistry and responses?

7. The action potential in the muscle cell membrane causes a rise in cytoplasmic Ca^{2+} levels. Where does this Ca^{2+} come from? How does it enter the cytoplasm? What must the muscle cell do when it is at rest to ensure that this signaling system will work?

8. Outline the entire pathway in chemical terms from wanting to move your arm to actually moving your arm. Tell this story using chemical and cellular language as you would for a traditional exam question. Then tell it in simpler terms as you would to a younger sister or brother. Use as many good analogies as you can.

----------STOP----------

System #4: How does a sperm cell tell an egg it has been fertilized?

We have studied 3 specific cases of signal transduction: 1) epinephrine bound to a liver cell receptor to tell the cell to put more glucose into the blood; 2) epinephrine bound to a heart cell receptor to tell the cell to contract harder; and 3) neurotransmitters bound to skeletal muscle receptors to tell the muscles to contract. Now we will examine a slightly different situation for signal transduction, fertilization.

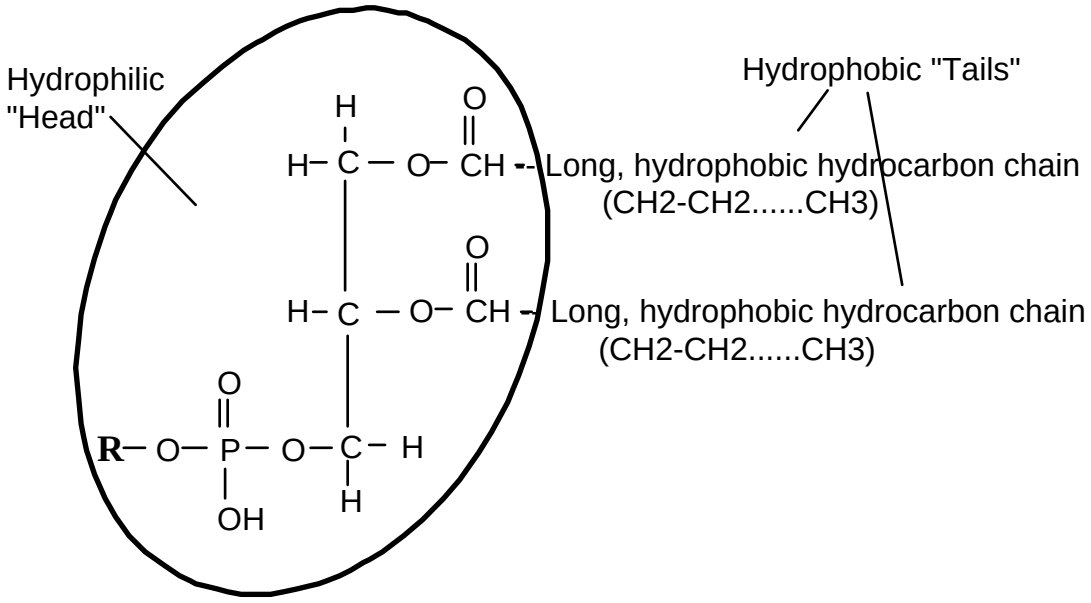
One thing is certain -- once nature develops a good system (of moving, of secreting, of transporting, of communicating) it keeps using it over and over again. The communication system used by an egg to sense fertilization is an ancient one. So ancient, in fact, that it developed before sea urchins, frogs, fish and mammals diverged from one another during evolution. How do we know this? Because all animal eggs use the same communication system. Odds are that evolution would not have produced the same communication system twice through random mutation and selection. Therefore, the system probably evolved before the split and all animals took this good idea with them as they diverged into different species. This communication system is called the **inositol triphosphate second messenger system**.

Focused Reading: p 287-8 " Two second messengers" stop at "Calcium ions are…"

p 288 Figure 15.13

The inositol triphosphate (IP_3) second messenger system uses the same kind of receptor-G protein-enzyme link that the cAMP system does. Except, this time, the enzyme linked to the G protein is not adenylyl cyclase, it is **phospholipase C**. Like most enzymes, this one tells you what it does. It cuts up a phospholipid.

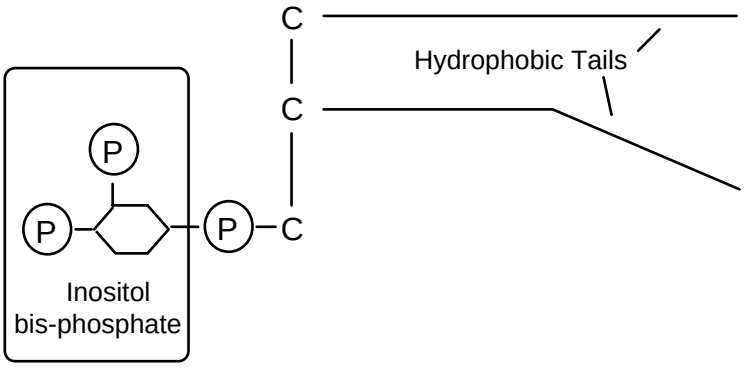
Let's talk a little more about phospholipids. You know that they make up the cell membrane and you know why. They have a hydrophilic "head" that dissolves in the water inside and outside the cell and they have hydrophobic tails that hate water and dissolve in each other in a hydrophobic bilayer. To refresh your memory:



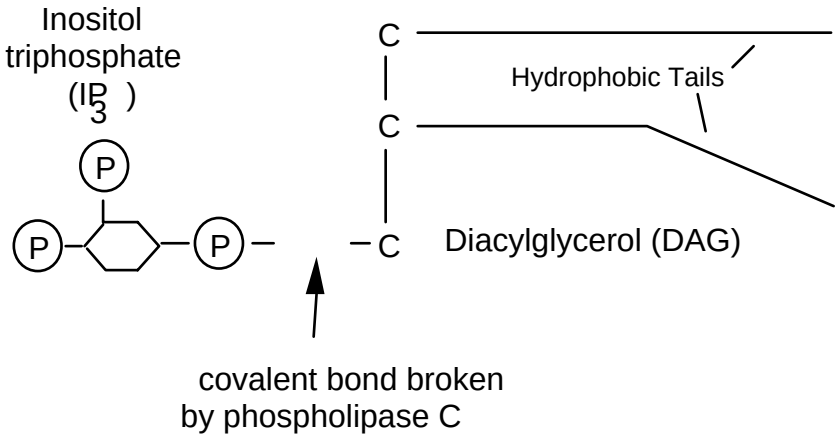
There are many different kinds of hydrophilic molecules that can be added to the phosphate on the phospholipid. These are added at the "R" in the diagram above. Regardless of what is added, these molecules are all highly polar and many of them are charged which greatly increases the hydrophilic nature of the "head" of the molecule. Some examples of molecules that are added to the phosphate group at "R" are serine, choline, and inositol. Phospholipids are named according to the molecule added to the phosphate. All phospholipids start with "phosphotidyl (blank)" and then the name of the added molecule fills in the blank. Thus, if serine were added, the phospholipid would be called **phosphotidylserine**. If choline were added, it would be called **phosphotidylcholine**. And if inositol were added, it would be called **phosphotidylinositol**.

Some phospholipids have **inositol bis-phosphate** added to their phosphate group. "Bis" means "two". So, inositol bis-phosphate is inositol with two phosphate groups on it. When inositol bis-phosphate is added to a membrane phospholipid, the resulting molecule is called **phosphotidylinositol bis-phosphate**. The abbreviation for this molecule is **PIP_2.** This is the substrate molecule for phospholipase C. PIP_2 can be diagrammed simply as follows:

Phosphotidylinositol bis-phosphate



PIP_2 sits in the <u>inner layer</u> of the plasma membrane's lipid bilayer. Like all other phospholipids in this layer, it has its hydrophobic tail embedded in the lipid bilayer and its "head" facing the cytoplasmic side of the membrane. When phospholipase C is activated by G protein, it cuts inositol off of PIP_2 in such a way that <u>all</u> the phosphates go with inositol and none remain on the lipid in the membrane. The products of this cleavage look like this:



The inositol with the three phosphates is called **inositol triphosphate (IP_3)**. This hydrophilic molecule floats away from the membrane into the cytoplasm where it will act as a <u>second messenger.</u> The other part of the molecule is called **diacylglycerol** (like triacylglycerol with two instead of three fatty acids chains.) This is abbreviated **DAG**. DAG stays embedded in the membrane, but nonetheless also acts as a <u>second messenger</u>. So the cleavage of PIP_2 by phospholipase C results in two cleavage products: IP_3 and DAG, both of which will now act as second messengers.

Study Questions

1. Describe or draw a simple diagram (like the ones presented above) of a triacylglycerol, a generic phospholipid, diacylglycerol, phosphotidylinositol bis-phosphate, inositol triphosphate. (If you need more information about lipids and phospholipids, see page 42-44 in your text)

2. Describe the pathway through which phospholipase C is activated.

3. Describe the enzymatic action of phospholipase C. What is the substrate for this enzyme and what are the cleavage products of the reaction? What general function do these cleavage products have in the cell?

We have all seen those classic films of sperm fertilizing an egg - the egg is surrounded by hundreds of sperm trying to penetrate the egg's plasma membrane. Why is it that only one sperm cell manages to fertilize an egg? With all those sperm cells trying to reach the same goal at the same time, you would think that at least two might enter the egg at about the same time (**polyspermy**). If this were to happen, the resulting zygote would be in trouble since it would have 3 haploid genomes (3 copies of each chromosome) instead of the normal 2 copies (we will cover this later in the Section II). An egg with 3 haploid genomes will probably not survive the process of development.

Evolution is a process of natural selection; natural selection allows those with favorable traits or abilities to reproduce. This results in advantageous traits or abilities being maintained in the population while less advantageous ones may disappear. Natural selection plays a part in all levels of molecular and cellular biology and fertilization is no exception. Any egg (i.e. organism that produces this egg) that has "learned" how to permit only one sperm to fertilize it will be more likely to produce a new individual that will have the same selective advantage its mother had which will in turn result in more successful matings for the mother's offspring.

So the question remains how has evolution (natural selection) given rise to an egg that permits only one sperm to fertilize it? Evolution is not a wasteful process. In fact, evolution was the first to recognize the importance of recycling. We have talked about G proteins that were coupled to receptors, which resulted in the production of cAMP as a second messenger. To "invent" a new second messenger system to facilitate signal transduction, evolution thought to herself, "How can I move across the plasma membrane the information that a sperm has arrived while not having to come up with a totally new mechanism?" The answer is beautiful in its similarity, or **homology**, to the cAMP messenger system but with a subtle twist to achieve a very different set of responses.

Focused Readings:
p 76-77 "Animal cells…" to end of chapter
p 734-36 "Sexual reproduction " stop at "Oogenesis…"
Figures 42.3, 42.4, 42.9
p 753-4 "Fertilization…" stop at "The sperm…"
Figures 43.1, 43.3, 43.4

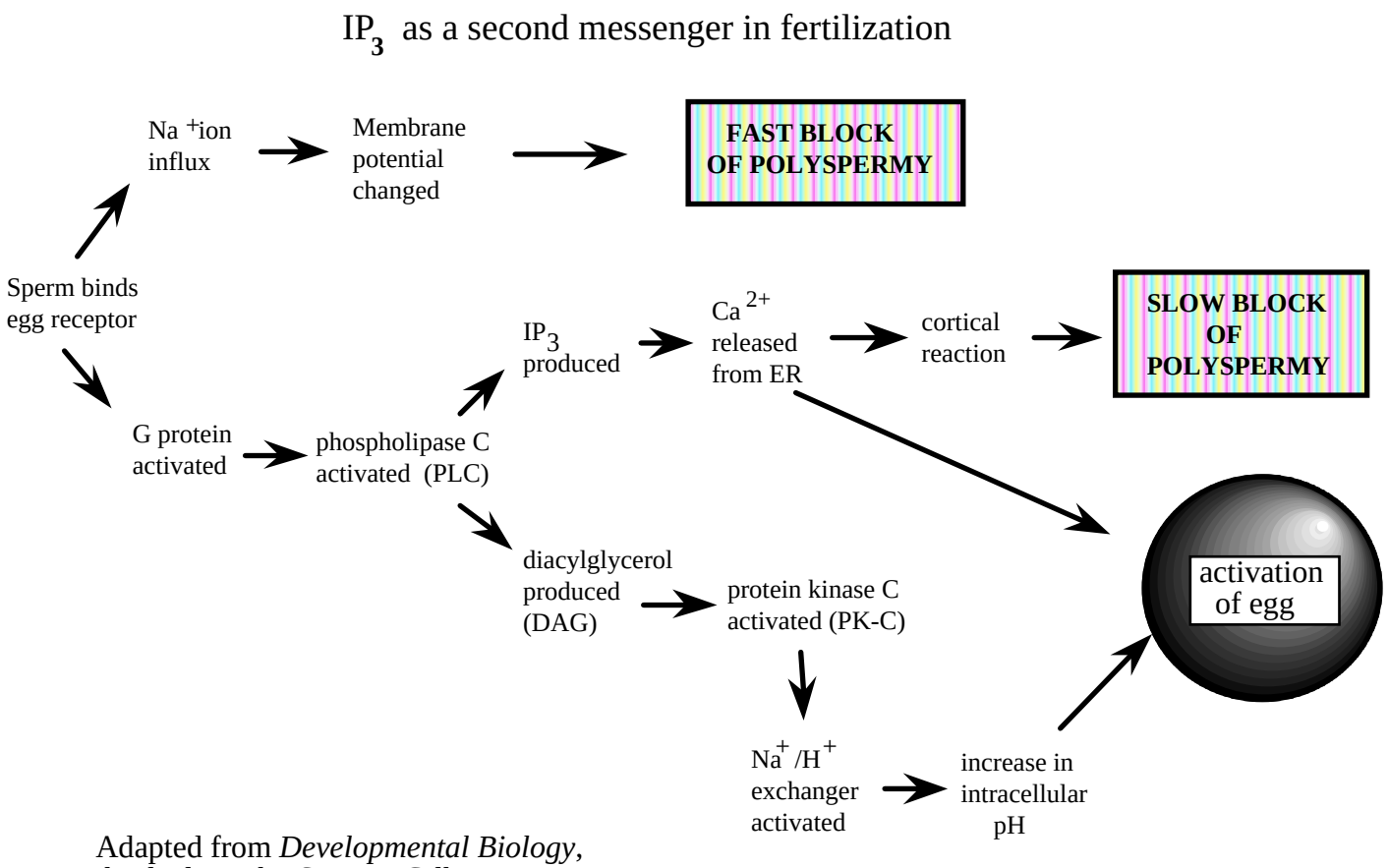
WWW Readings:
Immunofluorescence Labeling of ER
Movie of Calcium During Fertilization
Movie of Flash Animation:
"Mechanism for IP3 production and Ca2+ ion wave"
Purves6e, Ch43, Tutorial 43.1 Fertilization in the Sea Urchin

More links to help you visualize this process. They are all part of the same web site "Bill Wasserman's Developmental Biology Webpage" from Loyola University. Once at the site check out the following selections: http://www.luc.edu/depts/biology/dev.htm

Mammalian Oocyte Egg Ovulation: #2 Mammalian Egg
Sea Urchin Egg Fertilization: #4 Acrosome Animation
Egg activation: #3 Cortical Granule Exocytosis

An egg is just like any other cell in many ways. It has a plasma membrane, a nucleus, a Golgi apparatus, and an endoplasmic reticulum. The egg also has many unique features including the **vitelline layer** (also called the **z**ona **p**ellucida) which is an extracellular matrix (analogous to a plant cell wall) that contains many copies of a sperm-binding protein (called **ZP3**; the third protein discovered in the zona pellucida) which functions a receptor. ZP3 interacts with **bindin** proteins on the surface of sperm cells and initiate the **acrosomal reaction**. ZP3 is as specific as any other receptor we have studied; it will only bind ligands present on the surface of sperm from the same species as the egg. For example, rat sperm proteins will not bind to ZP3 on frog eggs. Interaction between ZP3 in the vitelline layer and ligands on the sperm head cause the two cells to fuse. Many different types of receptors in the sperm's plasma membrane trigger this fusion: some are protein kinases and others activate G proteins. The bottom line is the sperm is told by its receptors that it is timed to fuse with the egg.

Over time, evolution has selected eggs that have developed two separate mechanisms to prevent polyspermy, a **fast block** (an **electrical** barrier) and a **slow block** (a **physical** barrier).



Adapted from *Developmental Biology*, third edition by Scott F. Gilbert

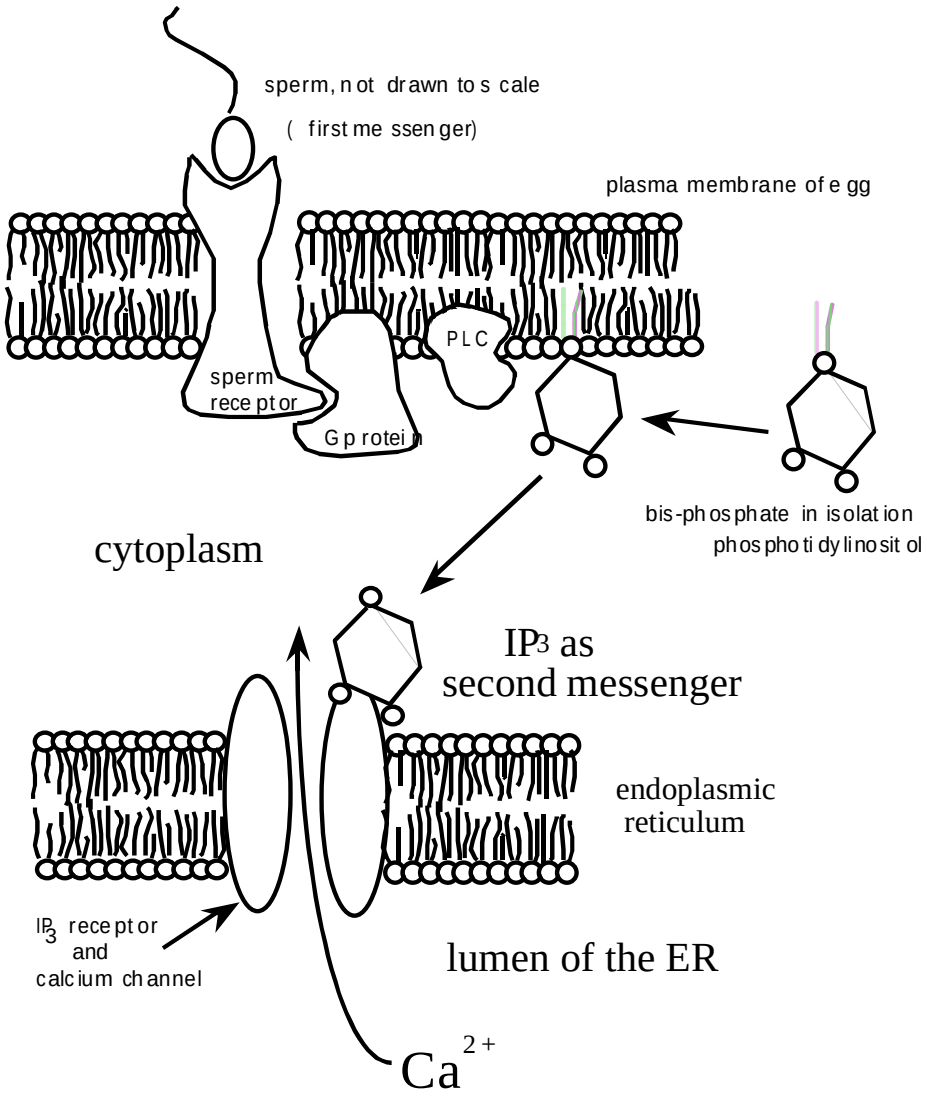
When the plasma membrane of the sperm first fuses with the plasma membrane of the egg, there is a change in the membrane potential of the egg cell. As we saw in muscles, egg cells have a resting potential of about -60 mV with a higher concentration of Na^+ ions outside the cell than inside. Fusion of egg and sperm membranes causes Na^+ channels in the plasma membrane of the egg to open. Although the exact gating mechanism for opening these Na^+ channels is unknown, the result is

predictable. Na^+ ions rush into the egg, down their concentration gradient, which changes the membrane potential from -60 mV to about +30 mV. For unknown reasons, this change in polarity prevents sperm from fusing and thus is a fast acting block to polyspermy.

The second block to polyspermy is where IP_3 is used but it is interesting that evolution has selected organisms that have two ways to block polyspermy. Can you imagine why you would need a second one, given the nature of the first? The second block to polyspermy will create a physical barrier to sperm entry rather than an electrical one.

If you were to examine an unfertilized egg in cross section, you would see lots of small vesicles, **cortical granules**, just below the plasma membrane. Inside the cortical granules are **proteases**, and **mucopolysaccharides**. During fertilization these cortical granules to fuse with the plasma membrane of the egg and, as we saw in neurotransmitter release, the contents of the vesicles is released into the adjacent extracellular space (exocytosis). The exocytosis of cortical granule contents causes two significant events to happen in close succession: 1) the proteases digest the proteins linking the vitelline envelope to the extracellular face of the egg's plasma membrane and probably disrupts the integrity of the unoccupied sperm receptors too; 2) the mucopolysaccharides increase the osmotic pressure in the small space between the vitelline envelope and the plasma membrane (we'll talk about osmotic pressure later) which makes water rush in which, like a hydraulic lift, causes the vitelline envelope to be pushed away from the plasma membrane. By pushing the vitelline envelope away from the egg's plasma membrane, a physical barrier has been created to prevent any more sperm from fusing with the egg.

Now that we know the cortical granules are capable of causing the physical block to polyspermy, we are still left wondering how the egg knows when to signal the cortical granules that one sperm has fused with the egg. This is where IP_3 comes into the picture. The sperm plasma membrane proteins interact with the sperm receptor in a manner similar to the cartoon (not drawn to scale) in the diagram below:



Adapted from *Developmental Biology*, third edition by Scott F. Gilbert

When the sperm receptor binds its ligand (ZP3), the receptor changes shape, activates the associated G protein, which stimulates phospholipase C, which cleaves phosphotidylinositol bis-phosphate into two parts: IP_3 and DAG. IP_3 is a second messenger that diffuses throughout the cytoplasm where it eventually bumps into the **IP_3 receptor** located in the ER membrane. The IP_3 receptor is a **homotetramer** (made of 4 identical subunits). The IP_3 receptor has a very high affinity for IP_3 and so IP_3 binds to its receptor and acts as an **allosteric modulator.** Each subunit has at least three allosteric binding sites; one IP_3 molecule and two calcium ions all have to bind to each subunit of the receptor. Calcium and IP_3 modulate the IP_3 receptor, which is also a **ligand-gated Ca^{2+} channel,** causing the normally closed channel to open. As we know, the ER is a rich source of Ca^{2+} ions. So, Ca^{2+} flows down its concentration gradient out of the ER into the cytoplasm. This elevates the concentration of Ca^{2+} ions in the cytoplasm, which is the signal, required to cause the cortical

granules to fuse with the plasma membrane. (You should reflect upon the number of similarities between this second messenger system and that used by neurons to secrete neurotransmitters.) As shown in your web reading this Ca^{2+} signal is propagated as wave, from the point of sperm penetration through the entire egg. The wave of Ca^{2+} creates a wave of cortical granule exocytosis that results in the entire egg being surrounded by a physical block to polyspermy. However, the wave of Ca^{2+} is not caused by a wave of IP_3. Instead, there is a recently discovered phenomenon called **Calcium Induced Calcium Release (CICR)** which is responsible for the wave of Ca^{2+}. The IP_3 created by phospholipase C causes just enough Ca^{2+} to be released from the ER to trigger CICR from adjacent Ca^{2+} channels in the ER. This is analogous to the way an action potential is propagated in a neuron but this time it is Ca^{2+} instead of Na^{+} ions. Exactly how this CICR works is an area of intense research. But what is clear is that in this system, Ca^{2+} has two functions: 1) to allow cortical granules to fuse; and 2) to spread information to adjacent areas that one sperm has entered the egg. However, as with all second messengers, we need a way to turn off the signal. When calcium levels reach a certain level (usually 1 - 10 seconds later), calcium ions cause the IP_3 gated channels to close. Therefore, the same ions that are used to open the channel also act to close it. The only difference is the concentration of ions. It seems likely that an additional allosteric site exists that has a lower affinity for calcium and this site is used to close the ion channel.

In summary, the sperm binds to its receptor; this initiates a chain reaction of enzymes (each of which can amplify the original single event) which results in the formation of the second messenger of IP_3. IP_3 binds to its receptor, causing it to open the Ca^{2+} channel so that Ca^{2+} floods into the cytoplasm (Ca^{2+} acts the third messenger), causing the cortical granules to dump their contents between the plasma membrane and the vitelline envelope, causing the vitelline envelope to rise up and create a physical block to additional sperm entering the egg.

> NEWS ITEM: Many researchers have tried to determine the identity of the ZP3 receptor and the results are inconclusive. So far, three proteins bind to ZP3, but it is not clear if all three are necessary. Rest assured, many people are trying to figure this out since there could be an effective birth control application once we know the complete story. (See *Developmental Biology*, 5th Edition, by Scott Gilbert.)
>
> Fertilin β, sperm plasma membrane protein, is required for the sperm and egg to adhere. A group in Research Triangle Park, NC collaborating with one at UC-Davis has identified that this protein is necessary for sperm binding AND for sperm motility. Without fertilin β the sperm never makes it out of the uterus to the oviduct (where the egg is). How these two functions are related is not known but it looks like another avenue for reproductive scientists. C.Cho et al. 1998 *Science* vol 281 p 1857-59.
>
> A group of collaborating scientists (Massachusetts, California, France and Mexico) has added to the ZP3 story. Their work demonstrates that ZP3 is involved in calcium regulation more than one time. When ZP3 signaling is initiated it triggers a very quick and transient opening of Ca2+ channels. If ZP3 signaling continues, the pathway activates a sustained Ca2+ influx mechanism and this sustained increase in Ca2+ drives the acrosome reaction. O'Toole, C.M.B. et al. 2000 *MBC* vol 11 p 1571-84.

Study Questions:

1) At which steps can the signal cascade be amplified and how does this amplification work?

2) Explain to a high school student the molecular events of the slow block to polyspermy.

3) Why does an egg need the second and slower block to polyspermy?

4) Explain how the egg uses a "third" messenger signal of Ca^{2+} twice.

5) In some of your focused reading, the text discusses how DAG is used as a second messenger. Exp

6) Compare and contrast: 1) a cardiac muscle's response to epinephrine, 2) depolarization leading to a neuron's secretion of neurotransmitters, and 3) an egg's response to fertilization.

7) List the similarities between a neuron communicating with a muscle and an egg trying to block polyspermy.

8) Explain how calcium is used to both open and close the IP_3 receptor.

9) How does cytoplasmic calcium return to resting levels?

10) How can the fertilization signal be deactivated?

----------STOP----------

NEWS ITEM: The molecular events in excitation-contraction of cardiac muscles have been visualized for normal and dysfunctional hearts. The key difference seems to be in the degree of CICR in the two situations. (*Science* Vol 276: 755. May 1997)

Many cells use the inositol triphosphate second messenger system. Here are a few examples:

The secretion of a digestive enzymes for carbohydrates by the pancreas (pancreatic amylase) in response to nervous system simulation.
The contraction of smooth muscle (involuntary muscle in internal organs and blood vessels) triggered by acetylcholine.
The secretion of insulin by pancreas in response to elevated plasma glucose levels
The secretion of histamine by mast cells when you have a cold or an allergy
The secretion of blood clotting factors by platelets when you are bleeding
The response of the immune system to bacterial invasions

OTHER COMMUNICATION SYSTEMS

It could be true that the majority of cells communicate through the 4 systems you have looked at here -- the cAMP second messenger system, the inositol triphosphate second messenger system, membrane voltage changes including action potentials and various method of producing elevated cytoplasmic Ca^{2+} concentrations. However, we have only scratched the surface in our knowledge of the cell and how it communicates so we will probably discover many additional ways that cells talk to one another. Here is a brief summary of some of the other systems of intracellular communication we now know something about:

The cyclic GMP second messenger system: Some cells use a second messenger system very much like the cAMP system where cyclic GMP is used instead of cAMP. cGMP is created by the enzyme **guanylyl cyclase** (analogous to adenylyl cyclase) which is activated by a G protein system. Probably

the most well investigated cGMP system is found in the photoreceptors (the rod cells) in the retina. In the dark, cGMP is bound to Na^+ channels in the cell membrane, keeping them open. When light strikes a rod cell, cGMP phosphodiesterase is activated, thus degrading cGMP to GMP. GMP disassociates from the Na^+ channel thus causing it to close. The opening and closing of this **ligand-gated Na^+ channel** causes voltage changes in the rod cell plasma membrane which are propagated toward the brain, thus allowing one to sense light.

> NEWS ITEM: Viagra, the 'wonder drug' for those suffering from erectile dysfunction is actually a phosphodiesterase inhibitor (it inhibits PDE5). Similar to the cAMP phosphodiesterase we learned about earlier, PDE 5's function is to breakdown cGMP into GMP. With Viagra around cGMP levels remain high and promote erection. Some not-so-well known Viagra facts are 1) it was originally developed to combat angina, 2) it is NOT an aphrodisiac and 3) it can cause distorted color vision. Why the vision change? Not because the users see the world through rose-colored glasses--because the eye uses cGMP as a 2nd messenger and Viagra also binds to PDE6, the phosphodiesterase found in the retina.

Stretch-activated ion channels: More properly called **mechanosensation**, the transformation of a physical stimulus to an electro-chemical signal is mediated by stretch-activated ion channels. These ion channels are responsible for our ability to hear, feel, and maintain our balance. These same mechanosensors enable our cells to "be aware" of their volume. These ion channels have been cloned recently, and will provide a great deal of understanding to this relatively unexplored area of sensations. (See Corey and García-Añoveros. *Science*. Vol. 273: 323-324. 19 July, 1996.)

Gap Junctions: Some cells communicate with one another directly, without the use of a chemical messenger. These cells are actually coupled to one another through proteins in their membranes called **gap junctions**. Gap junctions are like giant ion channels that allow small cytoplasmic molecules to pass directly from the cytoplasm of one cell to the cytoplasm of the adjoining cell. Heart muscle cells communicate this way, thus allowing the heart to contract as a unit. Many other cells communicate in this fashion as well.

Catalytic Receptors: Some receptors are enzymes themselves, and are therefore called catalytic receptors. An example of such a receptor is the receptor for insulin on muscle and fat tissue. When insulin binds to this receptor, it changes shape (sound familiar?) and this change in shape increases the enzymatic activity of the cytoplasmic tail of the receptor. The receptor then **autophosphorylates**, that is, it adds a phosphate to itself. Because this part of the molecule phosphorylates, it is called a kinase. And because it adds the phosphate at a tyrosine residue in its own structure (tyrosine is an amino acid), it is said to be a **tyrosine kinase** (an enzyme than phosphorylates tyrosine.) When insulin binds to its receptor, it causes a number of changes in the cells that include stimulating the transport of glucose into the cell and stimulating glycogenesis and the synthesis of triacylglycerol. Interestingly, several genes associated with the development of cancer (called **oncogenes**) have been shown to encode defective tyrosine kinase receptors. (We will cover this in detail in Unit IV.) For instance, the normal receptor allowing response of epidermal cells (skin cells) to the chemical messenger epidermal growth factor is a receptor with tyrosine kinase activity. The cancerous version of these proteins lacks the extracellular binding site for epidermal growth factor, but still has the tyrosine kinase part on its cytoplasmic tail. Without the binding site, the tyrosine kinase is always on, thus stimulating too much cell division.

Eicosinoids: These signaling molecules come in three varieties: prostaglandins**, leukotrienes,** and **thromboxanes**. Prostaglandins mediate pain and inflammation (aspirin works by inhibiting the

enzyme that produces prostaglandins). Leukotrienes mediate some of the immune aspects of inflammation. And thromboxanes facilitate blood clotting. Note that the eicosinoids are all involved in responses to injury. These molecules are actually derivatives of the fatty acid **arachidonic acid**. This makes them very unusual. They are produced by the cell membrane of injured or oxygen-starved cells and they mediate the inflammation, swelling, pain and blood clotting associated with injury.

Steroid Hormones: The hormones, neurotransmitters, and sperm cell proteins we have looked at in this unit are all hydrophilic -- they therefore cannot cross the hydrophobic cell membrane and must remain on the outside of the cell. However, steroid hormones (testosterone, estrogen, progesterone, cortisol, and aldosterone) are **lipids**. Therefore, they are freely soluble in the cell membrane and they cross into the cell. Their receptors are floating in the cytoplasm or the nucleus. The steroid hormones bind to their receptor, which then changes shape. The hormone-receptor complex then binds directly to control regions of genes in the chromosomes and causes these genes to be expressed (or stop being expressed.) These hormones tend to be slow acting and produce long-term changes.

> NEWS ITEM: A new family of about 100 genes has been discovered that function as human pheromone receptors. These receptors reside in the part of your nose that you might not know about called the vomeronasal organ. This organ is responsible for the perception of "odors" that we are not conscious of such as pheremones. Pheromones are usually fatty acids or steroids and their receptors appear to span the membrane 7 times and are linked to G-proteins. (see Science Vol. 278: 79. October 1997)

Fatty Acid-Based Signal Molecules: It turns out that plants are not as helpless as we vertebrates think. When corn is attacked by beet army-worm caterpillars, the injured plants release a mixture of chemicals called terpenoids which are fatty acids (same family as the long tails of phospholipids and DAG). These terpenoids are released into the air and attract a parasitic wasp which kills the armyworm caterpillars. Terpenoids also stimulate certain genes in the plants to fix the wound created by the caterpillars much the same way we produce scabs to seal wounds from possible infections. (See summary in *Science* Vol. 276: 912. May 1997)

Nitric Oxide: The cellular and molecular biology community is currently all abuzz about this newly discovered second messenger signaling system. Nitric oxide is a gas (not the same one the dentist gives you -- that's nitrous oxide.) This small molecule, which lasts only milliseconds inside a cell, nonetheless acts as a second messenger and triggers many interesting changes. A report in 1996 revealed that the levels of NO plays a role in the degree of symptoms when a person is infected with malaria.

> NEWS ITEM: The binding of oxygen to hemoglobin promotes the binding (allosteric modulation) of nitric oxide to a particular amino acid on the beta chain of hemoglobin. When oxygen is released from the hemoglobin molecule, the modulated hemoglobin changes shape. In this modulated but deoxygenated state, hemoglobin can cause blood vessels to become larger in diameter which results in increased blood flow. Therefore, NO increases the function of hemoglobin from simply a carrier of oxygen, to a modulator of blood flow so that areas of low oxygen will receive more blood. (*Science* Vol. 276: 2034. June 1997)

Study Questions:

1. In general, how do cells communicate? In answering this very big question in a manageable way, you cannot include very many details (although you might want to include a few examples) Rather, think about what central points you want to make. Think about this answer on many levels -- Explain it to me, explain it to a classmate in this class,

explain it to a Davidson student who has taken no biology since high school, explain it to your parents, and explain it to a child. Use good analogies when appropriate.

2. One of the basic tenets in cellular communication is that different cells respond in different ways to the same chemical signal. Using systems you have studied in this unit, give an example illustrating this point.

3. Second messenger systems have been studied extensively by cancer researchers because cancer cells ignore normal messages that tell them to stop dividing. Genes associated with the development of cancer are called **oncogenes** ("onco" means cancer, as in oncology.) One set of such genes called the *ras* genes (because they were discovered in a cancer called a rat sarcoma) code for the production of an abnormal G protein. The G protein has a slightly different amino acid sequence than the normal G protein. As a result, it cannot catalyze the cleavage of GTP to GDP by the G protein. Based on what you know about G proteins, explain how this abnormal G protein might produce uncontrolled growth in a cancer cell.

4. Over 70 different cellular protein kinases have been isolated and identified. What do all these kinases have in common? Choose three different protein kinases presented in this unit and compare and contrast their functions. What turns each of them on? What do each of them do? In what ways are these processes similar? In what ways are they different?

5. Myasthenia gravis is a disease, which produces a progressive weakening of skeletal muscles and ultimate paralysis. It is an autoimmune disease caused by the development of antibodies to the acetylcholine receptor. These antibodies bind to the receptor in such a way that they do not activate it, but they block the binding site for acetylcholine. (By the way, this is the same mechanism that the drug curare uses to produce paralysis.) Explain, in molecular and cellular terms, how this disease causes paralysis. What type of paralysis would result from this illness, flaccid (no contraction possible) or rigid (muscles permanently contracted)?

6. Certain types of "nerve gas" and pesticides act by blocking the action of **acetylcholinesterase** in the synaptic clefts and neuromuscular junctions. These agents produce paralysis. Explain, in molecular and cellular terms, how these agents produce paralysis. What type of paralysis would result from exposure to these agents, flaccid or rigid? Explain.

7. One of the most deadly poisons known is a toxin produced by the bacterium *Clostridium botulinum*, the organism that causes botulism. This toxin blocks the release of acetylcholine from nerve endings. How do you think this toxin kills you? Describe some of the symptoms you think would be produced by this toxin and explain how the blockage of acetylcholine secretion would produce such symptoms.

8. Summarize the role played by the cytoskeletal components in the systems you have studied so far.

9. While intercellular signaling systems differ in their details, they are all based on some common functions that are fundamentally important in all signaling systems. What do you think are the three or four phenomena that occur most consistently in cellular signaling systems and upon which cellular signaling is based?

SOME EXPERIMENTAL APPROACHES

The approaches and methods used to investigate the biology of cells and their communication processes are numerous and most are beyond the scope of this course. However, as a starting point, I have chosen a few basic methods upon which many others are based.

Microscopy -- The Direct Approach

Thanks to the Dutch lens grinders of the 17th century, we can see prokaryotic and eukaryotic cells simply by looking through a microscope. Because most animal cells are clear as are most of the parts of plant cells (only the chloroplasts and chromoplasts are colored), cells usually need a little help in order to be seen through the microscope. Without this help, they would be like small panes of glass -- present, but transparent. Several methods are available. The simplest is staining the cell to make it colored. Other methods allow the microscope to distinguish differences in structures due to their different abilities to defract light. For example, in **phase contrast microscopy** (we'll see this in lab), some structures will appear dark while others will appear light due to differences in defracted light. Finally, dyes that fluoresce when excited by light can be used to label organelles and molecular components of cells. These dyes are observed with a **fluorescence microscope** (See page 72 and 73 for examples).

Even with the most expensive optics available, the light microscope can only magnify about 1500 times. This is enough to allow one to see cells, but not enough to allow a clear view of most organelles and cellular inclusions. For that, you need a source of electromagnetic radiation that has a much shorter wavelength than light. In the 1950s, engineers perfected the **electron microscope** which uses electrons instead of light to produce images. This method is described on page 71 of your text. The transmission electron microscope allows the clear definition of cellular organelles and inclusions (such as cortical granules, microfilaments, etc.). Viruses can also be seen with this type of microscope. Using special methods, very large macromolecules can also be visualized (e.g. transport proteins in the cell membrane.) Some examples are found on pages 60 and 61

For fun (ok stop laughing) check out the Purves6e Ch4 Activity 4.3 'Know your techniques' #15 shows sickled blood cells—we'll be discussing the genetic mutation that causes this shape change in the next unit.

Isolating Living Cells for Experimentation -- Cell Culture

Most plant and animal cells can be kept alive for some time outside the host if they are maintained in conditions that mimic those of the body fluids. Cells are placed in **culture medium** which is a fluid designed to provide all the nutrients, salts, vitamins, etc. that the host normally provides in the right concentrations and at the right pH. If you can get cells to live in cell culture, you can do some pretty fancy experiments on them. For example, if you put muscle cells in culture medium that contains high levels of Ca^{2+}, nothing will happen because the living muscle cell can pump Ca^{2+} out of its cytoplasm as fast as it enters. However, if you then add a **Ca^{2+} ionophore** to the medium (an ionophore will insert itself into the cell membrane and create an artificial ion channel

which cannot be closed), the cell will contract. This indicates that high levels of intracellular Ca^{2+} trigger muscle contraction. By this approach, you could determine the concentration of Ca^{2+} necessary to elicit contraction. If you wanted to see that the concentration of an ion had actually changed inside a cell, you might use an **ion-sensitive dye** that will glow in the dark when it selectively binds to its ion.

Focused Reading: p 206 fig 11.9 research method
p313 fig 17.2
p 312-313 "Gel electrophoresis" stop at 'Recombinant…"
WWW Reading: Gel Electrophoresis Methodology

Isolation of Organelles, Cellular Inclusions, and Other Cell Parts

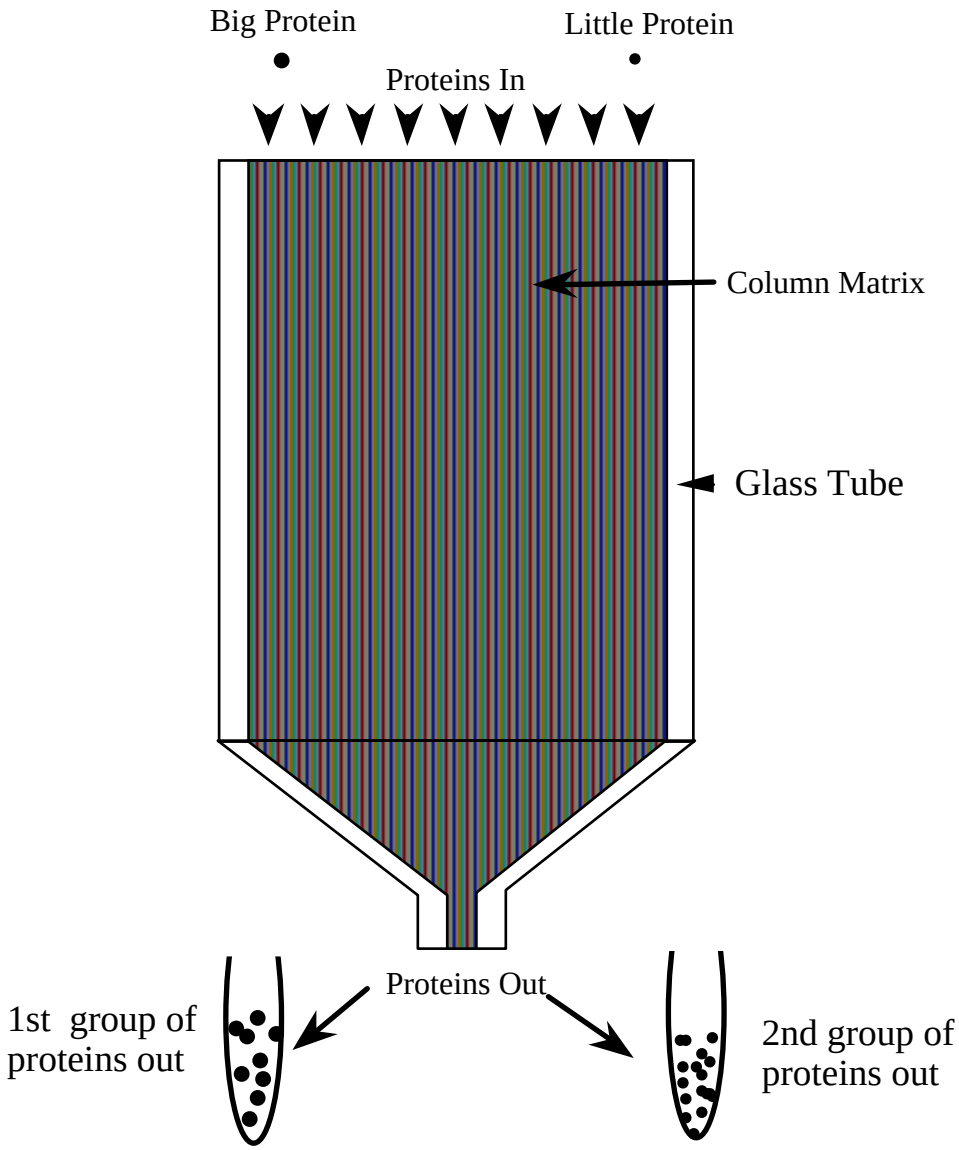
Sometimes it is beneficial to isolate part of a cell for study. For instance, if you are interested in a protein found only in the plasma membrane, it may be helpful to isolate the plasma membrane from the rest of the cell. Or if you are interested in ribosomes, you may wish to isolate them from the rest of the cell. All these cell parts are called **subcellular fractions** and they can be isolated using a method called **cell fractionation** using a **centrifuge** or an **ultracentrifuge**. A centrifuge 'spins' samples like a washing machine or the achine they use to train astronauts. The potential separation power of 'spinning' is seen in Figure 11.9 (p206). Density gradient centrifugation is used to separate pieces of DNA that have nucleotides that vary slightly in weight.

Isolation of Proteins by Molecular Sieves

Quite frequently, it is necessary to isolate a single protein from a cell. One method that is commonly used to do this is **gel electrophoresis**. In this method, a bunch of cells are homogenized to release all proteins. The cellular proteins are then usually dissolved in a detergent that covers them with negative charge. When these proteins are put in a gel (like a slab of Jello) and a voltage is placed across the gel (one end of the gel is made negative (the cathode) and the other end is made positive (the anode), the negatively charged proteins move toward the anode. Just like people in a thick forest, the smaller they are, the quicker they can move through the obstacle course of the gel to get to the anode. Thus, the smaller proteins move faster than the larger proteins and the proteins of the cell separate by size or molecular weight. This method, applied to DNA, is described on page 13 of your text and is very similar when applied to proteins.

If you want to study a protein further after it has been isolated, gel electrophoresis is not such a good method because detergent is very harsh on proteins and frequently destroys their native conformation during the separation process. A better method is one form of **chromatography** in which proteins are poured over a matrix in a glass tube (the tube length can range from 2 inches to 5 feet and the diameter from 0.25 inches to 3 feet.) The proteins are not treated to cover them with negative charge, as in electrophoresis, so they retain their native conformations. The proteins enter the matrix and, this time, the larger proteins get through the matrix first while the matrix retards the movement of smaller proteins so they come out last. This is because the matrix is made of small "beads" that contain tiny holes or channels, which the small proteins are small enough to enter, but the large proteins are too big to fit into. The small proteins spend a lot of time wandering around in these channels and it takes them a long time to get through the entire matrix. The large proteins cannot get into the channels so they continue through the tube on the outside of the beads (in the

space between the matrix particles). By taking this alternative route, they get to the bottom of the tube rapidly. Thus, the proteins are separated by size and maintain their native conformations and therefore can be used for further study. This type of chromatography can be diagrammed as follows:



----------STOP----------

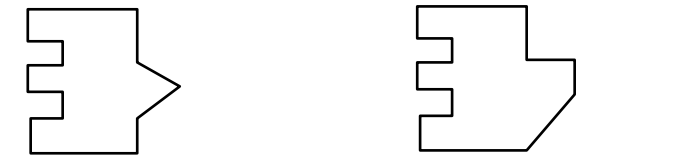
Focused Reading: p 363-4 " Antibodies share…" stop at "T cells… "
Fig 19.13

WWW Reading: Cartoon of Immunofluorescence Method

Identification of Proteins with Antibodies (Abs)

Since proteins do most of the meaningful work of living creatures, it is extremely important to biologists to be able to isolate and identify individual proteins. This can be done in a number of ways. One commonly used method involves the use of antibodies that bind to proteins with great specificity. When a foreign protein is injected into an animal (e.g. rabbit albumin into a mouse or goat insulin into a rabbit) the animal's immune system recognizes this foreign protein and interprets

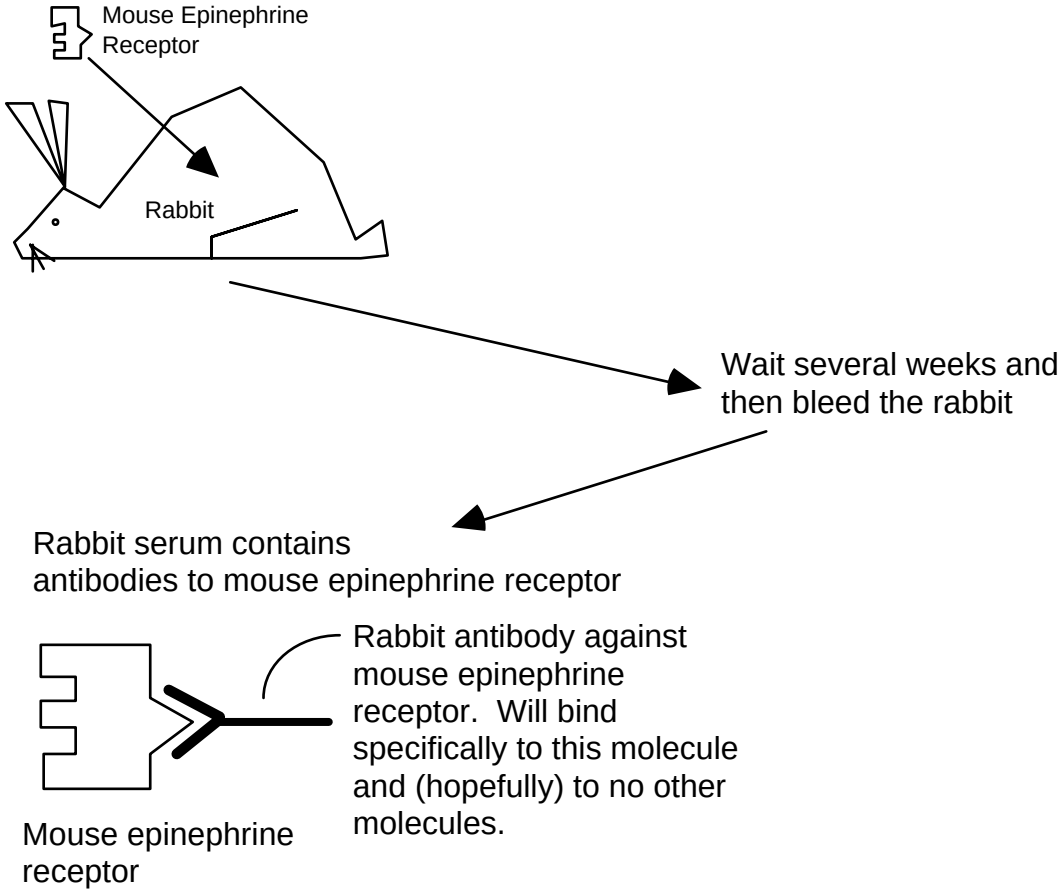
it as a microbial invader. (The immune system recognizes foreign molecular shapes whether they are harmful or not. Thus, you can get allergic reactions (an immune response) to pollen even though pollen can't harm you.) This immune response to the foreign molecule produces **antibodies** (which are proteins) which bind specifically to the foreign protein (called an **antigen**.) Antibodies have active sites, like enzymes, and the antigen is the **ligand** that binds there. The production of antibodies for research can be diagrammed as follows:



Mouse Epinephrine Receptor

Rabbit Epinephrine Receptor

The epinephrine receptors from these two species are slightly different in structure. Thus, mouse epinephrine will be seen as foreign by a rabbit and an antibody will be produced:



So, you can raise these specific antibodies against a protein you might be interested in studying and use the antibody as a probe for that protein since it will bind specifically to that protein and no other. You can probe for proteins ***in situ,*** which means that they are still in their normal location within the intact cell. The identification of proteins *in situ* using antibodies is called **immunocytochemistry** or **immunohistochemistry**. You can also remove the proteins from the cell, separate them by electrophoresis (see above), and then apply the antibody probe for the protein. This method is called an **immunoblot** (or a western blot in the vernacular).

Identification of Specific Proteins Through the Use of Radiolabelled Ligands

This method uses radioactivity to identify specific proteins. While there are many variations on this method, the basic idea is this. You buy or synthesize a ligand that contains a radioactive element. For instance, if you wanted to study the acetylcholine receptor, you would obtain **radiolabelled** acetylcholine. This acetylcholine could contain radioactive hydrogens (called **tritium**) or radioactive carbon (^{14}C) or an additional radioactive element (such as iodine - ^{125}I) could be added. These radioactive elements are **isotopes** of the non-radioactive elements. Isotopes are described on page 19 of your text.

These **radioligands** (in this case, radiolabelled acetylcholine) can be bound to various kinds of cells to determine whether they bear the ligand's receptor. For instance, if you wanted to know if liver cells have acetylcholine receptors in their membranes, you would incubate radiolabelled acetylcholine with liver cells. If they bind the ligand (i.e. if the cells become radioactive), then you can assume (if your experiment is properly controlled) that they are radioactive because they bound the ligand. This should mean that they have acetylcholine receptors. If they don't become radioactive, then they don't have acetylcholine receptors. You can also use this procedure to determine **the concentration or density** of a receptor in a membrane. Therefore, you can use this method to see if receptor densities change over time as you subject the cell to various treatments.

This method was used by Drs. Candice Pert and Sol Snyder in order to identify the receptors in the brain that bind (and respond to) opiates such as heroine. Through the use of this method (and others), we now know that we make internal or **endogenous** opiates called **endorphins** which reduce pain and may have other beneficial effects.

Molecular Models and Computer Graphics

One of the most exciting new methods in biology is the ability to build fairly accurate, complex 3-dimensional models of proteins based on computer analysis of data obtained by **x-ray crystallography**. This method is described on page 247 in your textbook. Because it is difficult to crystallize many important molecules, their 3-D structure at the atomic level (in their native conformation) remains illusive. However, if we learn enough about how amino acid sequence translates into 3-D structure, we may be able to predict (or teach a computer how to predict) the 3-D structure of a protein from its primary amino acid sequence. Because the amino acid sequence of proteins is becoming much easier to obtain (through the remarkable progress being made in molecular biology), this would be a tremendously important breakthrough and would give us new worlds of information about how living things function.

Focused Reading: p236 “induced mutations…” stop at “ Mutations are…”
p218 "One-gene, one-polypeptide hypothesis” stop at “DNA, RNA…”

p 263 Fig 12.1

WWW Reading: Movie of Microinjection

Use of Genetic Mutants

Because mutations are changes in the DNA that can alter the activity of one protein, they can be used to identify the protein responsible for a specific function. For instance, scientists have used genetic mutants to study the process of membrane traffic in the cell. Using mutant yeast, investigators have identified several mutant strains that each have one important protein altered. For instance, let's say Mutant strain #1 is missing Protein #1. Investigators find that this mutant strain cannot transport protein from the ER to the Golgi. Mutant #2 is missing Protein #2. This mutant strain cannot transport protein from the Golgi to the secretory vesicle. Thus, by identifying the protein that is missing and correlating it with the functional deficit in the cell, investigators can determine the proteins that are responsible for each step in a biological process. We will use genetic mutants to screen compounds to see if they are mutagens. We will perform this experiment (called the Ames test) later in the semester.

Microinjection

There is a very difficult, and labor intensive method to place a molecule of interest inside a particular cell and this method is called **microinjection**. As the name implies, you take a very small needle, usually made of a glass tube that has been pulled to a very fine point, attach the needle to a syringe, and inject a cell with a very small volume of solution which contains your favorite molecule. Typical molecules that are injected include DNA, RNA, antibodies, and proteins that have been purified from other cell types.

Study Questions:

1. Be able to describe each of the techniques outlined above.

2. If you had all of these methods available to you in the lab, how would you go about answering the following questions? Note: Just because a method is available does not mean it is the best approach to the problem. In each case, choose the method or methods that you think provide the most efficient route to an interesting and substantive answer.

   A. Do plant cells use cAMP second messenger systems?

   B. Is Ca^{2+} involved as an intracellular messenger in the secretion of saliva from the salivary glands?

   C. The microfilaments (actin and myosin) in vertebrate muscle cells are aligned in organized units which produce contraction as described by the sliding filament theory (outlined above). Are the microfilaments of the muscle cells of insects aligned in the same manner?

   D. Some forms of breast cancer are stimulated by estrogen (a female sex hormone). Do these breast cancer cells have a higher concentration of estrogen receptors than normal breast cells?

   E. Plant cells secrete the cell wall. Is the secretion of the cell wall constitutive or regulated?

F. What proteins mediate each of the steps that lead from ligand binding to cell division in fat cells?

G. Plants and animals both use the inositol triphosphate second messenger system which requires the use of phospholipase C. Is the phospholipase C used by plant cells similar in molecular weight and three-dimensional structure to the phospholipase C used by animal cells?

H. Does the Ca^{2+}pump in the SER membrane have the same molecular weight as the Ca^{2+} pump in the plasma membrane?

----------STOP----------

Unit II: Genetics

Brief Overview Reading: Chapters 3, 4, 9-12, 14 (Note: you have reviewed much of this already)

The earth is teeming with living things. We can easily see some of the larger organisms—trees, grass, flowers, weeds, cats, fish, squirrels, dogs, insects, spiders, snails, mushrooms, lichens. Other organisms are everywhere, in the air, in water, soil and on our skin, but are too small to see with the naked eye—bacteria, viruses, protists (single celled eukaryotes such as amoebae), and tiny plants and animals. Life is remarkable in its complexity and diversity, and yet it all boils down to a very simple idea—the instructions for making all this life are written in nucleic acids, usually DNA. Most organisms have a set of DNA that contains the instructions for making that creature. This DNA contains four "letters" in which these instructions are written—A, T, G, and C. The only difference between the code for a dog and the code for a geranium is in the order of those letters in the code. If you took the DNA from a human and rearranged the letters in the right way, you could produce an oak tree—arrange them slightly differently and you would have a bumble bee—arrange them again and you would have the instructions for making a bacterium. Acting through more than two billion years, the process of evolution has taken one basic idea—a molecular code that uses four letters—and used it over and over, in millions of combinations to produce a dazzling array of life forms.

As far as we know, we are the only creatures on the planet that have figured this out. The members of our species who get the credit for this discovery are James Watson and Francis Crick, although many others helped including Maurice Wilkins and Rosalind Franklin. Some believe Franklin was denied the Nobel Prize because of her gender but careful review of the facts will show that she was deceased at the time of the award and the prize is not given posthumously. Watson and Crick determined the 3D structure of DNA in 1952 and showed all of the human world (and any other species that could understand) that all of life is deeply united at the molecular level—indeed, we are all rearranged versions of one another.

The field of genetics is the study of how four bases make from aspen trees to zebras. Molecular geneticists study how the code is put together, how the code is translated into an actual living creature, and how the code is passed down from one generation to the next (dogs beget dogs, oak trees beget oak trees, and fish beget fish, although the offspring can be slightly different from the parents and from one another.)

In this Genetics Unit, we will look at the progress that has been made by researchers in understanding three inherited genetic diseases: **Cystic Fibrosis, Sickle Cell Disease** and **Huntington's Disease**. At the end of the Unit, we will also discuss some **sex-linked** genetic disorders. Many of the diseases that afflict humans have a genetic origin. Some diseases are caused exclusively by genetic defects. These include cystic fibrosis, Huntington's disease, phenylketonuria (PKU), Down's syndrome, Tay Sache's disease, sickle cell disease, muscular dystrophy, and hemophilia A. In other cases, such as cancer, one can inherit a genetic **predisposition** to a disease, but environmental factors also play a major role. Most disease conditions are probably in this category which certainly includes diabetes, hypertension (high blood pressure), and most forms of cancer.

Focused Reading:	p 332-3"hemoglobin" stop at 'receptors and transport'
	p 333-4 "Receptors and transport proteins" stop at "Structural proteins"
WWW reading:	For additional info when you have extra time
	Cystic Fibrosis Web Site
	Sickle Cell Disease Web Site

Huntington's Disease Web Site

The three diseases we will investigate in this Unit, cystic fibrosis (CF), sickle cell disease (SC) and Huntington's disease (HD), are caused exclusively by genetic defects; CF is the most common genetic disease in Americans of European descent, occurring in 1 out of every 2500 births. CF occurs with a frequency of 1 in 17,000 African Americans and with less frequency in other races. In the US, 1000 new cases are diagnosed each year, with 30,000 CF patients alive in 1996. Victims of cystic fibrosis accumulate thick mucus in the lungs and pancreas, produce elevated levels of very salty sweat and frequently develop cirrhosis of the liver. Digestion is disrupted in CF patients since pancreatic enzymes cannot reach the intestines. The mucus in the lungs makes breathing difficult and exhausting. This mucus is also attractive to microorganisms and therefore pneumonia is a constant threat in this disease - respiratory infections are the actual cause of death, not the thick mucus. Untreated children usually die by the age of 4 or 5 and the average life expectancy with medical care is 40 years.

SC is the most common genetic disease among African Americans afflicting 1 in 400 while 1 in 10 are carriers of the genetic trait. Most carriers are unaffected but some suffer from a mild form of the disorder (more about this later). Red blood cells are biconcave in shape (shaped like tiny doughnuts with a membrane across the hole) in non-affected individuals, but in this disease, they take on the shape of a crescent moon, or sickle, which causes several problems (see fig 12.17, p 235). The sickle-shaped cells tend to circulate more sluggishly in the body and clot as they pass through the tiny blood vessels of the tissues thus leading to tissue death and/or strokes. They are also destroyed more rapidly than normal red blood cells, which causes the symptoms of anemia—extreme fatigue, especially upon exertion.

HD is a fatal neurological disorder that causes severe mental and physical deterioration, uncontrollable muscle spasms, personality changes, and ultimately insanity. Perhaps the most troubling feature of this disorder is that the symptoms usually do not begin to appear until after the age of 40, after an individual has already had his or her children. Thus, until recently, people with this disease in their families have had to reproduce without knowing whether they have the disease and run the risk of transmitting it to their offspring.

The search for the causes and cures of these and other genetic disorders has been the goal of researchers for over 30 years. The recent revolution in genetics and molecular biology has dramatically improved our understanding of genetic diseases and greatly enhanced our ability to manipulate genetic systems to produce diagnostic tools and therapies.

In order to understand how these traits are pass on from one generation to the next, we need to understand the process of cell division in **somatic** cells (non-sex cells) and **gametes** (sex cells).

Focused Reading:	p 199 "DNA: …" stop at "The genetic material"
	p 157 "Eukaryotic cells divide…" stop at "Interphase and the control"
	p 160-5 "Mitosis: …" stop at "Reproduction by MEIOSIS…"
	Fig 9.6 & 9.8
WWW Reading:	Purves6e, Ch9, Tutorial 9.1 Mitosis
	Cartoon of Mitosis
	Movie of Mitosis

There is one rule that must not be broken for any cell to survive and function properly: a cell must maintain the right number of chromosomes at all times. This presents a problem for the average cell that is ready to divide. Lets say the cell has 23 pairs (it is **diploid**) of chromosomes and it wants to make two new cells. The first problem is how can a cell go from 1 X 46 to 2 X 46 chromosomes?

The obvious answer is that the cell must make 46 more chromosomes before it can divide. In its simplest form, that is all there is to **mitosis** - duplicate the DNA then divide. Of course any process as important and complicated as mitosis must progress in an orderly and stepwise fashion. The individual steps of mitosis are outlined in figure 9.8. You should be familiar with the major steps of mitosis (which should not include the cell cycle phase called interphase); 1) prophase; 2) metaphase; 3) anaphase; and 4) telophase (all 4 phases are reviewed in text p 160-165). Two points to note, 1) the text includes a 5th phase called 'prometaphase' and 2) mitosis does not include cytokinesis (but the two are closely associated).

Now that you have a handle on mitosis, we need to see what gametes do when they are formed. You know that to form a new individual by sexual reproduction, two gametes fuse to form a zygote. Since each gamete brings a set of chromosomes to **syngamy**, or fusing of gametes, we are faced with a mathematical dilemma. How can two cells contribute complete sets of chromosomes to a zygote without violating the cardinal rule of maintaining the proper chromosomal number? The answer is in **meiosis**.

Focused Reading:	p 735 Fig 42.4 --review
	p 204-211 "Meiosis…" stop at "Meiotic errors…"
	p 165-7 "Reproduction by meiosis…" stop at " Meiosis…"
	Figs 9.14 & 9.12
WWW Reading:	Purves6e, Ch9, Tutorial 9.1 Meiosis
	Movie of Meiosis

As you read, meiosis started off like mitosis with a diploid cell that replicates its chromosomes but instead of a single round of nuclear division, there were two rounds of nuclear division. This results in **haploid** cells that have only one copy of each chromosome (e.g. human egg and sperm have 23 chromosomes each). Therefore, when the two gametes combine their share of chromosomes, the zygote is back up to the proper (46 in humans) diploid or **2n** (2 copies of each chromosome) number of chromosomes. The important steps of meiosis are again well defined in the focused reading, and you should become familiar with them. But notice one other very important difference between mitosis and meiosis: chromosomes are not solid structures that cannot be modified but they can in fact switch parts with one another in a process referred to as **crossing over** (figure 9.16). This adds to the variation derived from independent assortment and provides a new source of individuality of each gamete, and ultimately the zygote and us.

Study Questions:

1. Be able to outline the major steps in mitosis and meiosis. Describe the major steps as if your your younger sibling has asked you why your eyes are the same color as your father's (and the inquisitive teen wants a really detailed answer).

2. Describe the significance of meiosis in relation to creating variation in the next generation.

> Dr. Alfred D. Hershey (as in the Hershey and Chase experiment) died at age 88 on May 22, 1997. He shared the 1969 Nobel Prize in physiology or medicine with two other researchers (Max Delbrück and Salvador Luria).

Now we know how cells inherit their DNA from the mother cell, and how haploid gametes are formed. In the last Unit, we saw how a sperm cell tells an egg it has been fertilized. Now we need to move on to the genetics, the pattern of inheritance. Genetics is a very logical discipline but the power

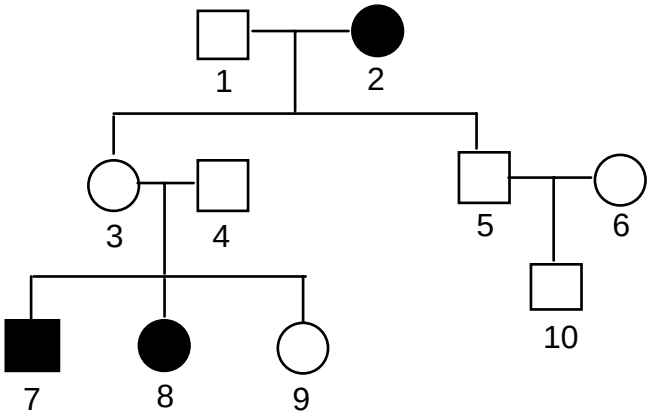
to genetics is numbers. The more progeny available for study, the easier it is to discern the pattern of inheritance. Unfortunately, genetic experiments with humans are not ethical or practical, since the generation times are so long. Given this inherent difficulty, it is amazing what has been learned about the genetics of human diseases.

Let's start by putting ourselves in the position of the first scientists who were interested in these diseases. Certainly one of the first thing people noticed about CF, SC and HD was that they run in families. Now because families usually live together and share a common environment, you cannot always conclude that something is genetic simply because it runs in families. Rather, you have to look closely at the inheritance pattern of the disease to see if it fits a classic genetic model of inheritance. For instance, coronary heart disease runs in families, but it does not fit a classic genetic model of inheritance. Therefore, we hypothesize that environmental factors also play a role in the development of this disease.

In looking for a classic genetic inheritance pattern in humans, the first thing you do is to research the disease occurrence in the family and draw a family **pedigree**. In drawing a pedigree, certain rules are followed.

1) Squares are used for males.
2) Circles represent females.
3) Non-affected individuals are blank while affected individuals are colored or patterned in some way. 4) Lines between a circle and square indicate a mating union (e.g. marriage) and all offspring of a mating union are drawn on the same level.
5) Individuals are numbered from top to bottom and from left to right.

Here is an example—a pedigree for a family with cystic fibrosis.



In this family, the woman (#2) in the first generation (grandma) had cystic fibrosis and yet survived long enough to have two children. Neither of her children (a girl and a boy -- numbers 3 & 5) had CF. Child #5 and individual #6 produced offspring #10, a normal, or **wild type**, child. Individual #3 and individual #4 had three children, two of whom (#7 and 8) have CF.

Study Questions:

1. Given information about a family, be able to draw a family pedigree that complies with standard rules.
1. Be able to interpret a pedigree drawn by standard rules.
2. Draw a pedigree for the cross that is outlined in figure 10.3 (p179).

----------STOP----------

What can we tell about the genetic inheritance of CF by looking at this pedigree? Well, in order to make sense of this pedigree, you have to understand a bit about the alternative ways by which genes can be inherited. To understand this, we have to go back 130 years to the Austro-Hungarian Empire and a Catholic monastery. Here a monk named **Gregor Mendel** was conducting breeding experiments with garden vegetables in an attempt to explain how genetic traits are inherited. His conclusions stand today as the foundation upon which modern genetics is built. Gregor Mendel defined the laws that govern the simple inheritance of traits and traits that are inherited in this straightforward manner are said to be **Mendelian traits** that obey the laws of **Mendelian genetics**

Focused Reading: p 177-88 "Mendel's work" stop at "Gene interactions…"

Study Questions:

1. Understand all the terms presented in bold face type in your reading assignment and be able to use them correctly in a description

2. Go back to the CF pedigree (on the previous page). In light of the concepts of Mendelian genetics and the information in this pedigree do you think that CF is a dominant, recessive or incompletely dominant trait? Explain.

3. Label the <u>generations</u> in this CF pedigree using Mendelian terminology (e.g. P, F_1, F_2).

4. What are the genotypes and phenotypes of each of the 10 people in the CF pedigree above? (Use proper <u>Mendelian</u> notation in assigning the genotypes.) In some cases, you will know a person's genotype and in other cases you will have incomplete information. Indicate this, and be able to explain the rationale you used to assign the various genotypes.

5. The mating of person #3 and #4 above represents the F1 of a **monohybrid cross** (refer to figure 10.4). Draw a Punnett square for this cross. (Use proper Mendelian notation here.) Does the actual mating outcome (two out of three children with CF) match the predicted outcome from the Punnett square? Why or why not? If they do not match, explain why this is the case.

6. Be able to solve genetics problems such as the following (From *Biology* by Villee, *et al.*):
   A. In peas, yellow seed color is dominant to green. State the colors of the offspring of the following crosses:
      1. homozygous yellow x green
      2. heterozygous yellow x green
      3. heterozygous yellow x homozygous yellow
      4. heterozygous yellow x heterozygous yellow

B. If two animals heterozygous at a single locus are mated and have 200 offspring, about how many would be expected to have the phenotype of the dominant allele?
C. Two long-winged flies were mated. The offspring included 77 flies with long wings and 24 with short wings. Is the short-winged condition dominant or recessive? What are the genotypes of the parents?
D. A blue-eyed man, both of whose parents were brown-eyed, married a brown-eyed woman whose father was blue-eyed and whose mother was brown-eyed. If eye color is inherited as a simple Mendelian trait (it actually is not), what are the genotypes of the individuals involved?
E. Outline a breeding procedure whereby a true-breeding strain of red cattle could be established from a roan (a blend of the incompletely dominant alleles for red and white) bull and a white cow.

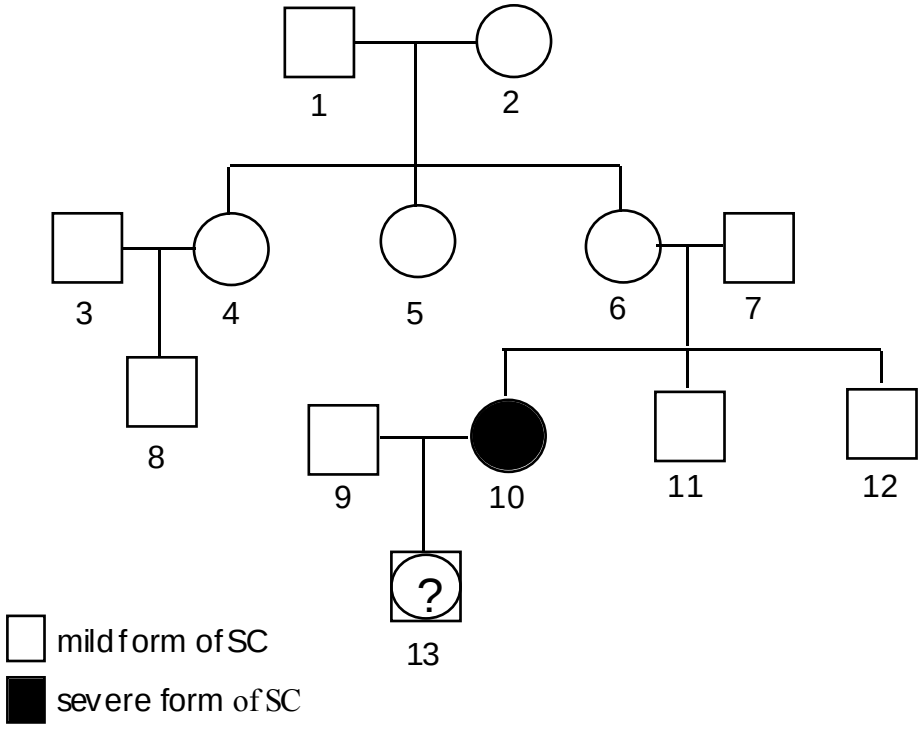
If you would like more practice, try the questions at the end of the chapter (p. 197-8).

NEWS ITEM: A collaboration between researchers at the Oregon State University and the University of Bristol (in the UK) have cloned the gene that encodes for the dwarf trait studied by Mendel. The gene is the last enzyme in a pathway that produces the plant hormone gibberellin. Without this hormone, the plant does not grow as tall. This is of more than historical interest. Plants that do not grow as tall often produce more seeds or fruit and are less likely to break and fall over since their stems are shorter. This is of considerable interest for genetic engineers that want to produce food crops that resist wind damage and produce more. (*Proc. Nat. Acad. Sci.* Vol. 94: 8907. 1997)

Focused Reading: p 185-188 "Alleles" stop at "Gene interactions"
P 189 "The environment" stop at end of page

When considering CF, an individual either expresses the phenotype (has two copies of the CF allele) or does not express the phenotype (has one copy of the mutant allele OR is homozygous wild type). This is the phenotypic expression pattern expected when the wild type allele is dominant over the CF allele. But now consider the pedigree for a family with members who have sickle cell disease (next page). Here we see individuals that have 'mild' cases of the disease. How can this be? Doesn't one allele 'win' over the other? Well, no. Some alleles show **incomplete dominance.** In these cases a heterozygous individual shows traits that are 'half way' between the homozygous possibilities. In sickle cells disease both incomplete dominance and **penetrance** come into play. Penetrance refers to the proportion of individuals that have a particular genotype that show the expected phenotype. The predicted phenotype of the mild form of anemia is not always seen in heterozygotes, so the mild form of the disease is said to be not fully penetrant. Environmental factors can affect 'seeing' the sickle cell phenotype. Heterozygous individuals may appear unaffected by SC except when faced with conditions of low oxygen, such as if they were to run a marathon or go to a vacation resort at a high altitude.

Here is a pedigree for a family with sickle cell disease:



Study Questions: All refer to pedigree on this page

1. Looking at the SC pedigree, Explain how you can tell that SC is an incompletely dominant trait.

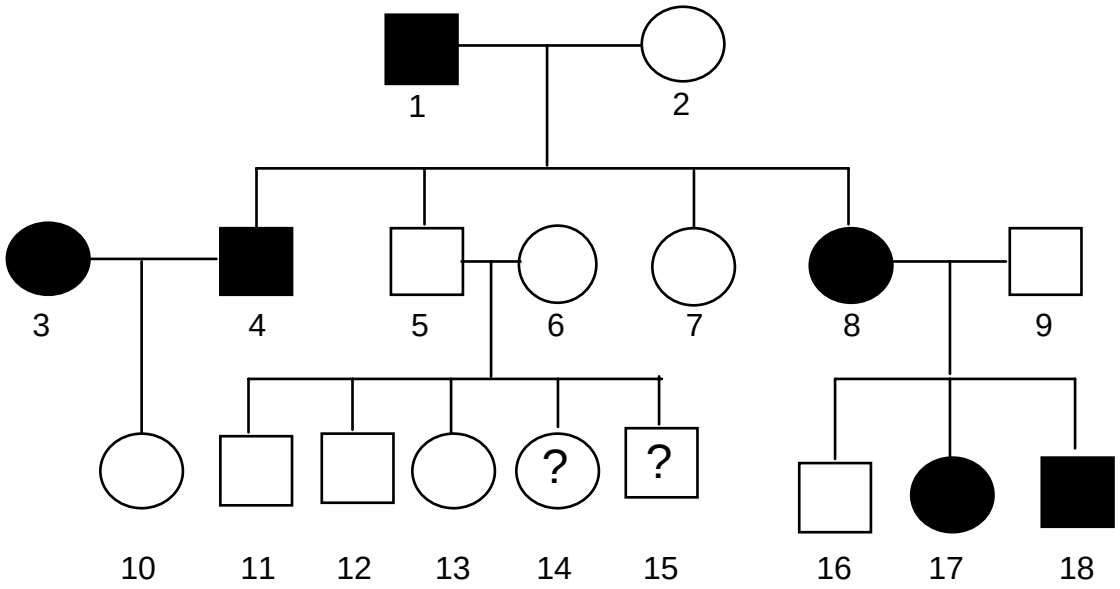
2. Label the generations in this SC pedigree using Mendelian terminology.

3. What are the genotypes and phenotypes for individuals 1-12 in the SC pedigree above? (Use proper Mendelian notation here.) In some cases, you will know a person's genotype and in other cases you will have incomplete information. Indicate this, and be able to explain the rationale you used to assign the various genotypes.

4. In Mendelian terms, what type of cross does the union of #6 and #7 above represent? (e.g. Monohybrid cross, test cross) Draw a Punnett square for this cross.

5. Individual #13 is still in the womb. For each of the following outcomes of this pregnancy:
   1) indicate the genotypes of the parents #9 and #10;
   2) What are the odds of these three outcomes?
      a) #13 is homozygous wild-type
      b) #13 is heterozygous
      c) #13 is homozygous disease

Here is a pedigree for a family with Huntington's disease.



Study Questions:

1. Looking at this pedigree, do you think that Huntington's disease (HD) is a dominant, recessive, or incompletely dominant trait? Explain.

2. Label the generations in this HD pedigree using Mendelian terminology.

3. What are the genotypes and phenotypes of each of the 18 people in the HD pedigree above? Individuals 14 and 15 are not yet old enough to determine whether or not they will get HD. What is your prediction about their disease status? What are their genotypes? Explain.

Many times people with genetic diseases in their family seek the advice of genetic counselors in trying to determine the probability that they will produce an offspring with the disease.

Focused Reading: p 339"Detecting human genetic variation" stop at paragraph beginning "Of the numerous…"
p 167 Fig 9.13 karyotyping

Study Questions:

1. Individuals #3 and #4 from the CF pedigree are considering having another baby and come to you as a genetic counselor. They want to know the chances that this baby would have CF. What will you tell them? (CF pedigree is on p 53)

2. What would you tell individuals #5 and 6 from the CF pedigree under the same circumstances?
   Assume that a person picked from the population at random has a 1 in 50 chance of being a carrier of a mutant CF allele.

3. What would you tell individuals 3 and 4 from the SC pedigree? Individuals 9 and 10? (SC pedigree is on p 55) Assume that a person picked from the population at random has a 1 in 100 chance of being a carrier of a mutant SC allele

4. What would you tell individuals 5 and 6 from the HD pedigree? Individuals 8 and 9?

5. A couple is planning to have children and comes to you to help them determine the chances that their children will have SC. Both parents have a very mild form of the disease.

   A. What is the probability that their first child will have SC (homozygous recessive)?
   B. What is the probability that their first child will be a carrier of SC or not have any SC alleles?
   C. If their first child has SC, what are the chances that their second child will have SC?
   D. If this couple has three children, what is the probability of all three children having full-blown SC?
   E. What is the probability that the first two children will have full-blown SC and the third one will be a carrier?
   F. What is the probability that one will have SC, one will be a heterozygote and one will be homozygous wild type?
   G. What is the probability that all three of the three children will be homozygous wild type?
   H. What is the probability that all 3 will be heterozygotes?

6. If couples from families with genetic disease decide to conceive and then want to know the genetic status of their fetus, what diagnostic tests are now available to them? Describe each test.

Answers to Questions 1 – 5:

1) 1 / 4
2) 1 / 50 x 1 /4 = 1/200
3) 1 / 100 x 1 /4 = 400
   1/100 x 1 /2 = 1/200
4) 0
   1/2
5) A 1 / 4
   B 3/4
   C 1 / 4
   D 1/64
   E 1/32
   F 3/16
   G 1/64
   H 1/8

So far, through pedigree analysis of the afflicted families, we know that cystic fibrosis is a recessive trait, Huntington's disease is a dominant trait and sickle cell disease usually behaves as a recessive trait (heterozygotes are asymptomatic [have no symptoms]) but sometimes SC behaves as an incompletely dominant trait (when the heterozygotes have a mild form of the disease.) What does all this actually mean at the molecular level? What does it mean to have a "dominant trait" or a "dominant allele"? How do alleles dominate one another?

In order to examine this question, we have to know what genes actually do, what they actually are. As you know from the previous Unit, your life is embodied in your structure (mostly proteins and fat) and your chemical reactions (each one catalyzed by an enzyme which is a protein). Your proteins control your life, and your genes control your proteins. The simplest definition of a gene (one that is outmoded, but a good place to start) is that a gene is a segment of DNA which encodes one protein. This statement is called the **one gene-one polypeptide** theory and it is still basically sound although we now know that the story is much more complicated than this statement suggests.

Genes encode proteins, that is, they contain the instructions that the cell can "read" in order to be able to make all the proteins it needs to live. We know from Mendelian genetics that we inherit two alleles for each gene. If we use the three genetic diseases we have introduced above as examples, we can (and investigators do) begin speculating about the genes that might be involved. In cystic fibrosis, you have too much thick mucus in the lungs and pancreas. There must be genes that encode proteins that prevent it from thickening. These genes could be involved in the production of mucus, the secretion of mucus, the control of mucus production and secretion, the movement of water into and out of the lungs and pancreas (since mucus become thicker when water is removed), etc. In the first part of this discussion, I will refer to this gene and the "mucus gene" and its protein as the "mucus protein" even though this description doesn't explain the high salt concentration in sweat or the liver cirrhosis. Nevertheless, it gives us a common language with which to refer to the normal gene that, when mutated, causes cystic fibrosis.

Because CF is a recessive disease, it is a good bet that the disease allele fails to encode a functional protein. In the case of a recessive disease, heterozygotes (carriers) do not have the disease because their one wild-type allele is enough to allow them to make all the functional protein they need. The second allele is redundant. But homozygotes for the disease have no wild-type alleles, no wild-type proteins, and they get the disease. So, in the case of a recessive disease, we are usually looking for a gene that does not encode for a functional protein.

In the case of sickle cell disease, the phenotype is sometimes incompletely dominantly expressed and sometimes expressed as a recessive trait. However, at the molecular level, SC is always codominantly expressed. This usually means, as in the case of recessive genetic disease, that the disease allele does not encode a functional protein. However, in the case of incompletely dominant expression, the normal allele in a heterozygote cannot fully compensate for the loss of protein caused by the disease allele. SC heterozygotes have some wildtype and some SC form hemoglobin in their red blood cells and thus experience some mild sickling in those cells. While these cells are usually able to function properly and are destroyed at a normal rate, sometimes under extreme conditions (heavy aerobic exercise, high altitudes) they function poorly and produce mild symptoms of SC. Thus, in this case, the trait is incompletely dominant. In a heterozygote both wildtype and SC hemoglobin are made but the severity of symptoms in the heterozygote varies widely depending on environmental conditions.

Because the symptoms of Huntington's disease involve many brain centers, a gene that has wide ranging effects on the function of the nervous system must cause the disease. Because Huntington's is a dominant trait, we would look for a gene that makes too much of its protein or makes a form of the protein that is hyperactive. When the disease gene is present, it causes its protein to be too active or in too high a concentration, but remember that onset of the disease comes around age 40. Regardless of the presence of the normal allele, the person has too much of an enzyme or structural protein. In the delicately balanced living system, having too much of something is frequently just as bad as not having enough.

Study Question

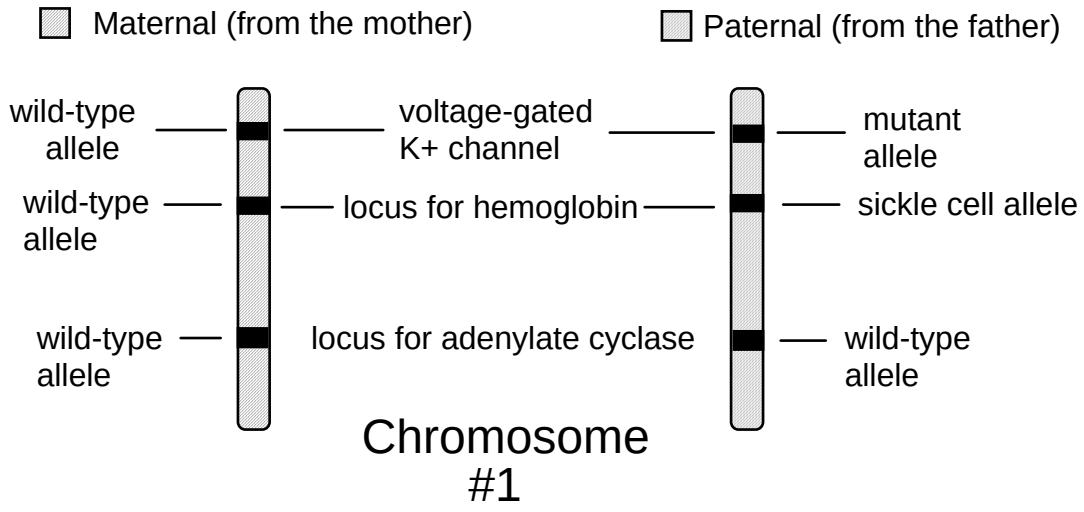
1. Explain how traits wind up being recessive, incompletely dominant or dominant based on the type of defect produced at the level of the protein. Give examples for each. (Do not use CF, SC or HD as examples here. Your examples need not be diseases. They can be normal traits.)

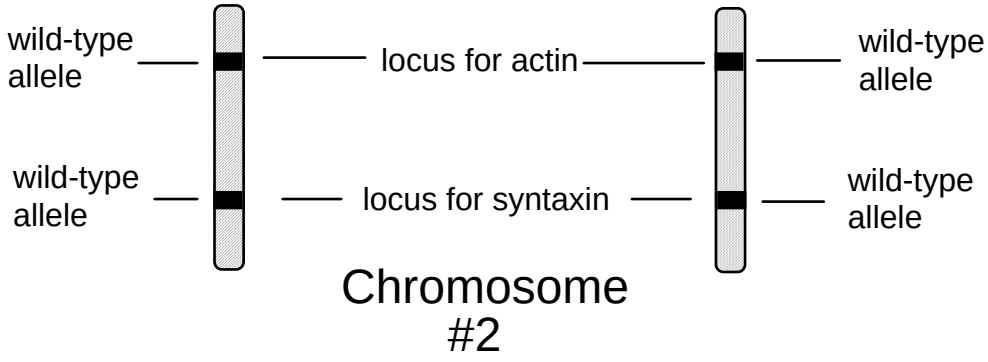
----------STOP----------

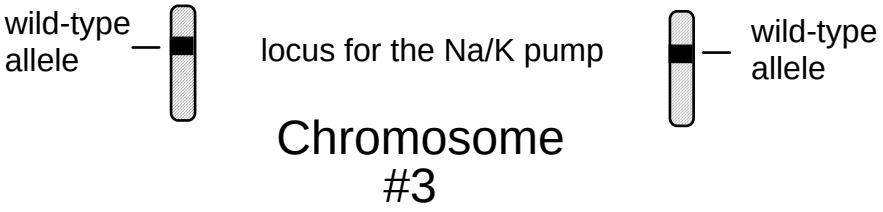
We need to stop and look at how wild-type genes produce wild-type proteins. Genes don't exist as individual strands of DNA, but rather, they sit one after another in long complexes called **chromosomes**. Chromosomes contain the DNA encoding enzymes as well as the proteins that are involved in packaging the chromosome (so it fits into the nucleus), and in the control of **gene expression**. Gene expression is the term for the process where the genetic blueprint of DNA is actually converted into a functional protein. Bacteria have one circular chromosome and eukaryotes have multiple linear chromosomes. Each species has a certain number of chromosomes, and humans have 46. However, as you know, each trait is encoded at a particular locus at which we inherit two alleles, one from our mothers and one from our fathers. Organisms that have two alleles for each locus or trait are said to be **diploid**. Humans are diploid and, therefore, their 46 chromosomes actually come in 23 pairs -- 23 pairs of **homologous** chromosomes.

The diagram on the next page is an illustration of the organization of genes on three homologous pairs of chromosomes. (Note: These loci are probably not on the correct chromosomes—this is simply an illustration.)

You can see from the diagram that <u>loci are always identical on homologous chromosomes</u>. Loci are like file folders. You have two file folders for a voltage-gated K+ channel; one on your maternal chromosome 1 and one on your paternal chromosome 1. The actual file (instructions) you store in this folder, however, can be quite different. The maternal voltage-gated K+ channel locus contains the instructions for producing a wild-type channel, while the paternal voltage-gated K+ channel locus contains the instructions for producing a non-functional channel. Therefore, this organism is **heterozygous** for the voltage-gated K+ channel. It has a heterozygous **genotype** at that locus. The **phenotype** that results from the expression of these alleles will depend on whether the alleles are dominant, recessive, or codominant, to one another.







A number of loci have been added to this genetic map for the sake of illustration and hopefully to alter some misconceptions. We concentrate a lot in genetics on the loci that produce individual differences in: height; eye, skin and hair color; disease states; etc. However, the vast majority of loci have only one allelic alternative in the species, they are **monomorphic**. For instance, in humans, there is only one allele for the IP_3 receptor, insulin (a hormone), collagen (the fibrous component of bones, tendons, and ligaments), keratin (hair and nails), acetylcholinesterase (the enzyme that destroys acetylcholine in the synapse or neuromuscular junction), etc, etc, etc. The vast majority of human proteins are encoded by an allele that all humans share; there is no variation from person to person. Therefore, we are almost totally homozygous. Loci at which there are a number of alternatives are the exception and are called **polymorphic loci** and the traits encoded at these loci are called **polymorphic traits**. Most traits are not polymorphic, but many interesting ones are, including all the features that make us different from one another.

Study Questions:

1. Describe the organization of genes along chromosomes and the concept of homology.

2. What is a genetic locus? An allele?

3. What does it mean when a trait is polymorphic? Give an example (not given above) of a polymorphic trait. Give an example (not given above) of a monomorphic trait.

So, to return to our tale, a person with cystic fibrosis would have two defective genes at the locus that controls mucus production in the lungs and pancreas. One defective "mucus gene" would be on the maternal chromosome (the person inherited this chromosome from his/her mother) and the other defective gene is on the paternal chromosome (the person inherited this chromosome from his/her father). A person with sickle cell disease would have two defective alleles at the locus controlling some aspect(s) of the red blood cell's shape. A person with Huntington's disease would have one defective allele at the locus controlling an important brain protein. This allele could be on the maternal or paternal chromosome. (Note: A person with HD could have two defective alleles, but because the disease is so rare, it is highly unlikely that two people with HD would mate, a requirement for producing a homozygous HD offspring.)

What is defective about these genes? What can a normal gene do that these disease genes can't do? In order to address this important question, we have to understand what genes do normally. Somehow, the instructions for making a protein have to be encoded in the DNA molecule in such a way that they can be translated into protein by the cell.

Overview reading	p 205 Fig 11.7
Focused Reading:	p 202-9 "The structure of DNA" stop at "The mechanism of..."
	Fig.s 11.8 through 11.10
	p 209-214 "The mechanism of DNA replication" stop at "Practical applications…"
	p 220-1 "RNA differs from DNA" stop at "RNA viruses…"
	p 223-5 "The genetic code" stop at "Preparation for…"
WWWeb Reading:	http://www.umass.edu/microbio/chime/dna/index.htm look at #3 and #5 DNA polymerase 1 and 2
	Purves6e, Ch11, Tutorial 11.1 DNA Replication

The DNA molecule is "written" in a code that has four "letters". The four nucleotides 'letters' in DNA are guanine (G), adenine (A), thymine (T) and uracil (U). In general terms the nucleotides are also called bases. In the DNA code three bases in a row equal a 'word' known as a **codon**, and each codon encodes a single amino acid. Following this through it is the base sequence of DNA that determines the amino acid sequence of the protein. Because amino acid sequence determines native conformation and native conformation determines function, the nucleotide sequence controls all living processes and structures.

Study Questions:

1. In a basic outline form, describe and/or draw the structure of DNA. What chemical groups does DNA contain and how are they arranged in the molecule?

2. Many times, DNA and RNA are described as having a 3' and 5' end. Explain what this means in terms of the structure of the molecules.

3. How is DNA transcribed into RNA? Where in the cell does this process occur?

4. Be sure you understand how to interpret the genetic code in Figure 12.5 (p 224). Given the base sequence of DNA or mRNA, be able to give the amino acid sequence of the resulting protein.

5. What proteins are involved in DNA synthesis and what are their roles during this process?

At this point, you should be able to come up with one hypothesis about what is wrong with the CF, SC, and HD genes. Their nucleotide sequences may be incorrect (i.e. contain some typos). Changes in the nucleotide sequence of DNA are called **mutations**. A number of different mutations could be interfering with the function of these genes.

Focused Reading: p 233-7 "Point mutations..." stop at "Some chemicals…"

You can see by studying the genetic code on page 224 that mutations in the third base of the codon frequently produce no change at all in the amino acid encoded by that codon. For instance, if the mRNA codon CCU were changed to CCC or CCA or CCG, it would still encode the amino acid proline. Thus, some point mutations have no impact at all on protein structure and function. However, some point mutations can make a very big difference in the function of proteins. By substituting one base for another in the DNA, you can change the amino acid at that position in the resulting protein. Look at the genetic code on page 224 and see which mutations would make such a difference. For instance, the code for serine (Ser) is UCG (there are actually 6 codons for Ser) while the code for tryptophan (Trp) is UGG. By changing "C" to "G", you can change the amino acid at that position in the protein. Now look on page 37 at the R groups of the amino acids. Serine's R group contains an OH group which means it is polar. Tryptophan has a large hydrophobic and non-polar R group. These two amino acids would behave differently in water, and thus this mutation would cause a slight alteration in the 3-dimensional shape of the protein. Depending on the exact location of this mutation, the protein may or may not be significantly altered in its shape.

Go back to p 224. The code for aspartate (Asp) is GAU while the code for glutamate (Glu) is GAA. If U were changed to A, glutamate would be put into a protein where aspartate should have been. Now go back to page 37 and look at the R groups of these molecules. Both R groups are organic acids, both are negatively charged. Therefore, this mutation probably would not have as great an effect on protein structure since glutamate and aspartate would behave very similarly in an aqueous environment.

Mutations that cause a change in the amino acid sequence of proteins are called **missense mutations**. The ultimate effect of such a mutation on the function of the affected protein, as you can see, depends on the type of amino acid substitution the mutation produces and the position of the amino acid substitution. As you know, enzymes, receptors, transporters and most other functional proteins have **active sites**, i.e. areas on the protein molecule that actually come into contact with important ligands, e.g. substrates, hormones, neurotransmitters, transported nutrients, etc. In addition, proteins frequently have **allosteric sites** at which they are regulated, **ATP or GTP binding sites**, and/or **phosphorylation sites** at which energy is transferred and the protein is regulated.

Amino acid substitutions at these important sites have a far greater impact on the protein molecule than do mutations that are in the framework or scaffolding areas. For instance, the change from glutamate to aspartate would probably cause no change in function if it occurred in a framework region of the protein. However, if it occurred at an active or regulatory site, it may dramatically alter the protein's function since aspartate is a smaller molecule than glutamate and would alter the topology of the surface of the active site that is so critical to specific binding. (A slightly bigger or smaller bump at one spot in the binding site may make specific binding to the normal ligand inadequate or impossible.)

A missense mutation is very likely to be the cause of a disease if the protein product is still present, but functioning poorly. However, if the protein is simply not present, we may be dealing with a **nonsense mutation**, or an **insertion,** or **deletion mutation** that has caused a **frameshift**. In either case, no protein is made at all.

HD is dominant; therefore we suspect that the protein encoded by the mutant gene is hyperactive. We might hypothesize at this point that a missense mutation in the active site increased the affinity of this molecule for its ligand. Or, possibly (and more likely), a missense mutation might have destroyed an allosteric site, making it impossible for an allosteric modulator to turn the protein off. Thus, the protein continues to function at a high rate at all times, producing too much of something that causes the disease. Conversely, it does not seem likely that a nonsense mutation is responsible for HD.

In addition to environmental agents causing mutations (irradiation, some chemicals, and viruses), the genetic material itself is constantly changing in ways that may cause mutations. For instance, genes or parts of genes can be duplicated (**gene amplification**), **methylated** (this permanently turns the gene off making it unable to be expressed), **rearranged**, or **transposed** (moved to another chromosome) [fig 14.5]. Then of course, our cells can make mistakes in DNA replication which can lead to mutations too (page 212-4 "DNA proofreading..."). Any of these natural changes may induce a mutation that destroys or amplifies a protein's function.

Study Questions:

1. Describe the effect of a single point mutation on protein structure and function. What types of point mutations are the most harmful? The least harmful? Explain. What two factors play a major role in determining the impact of a mutation on protein function? Explain.

2. Given the genetic code and the R groups of the amino acids, be able to develop a reasonable hypothesis about the effect of a given mutation on protein function.

3. Nonsense and frameshift mutations virtually always destroy the gene's ability to produce a product. Explain why this is so.

4. Explain how a missense mutation may increase the activity of a protein product.

5. Describe changes that occur in the DNA (without external mutagens) that may lead to the development of a non-functional or hyper-functional gene.

6. Explain why the five base pair insertion mutation described in the NEWS ITEM back on page 29 would cause 300 amino acids to be deleted.

NEWS ITEM: A very new classification of mutations has been discovered recently. This mutation does not happen at the DNA level, but at the mRNA level. It appears that the RNA polymerase makes certain mistakes frequently, such as reading the DNA sequence GAGAG and producing an mRNA that is only GAG, a two base deletion. This new form of mutation has been discovered in the brains people with Down's syndrome and Alzheimer's disease. (*Science* Vol. 279: 174. January 1998)

7. Would the form of mutation described in the News Item be passed on from one generation to the next? (This is a trick question so think about two possible answers.) To which category of DNA mutations is this mRNA mutation most similar?

----------STOP----------

When we talk about mutations, it is a common misconception that we are always talking about changes in the DNA that occur in the individual bearing the trait. This is not the case, and it is important that you understand this point. Mutations can occur in this manner, in which case, they are called **new mutations**. Some diseases, especially some forms of cancer (e.g. skin cancer) are thought to be enhanced by new mutations within individuals. However, the classic genetic diseases are caused by mutations that occurred hundreds or even thousands of years ago in an ancestor and are transmitted through inheritance to the individual with the disease. Thus, even though the disease was originally caused by a new mutation, it occurs in individuals as an inherited trait. For this reason, classic genetic diseases are sometimes referred to as **inherited diseases** to distinguish them from those that are caused by new mutations in the afflicted individuals.

WWW Reading: SRY paper
Human SRY binding to DNA
Focused Reading: p 215 Fig 11.20 Sequencing DNA

At this WWW site, you will find a virtual reprint of an article that illustrates how important each and every nucleotide is. A Japanese couple has had problems conceiving a child and both of them go to a fertility specialist for some advice. This woman a mutation with dramatic system wide phenotypic consequences. The paper discusses a person who has a mutation in the SRY gene which is a gene located on the Y chromosome. A functional copy of SRY is required for embryos to develop as males rather than females.

Scott Gilbert's Developmental Biology text gives an overview of how the SRY gene was identified. You can check it out at http://www.devbio.com/chap17/link1702.shtml

Study Questions:

1. What were the clinical symptoms of the woman described in this paper? Which sex chromosomes did she have?

2. What kind of mutation(s) did she have in her SRY gene?

3. Do you think she inherited this mutation or do you think it is a new mutation in her?

4. Be able to explain to your non-science friends why this woman was infertile.

5. What would happen to her if she wanted to compete in the Olympics and was subjected to a

karyotype analysis?

To be precise, all of our physical traits originated as new mutations that were passed down to succeeding generations. This is one of the major tenets of the theory of evolution—new mutations arise spontaneously all the time. These mutations are either advantageous to the organism (the 'mutant' organism lives and successfully transmits these genes to their offspring), disadvantageous to the organism (the 'mutant' individual is less successful or unsuccessful in passing on these traits) or neutral (the mutation is of no consequence to survival, in the current environment—it just gets passed along to the next generation). Thus, as mutations occur and provide advantage to the organisms bearing the mutations, they are **selected** by the environment (a process called **natural selection**) and they eventually become a standard trait of the species as more and more individuals who bear this trait out-compete individuals who lack the trait.

A theory from the tale of human evolution should illustrate this point. Humans (*Homo sapiens*) first arose in Africa from lower primates that were covered with thick body hair. Humans began to lose their thick body hair due to an advantageous mutation. (What is so advantageous about this remains a topic of debate.) Upon the loss of thick hair, the skin became more exposed to the harmful ultraviolet radiation in sunlight. These high-energy rays can mutate thymidine bases and break DNA, causing a mutation and skin cancer. (By the way, this is still the case—exposure of human skin, especially the lighter skin colors, can break DNA and cause skin cancer.) These early, thin-haired humans had to rely on the expression of genes that control the enzymes that make **melanin**, the dark pigment of skin. Having higher production of skin melanin and turned the skin a dark color. These individuals didn't get skin cancer as often because their dark skin pigment blocked the penetration of UV light. They were healthier and more able to reproduce and raise offspring to maturity. These dark skinned individuals therefore became the wildtype phenotype in the population. Their pale counterparts represented spontaneous mutations in genes that caused less melanin production. Since the pale skinned individual was more susceptible to UV light damage slowly over generations, dark skin came to be the dominant trait of the human species which is what we see today in Africans.

It should be noted that mutations occur all the time (on average, one mutation per $10^{10\text{-}12}$ bases of DNA). For instance, while some early humans had mutations that increased melanin production in the skin, others had mutations that decreased melanin production, eliminated vital blood proteins, incapacitated vital liver enzymes, destroyed the pigments in the retina that produce color vision, etc. None of these mutations survived in humans because they are not advantageous to the individual and thus do not enhance survival and reproduction.

Your body contains some new mutations which developed in the egg and/or sperm that joined to produce you, or in the cells of your body during development *in utero*, or after you were born. As you know from the discussion above, these mutations can cause a variety of protein changes ranging from no change to complete destruction. You may think that your presence on the planet means that none of these mutations is harmful in any significant way. However, it is quite possible that you do harbor at least one lethal new mutation (destroying an absolutely essential protein), but you are protected from its effects by being **diploid**. One of the tremendous benefits of being diploid is that you can have lethal or harmful mutations in a gene and frequently it won't kill you or harm you because the other allele is wild type and compensates for the deficient allele. You have built in genetic redundancy that safeguards you against mutations. Big, multicellular creatures such as ourselves that take a lot of energy to produce are virtually always diploid which provides such an enormous adaptive advantage.

The presence of a potentially lethal or harmful new mutation makes you a carrier of a defective gene. If you mate with someone who is a carrier of a mutation in the same gene, you stand a 25% chance of

producing an offspring with two mutant alleles at that locus—that child would have a diseased phenotype. Because mutations occur spontaneously (i.e. randomly) in the DNA, it is extremely unlikely that you would pick a mating partner with exactly the same genetic mutation that arose spontaneously in you. However, because mutations are passed down to offspring, they run in families. This is why it is a very bad idea for close relatives to form mating unions. [For instance, if a spontaneous mutation occurred in grandma, she would pass this down to half of her children, who would in turn pass it down to half of their children. If these first cousins married, they would have a dramatically increased probability of producing an offspring with two bad genes, a homozygous individual with serious or lethal genetic problems. Most countries have laws or traditions against such incestuous relationships. One can only speculate about the origin of such traditions.]

If mutations have to confer an adaptive advantage in order to be selected, how then do disease alleles manage to stay in the human population and get passed down from generation to generation? Recessive disease genes get passed down because individuals can be carriers without actually having the disease. Thus, heterozygous individuals are just as healthy and able to reproduce as homozygous "normal" or wild type individuals and the defective genes get passed down. The situation is different with incompletely dominant or dominant traits. If the disease trait interferes with health and reproduction, it should be slowly weeded from the population since anyone with a single diseased allele is not as fit to compete for survival and reproduction. Most classic genetic diseases, therefore, are recessive—not dominant. Exceptions are those diseases that afflict individuals after they have reproduced, such as most cancers and Huntington's disease.

Study Questions:

1. Explain the role of new mutations in evolution.

2. Explain the difference between a new mutation and an inherited mutation. Give examples.

3. In animal and plant breeding, the concept of hybrid vigor is used to explain why hybrid (heterozygous) organisms are heartier than inbred (homozygous) individuals. Explain why this is so.

4. Most genetic diseases are recessive. Explain why this is the case. If maladaptive mutations are selected against, how do dominant and recessive inherited diseases remain in the population despite their detrimental effect on health?

Let's return to our study of the cause of these genetic diseases. Mutations in actual structural genes may be responsible for producing CF, SC and HD. However, in order to develop a more complete understanding of potential genetic flaws, we have to look a bit more closely at the process of **gene expression**. (Gene expression is the process through which the genetic code is used to produce a functional protein→ going from DNA to RNA to protein) In the following discussion we will explore the possible sources of the genetic defects that cause the classic genetic diseases.

Focused Reading:
Fig 12.4 p 222
p 222-3 "Transcription:…" stop at bottom of page
p 265-6 "The structures of…" stop at bottom of page
p 268-70 "RNA processing" stop at "Contrasting eukaryotes…"
p 376 "Posttranscriptional control" stop at "Translational…"

A number of steps comprise the process of transcription. A defect at any one step would interfere with the production of an accurate mRNA. Without accurate mRNA, accurate proteins cannot be produced and genetic disease may occur.

So, in revisiting the three diseases in question, what might be causing the problem with the disease alleles other than a direct mutation in a structural gene? Well, you could have a mutation in a gene that encoded any of the proteins that are required for transcription (e.g. RNA polymerase or transcription factors), RNA processing (splicing, adding a cap or poly-A tail), or transporting the mRNA from the nucleus to the cytoplasm. However, if this were the case, the cell could make no proteins since all proteins use the same polymerases, transcription factors, spliceosomes, processing enzymes and transport proteins. The cell wouldn't exist (this would be a **lethal mutation**), so this is an unlikely hypothesis.

Alternatively, the faulty gene might contain a mutation in its **promoter**. This region normally controls the expression of the gene so that it is expressed in the appropriate cells (lungs, pancreas, liver, and sweat glands) and not expressed in incorrect cells (brain, bones, and kidneys.) The promoter is a sequence of DNA immediately "**upstream**" from the structural gene that is recognized by RNA polymerase and by molecules that specifically control the expression of this gene. Thus, a mutation in the promoter that changed this recognition area might cause the gene to be expressed too much (the promoter is "on" too often allowing too much transcription); too little (the promoter is not "on" enough allowing too little transcription) or not at all (promoter is non-functional and RNA polymerase cannot bind to it). Alternatively, the mutated and defective gene might be in a region called an **enhancer**. As its name implies an enhancer is a segment of DNA that enhances the expression of the gene. The unexpected thing about enhancers is that they can occur several thousand bases (**kilobases**) away from the actual gene and can also be found in introns. A defect in an enhancer may cause a gene to be expressed too infrequently or too frequently.

A third alternative involves a defect in the **introns** of the gene. In order to be successfully spliced out of the primary transcript to form mRNA, introns must contain base sequences that are recognized by the spliceosome and used to determine where the mRNA should spliced. If a mutation occurred in these recognition areas of the intron, correct splicing may not occur in which case accurate mRNA would not be formed and an accurate protein could be made.

Finally, we could hypothesize that a mutation occurred which made the mRNA more or less susceptible to enzymatic degradation in the cytoplasm. If the mRNA remains intact longer than normal, more protein than normal could be made. Likewise, if the mRNA is degraded too quickly, less protein than normal could be made. Thus, the amount of protein may be altered, producing a disease state. This is a viable hypothesis since the signals for degradation of each mRNA are probably at least partially inherent in the mRNA molecule itself and thus specific to this one gene.

Thus, a mutation need not be in the coding portion of the gene (the **exons**) in order to cause a genetic defect. It can also be in any of the genetic elements that control the transcription of the gene, the splicing of the primary transcript into mRNA, or the transport of the mRNA out of the nucleus into the cytoplasm.

Study Questions:

1. What types of mutations may affect protein function besides those within the structural gene? Explain how these mutations produce these changes.

2. Many proteins are involved in gene transcription. Some of them are likely candidates in the quest for the causes of genetic disease, and others are not. Which of these proteins are unlikely to be the cause of any of the classic genetic diseases and why?

3. Describe the role of each of these components in transcription and mRNA processing:
   A. RNA polymerase II
   B. The promoter
   C. The spliceosome
   D. snRNPs
   E. The mRNA transport proteins (in the nuclear pore)
   F. Introns and Exons
   G. Enhancer
   H. Transcription factors

4. What is SRY, what is its function?

In addition to genetic defects in the proteins that control transcription, RNA processing and mRNA transport, genetic diseases may be caused by defects in the proteins that control translation.

Focused Reading: p 225-30 "Preparation for…" stop at "Regulation of translation"
p 231 "Posttranslational events…"
p 276-7 "Translational and post translational control" to end
p 64-5 "The endomembrane system" stop at "The Golgi..."

WWW Reading: Immunofluorescence labeling of the ER
Hemoglobin (RasMol Image)

A defect in translation and post-translational processing may be responsible for causing CF, SC or HD, although it is much more difficult to develop a viable hypothesis about these processes. We could hypothesize, for instance, that a disease was caused by a defect in any of the genes that control the proteins of translation (ribosomal proteins, initiation factors, elongation factors, enzymes such as peptidyl transferase, etc.) However, as in the case of transcription, all proteins are made using the same set of translational proteins and if a defect existed in any of these important molecules, the mutation would be lethal and the cell would not exist.
Another hypothesis could be a defect in the genes that encode tRNA or rRNA. If, for instance, the tRNA that binds to the amino acid alanine were defective, alanine could not be activated and could not be incorporated into proteins thus leading to defects. Again, however, this would affect all proteins of the cell, and would be a lethal mutation.

The defect could be in the enzymes that perform post-translational modifications such as glycosylation, sulfhydryl bond formation, chain cleavage, etc. Again, these are "global" or "house keeping" enzymes that modify all proteins and one would expect to see widespread protein abnormalities if such a mutation existed.

Study Questions:

1. Gene expression is a highly energetic process requiring the expenditure of significant amounts of ATP and GTP. Describe the expenditure of energy (ATP and GTP) during transcription and translation. How is the energy expended? Which parts of the process

require the expenditure of energy?

2. Describe the steps of translation.

3. How are proteins altered during post-translational modification?

4. Some genes encode 'processing proteins' that control the translation and post-translational processing steps of gene expression. Explain why it is unlikely that CF, SC and HD are caused by a defect in a 'processing protein' gene.

The defect in some genetic diseases may cause the protein to get "lost" in the cell after it is made.

Focused Reading: p 65-6 "The Golgi…" stop at "Lysosomes…"
p 230-3 "Polysome formation…" Stop at "Mutations…"

Secreted and membrane-bound proteins require the presence of a **signal sequence** for transport into the ER. The signal sequence is a stretch of amino acids in the protein that act like a 'zip code' telling the cell where the protein belongs. If a protein is supposed to be membrane-bound or secreted, a defective form that causes a disease may contain a mutation in its signal sequence. In this case, the protein could be made, but it would never get to the appropriate area of the cell to be used. Proteins going to the ER are not the only ones that use signal sequences other proteins contain different 'zip codes' which instructs the cell to send the protein to the mitochondrion, the nucleus, or the chloroplast.

Study Questions:

1. Describe the process by which secreted and membrane-bound proteins get from cytoplasmic ribosomes into the ER. What role does the signal sequence play in this process?

2. Explain how a mutation in a gene's signal sequence could produce a genetic disease.

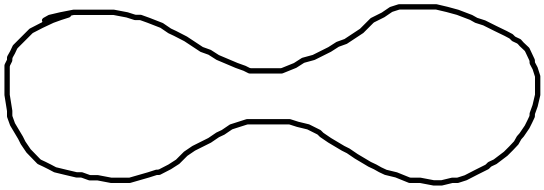
----------STOP----------

While we do not understand the cell or chromosomes well enough to speculate about all the possible mutations that may cause genetic diseases, the preceding discussion certainly gives you an idea about the complexity of genetic systems and the incredible number of steps involved in producing a normal protein. It is nothing short of a miracle that we exist, given all the reactions that have to work exactly right in order for us to produce one gene product, not to mention the products of all 100,000 genes.

We have explored many of the possible proteins that may be mutated to cause genetic disease. Now let's see what we know about the causes of CF, SC and HD and how investigators acquired this knowledge.

Certainly, if you want to know how to cure a genetic disease, it would be very helpful if you could find out which protein is defective and how it is defective. In the case of sickle cell disease, this was a relatively easy process. Because the disease symptoms produce disease and because you can actually see the sickled red blood cells (**RBCs**) under a microscope, it seemed very likely that the

defect is in a molecule in the RBC, possibly a protein that controls the shape of the RBCs. RBCs are normally shaped like this:



This shape is called a **biconcave disk**. Because of the thermodynamic properties of phospholipid bilayers, the most thermodynamically stable shape for a cell is a sphere. Like soap bubbles, if you don't do something special, a cell will always assume a spherical shape. So, in order to maintain the RBC in this odd biconcave shape, the cell has to distort and support the membrane with proteins. One such protein is called **spectrin** and it lies immediately under the cell membrane and holds it in its unusual shape. So, the sickle cell disease mutation could be in the gene that controls the production of spectrin.

However, investigators noticed that the red blood cells were not always sickle shaped. They only became sickle shaped when oxygen levels were low, as in the veins (as opposed to the arteries). The molecule that carries oxygen in the RBC and which changes shape when it binds to oxygen is called **hemoglobin** (see page 41, figure 3.7). Hemoglobin is a molecule much like chlorophyll (we will talk about this more in Unit III) with a porphyrin ring structure containing an atom of iron (in hemoglobin) instead of magnesium (in chlorophyll—see 141, figure 8.9 to see what a metal-bearing porphyrin ring looks like.) It is the iron atom that actually binds the oxygen. RBCs are really bags of hemoglobin—over 90% of their protein content is hemoglobin. Investigators were quick to suspect that the genetic defect may be in the hemoglobin molecule.

Hemoglobin can be isolated from RBC very easily. The red cells are burst open, usually by osmotic pressure. Are they burst by tiny molecular cherry bombs? No, RBCs are put into pure water, which has a very high osmotic pressure. Because of all the proteins, nutrients, and ions dissolved in its cytoplasm, the osmotic pressure inside RBCs is low. Water, therefore, moves into the red blood cell. All that water makes the RBC swell until it bursts, freeing all of its hemoglobin. This process is called **hemolysis** (Heme = red blood cells; lysis = slicing open or cleaving.)

The hemoglobin can then be purified by a number of processes including **column chromatography** (described in Unit 1.) Hemoglobin will be separated from the other proteins in the red blood cell because it moves through the column at its own specific rate. Other proteins will move through faster or slower and thus separation will occur. SC hemoglobin and normal hemoglobin can also be compared using **electrophoresis** (also described in Unit 1). If they move at different rates in the electrical field, they are different sizes. In this case column chromatography was not sensitive enough to detect a change in hemoglobin. Gel electrophoresis detected a slight change in mobility. This change indicated that wildtype and SC hemoglobin had some difference in their size and that scientists should look more closely in this direction. (Note they had to try two approaches that seem to 'do the same thing' [separate by size] before they saw the difference in the hemoglobin—sensitivity counts).

The approach that got at the difference was to determine the **amino acid sequence** of the proteins to see if a mutation has produced a change that could lead to an alteration in function. Each hemoglobin molecule is composed of four chains or **subunits** (the complete and functional molecule has a four-subunit **quaternary structure**) two alpha chains (each 141 amino acids long) and two

beta chains (each 146 amino acids long). These four chains, each containing a porphyrin ring and an atom of iron, interact with one another forming a very large hemoglobin molecule that can bind to four molecules of oxygen, one at each iron atom. Determining the complete amino sequence of each chain was a time-consuming, labor-intensive and tedious process. While you certainly don't have to understand the details of this process, the following brief discussion should give you an idea of some of the technical difficulties involved in this process.

Focused Reading: p 901 Table 50.3; focus on enzymes digesting proteins or peptides

Amino acid sequencing relies on the use of analytic chemical methods to identify amino acids after the digestion of the protein with enzymes that cleave peptide bonds (peptidases or proteases) as specific sites. For instance, if you subject a protein to **carboxypeptidase**, the enzyme will cleave off the last amino acid—the amino acid at the carboxyl terminus of the protein. (The first amino acid translated always has a free amino group (amino=NH) so that end of the protein is called **N-terminus**. At the other end of the chain, the last amino acid always has a free carboxyl group so it is called the **C-terminus**. See page 38, Figure 3.4 for an illustration.) Trypsin will cleave on the carboxyl side of lysine or arginine; chymotrypsin will cleave on the carboxyl side of phenylalanine, tryptophan or tyrosine; etc. In addition, various chemical processes can be used to tag or label the C- or N-terminus amino acid or other specific amino acids so they can be identified by analytical procedures. For instance, you could radioactively tag the C-terminal amino acid and then subject the peptide to carbaminopeptidase. The analysis of the liberated amino acids should show only one amino acid bearing the tag. This is the amino acid at the C-terminus end of the protein. You could then subject the protein to carbaminopeptidase for a short period of time and then tag the C-terminus. The tagged amino acids in this analysis should be the second, third and/or fourth amino acids (the enzyme just keeps cleaving amino acids, one after another, off the C-terminus end.) While the procedure is very complicated, this brief example may give you an idea of some of the technical manipulation that is involved. Needless to say, amino acid sequencing is not something a laboratory attempts unless it has a very good reason to do so.

Focused Reading: p 39-42 "The primary structure…" stop at "Chaperonins…"
p 876 Fig 49.11 note how RBCs must fit through capillaries
p 234 Paragraph beginning "In contrast to silent…" stop at "Nonsense mutations, another type of mutation…"

WWW Reading: Hemoglobin Mutagenesis Page

Wild-type hemoglobin (called **hemoglobin A** was sequenced in the 1950s in Germany and the United States. Hemoglobin from a sickle cell disease patient (now called **hemoglobin S**) was found to be absolutely identical in amino acid sequence except for a single difference at position #6 on the beta chain (6 amino acids from the N-terminus). Hemoglobin A has a glutamic acid at position #6 while hemoglobin S has a valine at this position. If you look at the genetic code on page 224, you see that the difference between the code for glutamate (GAG) and valine (GUG) is a single base in the middle of the codon. By changing the sequence of the codon for glutamate from A to U, valine is put at position 6 instead of glutamate. You can see on page 37 that glutamate (glutamic acid) has a negatively charged organic acid in its R group while valine has a non-polar hydrocarbon. The switch from a charged to a non-polar R group changes the three-dimensional shape of the molecule enough to alter its shape. The shape change is in a critical enough point to change the function of the protein, thus causing the sickle cell disease.

Note, again, that people with SC inherit this mutation from their parents—it does not occur spontaneously in SC patients. The original mutation occurred thousands of years ago. In fact, this mutation appears to confer some adaptive advantage to heterozygotes. Malaria is a dangerous and widespread disease, especially in Africa. This disease is caused by a protozoan that spends part of its life cycle in the RBC. SC heterozygotes are resistant to this phase of the disease and are therefore somewhat more protected from malaria than are normal individual. Thus, despite its harmful effect in homozygotes, the SC gene may in fact have been an adaptive trait for Africans (in Africa) and naturally selected in heterozygotes. This helps explain why SC is so prevalent in African Americans. (This also provides an example of why mutations are not inherently 'good' or 'bad'—it depends on the environment that the organism is in.)

> News Item: Genetic mutations are not the only way to make RBCs less effective. Exposure to carbon monoxide (CO) inhibits the hemoglobin in RBCs from binding oxygen. The cells get through the blood vessels but have nothing to deliver (No oxygen—cells die. CO is why car exhaust is poisonous). A group at the European Molecular Biology Laboratory (EMBL) has used X-ray crystallography and molecular modeling to visualize the protein and determine cellular mechanisms that block CO binding. Their work on sperm whale myoglobin indicates that CO can only bind after two α helices shift position slightly. How hemoglobin can tell the difference between oxygen and CO is not yet known but, as cities get bigger and we all must drive SUVs, its a good research direction to give 'air time' to. G.S. Kachalova 1999 *Science* vol 284 p463-6.

Study Questions

1. Describe the process by which red blood cells are lysed by osmotic pressure. Explain why water moves into the cell under these experimental conditions.

2. What approach was taken to determine the cause of sickle cell disease?

3. What specific genetic defect causes sickle cell disease?

4. Describe the selective pressure that may have actually enhanced the presence of the SC allele in the African and African American populations.

5. Why is glutamic acid the 6^{th} amino acid if it is encoded by the 7^{th} codon?

> NEWS ITEM: The FDA is about to approve a drug that has been used to treat cancer for over 30 years for the treatment of SC. The drug is called hydroxyurea and it has the ability to activate the transcription of a gene that is normally silent. The gene being activated encodes for a form of hemoglobin that we produced while we are embryos, but is silent forever once we are born. These fetal hemoglobins work just as well as adult hemoglobin and so it should work as a good alternative for those with SC. (see article by Robert Finn in *The Scientist* February 16, 1998. Pg 9.)

6. Is the hydroxyurea treatment considered a cure? Will those being treated still be at risk of having children with SC?

> **News Item:** In December 1999 the Associtaed Press reported the success of a new cure for Sickle Cell Anemia. A thirteen year old suffering from SC was treated by introducing stem cells from the umbilical cord of an unrelated infant who did not have SC. (Stem cells are undifferentiated cells found in bone marrow that develop to produce red blood cells.) The transplant, performed Dec 11, 1998, is the first time unrelated cord blood has been used to treat sickle cell anemia and is much less painful than bone marrow

transplants that have been used in the past. His treatment was to provide him with a self-renewing source of healthy red blood cells (the stem cells). After one year the cord cells have taken hold in the boy's bone marrow and are making healthy blood cells so the doctors have declared the child 'cured'. Do you consider this a cure? If he should have children would they be at risk of having SC?

The sickle cell disease puzzle was solved relatively early because a cellular defect was visible through the microscope and the probable protein affected by SC was fairly easy to deduce. Unfortunately, the overwhelming majority of inherited genetic diseases are much more difficult to investigate. In the case of cystic fibrosis, all investigators knew for many years was that the disease altered the way in which mucus is handled by the lungs and pancreas. Patients suffered from pneumonia, loss of digestive enzymes, liver cirrhosis, production of profuse sweat with a high salt content and, in some cases, sterility. This mixture of symptoms doesn't immediately point to a culprit. We have been referring to the CF gene as a "mucus gene", but that doesn't explain the all the symptoms of the disease.

A real breakthrough in CF research came in 1984 from a lab investigating the differences between respiratory cells from CF patients and wild-type individuals. This group tested the ability of respiratory cells to respond to second messengers. Wild-type respiratory cells pump Cl^- into extracellular spaces in response to the activation of the cAMP second messenger system. To review, the cAMP second messenger activates cAMP-dependent protein kinase which, in this case, presumably phosphorylates the Cl^- pump and increases the rate at which it pumps Cl^- from the cytoplasm to the extracellular space. Because Cl^- exerts osmotic pressure, water follows the Cl^- and moves outside the cell in response to the cAMP signal.

This lab group (Sato and Sato) found that respiratory cells from CF patients were not able to pump Cl^- in response to cAMP activation. They asked whether this might be because cAMP cannot activate cAMP-dependent protein kinase, and they found that the protein kinase <u>does</u> become activated, but it does not activate any Cl^- pumping action. While we always have to be wary of jumping to conclusions that are insufficiently supported by the data, this was a very exciting finding since it correlates with several of the disease symptoms. In the lungs and pancreas, if Cl^- cannot be pumped into the breathing tubes (bronchi) of the lungs or secretory ducts of the pancreas, water will <u>not</u> follow and the mucus normally found on these internal surfaces will remain thick and dry. This condition will harbor bacteria in the lungs causing pneumonia and will block the passage of digestive enzymes from the pancreas to the intestine.

It looked at this point as though the CF gene produced a defect in the Cl^- pump in the membranes of respiratory cells and possibly the cells of the pancreas. Because of the difficulty in working with membrane-bound proteins, and because of the availability of new techniques in molecular biology, the next steps in the solution to this genetic disease came not from cell biology or genetics, but from molecular biology.

The nucleotide sequence of an isolated gene can be determined much more easily than amino acid sequences can be determined. This allows investigators to work backward using the genetic code to determine the amino acid sequence of the protein. Sometimes this amino acid sequence gives a clue about the protein's function. For instance, membrane-bound proteins tend to have alternate stretches of hydrophobic amino acids with stretches of hydrophilic amino acids. While this pattern does not necessarily mean that it is a membrane protein, it gives investigators a clue about where to look. Therefore, the hunt was on for the CF gene. Once the CF gene cold be found, investigators would use the gene to determine the structure of the protein involved, and then use the protein to determine the cell biology that is actually causing the disease.

The human genome (the sum of all of the DNA in all 23 pair of human chromosomes) contains about 6×10^9 base pairs and about 100,000 functional genes. (Over 98% of the genome is non-

coding sequences!) So locating a single gene in this gigantic mass of DNA is like looking for a needle in a haystack of DNA, but even the haystack is too small to see! Investigators working on genetic diseases are trying to find these needles by some very ingenious techniques we will describe below.

As an interesting aside, the US government has funded an enormous scientific enterprise called the **Human Genome Project**. The project was first headed by James Watson (of Watson and Crick fame) and is an internationally coordinated effort to identify the base sequence of the entire human genome in about 15 years. Recently a 'rough draft' of the human genome was completed. Estimates place the cost of this project at about $3-5 billion (about one dollar per base pair). The ethics of this project are currently of being widely discussed. The knowledge of the entire base sequence of the human genome will give scientists tremendous power to manipulate the genetics of the human species. We have already seen a small glimpse of this power in the ability to detect genetic abnormalities before birth through amniocentesis. Many couples have chosen to abort fetuses when they find that they have **Down Syndrome** (called **trisomy 21** because it is caused by the presence of an extra chromosome #21 (three instead of the normal two). [The detection of this abnormality does not require the techniques of molecular biology.] One can only wonder what parents will do when many, many more genetic diseases and traits can be diagnosed *in utero*. What if the fetus has genes that predispose it to cancer, to heart disease, to homosexuality, to baldness, to being overweight? Further, *in vitro* fertilization now allows the predetermination of genetic traits. Egg and sperm can be joined in a petri dish producing embryos whose genetic traits can be screened before they are implanted in the woman. As we gain more and more knowledge about the human genome, more and more traits will be screenable. The "correct" embryos can then be implanted in the woman's uterus and the "defective" embryos discarded (similar to polar body diagnosis described on p 385 bottom of first column). See any ethical issues here?

Obviously, this raises significant questions about what we mean by "normal" and "defective." One could hold the view (and many in the disabilities movement do) that we abort Down Syndrome fetuses because, as a society, we place far too much emphasis on physical and mental perfection. As was the case with nuclear technology in the 1940s, the knowledge we gain through the Human Genome Project will test our wisdom as a society in unprecedented ways.

Study Questions:

1. Explain the approach taken by Sato and Sato that identified a defect in the respiratory cells of CF patients. Describe this molecular defect.

2. What are the goals of the Human Genome Project? How does this approach differ from the approach taken by investigators studying individual genetic diseases? Briefly discuss some of the ethical and economical issues raised by the Human Genome Project.

----------STOP----------

Investigators working on specific diseases usually begin to identify and isolate the disease gene by trying to determine the rough location on a chromosome of the gene so they can limit their search to part of a chromosome rather than the entire genome. As a beginning, investigators try to determine which one of the 23 pairs of homologous chromosomes bears the locus for the disease gene and its normal allele. In order to understand how investigators determine this, we need to look at the phenomenon of **linkage**.

Focused Reading: p 180-186 "Mendel's experiments…" stop at "Many genes..."

p 190-192 "Genes and chromosomes" stop at "Sex is determined…"
Fig 10.23 & 10.24

Genes that are on different chromosomes are passed down to offspring through independent assortment as described by Gregor Mendel. Here is an example. Let's say that the locus controlling CF is on chromosome #10 and the locus controlling some other polymorphic trait, let's say blood group, is on chromosome #3. For the CF locus, you have two alternatives. The allele can be wild type or CF. As you learned from this reading assignment, we now use a more modern terminology to express these alleles. In Mendel's notation, the dominant allele had a capital letter and the recessive a low-case letter. The letter was determined by the dominant trait (e.g. green (G) and yellow (g) -- green is dominant to yellow.) However, because the recessive trait (i.e. yellow) is usually the one that is under investigation as an interesting mutation, this notation isn't very helpful. Thomas Hunt Morgan devised a system of notation in which the mutant allele is designated by italicized letters, and the wild-type allele is designated by the mutant letters with a "+" superscripted. If the mutant allele is recessive, it begins with a lower-case letter, if dominant, with an upper-case letter. In the case of CF, we could use *cf* to designate the mutant (disease causing), recessive allele that causes CF. Given this nomenclature, you could have the following genotypes at the locus in question:

$cf^+\ cf^+$	wild-type
$cf^+\ cf$	heterozygous carrier
$cf\ \ cf$	homozygous recessive, disease phenotype

For blood groups, you can be phenotypically A, B, AB, or O. A and B are codominantly inherited while O is recessive. Because all three blood types are caused by naturally occurring, wild-type alleles, we can designate the wt alleles A^+, B^+, and o^+. The possible phenotypes and their corresponding genotypes are listed below.

Phenotypes	Genotypes
A	A^+A^+ or A^+o^+
B	B^+B^+ or B^+o^+
AB	A^+B^+
O	o^+o^+

Now, if CF and blood groups are on different chromosomes, these traits will be independently assorted when they are passed down to the next generation. Here is an example. Let's say Maria is blood type AB and is a carrier for CF.

Maria's genotype is:

$$A^+\ B^+\ cf^+\ cf$$

Louis is blood type O has CF.

Louis' genotype is:

$$o^+\ o^+\ cf\ cf$$

According to Mendel's first law, the alleles at each locus segregate independently of one another when gametes are formed. Therefore, in Maria's case each egg receives one blood group allele and one CF allele. If the alleles are on different chromosomes, then they are **not linked** and they assort independently into the gametes. That means that four types of eggs will be produced:

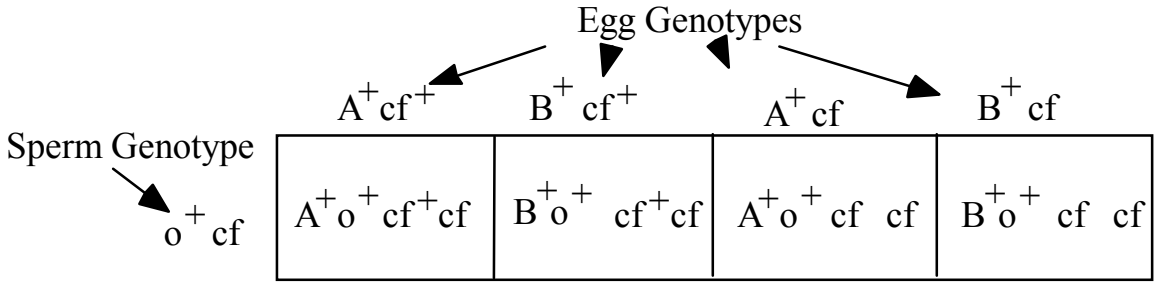
Egg type 1: $A^+ cf^+$

Egg type 2: $B^+ cf^+$

Egg type 3: $A^+ cf$

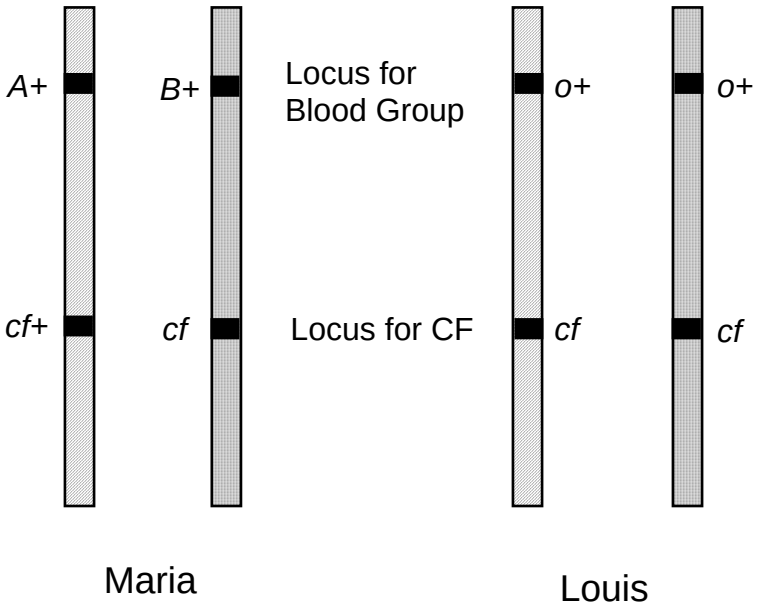
Egg type 4 $B^+ cf$

Louis's alleles also segregate independently during meiosis, but because he is homozygous at both loci, all of his sperm would get one o^+ and one *cf*. If Maria and Louis should produce offspring (and this is fairly unlikely in this case since CF causes infertility in males, but let's say Louis is an exception to the rule), this is what the Punnett square would look like:



This is a classic Mendelian test cross in which a dihybrid is mated with a homozygous recessive individual. If blood group and CF are on different chromosomes, there are four possibilities for the children: carriers of CF with blood type A or B and afflicted individuals with blood type A or B. All possibilities are equally probable. If Maria and Louis were elm trees producing thousands of offspring, about 25% of the offspring would be in each category. (Not that elm trees have blood or get CF, but you get the point.)

Now, let's say that CF and blood groups are on the same chromosome—that they are **linked**. Here is a picture of what this chromosome (homologous pair) might look like in Maria and Louis:



Because A^+ is linked to cf^+, the two alleles go together (assort together) into the gametes. Likewise, because B^+ is linked to *cf*, these two alleles assort together. Thus, if Maria and Louis have children under these circumstances, this is what the Punnett square would look like:

Sperm Genotype ↘ / Egg Genotypes	A^+cf^+	B^+ cf
o^+ cf	$A^+o^+cf^+cf$	B^+o^+ cf cf

In this case, there are only two alternatives for the offspring. They are <u>either</u> 1) blood type A and a carrier <u>or</u> 2) blood type B and afflicted with the disease.

Now, here's the deal. You gather up all of the information you have on many many families that have members suffering from CF and determine the blood type of each member (afflicted or not). By analyzing this information you can see if the traits follow the pattern on the last page (4 possible combinations) or the pattern on this one (a child that is blood type B always has the disease). If blood type B is always inherited with cf then the two loci are 'linked' on the same chromosome. If you know which chromosome carried the bloodtype gene you now know that that chromosome carries the CF locus (the same one). If inheritance patterns follow the example on the previous page you know that CF and blood type are not on the same chromosome so, in your search for a chromosomal location you have eliminated one and only have 21 left to go. (21 because you also know that CF is not on the sex chromosome because the disease is not sex-linked—that is, it occurs in males and females in approximately equal numbers. More on this later.)

Study Questions:

1. Explain the Law of Independent Assortment. What exactly does this law tell us about genetic inheritance?

2. Understand and be able to use Morgan's genetic notation.

3. Be able to predict the genotypic and phenotypic frequencies for dihybrid crosses and dihybrid test crosses in situations where the loci are linked and unlinked.

4. Be able to solve genetics problems such as the ones below (from Biology by Villee, et al)

   A. In rabbits, spotted coat (S) is dominant to solid coat (S^+) and black (B) is dominant to brown (B^+). A brown spotted rabbit is mated to a solid black one, and all the offspring are black and spotted. What are the genotypes of the parents? What would be the appearance of the F_2 generation if two of these F_1 black spotted rabbits were mated?
   B. The long hair of Persian cats is recessive to the short hair of Siamese cats, but the black coat color of Persians is dominant to the brown-and-tan coat color of Siamese. If a pure black, longhaired Persian is mated to a pure brown-and-tan, shorthaired Siamese, what will be the appearance of the F_1 offspring? If two of these F_1 cats are mated, what are the chances that a longhaired, brown-and-tan cat will be produced in the F_2 generation?
   C. What kinds of diploid matings result in the following phenotypic ratios?
      3:1 1:1 9:3:3:1 1:1:1:1

5. Given information about the chromosomal location of one trait, be able to devise a genetic cross that will allow you to determine if a second trait is also encoded on that same chromosome.

6. Given data from a linkage experiment such as the one presented above or the one you devised in question #5, be able to interpret the data to assess whether or not the traits are linked.

----------STOP----------

Well, as is usually the case, CF is not linked to something as obvious and easy to detect as the ABO blood group. However, it is linked to something almost as good—a RFLP (pronounced "rif-lip").

Overview reading Chapter 17
Focused Reading: p 315 "Plasmid as vector" (3 paragraphs)
p 312-13 "Restriction endonucleases…" stop at "Recombinant DNA…"
Fig 17.2
p 337-8 "DNA marker…" stop at "Human gene…"
Fig 18.7
WWW Reading: Cartoon of Southern Blot Method
Real Southern Blot

RFLPs can be thought of as genetically inherited traits like brown eyes and dark skin. Polymorphic traits, such as eye color, skin color and RFLPs allow investigators to follow genes on a chromosome. As in the hypothetical case of CF being linked to blood groups, you can tell CF is on the same chromosome as blood groups because both loci are inherited together, they are linked. "A" followed the *cf*+ gene and "B" followed the *cf* gene. Without different allelic alternatives to follow, you can't

do genetic analysis. The problem is, as mentioned earlier most human traits are not polymorphic. For most proteins, every human has exactly the same alleles as every other human. So finding polymorphic traits that can be easily detected has been a tremendous problem and barrier to progress in genetics. Our problems have been solved by the discovery of RFLPs, thanks to the 98% of the DNA in our chromosomes that is non-coding DNA.

Although, 98% of the DNA in the genome does not encode functional proteins, these base sequences are passed on from generation to generation. You inherit your non-coding DNA from your parents with the same degree of accuracy as you do your functional genes. Mutations can occur in these non-coding sequences (just as they can in functional genes) and these mutations are then passed on to offspring. As far as we know, mutations in these non-coding areas do not matter much to the survival of the organism, so they are not selected against and tend to stay in the gene pool.

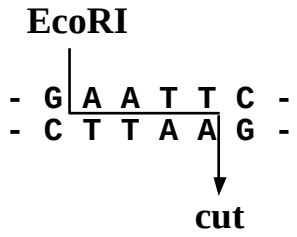
Because these non-coding areas do not code for a protein, we cannot analyze them by looking at the amino acid sequence or the function of the proteins they produce. Rather, if we want to analyze these non-coding regions, we have to look at their nucleotide sequence.

In order to establish the presence of a RFLP on a chromosome, or segment of chromosome, you have to have a way of labeling certain DNA or RNA sequences so that they can be seen with the naked eye. You do this with a **probe** that traditionally was radioactive (because it contains radioactive phosphorus in its phosphate groups) and complementary in nucleotide sequence to some chosen sequence of bases. [Currently, most researchers are switching to non-radioactive probes since they are cheaper, more sensitive, and safer.] Most probes are pieces of DNA isolated from other species. For example, if we wanted to clone the human version of the glycogen synthase gene, we might use the previously cloned mouse version of the same gene as our probe. Since they have a highly conserved structure and function, we assume that the nucleotide sequences for the two genes would also be conserved. Another type of probe is called an **oligonucleotide** (oligo- means a polymer of unspecified length; nucleotides are what get polymerized). Oligonucleotides are short stretches of single-stranded DNA, which are synthesized by a machine called a nucleic acid synthesizer. On this instrument is a 4 letter keyboard so you can type in the sequence you want and the synthesizer makes millions of copies of the short nucleic acid chain with the base sequence you typed. The machine is loaded with dATP, dGTP, dTTP and dCTP, and in this case the synthesizer is programmed to create an oligo with the sequence 5' AATTCCGGTGGCATTACT 3'. (Note: by convention, DNA sequences are always written with the 5' end on the left, but where indicated in this illustration, we have written some sequences backwards, 3' to 5'.) This oligo is then made radioactive by using a kinase to add a ^{32}P phosphate to its 5' end. The radioactive oligo is now ready to be a **probe** and will bind (by complementary base pair bonding) with the DNA sequence 3'TTAAGGCCACCGTAATGA5' which becomes our DNA **marker**. It's a stretch of DNA that we can always label or mark with our radioactive probe and follow in a family pedigree.

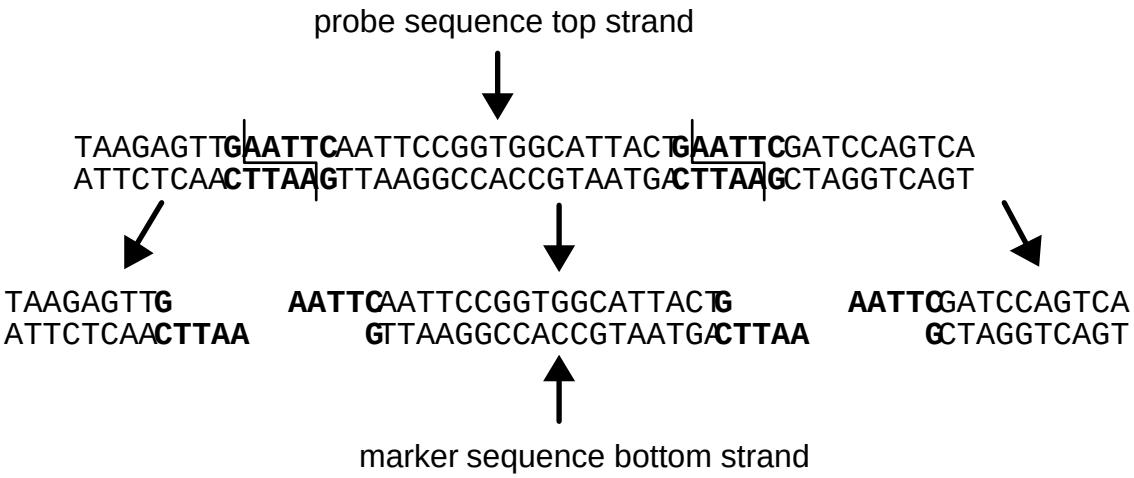
Now let's look for RFLPs. In order to do this, we have to get DNA from many different individuals since we are looking for a <u>polymorphism or genetic variability between individuals</u>. Let's say we get DNA from Jack and Jill for starters. To get a complete set of chromosomes from a person, you simply have to take any cell from their body that has a nucleus. Every nucleated cell of the body (all 50-70 trillion of them!) contains a complete set of chromosomes. This is called **genomic DNA**—at the genetic level all of your cells are equivalent even though they have quite different phenotypes. The genes found in your DNA are expressed differently in different cells so that you wind up with liver cells that look and act differently than hair follicle cells.

In humans, the white blood cell, or **leukocyte**, is a popular source of DNA for analysis since sampling merely requires drawing blood. You then incubate the DNA from the chromosomes with a **restriction enzyme.** Let's say you choose the restriction enzyme **EcoRI** (pronounced eco-are-one

and named after the *E. coli* bacterium from which it was isolated. This was the first restriction enzyme ever discovered and was called restriction enzyme #1, or RI. This discovery was worth a Nobel Prize.) This restriction enzyme recognizes the following base sequence and every time EcoRI sees GAATTC, the enzyme makes the following cut:



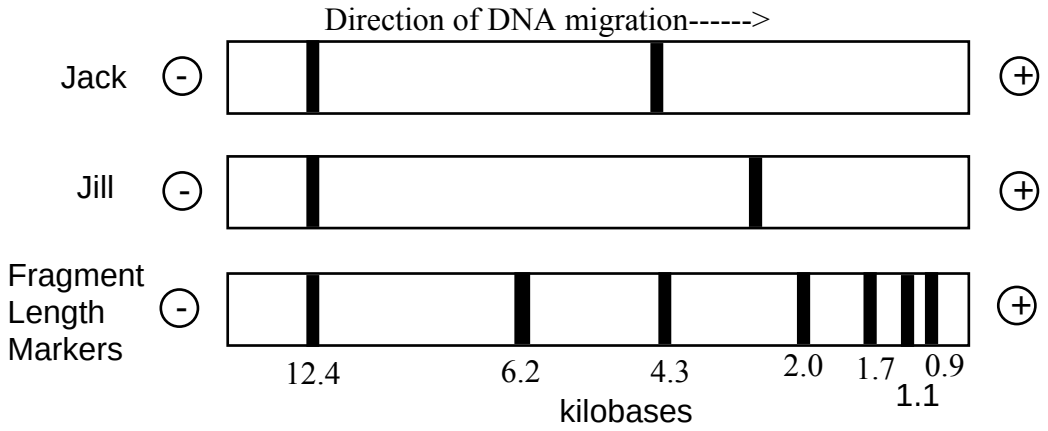
Every time this sequence (GAATTC) appears in Jack and Jill's DNA, this enzyme will make this cut, as shown below:



Thus, a restriction enzyme cuts the large, chromosomal DNA into small fragments (called **restriction fragments** because they are created by restriction enzymes) that can then be electrophoresed and separated by size. Because the restriction enzyme digestion of the entire genomic DNA creates millions of restriction fragments of different sizes, the bands of this electrophoretic separation are so numerous that it looks like a continuous smear all the way up and down the gel. This doesn't help much, so we have to use our radioactive probe.

The DNA separated by electrophoresis is transferred to a piece of special filter paper (usually researchers use nitrocellulose) and the DNA binds to the nitrocellulose so the immobilized DNA can incubate with a probe that is floating in a solution that bathes the nitrocellulose. This process is called **Southern blotting**—you have created a **Southern** blot by transferring DNA from an electrophoretic gel onto nitrocellulose. The words are big but the process is very simple. Basically, if you can make Jello™ and handle paper towels, you can do this. The DNA is also **denatured** during this process. Denaturing DNA is a bit different from denaturing protein. When you denature DNA, you unzip the double helix and convert the molecule into two single strands. You then apply your radioactive probe, allow the probe to bind, or **hybridize**, to its complementary sequence, wash the blot to remove unbound probe, and see where the radioactivity is. In order to see this radioactivity, you have to place an X-ray film over the DNA and give it time to be exposed by the emissions of the radioactive phosphorus. Everywhere the probe has bound, the film will be exposed and turn black. This process is called **autoradiography**. This whole process is diagrammed and explained in more detail on in your web reading.

So, here are the Southern blots from Jack and Jill after they have been hybridized with the radioactive probe and the resulting blot is exposed to X-ray film. In a real blot, you would not see the outlined lanes for each sample of DNA. We have outlined each lane for illustration purposes.



When restriction fragments are electrophoresed, molecular weight, or fragment length, markers are electrophoresed at the same time. These markers are DNA fragments of known length. Their lengths are measured in **kilobases** (1000 bases to a kilobase) or **kb**. By running these markers along with the restriction fragments, you can estimate the length of the restriction fragments in your sample.

The restriction enzyme EcoRI has digested Jack's DNA into many, many fragments, two of which contained the marker sequence 3'TTAAGGCCACCGTAATGA5'. For the sake of clarity, let's call these Jack 1 (on the left) and Jack 2 (right). In Jill's case, EcoRI created many, many fragments, two of which contained the same marker sequence. We'll call these bands Jill 1 (left) and Jill 2 (right).

If we focus only on the restriction fragments that bear the marker (the only ones we can see in a Southern blot), Jack 1 and Jill 1 are the same length (about 12.4 kb). For one of their two chromosomes, the DNA carried by Jack and Jill are probably identical at this locus. However, their other chromosome resulted in different size restriction fragments hybridizing with the radioactive probe. Here is an illustration (with DNA written backwards, 3' to 5') with the numbers in the parentheses being hypothetical distances between the given sequences:

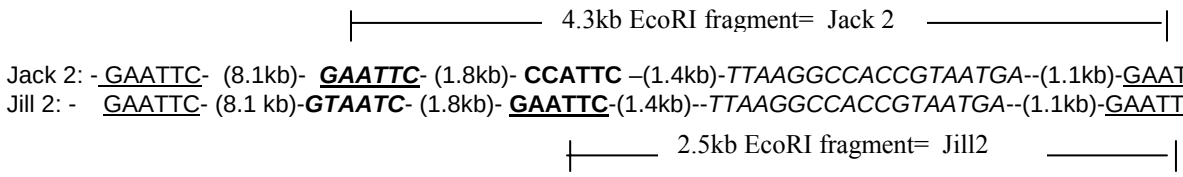
Jack1 or Jill 1: -GAATTC---(11.3 kb)---TTAAGGCCACCGTAATGA---(1.1 kb)---GAATTC-

These fragments are flanked by two restriction target sites for EcoRI and contain the marker sequence. While we cannot say that Jack 1 and Jill 1 are identical (they may differ in the bases within parentheses above), we do know that they both have the marker nucleotides (the probe) and they both are flanked by the target site for the restriction enzyme EcoRI.

But Jack has 2 bands indicating that the probe hybridized with 2 <u>different</u> size restriction fragments of DNA. For this to have happened Jack's 2 copies of this chromosome must not be identical, the copy of the chromosome containing the region we call Jack 2 must contain another EcoRI site. The same is true for Jill. For Jill2 to exist there must be another EcoRI site in this region that puts the 'probe-containing piece' in a 2.5kb piece of DNA.

Comparing Jack 2 and Jill 2 we see that these 2 bands are not the same size. Jack2 Is 4.3kb while Jill2 is 2.5kb. Remember, both of these fragments <u>must</u> be flanked by EcoRI sites and contain the

marker sequence. Because they are different lengths in Jack and Jill's blots, they represent differences in the DNA we call **RFLPs** (**restriction fragment length polymorphisms**). To understand what this means, let's look at one possible scenario that would produce this RFLP.



(# kb=length of DNA between regions—not shown to save space)

In this case, Jack has a 4.3 kb fragment bearing the marker sequence and flanked by two EcoRI sites (underlined). About 8.1 kb downstream from the first restriction site, there is an EcoRI site not found on Jack's other chromosome (the one that gave 12.4kb Jack1). EcoRI 'sees' this recognition site and cuts Jack's fragment into a 4.3 kb length. This piece of DNA contains the marker sequence (in italics) so it hybridizes with the probe and is observed on the autoradiograph. However, Jill inherited a slightly different sequence in this part of her DNA. In this copy of the chromosome she did not inherited the EcoRI site Jack2 has but instead has a sequence 9.9 kb downstream from the first EcoRI target site in which there exists "**GAATTC**". This is the target sequence for EcoRI, and the enzyme will cut Jill's DNA at that site. The digestion of Jill's DNA will produce a 9.9 kb fragment that does not have the marker sequence (so it will not be observed on the autoradiograph), and a 2.5 kb fragment that does contain the marker sequence.

You should understand that this explanation is hypothetical. We usually cannot deduce this much detail from Southern blot data, but something like this happens. We do know that the EcoRI sites that produced Jack's blot were slightly altered in Jill's DNA. She inherited different DNA sequences than Jack did (analogous to different alleles) and this constitutes a RFLP. Different people will demonstrate this particular RFLP if their DNA is digested with EcoRI and probed with the 5'AATTCCGGTGGAT TACT3' probe. This type of RFLP analysis can be used to produce a "**DNA fingerprint**" which can be used as a very accurate form of identification in forensics. (We will perform a different kind of "DNA fingerprinting" during the last two weeks of lab for a preview, see p 328-29.) RFLPs are so polymorphic in the human population that the chances are virtually zero that you would produce an identical DNA fingerprint to anyone else on the planet (except an identical twin) if you use several different RFLPs (i.e. different combinations of restriction enzymes and probes).

WWW: Reading: Genotyping with RFLPs

Study Questions

1. What is a restriction enzyme? Where do they come from and what do they do?

2. What are restriction fragments? Explain the process of electrophoresis. When restriction fragments are electrophoresed they produce a banding pattern. Why? Be able to interpret the band pattern produced by such a technique.

3. Why are fragment length markers run along with sample DNA in electrophoresis experiments?

4. What is a kb? A Mb? What do these terms mean?

5. Explain how a Southern blot is performed. What types of information can you get from a Southern blot that you cannot get from simply electrophoresing a sample?

6. Explain the process of autoradiography. How is this used in the Southern blot?

7. Explain as clearly as you can what a RFLP is. What does the acronym stand for? What is a probe made of and what does it do? The discovery of RFLPs has revolutionized molecular genetics. Why are RFLPs an important tool in genetic analysis?

8. Explain the two parts or components that are required to define a RFLP. In other words, if I told you that investigators had identified a RFLP called DC28036, what information would you expect to get in the published article about this RFLP?

9. What is an oligonucleotide and how is it made? How are oligonucleotides used in the characterization of RFLPs?

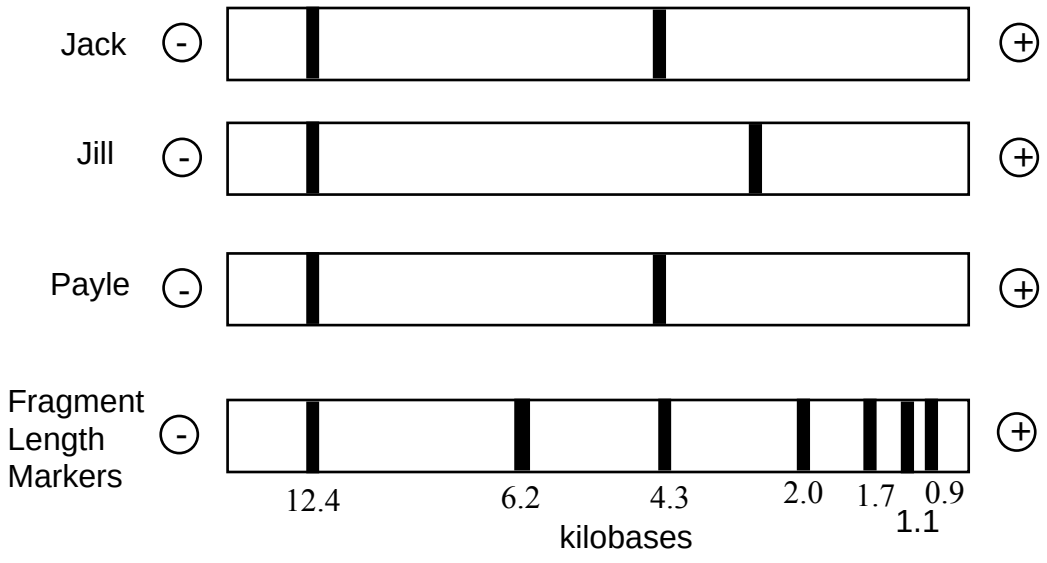
10. In recent years, DNA fingerprinting has become the basis for conviction in criminal trials. If you were called as an expert witness to explain DNA fingerprinting to a jury, what would you tell them?

11. How are RFLPs related to the process of DNA fingerprinting?

Optional WWW Reading: RFLPs Summary and test your knowledge.

----------STOP----------

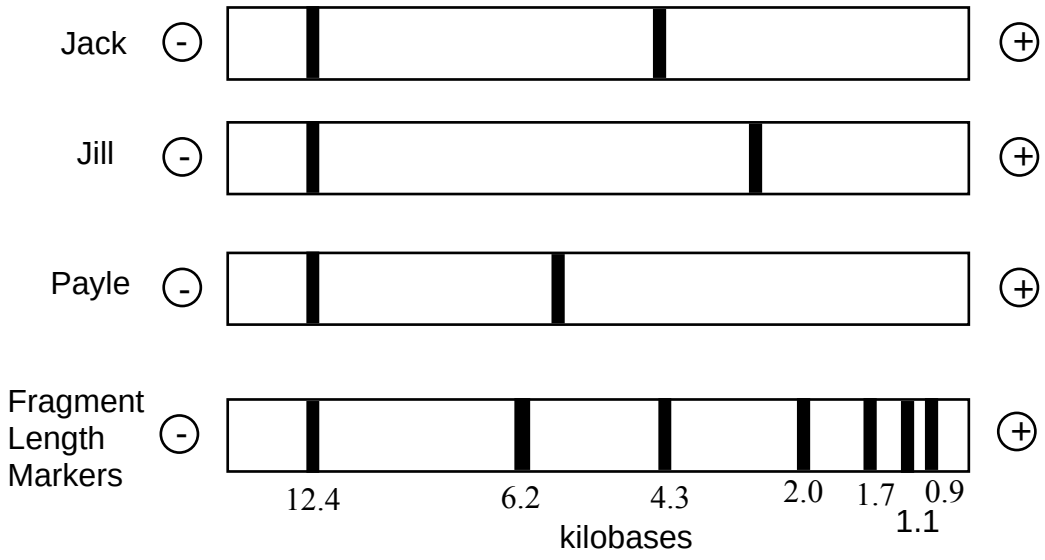
In addition to identifying individuals, RFLPs are passed on to children just as alleles are passed on. To illustrate the power of this multigenerational analysis of RFLPs, let's say that Jack and Jill have a child together. We'll name the child Payle. Let's say we did the same genetic analysis to Payle that we did to Jack and Jill, and this is what we found:



In analyzing these gels, remember that the marker sequence (3'TTAAGGCCACCGTAATGA5') can't simply disappear (except through new mutation and we will assume here that new mutations have not happened.) Jack has two copies of the marker sequence (Jack 1 and Jack 2) Jill has two (Jill 1 and Jill 2). Payle inherited two copies of the marker sequence too. It appears that he inherited Jack 2 and Jill 1 and he did not inherit Jill 2 or Jack 1.

How can we say that he inherited Jill 1 but not Jack 1? Don't parents have to pass their genes on to their offspring? And how is it that Payle didn't inherit Jill 2? We said above that this marker sequence couldn't simply disappear. Well, remember that both Jack and Jill are diploid organisms that produce haploid gametes, which means they pass only half their chromosomes to their offspring. Because Payle had to fetch something from his mother and he did not inherit Jill 2, he had to inherit Jill 1. Likewise, because Payle inherited Jack 2, he could not also inherit Jack 1 since Payle can only get one copy from each parent.

This is one example of the many things you can determine by analyzing family RFLPs. You can tell whether an offspring is actually the child of a couple. Let's say that the RFLP analysis went like this:



It is of little concern that Payle did not inherit Jack 2. Payle could have inherited Jack 1. But, how did Payle get Payle 2, which is not present in either "parent"? He didn't inherit it from Jill—she doesn't have such a fragment, and he didn't inherit it from Jack—he doesn't either. So, the possible

conclusions are: 1) Payle has a new mutation in his DNA; 2) Jack is not the father; or 3) Jill is not his mother (which is unlikely if she gave birth to Payle).

We have analyzed only one RFLP here, but in real paternity cases, several RFLPs are analyzed. Even if one new band in the offspring is due to new mutation, the chances are infinitely small that all new bands are due to new mutations. Therefore, RFLP paternity testing is extremely sensitive and reliable.

It should be noted before moving on that the process of finding a RFLP has been greatly over simplified in these examples. Investigators have to test thousands of probes and scores of restriction enzymes in order to produce the kind of neat package presented here. It is a labor-intensive process, but once the system is set up, it is an extraordinarily powerful and reproducible tool in genetic analysis.

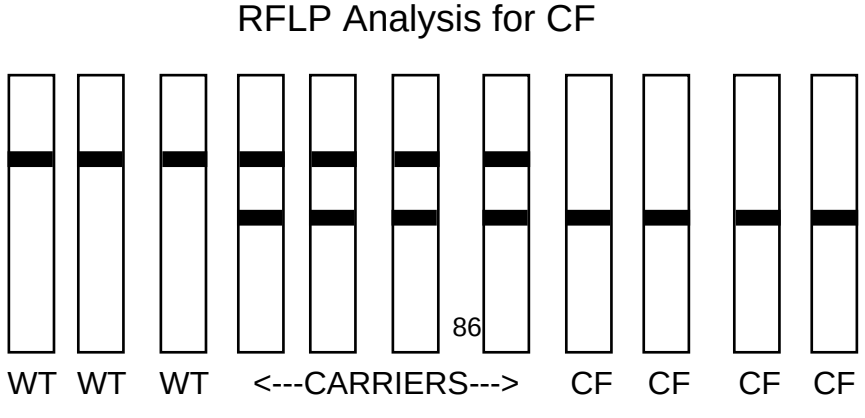
> NEWS ITEM: Ever wonder what makes a 'Chablis' a 'Chablis' and not a 'Chardonnay'? Did all those grapes start out in France or did invaders of long ago bring along their favorites? "Paternity testing" has now been used to trace the lineage of certain cultivars (varieties) of wine grapes. By examining the DNA at 32 different loci scientists have determined that your parents' favorite 'Chardonnay' and 'Melon' may be offspring of the same grape parents. Bowers, J. et al. 1999 *Science* vol 285 p 1562-3.

Study Questions

1. Be able to interpret a multigenerational RFLP analysis. Be able to explain how the analysis does or does not support the assertion that the child is, in fact, the offspring of these parents. Be able to interpret such an analysis to determine which RFLPs represent a heterozygous trait in the parents.
2. Read p328-9 "DNA Fingerprinting…" where VNTR's are discussed. What is a VNTR? How is it similar to a RFLP? How is it different? (NOTE: you are determining your own VNTR pattern for the D1S80 locus in lab)

Lap-Chee Tsui, John R. Riordan and Francis Collins determined that the CF gene was on Chromosome #7 by finding that it was linked to a RFLP that was located on that chromosome. To do this, they gathered DNA from hundreds of families—families without any CF history as well as families afflicted with the disease. They isolated DNA from carriers (the parents of afflicted individuals) as well as CF patients. They looked for linkage between the presence of CF and all the RFLPs they could generate, and they found a linkage between CF and two markers on Chromosome #7. Below is a simplified version of their Southern blot data.

[It should be noted here that many different restriction enzymes and probes are used to do RFLP analysis. The important element in this approach is that once you have identified a RFLP using a certain restriction enzyme and a probe, you must use the same restriction enzyme and the same probe to look for that particular RFLP in everyone.]



Individual 1 2 3 4 5 6 7 8 9 10 11

Individuals 1-3 are from families without CF; individuals 4-7 are carriers (parents of a CF patient but without the disease themselves) and individuals 8-11 are CF patients. Notice that the top band is present in homozygous wt individuals and in carriers but never in CF patients. The bottom band is present in CF patients and carriers, but never in homozygous dominant individuals. The top band contains the marker sequence linked to (inherited with) the wild-type allele. This allele is the only one present in wild-type individual (homozygotes). The bottom band is linked to (inherited with) the disease allele (the CF causing allele) and is the only allele present in CF patients (homozygotes.) Carriers (heterozygotes) have both bands. The RFLP represented by the top band is known to be located on Chromosome #7. (That is, if you digested each chromosome (1 through 22 plus X and Y) individually with the restriction enzyme used in this analysis and applied the probe used in this analysis, only chromosome #7 would give you bands at the positions shown above. If the wild-type gene is on chromosome #7, the disease gene must be there as well.)

It is important to note that, while the restriction fragment with the marker sequence may also contain the CF gene, this need not be the case. All this analysis shows is that CF and this RFLP assort together and are inherited together—they are linked (are neighbors) on the same chromosome. Linkage will be inhereitted within a family so if members of a family have CF, linkage of a consistently identifiable RFLP with the CF gene makes it possible to determine whether someone were a carrier, wild-type, or an afflicted individual using a Southern blot. In other words, the Southern blot can be used to **diagnose** the disease state. For example, if the Southern blot above were performed on a person of unknown disease status and the blot looked like that of Individual # 1 (one higher band), the person would be homozygous wild-type. If the blot looked like that of Individual #4 (two bands), the person would be a carrier. If the blot looked like that of Individual #8 (one lower band) the person would have the disease. This kind of diagnosis can be used to determine whether individuals are carriers or even *in utero* to determine the genetic status of a fetus. This has been a real boon to genetic counselors. Before this test was available, they could only estimate from pedigrees whether or not an individual was a carrier. Now, they can be more certain and offer the family more realistic information on probability of inheritance.

The CF gene was initially found to be linked to two RFLPs on chromosome #7. The next step in the isolation of this gene was to try to pinpoint the location of the gene on the chromosome so that its base sequence could be determined. There are on average 130 million base pairs on each human chromosome. This many base pairs cannot be sequenced easily. One has to work with a more manageable unit, a much smaller segment of DNA. It is much faster to try to pinpoint the general location of the gene on the chromosome, and then sequence the DNA in that specific area. Once the location has been determined, the gene's sequence can be determined.

So, how do you locate a gene on a chromosome? In order to understand how this is done, you have to know something about a process that occurs naturally during inheritance called **recombination**.

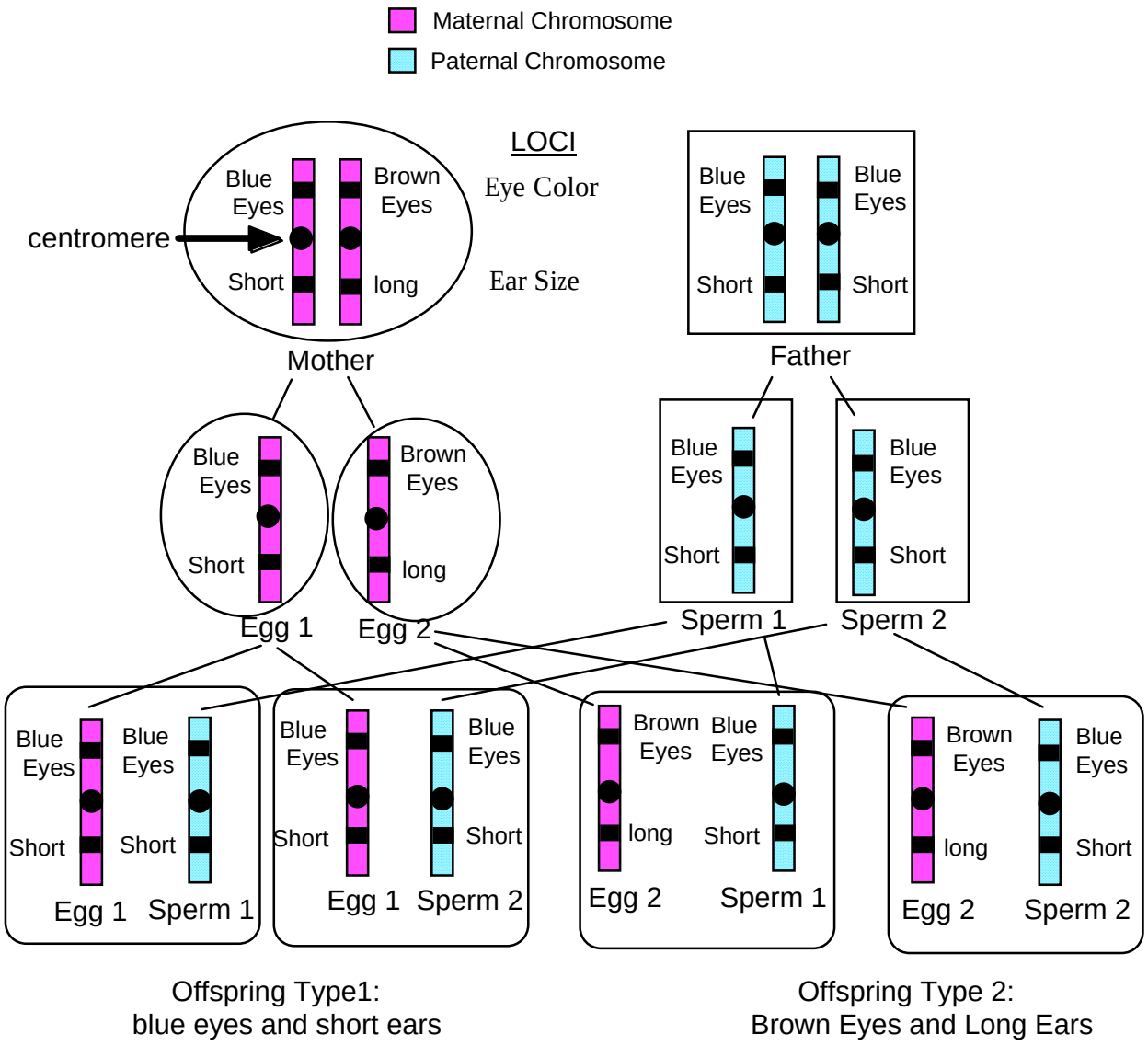
Focused Reading: p 180-2 "Mendel's first law…" stop at "Punnett squares…"

Meiosis Review
p 167-72 "Meiosis:…" stop at "Meiotic errors:…"
Fig on 168-9 Fig 10.19
p 190-2 "Genes and chromosomes" stop at "Geneticists make maps…"
p 341 "There are several ways…" stop at bottom of page
Fig 18.11 and 18.12
p 192 "Geneticists make maps" (2 paragraphs) and Fig 10.22

Fig 10.23 & 10.24

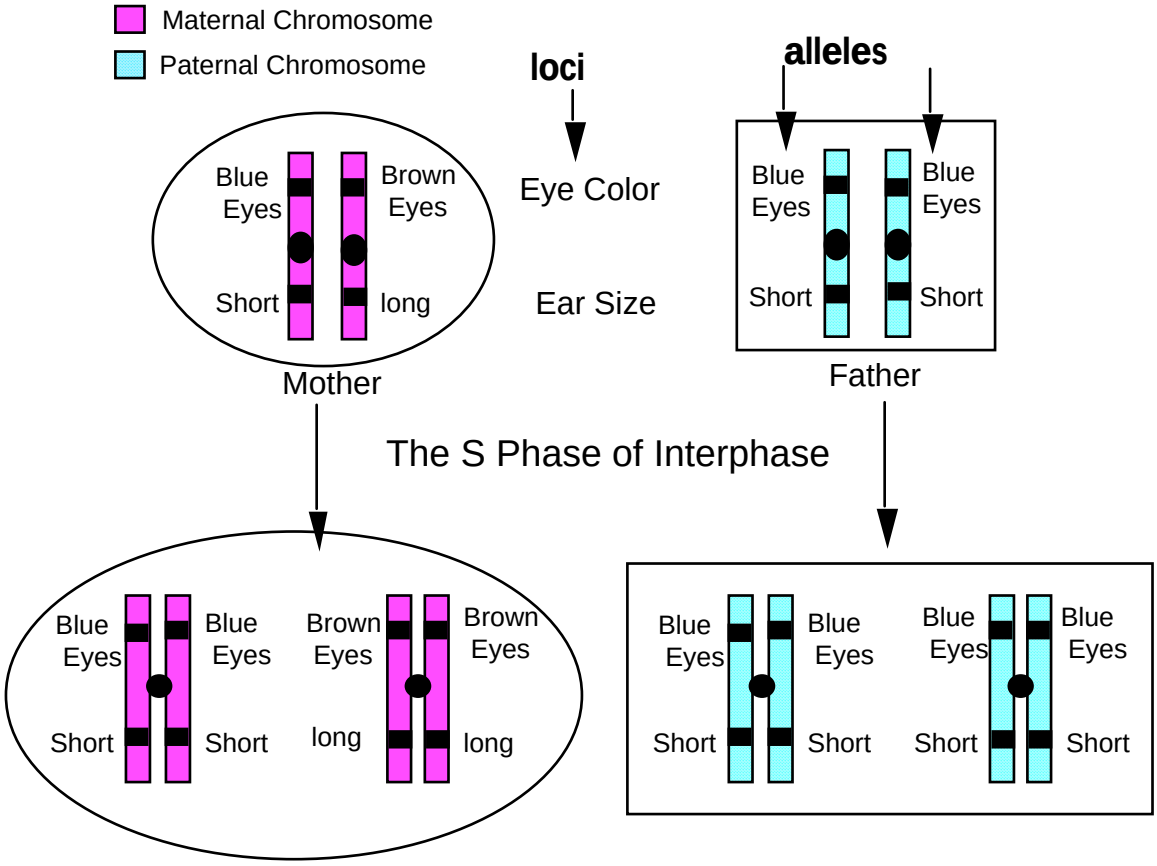
WWW Reading: Karyotyping

The homologous chromosomes segregate during meiosis and are independently assorted into the gametes. Thus, at your own fertilization, you received chromosome 1-23 from your mother and chromosome 1-23 from your father. Thus, you have two of each chromosome—homologous pairs. You inherit your genes in these chromosome "packages". Each chromosome is a long line of genes. Here's an illustration in which brown eyes are dominant to blue, and tall is dominant to short. This illustrates the inheritance of only one chromosome. This happens to all 23 pairs during inheritance.

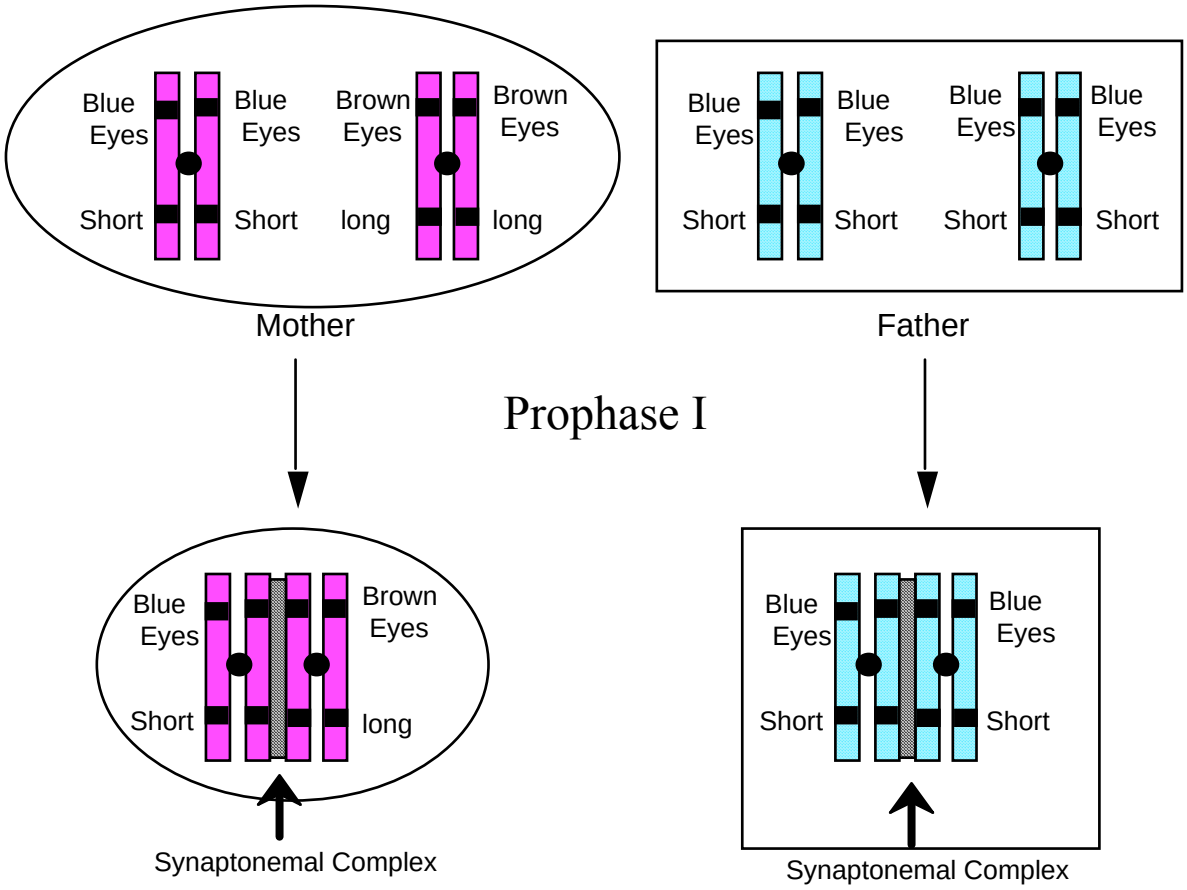


In this example, long stays with brown eyes and short stays with blue eyes. Therefore, you cannot get an offspring from this union that has long ears and has blue eyes, or that has short ears and has brown eyes. So, if you wanted that combination in your offspring, you would be out of luck. Brown is linked to long and blue is linked to short forever and ever and ever. (We'll modify that statement later.)

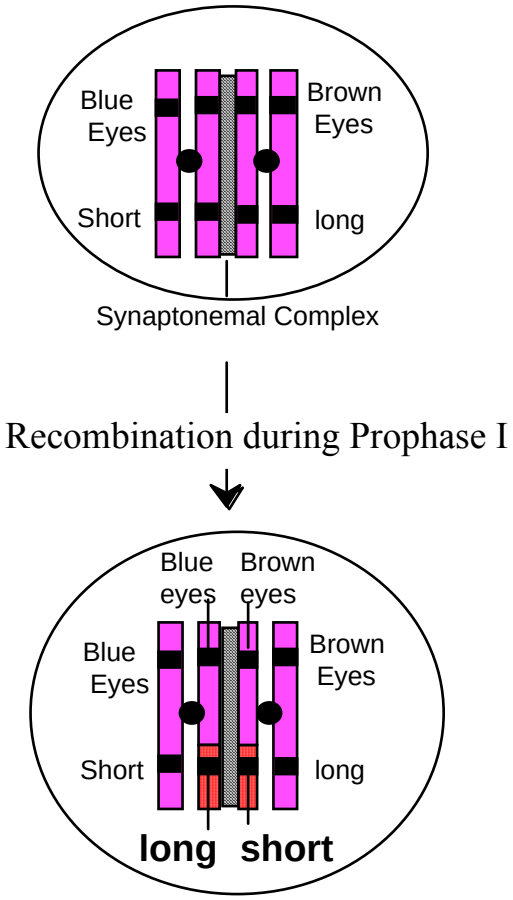
Well, as you know from your reading, genes and chromosomes are not nearly that rigid and immutable; they tend to exchange pieces when eggs and sperm are produced. In the case above, then, when the mother created her eggs, she produced two types: Egg Type 1 and Egg Type 2. However, in actuality, such a woman could produce four types of eggs because this homologous pair might undergo **recombination** during meiosis. During the S phase of interphase, an identical copy of the DNA is made. Thus, each chromosome goes from being a single linear molecule to a double molecule as follows:



Each chromosome makes an exact copy of itself. The copies are attached to one another by the **centromere.** Each half of this double chromosome is now called a chromatid (remember Fig9.6 p160?). The homologous pairs, (which have been ignoring one another in the cell up to this point), find each other and join together, or **synapse**, through a protein complex called the **synaptonemal complex** as follows:



This process, where the homologues find each other and bind is called **synapsis** and it produces a bundle of four chromatids called a **tetrad**. Enzymes called **recombinases** reside in the synaptonemal complex and these enzymes can cut chromosomes and swap pieces in the process of **recombination**. The inner two chromatids in the tetrad (the ones bound by the synaptonemal complex) might swap segments through this process. Thus, after this process, the mother's chromatids would look like this:



The two outer chromatids are the original ones or the **parental chromatids**. However, because of the recombination event, the two inner chromatids are now different from anything that existed in the mother. They are called **recombinant chromatids**. As the mother puts each of these chromatids in different eggs, some eggs will get chromatids in which blue eyes are linked to tallness, and brown eyes to shortness. Four different types of offspring would result from this union: blue eyes and short, brown eyes and tall, blue eyes and tall and brown eyes and short.

Note here that recombination happens in the father as well when he produces sperm, but because he is homozygous at both of these loci, recombination does not produce any new combinations. He still can produce only one kind of chromatid—blue eyes and short. Also note that recombination can happen on all 4 chromatids, not just the inner two chromatids as shown in this simplified diagram.

This feature of meiosis and inheritance was discovered by Thomas Hunt Morgan and is used by nature as a way to increase the diversity in a population, thus giving natural selection a greater variety of organisms to work on. Nature was recombining its chromatids long, long before humans ever populated the earth. Determining the location of genes on chromosomes is called **chromosome mapping** and it relies on a discovery that TH Morgan made about recombination: That the frequency of recombination between two loci is proportional to the distance between the two loci on the chromosome. That is, if two loci are very far apart on a chromosome (say at opposite ends), then recombination is very likely to occur at a point between these two loci, thus moving their alleles to

homologous chromatids. Conversely, if two loci are very close together on a chromosome, it is very unlikely that recombination will occur in the tiny stretch of chromosome between them and thus they are likely not to have their alleles separated on different but homologous chromatids.

Understand? Good. But how does this allow you to map genes or RFLPs on a chromosome? Well, if you had a way to measure the frequency of recombination between two loci, you could determine how far apart they are on a chromosome. In order to do this, geneticists have defined the distance on a chromosome called a **map unit**. A map unit is the distance that corresponds to a recombination frequency of 1%. Thus, if recombination occurs between two loci 12% of the time, these two loci are 12 map units apart on the chromosome. This doesn't tell you how may kilobases apart they are, but it does give you an approximate distance to use as a starting point.

You can tell how far apart 3 loci are if you use three loci at a time in your analysis. For example, let's say you know that Statesville, Davidson and Charlotte are all located on the same perfectly straight highway. Statesville is 20 miles from Davidson, and Charlotte is 50 miles from Statesville. If I asked you to draw a map of these cities, you would have two alternatives:

Charlotte --- 50 miles --- Statesville --- 20 miles --- Davidson
or Statesville --- 20 miles --- Davidson --- 30 miles --- Charlotte

In order to choose between these two alternatives, you have to know the distance between Davidson and Charlotte. If it's 70 miles, then the first map is correct. If it's 30 miles, then the second one is correct.

This is exactly how you map genes on a chromosome. You take three points, three loci, and you find out how far apart each of the pairs of loci is by determining the recombination frequency between each pair, and then you map them. Such a map is called a **genetic linkage map** because it relies on the properties of linkage to determine map distances.

Study Questions

1. Describe the methods used to isolate individual chromosomes. Why is this an important component in the process of mapping genes?

2. What is a tetrad? How do the chromatids in a tetrad assort? (That is, how many and which ones go into each egg or sperm cell?)

3. What is recombination? When does it normally occur? What are the genetic consequences of recombination?

4. Linkage analysis is based on the idea that recombination frequency is proportional to the distance between loci. Explain what this means.

5. Given genetic data, be able to construct a genetic linkage map.

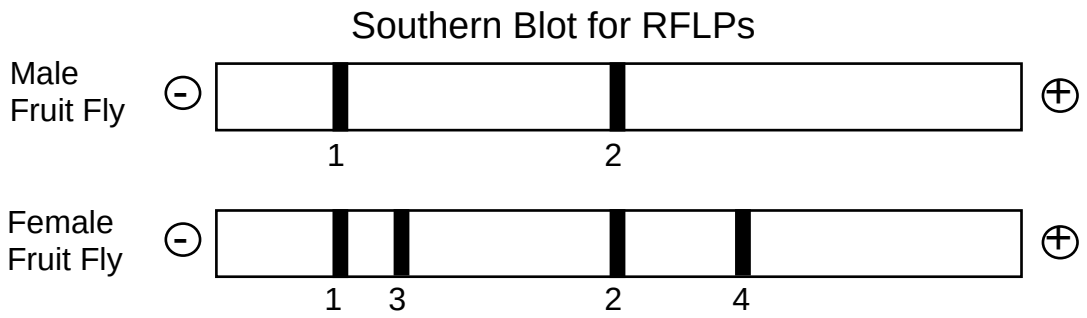
----------STOP----------

Focused reading p192 "Geneticists make maps"
Fig 10.22, 10.23 & 10.24

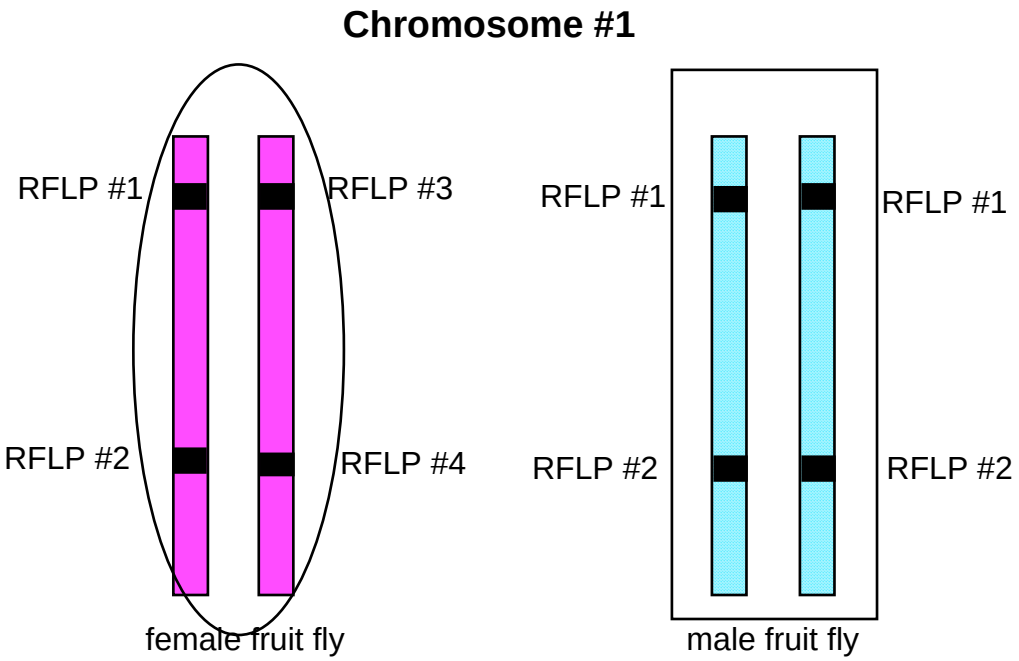
WWWeb reading: For further help mapping genes: http://www.whfreeman/purves6e Click on the "Math for Life" link and read Topic 7.1 Mapping Genes. Contains 'how to' as well as practice problems' (and solutions)

But how do you determine recombination frequency? You have to be able to detect the alleles and follow them as they are inherited. In the example above, this was fairly easy—you can see eye color and height so you can follow the alleles. Even though its a bit more technical, following RFLPs and disease states allows you to determine recombination frequencies and thereby determine map distances.

If we were trying to use RFLPs to develop a linkage map of a chromosome in a organism that produces many, many offspring—say *Drosophila*, it would be relatively easy to do so. In diploid organisms, the simplest way to map chromosomes is to do a **dihybrid test cross** (a heterozygote by a homozygous recessive.) Here is an example of how this would go:



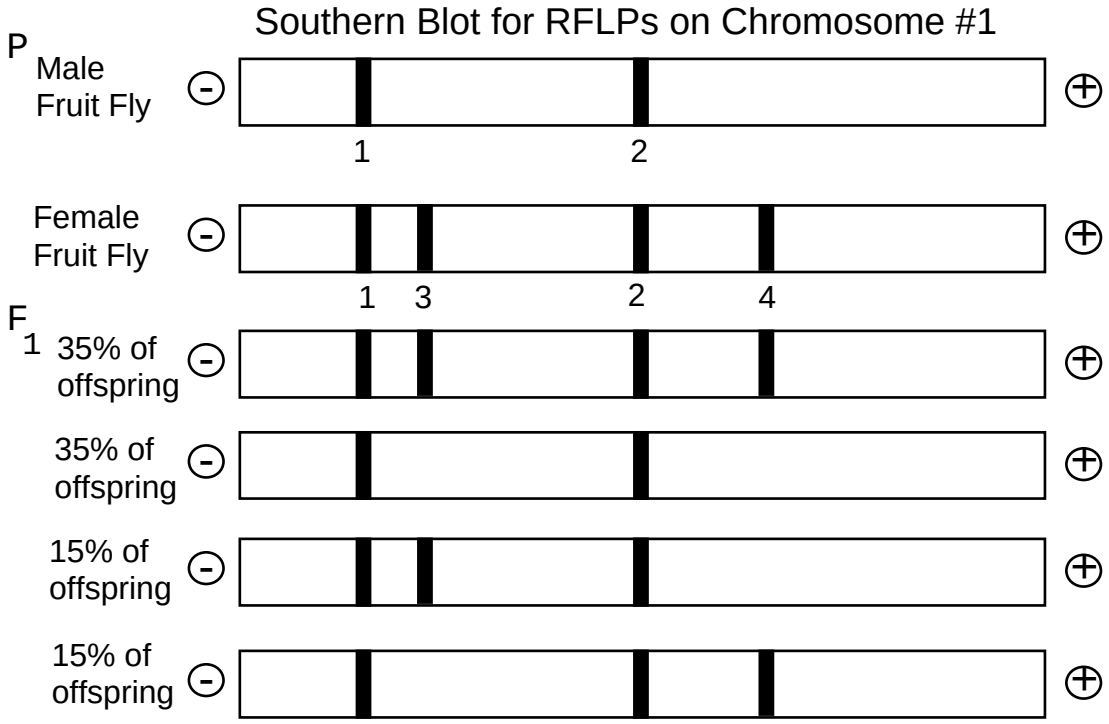
Let's say we are looking at Chromosome #1 of the fruit fly. You obtain chromosome #1, digest it with a known restriction enzyme and probe it with two different radioactive probes, and you get the above Southern blot. The male fruit fly has two RFLPs on chromosome #1 whereas the female has four—she shares two with the male (1 and 2) and has two that she does not share with the male (3 and 4). Thus, the male is homozygous for these two RFLPs and the female is a heterozygote. While we don't know exactly which RFLPs correspond to which loci, chromosome #1 in these flies might look something like this:



As it is drawn, band 1 and 3 on the female fly's Southern blot are alleles of the same locus, and bands 2 and 4 are alleles of a second locus. Thus, this female fly is a heterozygote at both loci. The male has identical alleles at the first and second loci; thus he is a homozygous at both loci. Now when this female fly creates her eggs, she will make four different kinds of chromatids (eggs) from chromosome #1.

Type of Chromatid	Alleles
Parental	RFLP 1 and 2
Parental	RFLP 3 and 4
Recombinant	RFLP 1 and 4
Recombinant	RFLP 3 and 2

Genetic linkage mapping is based on the idea that the frequency with which the recombinant chromatids occur is proportional to the distance between the two loci. Let's say you mate this female and male fly. You do a Southern blot on the offspring and obtain the following data:



Because the male fly is a homozygote, he always passes on RFLP 1 and 2. Thus, all of the F_1 offspring have RFLP 1 and 2. 35% of the offspring also inherited RFLP 3 and 4. Thus, they received a chromatid bearing 1 and 2 from their father and a chromatid bearing 3 and 4 from their mother. This chromatid from their mother is a parental chromatid and thus, these flies are not the products of recombination. Likewise, 35% of the offspring inherited two copies of RFLP 1 and 2. Thus, they received 1 and 2 from their father and 1 and 2 from their mother. Again, they inherited a parental chromatid from their mother and are not the products of recombination.

15% of the offspring inherited RFLP 1 and 2 from their father and RFLP 3 and 2 from their mother. A chromatid bearing RFLP 3 and 2 is a recombinant chromatid, and thus, these offspring are the products of recombination. Likewise, 15% of the offspring inherited RFLP 1 and 2 from their father and 1 and 4 from their mother. A chromatid bearing RFLP 1 and 4 is a recombinant chromatid, and thus these flies are the products of recombination. Thus, in the above example, 30% of the offspring are the products of recombination. Thus, the **recombination frequency** between these two loci on Chromosome #1 is 30%, which represents **30 map units.** If 10% of the offspring had been recombinant forms, then these two loci would be 10 map units apart (3 times closer together than if they were 30 map units apart.)

It is a relatively simple task to produce a linkage map of an organism that has many, many offspring. However, this is much harder in humans. You can't do recombination frequencies in a single family. Rather, you have to look at an entire population and determine recombination frequencies there. Thus, you have to gather many, many samples and run many, many Southern blots.If you remember, in the case of CF, the disease was shown to be linked to two markers on chromosome #7. These markers have names (everything in biology has a name!) -- they are called *MET* and *D7S8*. The lab that characterized the marker chooses the names and they can mean almost anything, so don't try to look for a scheme to these names—there is none. You can think of them as human names. You name your kid Met or D7S8 and that's the name that identifies that individual.

If you do the kind of RFLP analysis outlined above and you look for linkage to the disease at the same time, you can determine the order of *MET, D7S8,* and *the CF gene* on chromosome #7. Consider the following simplified and hypothetical data

Loci Analyzed	Recombination Frequencies
MET and *D7S8*	10%
MET and cystic fibrosis	4%
D7S8 and cystic fibrosis	6%

Just as in the example of the three Cities above, you now can determine the order of these alleles on chromosome #7. The only map that works for all the data is:

------*MET* ----4 map units ---CF----------6 map units---------*D7S8*-----

Thus, investigators were able to determine that *MET* and *D7S8* flanked the cystic fibrosis gene. This is important information because it defines the location of the CF gene on chromosome #7. We now know that the gene is somewhere between *MET* and *D7S8* and both of these markers are identifiable by the presence of restriction target sites and marker sequences. Investigators continued to look for RFLPs in this region of chromosome #7 using many different restriction enzymes and many different markers. They found 2 more RFLPs that mapped between *MET* and *D7S8*. Recombination frequency analysis and Southern blots using pulse-field electrophoresis determined the order of and distance between these RFLPs to be:

MET ---- 500 kb ----------CF---------- 980 kb ----------*D7S8*

The total distance between *MET* and *D7S8* was determined to be 1480 kb or about 1.5 **Mb** (million base pairs.)

Study Questions

1. Understand how RFLPs can be used to locate genes. Be able to interpret a Southern blot to determine which RFLPs are linked to a disease gene.

2. Increasingly, RFLP analysis is being used to diagnose the presence of carrier status or the genetic status of fetuses by amniocentesis. If you were a genetic counselor, how would you explain this process to someone who wanted to understand how her disease status would be determined?

3. How can investigators determine which RFLPs are on which chromosomes? How are individual chromosomes obtained?

4. If you were the technician performing the diagnostic test to determine if someone were a carrier of CF, what controls would you run? Whose DNA would you sample?

5. What is a map unit?

6. Explain how linkage maps are created.

7. Be able to map a DNA segment given the outcome of dihybrid testcrosses.

8. Be able to map a DNA segment given the outcome of a Southern blot analysis of RFLPs resulting from a dihybrid cross.

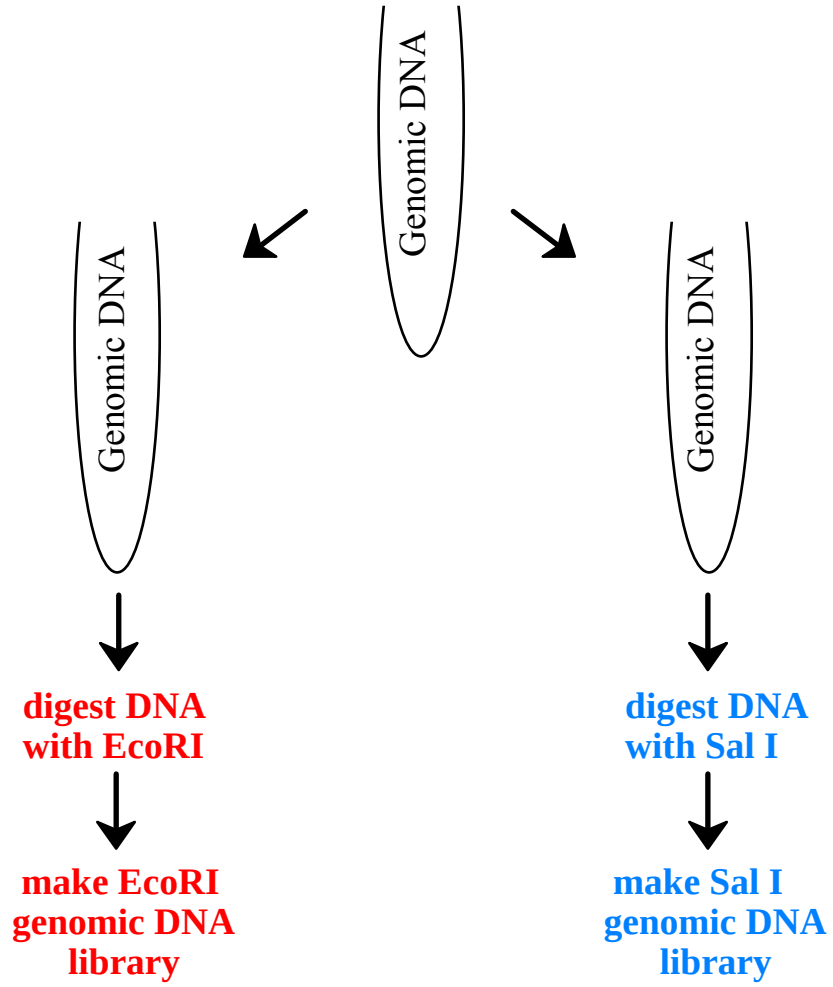
The search for the CF gene had been dramatically narrowed by linkage analysis of the RFLPs on chromosome #7. Investigators knew that the CF gene was somewhere within a defined 1.5 Mb segment. So, what now? Linkage analysis won't help you any more because the distances between loci in this region are so small that recombinant doesn't occur often enough to be detected. So, investigators had to turn to a different technique called **positional cloning** or **chromosome walking.**

Focused Reading: p 337-8 "DNA markers…" stop at "Human gene mutations…"
p 394-5 paragraph beginning "In physical mapping of the chromosome…" stop at "The genomes of several organisms…" (note HGP stands for Human Genome Project—p393)
P 348-50 "Sequencing of the…" stop at " The human genome…"

Let's pause for a moment and look at the theory behind DNA marker sequences a bit more closely. Ideally, a DNA marker sequence would appear only <u>once</u> in the entire genome. Ideally, a given probe should be able to identify one and only one inherited marker sequence—this inherited sequence would then be unique in the genome—like the gene for insulin or the gene for cytochrome c.

How long does a probe have to be to meet this criterion? Well, if there are 6×10^9 base pairs in diploid human genome, a base sequence should be long enough to have a probability of existing at the frequency of 1 in 6 billion. How long is that? Well, what are the chances that a given base sequence starts with "A"? The answer: 1 chance in 4 since there are four bases (we'll assume each is equally probable, although that does vary a bit in different species.) If "A" is the first letter of our sequence, what are the chances that the next letter is "C"? Again, 1 in 4. But the chances of having a base sequence "A" followed by "C" is the product of the probabilities of each letter, 1/4 x 1/4 or 1/16. Well, with 6 billion base pairs, if the chances of "AC" occurring are 1 in 16, you are going to have millions of "AC" combinations in the genome (6 billion ÷ 16).But, let's keep going. What are the chances of having the base sequence "ACC"? $(1/4)^3$ or 1/4 x 1/4 x 1/4 = 1 in 64. The real question we want to ask is coming into focus. To what power do you have to raise 1/4 to get a chance of around 1 in 6 billion? The answer is 14; that is $(1/4)^{14}$ = about 3×10^{-9}. So if you had a marker sequence 14 bases long, the chances are that it is one of a kind in the genome. However, due to practical considerations, like the effects of temperature and salt concentrations on hybridization of complementary sequences, probes are usually in the 20 to 40 base range.

What we want to do is clone the CF gene that is hidden somewhere in a 1.5 Mb piece of DNA which is flanked by two RFLPs (MET and D7S8). We have probes for the RFLPs which are located at the two ends of the DNA, but we do not have any probes for CF so we will have to walk from one end to the other and look for CF as we walk. This process is called **chromosomal walking** and it allows us to clone an uncloned gene that is hiding near an RFLP. The term "walking" is a nice metaphor since we must use two feet in order to walk. Likewise, we must use two separate genomic DNA libraries; each made from the same genomic DNA but using different restriction enzymes. In our example, we will use the enzymes EcoRI and Sal I:



While the procedures of chromosome walking (and jumping) are a bit complicated, they are based on a fairly simple idea. Let's say you have the segment of DNA that contains the CF gene—the 1.5 Mb segment of chromosome #7. First, you pull a restriction enzyme out of the freezer. Let's say EcoRI cuts your 1.5 Mb DNA segment into 6 restriction fragments that look like this after gel electrophoresis:

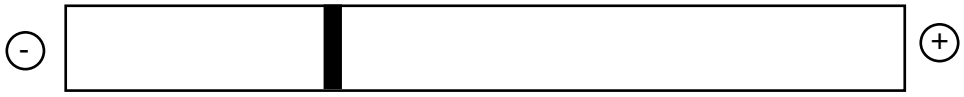
DNA fragments on a gel

Above, you see the DNA bands in the gel, not a Southern blot. From this experiment you learn that EcoRI can cut the 1.5 Mb portion of the chromosome into six fragments. Remember that DNA

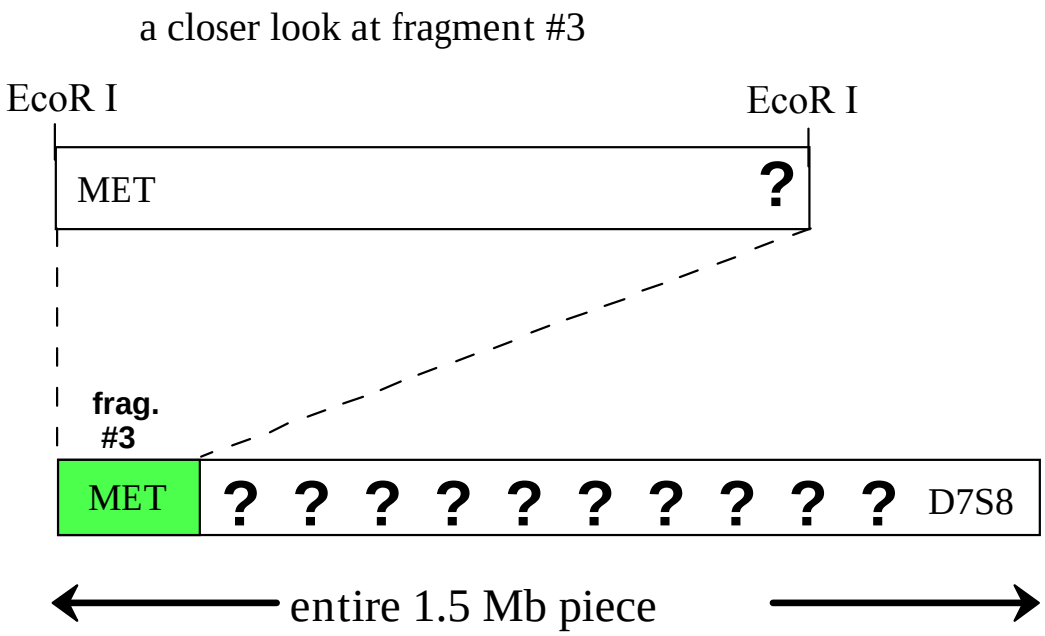
separates on the basis of size and the arrangement of band on the gel does not have anything to do with the linear order of the fragments in a chromosome. To 'walk' between two chromosomal markers you need to determine the correct linear order of these fragments. Fortunately you have probes—those RFLPs MET and D7S8-- and you decide to begin looking for the MET RFLP. Because MET is a specific piece of DNA one (and only one) of these fragments should contain the marker for MET. You perform a Southern blot by transferring these restriction fragments onto nitrocellulose and probing with the MET probe sequence (for the sake of simplicity this probe is CCCCCCCCCCCCCC, thus it would recognize the marker sequence GGGGGGGGGGGGGG in the DNA).

Your original gel with all of the restriction fragments looked like this:

And your Southern blot using the MET probe looks like this:

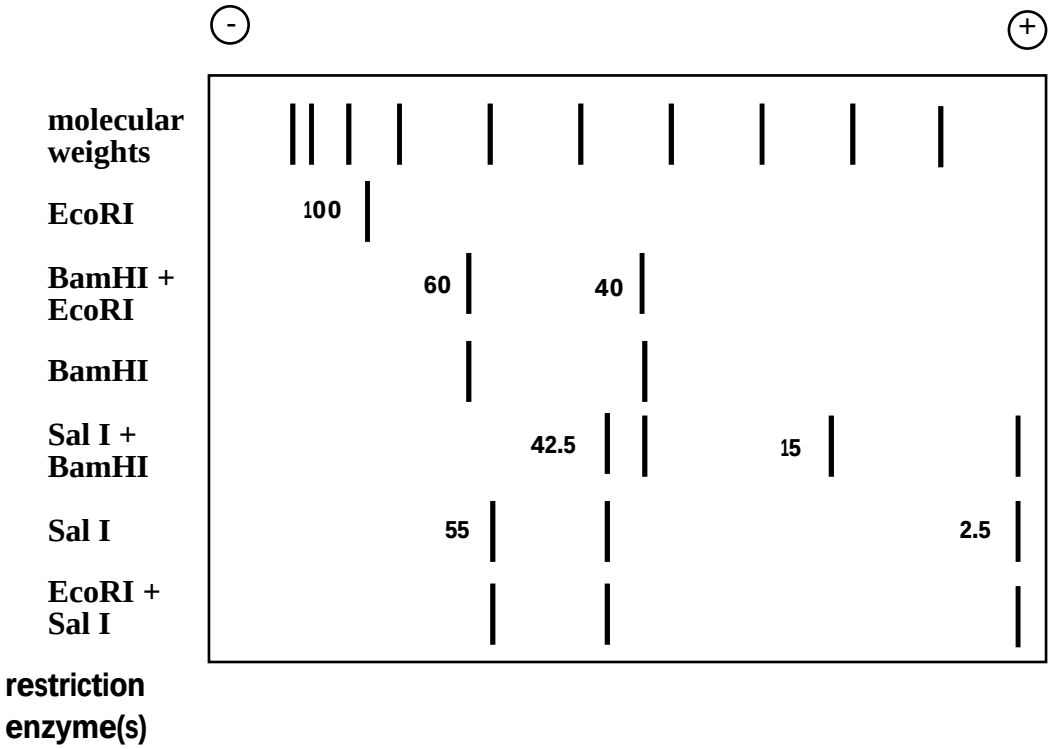


This result indicates that the third band from the top of the gel contains the MET marker (the 'top' of the gel is the end you initially add the DNA to and is always labeled "-" because it is closest to the anode in the electrophoresis unit). Now you know that one end (the MET end) of the 1.5 Mb fragment is in EcoRI fragment #3. You will need to clone this EcoRI restriction fragment from your EcoRI genomic DNA library (see p 359-60 and Fig 16.9). Once you have isolated fragment #3 from the library, you are ready to take your first step in chromosomal walking. The goal is to walk down the chromosome from MET to D7S8. In order to walk, you need to know which direction to take your next step; you want to walk towards D7S8 and not away from it. Therefore, you need to produce a restriction map of fragment #3 so you can isolate a piece of DNA from fragment #3 that is on the opposite end of fragment #3 from MET. We want to isolate the portion of fragment #3 that is labeled with a "?" in the diagram below (the location of fragment #3 has been put onto the 1.5 Mb DNA so you can see that we have identified only 1 of the 6 EcoRI fragments).

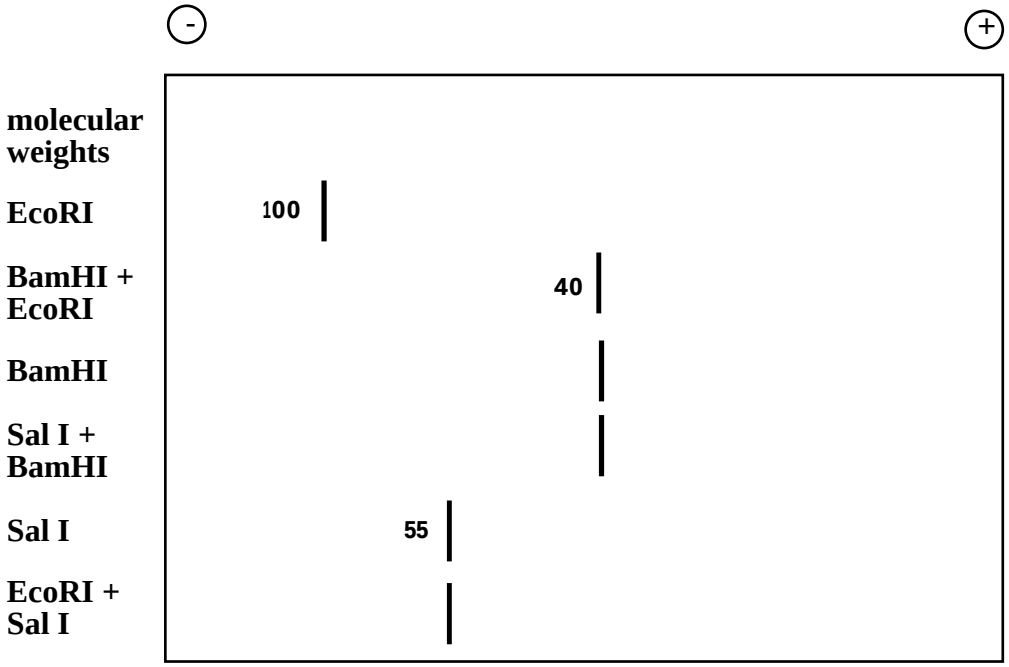


Further characterization of fragment #3 will require you to use several restriction enzymes to digest <u>just</u> fragment #3. You digest fragment #3 with each of 3 enzymes, alone and in combinations, and run them all on one gel. The enzymes you choose are EcoRI, SalI (the enzyme you made your other genomic library from) and BamHI (another enzyme to help with ordering). This experiment will help in two ways: it lets you to place the EcoRI fragments in order and it allows you to identify the SalI fragment that is furthest away from the MET marker so that you can continue 'walking'. The DNA banding pattern from your restriction digests might look like this (with molecular weights indicated to facilitate mapping):

Agarose gel of DNA fragments that result from digesting Fragment #3 with the enzymes indicated



If this gel were blotted and probed with MET, the resulting autoradiograph might look like this:



"DNA FRAGMENTS" TO USE WHEN PRACTICING RESTRICTION MAPPING:

The boxes below were drawn to scale using the information found in the agarose gel seen on page 93. The shading of each box indicates the restriction enzyme used to obtain the fragment and only the single restriction enzyme digest information is included. Cut out the strips and use the information from the agarose gel (single and double restriction enzyme digests) and Southern blot to determine the order of the restriction sites in Fragment #3.

Fragment #3 (100kb)

EcoRI (100kb)

BamHI (60kb)

BamHI (40kb)

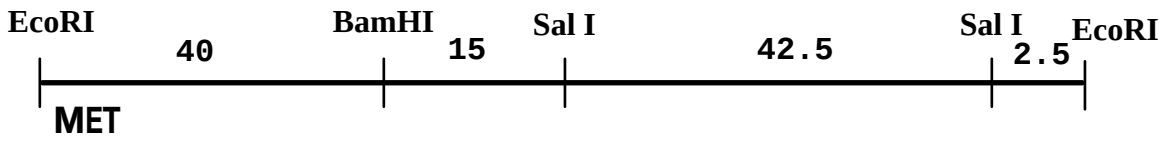
Sal1 (55kb)

Sal1 (42.5kb)

Sal I (2.5kb)

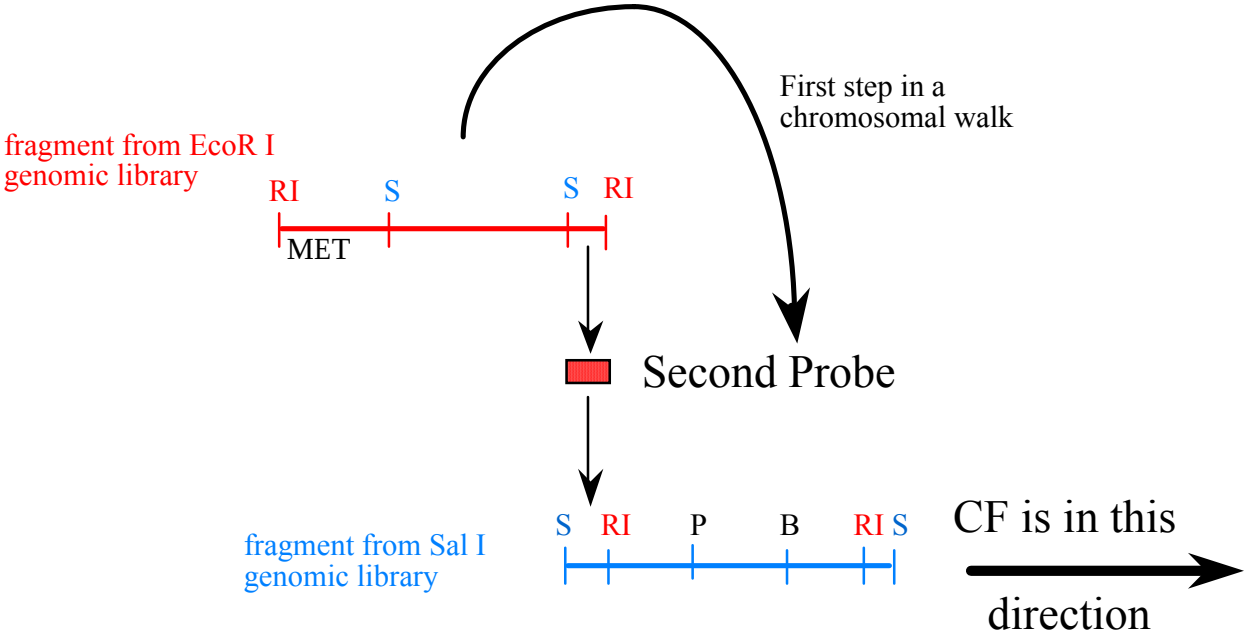
Again each of the bands seen on the autoradiograph indicates the DNA fragment that contains the 14 base pair long probe sequence (and ONLY band that contains the probe sequence). When you look at these data, you can start to figure out the **restriction map** for fragment #3. A restriction map is like a road map with several cities (i.e. restriction sites) in a row connected by different lengths of highway (DNA). This type of map is called a restriction map because it locates the restriction sites relative to each other (similar to a linkage map). The restriction map for fragment #3 might look like this:

(note this is not drawn to scale)

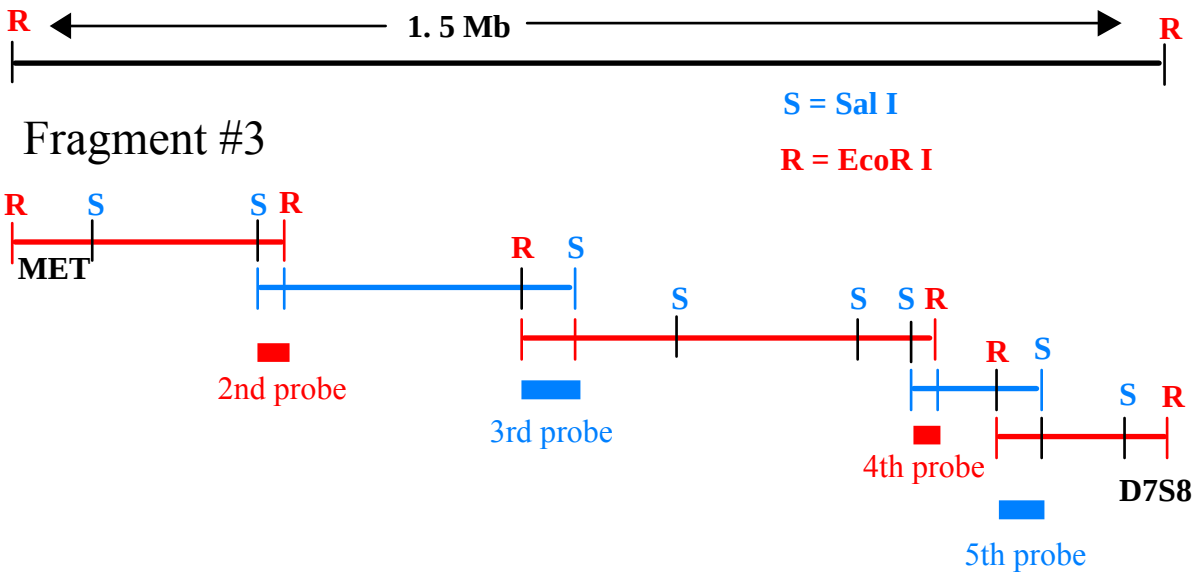


See if you can reconstruct the restriction map using the information found on the agarose gel of digested DNA and the Southern blot. The next page contains a set of "DNA fragments" (boxes drawn on the page). Cut out the fragments and line them up on the template for Fragment #3. Remember that the gel lanes containing DNA digested with two enzymes will give clues to help you determine how the fragments overlap. (This is actually study question #2 below) Can you determine the fragment order definitively?

The next task is to isolate the 2.5 kb Sal I - EcoRI fragment and use this 2.5 kb fragments as your second probe the Sal I library to isolate the piece of DNA along your walk. The probe 2.5 kb piece will be your second probe because it is the DNA fragment furthest from the MET marker and therefore must be closer to CF. When you isolate a piece of genomic DNA from the Sal I library, it will overlap some with the EcoRI fragment #3 but then extend further to the right because probe #2 is bounded on the left by a Sal I site. You know any new fragment that has Sal I sites on both ends and binds to the second probe will extend towards the right (in the direction of CF and D7S8):



When you find which Sal I fragment binds to probe #2, you figure out its restriction map the same way we did for EcoRI fragment #3. This process continues until you reach the other end of the 1.5 Mb fragment (which is defined by the D7S8 marker). The final product is a series of overlapping fragments covering the entire 1.5 Mb piece of DNA and each of these fragments has been restriction mapped. (You may begin to realize that this is a huge project taking lots of time, people, and money.) The final restriction map, and the overlapping fragments, might look like this:



Using this technique, you can "walk" down the chromosome, identifying unique marker sites and restriction target sites as you go, and establishing distances between the restriction sites. If you keep at this until you reach the D7S8 marker at the far end of the DNA segment, you will have produced a complete **restriction map** of the 1.5 Mb DNA segment that contains the CF gene. A good restriction map will identify a marker sequence or a restriction site every 5 to 20 kb all along the entire 1.5 Mb segment.

Study Questions

1. Explain in general terms how chromosome walking is done.

2. Given the kind of data presented on page 93-94 above, be able to construct a simple restriction map. (Cut out the paper DNA fragments on page 95 and use them fragments to help you.)

----------STOP----------

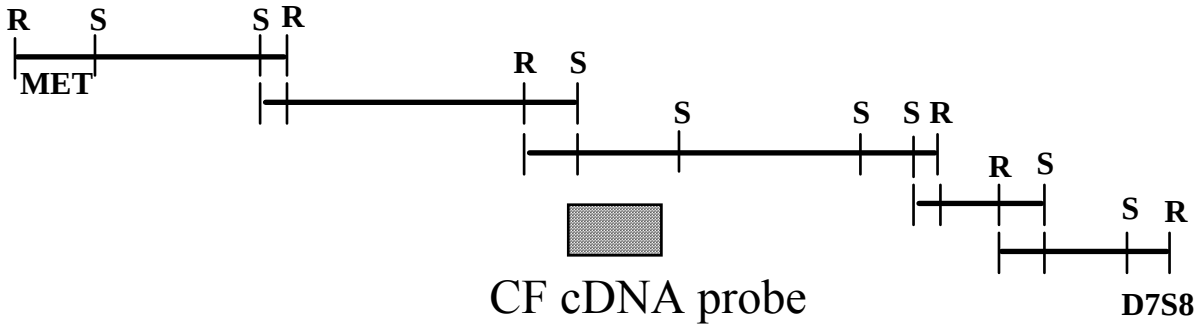
Focused Reading: p 355 "Genes can be cloned…" to bottom of page
p 360 "A DNA copy of mRNA can be made" (2 paragraphs)
p 361 Fig 16.10

We now have a restriction map of the segment of DNA containing the CF gene. How does this allow us to isolate the CF gene? We know that all cells contain all genes how has the restriction map helped us narrow our search? Well, all cells contain all genes but each cell; type (liver, retina, and muscle) uses only specific parts of the genome. If we look at cells that use our gene we can narrow our search. In this case we go to the cells that actually make the wild-type version of the CF protein and isolate mRNA from these cells. Because we know that CF patients have problems in their lungs, pancreas and sweat glands, these cells are a good place to start. Investigators took these cells from wild-type individuals and isolated the mRNA from these cells. If these wild-type cells make the wild-type version of the CF protein, they must contain mRNA for this protein (they must

'use' the gene). After isolating the mRNA from these cells, investigators made radioactive probes that were complementary to this mRNA by incubating the mRNA with radioactive nucleotides and an enzyme called **reverse transcriptase**. Reverse transcriptase, as the name implies, does transcription in reverse. It uses RNA as a template to create a complementary strand of DNA (**cDNA**), so reverse transcriptase is a kind of DNA polymerase too. We will talk more about this later when we discuss HIV.

Because mRNA has already had its introns spliced out, it contains only exons of the functional gene. When cDNA is made from lung tissue, every mRNA made in lung tissue (e.g. cAMP dependent kinase, actin, elastin, spectrin, Na/K pumps, ion channels, calcium pumps, and the Cl^- channel) is converted into cDNA. Conversely mRNAs that are specific to other tissues (like opsin in the eye) will not be included. The cDNA present will be a mixed bag of probes that should contain some CF cDNA. It is unlikely that our 1.5 Mb fragment contains more than one gene since 98% of our chromosomes do not contain any genes. Therefore since we know that the 1.5 Mb chromosomal fragment has the CF gene in it (by RFLP mapping), a cDNA probe from lung that hybridizes with part of our 1.5 Mb it will identify the piece of DNA that contains the CF gene. Since it is technically possible that there are two genes in this 1.5 Mb area of the DNA from chromosome #7 (one is the CF gene and the other some other lung expressed gene) we will still have to prove that we have the right gene. We'll get to that in a paragraph or two.

First let's go back to the mix of lung cDNA probes. The mixture of all the cDNAs from the tissue (e.g. lungs) is made radioactive and this "hot" cDNA is used as a probe to see if any of the cDNA binds to the 1.5 Mb area of DNA. After performing a lot of Southern blots you might deduce that the cDNA binds to the DNA in the area of the checkered box as indicated below:



This is a simplified and hypothetical example. The real cDNA probe used to find the CF gene was synthesized from sweat glands, since those are easier to obtain than lungs, and was 6129 bases long. The restriction map of the 1.5 Mb segment on chromosome #7 was much more complicated than this diagram. Nevertheless, the underlying approach is exactly the same.

Study Questions

1. What is cDNA and how is it made?

2. How was cDNA used to actually pinpoint the location of a gene on a restriction map?

3. Given the kind of data presented above, be able to pinpoint the location of a gene on a restriction map.

4. Sometimes investigators probe DNA segments with cDNA and identify the wrong gene. Why does this mistake occur? Why isn't this system of gene identification foolproof?

Using this cDNA probe method, investigators determined that the gene encoding the normal allele at the CF locus is 250 kb long (huge!) and contains 27 exons. After mRNA processing, the 250,000 bases in the gene are reduced to 6129 bases in the final mRNA. (This means that 243,871 bases in the gene are in introns.) This mRNA is translated into a protein that is 1480 amino acids long with a molecular weight of 168,138 daltons (168 **kilodaltons** or **kD**). Having pinpointed the gene and knowing which restriction target sites flank the gene, this very specific piece of DNA can be isolated, **cloned** (or copied) and **sequenced**.

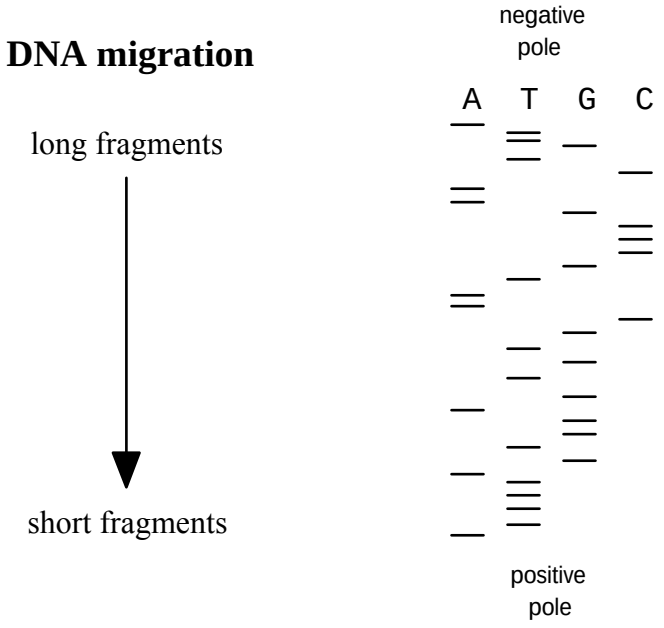
Review and Focused Reading: p 2.6-216 "DNA replication" to end of Chapter
p 215 Fig 11.20

Study Questions:

1. Describe the natural process of DNA replication. What proteins are involved in the process? What role does the primer play in this process? What is the primer made of?
2. Why is DNA replication called "semi-conservative?" What is conservative about it? What is "semi" about it?

3. Explain the process of DNA sequencing. Why are dideoxynucleotides used in this process?

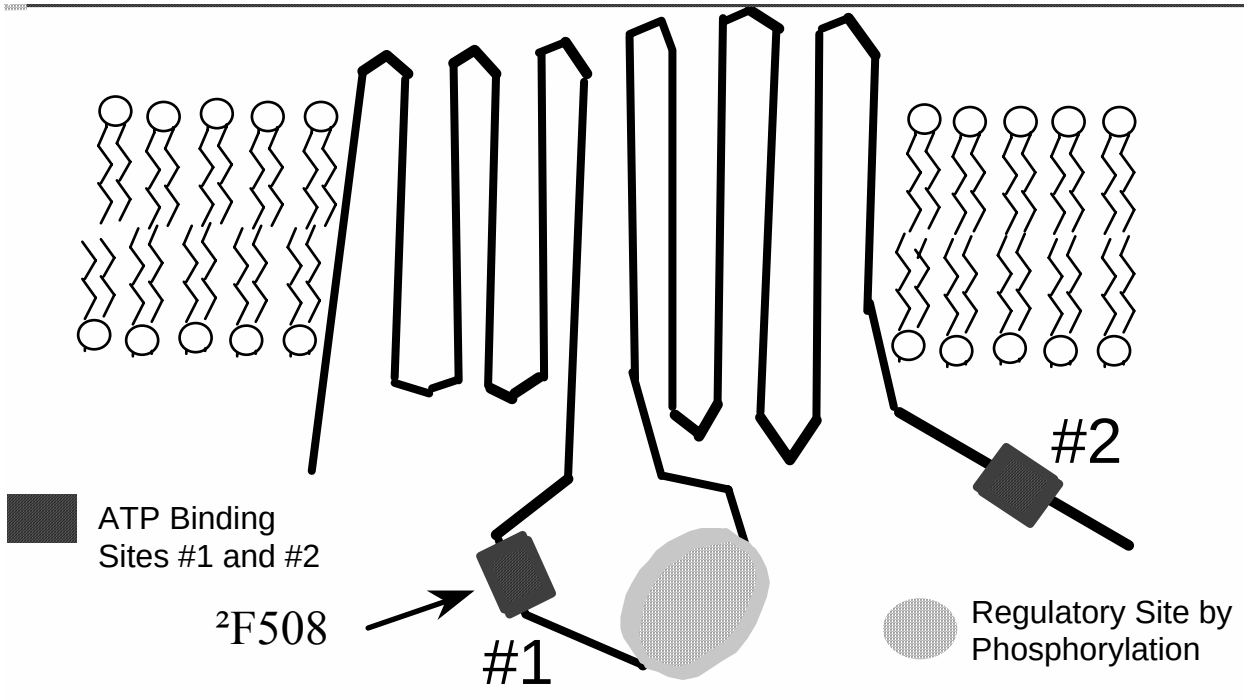
4. Be able to interpret a Sanger sequencing gel to give the correct base sequence of a DNA segment (with the correct 5' to 3' orientation.)



The sequence of bases in the 27 exons of the gene at the CF locus was determined by DNA sequencing. Once the base sequence of these exons was identified, the amino acid sequence of the wild-type protein was deduced using the genetic code (on page 224) Investigators noted that, in this protein, long stretches of hydrophobic amino acids alternated with long stretches of hydrophilic amino acids. This pattern of amino acid distribution is consistent with an integral membrane protein. Also, the amino acid sequence of this protein had a pattern that was similar to several ion channels whose encoding DNA had been sequenced already (i.e. there was some **homology**). Now, investigators performed the crucial test—they needed to establish that some of the DNA bases in this gene are different in CF patients than they are in wild-type individuals. Remember there is still a slim possibility that this gene could actually encode some other protein made by sweat gland cells, investigators had to establish that this gene is altered in CF patients to support their hypothesis that this gene product is involved in causing cystic fibrosis. They used the wildtype cDNA as a probe to isolate cDNAs from CF patients and they sequenced these cDNAs. After comparing the DNA sequence from the wildtype gene to the sequences of the same gene in people having CF they found that in 70% of CF patients one codon was deleted from an exon in this gene. The missing codon encoded amino acid #508, which is a phenylalanine in the wild-type gene. The shorthand abbreviation for phenylalanine is "F". Thus, this mutation is called ΔF508 -- a deletion (Δ) of phenylalanine (F) at position 508.

So, it appears as though investigators have found the gene that causes CF, at least in 70% of the cases. Unfortunately, the remaining 30% of cases are caused by over 600 different mutations in the CF gene—a very difficult basis for finding a common cure. Approximately 4% of CF alleles contain nonsense mutations at different codons.

The next step in the process was to try to figure out what this protein does and how the ΔF508 mutation keeps it from doing its job. Computer assisted analysis can produce a likely three-dimensional structure, or **topology**, of a protein from its amino acid sequence by predicting common protein folding patterns, or **motifs**, based upon what is known about homologous proteins. For instance, given the position of polar and non-polar R groups, we can predict which domains probably form an **alpha helix** (like a corkscrew) or a **ß pleated sheet** (like corrugated cardboard), or if this protein is embedded in the membrane. Computer assisted prediction of protein conformation is a rapidly growing field but predictions for large proteins are still fairly crude. Nevertheless, the analysis of the wild-type version of the CF protein clearly predicted that it was a 12-pass integral membrane protein. A 12-pass protein zigzags back and forth through the membrane 12 times like so:



The nucleotide sequence in two cytoplasmic areas are predicted to be ATP-binding sites and sites needed for regulation of the protein by ATP binding and hydrolysis. This structure, with sites for ATP binding, is typical of ion pumps and ion channels and is consistent with the hypothesis that this gene encodes a Cl^- ion channel. The regulatory domain ® can be phosphorylated by a cAMP-dependent protein kinase (sound familiar?). When R gets phosphorylated, then the gate is opened to allow Cl^- ions to move out of the cells.

The early evidence that lung cells from CF patients cannot export Cl^- when cAMP levels rise correlated very well with the protein structural information acquired through molecular, or DNA, methods. When mutated, this integral membrane protein causes CF, therefore it was given the name **CFTR—Cystic Fibrosis Transmembrane Conductance Regulator**. (The "conductance" being referred to here is chloride ion conductance.) This is a fairly vague name, but good scientists hate to jump to conclusions with preliminary evidence. No one wants to be the person who named this protein the cystic fibrosis ATP-dependent chloride ion pump only to find out a few years from now that it isn't a chloride ion pump at all. When something appears in print for all eternity, better cautious than wrong.

At this point, we need to figure out why a chloride ion channel would make the mucus in lungs more viscous, and all the other problems associated with CF. In order to understand this, we need to understand **osmosis**.

Focused Reading: p 86-87 "Osmosis is passive…" stop at "Diffusion may be…"
p 87 Fig 5.8

Unlike sodium or calcium, water is not a leader but a follower -- a lamb in a world of Marys. Think of ions as Mary (as in Mary had a little lamb). Wherever the ions go, the water is sure to follow. All cells have to control the amount of water in their cytoplasm in order to survive. This is

most obvious in plants that do not get enough water and begin to wilt. Cells have to move water to maintain their cell volume and internal pressure, but they cannot actually bind water and move it. Likewise, animal cells and their secretions need to have a balance of water and salt. So they rely on the process of osmosis to move water. If chloride ions cannot leave the cell and enter the mucus, the mucus does not have enough ionic strength to pull more water out of the cells, and the mucus is left as a sticky paste.

NEWS ITEM: Having too much water in mucus causes as much trouble as having too little. A rare genetic disorder called pseudohypoaldosteronism I (PHA) causes fluid buildup in the airways of the lungs. The fluid causes wheezing and infection but, fortunately the condition is usually outgrown with time. The cause? A defective epithelial sodium channel that can't pump sodium out of the cell. Using what you know about osmosis, why would this result in fluid in the airways? Why might these people be able to 'outgrow' their problem? (The first question you should be able to answer, the second requires speculation) Reviewed in Dorrell, S. 1999 *Molec Med Today* vol 5 p 462.

Study Questions:

1. Explain the process of osmosis. What is producing the force that moves water during osmosis? In what way is the process of osmosis an example of the concept expressed by the 2nd law of thermodynamics?

2. While the movement of water across cell membranes cannot be directly controlled, it can be indirectly controlled. Explain how the transport of water is controlled. Explain how this process may ultimately rely on ATP as a source of energy.

3. What is osmotic pressure? What makes a solution hypotonic? Hypertonic? Isotonic? Understand the direction of movement of water under different conditions of osmotic pressure (See Fig 5.8 p 87).

Now, back to our understanding of CF. Where does the ΔF508 mutation appear in the CFTR? It is in the first ATP-binding site. Ah ha!! Good place for a mutation that seriously impairs protein function. One hypothesis would be that maybe this protein can't bind ATP and therefore can't get any energy to move Cl^-. Cl^- cannot move from the cells into the airways of the lungs and pancreatic ducts. The water, which would have normally followed the Cl- by osmotic pressure, does not enter the mucus so the mucus becomes thick. You get cirrhosis because some other product (bile?) requires this dilution effect as well and, when it doesn't happen, this dry product clogs the liver ducts causing cirrhosis. And finally, the sweat glands cannot move Cl^- into the sweat, water does not follow, and therefore the sweat remains highly concentrated with Na ions.Simple, right? Well, a cardinal rule in science is this: An explanation can make perfect sense, be flawless in its logic, and be dead wrong. So, let's not jump to any conclusions prematurely—this is only one hypothesis. We need to see if experimental evidence about the role of the CFTR in cells supports this hypothesis or if another hypothesis is more plausible.

Study Questions:

1. Draw the hypothetical structure of the CFTR protein and explain each of the significant features of the protein. From what experimental evidence and methods is this structure derived?

2. In what portion of the CFTR protein is the ΔF508 mutation located? Given the location of this mutation, describe the most straightforward hypothesis explaining the failure of this protein to successfully move Cl^-.

WWW Reading: *in situ* methodology

You could hypothesize that the protein is in the membrane, but cannot function properly because it cannot bind ATP or because it cannot cleave ATP to ADP or because it cannot be phosphorylated by cAMP-dependent protein kinase. Studies on the normal version of CFTR protein show that phosphorylation by protein kinase A is also a requirement for Cl^- movement. Thus, the mutation may make this phosphorylation event impossible.

These questions can be approached in several ways. For instance, you could hypothesize that the mutation in the CFTR gene keeps it from being transcribed into mRNA. To approach this question, you would perform **in situ hybridization** on the usual tissues from a CF patient. If you did not find mRNA for CFTR, you could conclude that the mutation caused a problem in the creation or stability of mRNA. Alternatively, if you found normal levels of CFTR mRNA in CF patients, you could hypothesize that the mutation keeps the protein from being translated or properly targeted within the cell. You could use immunohistochemistry to look for the protein on the cells of CF patients. The absence of the CFTR protein would mean a defect in translation or post-translational processing or transport.

Investigators looked for CFTR mRNA with the procedure called **in situ hybridization.** *In situ* means in the normal location (in this case in the intact cell), and, as with all DNA probes, the probe hybridizes to its complementary sequence. In the case of *in situ hybridization*, the target is mRNA within the cell's cytoplasm. For these studies, they took radioactive CFTR cDNA and used it as a probe directly in lung tissue. All cells containing mRNA for CFTR will become radioactive when the cDNA hybridizes to the mRNA. Cells not expressing this mRNA will not become radioactive (because the probe had nothing to bind with). These studies showed high expression of the mRNA in pancreas, sweat glands, salivary glands, intestine and reproductive tract and lower expression in respiratory tissue. So, this study demonstrated that CFTR mRNA exists everywhere there are clinical symptoms.

Does this support our hypothesis above that CFTR is the CF protein? Well, it is certainly accepted by the scientific community. However, you will note from this discussion that you can never be absolutely sure you are right. "Proof" in science is based on evidence—sometimes solid, sometimes shaky—but only evidence. No one ever comes along to say, "You've solved it! You're right!" The best that happens is that you and other scientists base many, many experiments on your theory and it always holds up. That's as close as we come to having scientific "proof".

So, even though the mutated version of the CFTR protein is pretty much accepted as cause of CF, much controversy still remains about what wild-type CFTR actually does and how the mutation keeps it from doing its job (the localization studies did not address that part of the hypothesis).

Several approaches can be taken in order to try to determine the function of a protein once its gene has been identified and isolated. If you remember, wild-type respiratory cells will pump Cl^- to the outside when intracellular cAMP levels rise. Respiratory cells from CF patients cannot do this. One standard approach, then, is to **transfect** respiratory cells from CF patients with the dominant wildtype CFTR gene (isolated from a wild-type individual). In this process, the functional gene is transferred into the CF cell to see if this gene can restore the wild-type condition.

There are several ways to do this. You first need to connect the cDNA that contains the CFTR to an appropriate promoter. This promoter need not be the CFTR promoter; rather it could be a promoter for a gene that is turned on by some easily controlled environmental event. For instance, the protein hormone insulin is produced when blood glucose levels are high (insulin lowers the blood glucose levels.) Therefore, the insulin promoter promotes gene expression in response to high glucose concentrations in the fluid bathing the cell. If you put the insulin promoter upstream from the CFTR gene, this gene will be expressed in response to high blood glucose levels. Figure 17.13 (p323) contains a diagram of an **expression vector**—a plasmid that allows you to express a foreign gene.

The CFTR cDNA with its artificial promoter is incubated with CF respiratory cells in tissue culture. Under certain conditions, the cells will take up DNA and begin to express this foreign gene as if it were their own. The transferred gene is called a **transgene** and the cell containing the transgene is called a **transgenic cell.** The process by which transgenes are put into eukaryotic cells is called **transfection**.

When you transfect CF respiratory cells with the CFTR transgene, these cells are restored to wild-type function (i.e. when intracellular cAMP levels rises, they move Cl^- across their plasma membranes at normal rates). This is pretty good evidence that this gene encodes a CF protein that moves Cl^- in response to a cAMP signal. Cl^- movement requires ATP because ATP is a ligand and CFTR is a ligand-gated ion channel. However, the inability to bind ATP is **NOT** why ΔF508 causes CF.

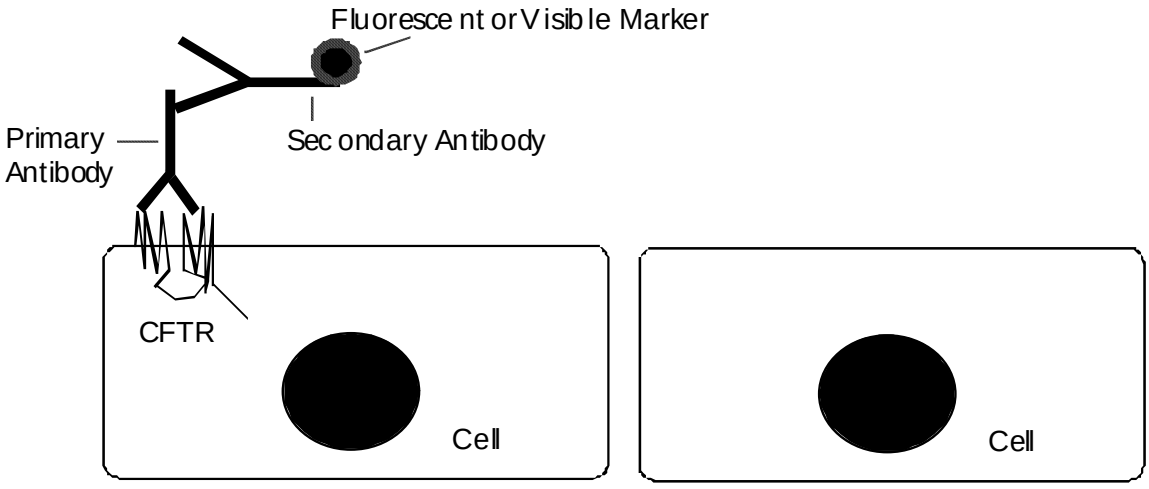
WWW Reading: Immunofluorescence Methodology

Okay. So if our initial hypothesis is not supported, as good scientists, we have to modify it. Another possibility is that the ΔF508 mutation effects expression of the **gene product** (the CFTR protein) in cells that are affected in CF—respiratory, pancreatic, hepatic (liver) and sweat gland cells. One approach to studying the protein localization is called **immunocytochemistry**. In previous approaches we used nucleotide probes (DNA or cDNA) to detect nucleotides (DNA or mRNA). In this approach you need a way to 'see' or detect a protein. To do this you must inject your protein of interest (i.e. CTFR) into an animal (like a mouse, rabbit, rat, goat, etc.). Because it is a human protein, parts of its structure will be foreign to this animal. Immune systems react to any protein shape that is not "self", and the animal will react to this "foreign" shape by producing an **antibody**. Antibodies are proteins with specific binding sites for foreign shapes. These foreign shapes are a kind of ligand called an **antigen**. Thus, antibodies bind antigens like enzymes bind substrates; like receptors bind hormones; like transport proteins bind transported substances; etc (see a pattern here?). This binding is **specific**—just as in the case of all these other proteins, an **anti-CFTR antibody** will bind to CFTR and only CFTR.

To detect CFTR in cells then, you bathe the cells in a solution containing anti-CFTR antibody. The antibody will bind to CFTR wherever it is located in the cell. This antibody is called the **primary antibody** in the immunocytochemistry (see the diagram below).

Now you have tagged the CFTR with the primary antibody and you need to provide a way to 'see' that tag. So, you then apply a **secondary antibody**—one that has been produced 1) to recognize the primary antibody and 2) has been covalently bonded to a fluorescent tag that can be seen under the microscope or by a machine. For instance, if the primary antibody was produced in a mouse, the secondary antibody would made by injecting mouse antibodies into a goat (to make an anti-antibody) and then chemically binding the goat anti-mouse antibody to a fluorescent dye. This secondary antibody is incubated with the cells from above. Every place the antigen (CFTR) exists,

the primary antibody binds and then the secondary antibody binds to the primary antibody making the area colored or fluorescent. Here's a picture:



In this example, the cell on the left bearing the CFTR protein will become fluorescent during this procedure while the cell on the right will not. Thus, you can determine the presence of CFTR and, in some versions of this technique, you can determine the density of the protein in the membrane, and precise subcellular localization.

When investigators used **immunohistochemistry** to look for CFTR in the wild-type tissues, they found the protein expressed in high concentration in the pancreas, sweat glands, salivary glands, intestine, and reproductive tract, and lower levels of expression in the respiratory tract. However, in patients with ΔF508, all of the CFTR was trapped in the ER. Somehow, this mutation causes the CFTR to be inappropriately sorted - it never reaches the plasma membrane and this is the cause of 70% of all CF cases.

Study Questions

1. What is the cause of CF in patients with the ΔF508 mutation?

2. Describe the process of transfection, immunocytochemistry and *in situ* hybridization. How have these approaches been used in CF research?

3. We know a great deal about the CF protein, but much remains to be discovered. If the editor of the prestigious scientific journal *Science* called you and asked what were the three most compelling questions remaining about this protein, what would you tell him? [Note, he would most certainly want you to explain your rationale for these choices.]

4. If CF causes cells to die and release their contents, why would a physician prescribe DNAase to reduce the viscosity of the mucus?

NEWS ITEM: Just because CFTR has its function at the plasma membrane does not mean that it is always located there. Some channels (like the GLUT4 channel involved in glucose uptake) spend most of their 'lives' in vesicles inside the cell and are only placed in the plasma membrane when they are needed (why have a 'hole' in the cell if there is no reason for it!). Since Cl- secretion by CFTR is activated by cAMP, researchers at Dartmouth Medical School examined whether cAMP changes the localization of CFTR or if it simply turns the channel 'on'. To watch CFTR they made a DNA construct that would code for CFTR attached to the green fluorescent protein (GFP). GFP glows so anywhere this CTFR-GFP was found researchers could see it glowing under the microscope. The conclusion: in the cells tested cAMP acted like a switch to open the channel already located in the plasma membrane not like a moving van that got the channel there in the first place. B.D. Moyer et al. 1998 JBC vol 273 21759-68.
----------STOP----------

Focused Reading: p 347-8 "Gene therapy…" stop at "Sequencing…"
p 315-17 "Vectors carry…" stop at "Genetic markers identify…"

We now know that the binding of ATP at site #1 converts the channel from a locked mode to an unlocked mode, but this does not open the channel. ATP binding at site #2 open the channel, but only if there is ATP already bound to site #1, and the R domain is phosphorylated. This may seem complicated but this is a simplified version of a process we don't fully understand. It will get more complex each year. (Aren't you glad you didn't put off Bio111 until next year!)

The hope in all of this, of course, is for a cure to cystic fibrosis. Because it is a genetic disease, it could theoretically be cured if a "good" CFTR gene were delivered to the cells of the CF patient in such a way that it could express a normal protein. Such an approach is called **gene therapy**. Because the most life-threatening symptoms of the disease occur in the respiratory system, such a gene could possibly be delivered in an inhalant aerosol spray. Several DNA delivery systems are being currently investigated including retroviruses, adenoviruses, liposomes and DNA-protein complexes. As we will discuss in Unit IV, viruses function by entering living cells and expressing their genes using the cell's protein manufacturing system. If the disease-causing genes from a virus are removed and a functional CFTR gene added, these viruses could enter the respiratory cells and begin expressing the CFTR gene. Such a "carrier" of a gene is called a **vector**. **Liposomes**, small spheres of phospholipid, are another way to apply gene therapy. By loading a functional CFTR gene onto a liposome and then spraying it into the respiratory tract, it may be taken up by respiratory cells (the cell membrane will fuse with the liposome as in the processes of endocytosis) and may be expressed as a normal gene product.

Now all this sounds really straightforward, but it is a long and journey from an idea to the finished product. We don't know, for instance, if any of these genes will actually be expressed once they are inside the respiratory cells. In addition, Francis Collins has defined a number of other questions that must be addressed before a viable therapy is available (*Science* Vol 256, p 778-779):

1. What are the relevant cells to treat? The respiratory tract is full of all kinds of different cells. Which ones are the best ones to treat in gene therapy?

2. What fraction of the responsible cell types must be corrected to achieve clinical benefit? Certainly one would not have to correct the CF defect in every single cell in the lungs in order to reach an acceptable level of health. How many cells do you have to treat?

3. Is over expression of CFTR toxic? One problem with transgenes is that they do wind up at the CF locus of the person's chromosome number 7 and therefore are not subject to the normal genetic control systems of the promoter that function at the level of the chromosome. Over expression --- unregulated expression—is a constant threat in gene therapy. Would such a thing be toxic to the individual?

4. How long will expression persist? Even if you can get these transgenes to be expressed, will they continue to be expressed indefinitely? Transgenes vary widely in their level of stability. Some function only very briefly, some function for the life of the cell. How will these respiratory transgenes behave?

5. Will the immune system intervene? As we discussed earlier, the immune system will respond to anything that is not "self". If the CFTR protein is <u>not</u> expressed in a particular CF patient, it may be seen as "foreign" by the immune system. Thus, its sudden expression could cause an immune reaction that destroyed the respiratory cells. This process is called **autoimmune disease**.

6. Can safety be insured? This is always a question with bioengineered organisms such as the viral vectors in this approach. Will they "get loose" (especially if it is delivered in an aerosol) and infect everybody, thus transfecting normal individuals with the CFTR gene? And if this happened, would this be dangerous?

Study Questions:

1. Explain the approaches that are currently being tested in gene therapy for CF.

2. What are some technical barriers which must be solved before an effective gene therapy for CF becomes available?

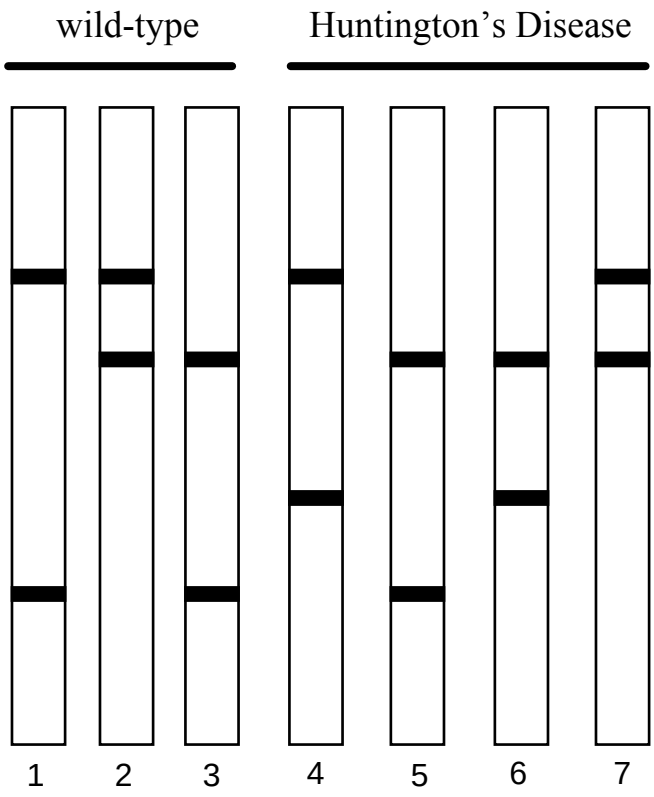
Due to the problems associated with gene therapy, researchers are still looking for conventional means for treating CF. Recent efforts have focused on the salt concentration in the lungs of CF patients. As you should remember from lab (isocitrate dehydrogenase (IDH) experiments), proteins do not work well in high salt environments. When CF and wild-type lung epithelial cells were grown in culture and incubated with the bacteria most commonly found in CF infections, the wild-type cells were able to kill the bacteria while CF cells could not. When salt was added to the wild-type cells, they were no longer able to kill the bacteria and when the salt was reduced for CF cells, the bacteria were killed. This suggested that lung epithelial cells secrete a bactericide that is salt-sensitive. Therefore, researchers began to look for other ion channels located in the plasma membranes of lung epithelia. Their rationale was to increase the secretion of Cl^- ions, which would draw water into the mucus, dilute the salt concentration and allow the lung's naturally produced bactericide to function. This "alternative" Cl^- ion channel has been found. It is a calcium-activated chloride-ion channel which can be stimulated to open when ATP or UTP is administered to the <u>outside</u> of cells. This breakthrough has lead to the first clinical trials in which CF patients have been given aerosolized UTP (UTP had prior FDA approval while ATP did not). Patients treated with UTP are able to clear their lungs better and over time, it is hoped they will have fewer infections. Meanwhile, the search is

on for the bactericide in hopes that this could be given directly to CF patients in addition to UTP treatment. (This work is being headed by Dr. Michael Welsh at the University of Iowa and was published in April 19, 1996 issue of *Cell*.)

NEWS ITEM: Gerald Pier and his colleagues at Harvard and UNC-Chapel Hill have determined that the bacterium *Pseudomonas aeruginosa* (a cause of chronic lung infection in CF patients) binds to CFTR in lung cells. In wild-type cells, the bacteria bind to the CFTR and are internalized by phagocytosis and killed. In patients with ΔF508, the bacteria are not internalized and killed, which can permit the bacteria to live and reproduce in the lungs. Therefore, CF patients are hypersusceptible to infection by *P. aeruginosa.* (See Pier *et al., Science*. Vol. 271: 64-67. 5 January, 1996.)

Some CF patients suffer from thick mucus and also show altered fatty acid levels in their cellular membranes. Juan Alvarez and Steven Freemen (Harvard Med School and Beth Isreal Deaconess Med Center) created transgenic mice that still had mutant CFTR but corrected the lipid biosynthesis problem. Amazingly these mice showed none of the pathology (symptoms) associated with CF! This work may point the way to CF treatments through treating patients with high levels of particular fatty acids. Reviewed by Greener, M. 2000 *Molec med today* vol6 p 47-49.

As the last part of this unit, we will look briefly at the quest for the gene that causes Huntington's disease. While the fine points vary from the CF story, the approach to identifying the gene was essentially the same. However, HD investigators did not have a protein candidate early on, as in the case of CF. In fact, the protein that causes HD is still a complete mystery. Nevertheless, in March of 1993, the HD gene was finally identified and cloned. HD is a dominant trait and less common than CF in the human population. Thus, the odds of finding a person who is homozygous for the disease are very low. As a researcher, why would you want to find a homozygote to include in your genetic analysis? Consider this, if you have only homozygous wild-types and heterozygotes to analyze for RFLP linkage, you might get the following Southern blot:



Individuals 1-3 are wild type and 4-7 are HD sufferers. Unfortunately, there are no obvious bands that correlate 100% with the disease. It would be much easier to identify if one person, with offspring, had two affected HD alleles. You could then identify the band that is passed on to subsequent generations and co-segregates with the disease.

A real breakthrough in HD research came when Dr. Nancy Wexler, a clinical psychologist and daughter of an HD afflicted parent, found an area in Venezuela where the incidence of the disease was very high (maintained through intermarriage within the town.) Wexler recommended that investigators study the inheritance pattern of the disease in this group. Investigators arrived at the town in 1979 to try to find a homozygote for HD. Little did they know, they were about to encounter an enormous extended family of over 10,000 individuals, all with HD or related to someone with HD. This is certainly the richest source of familial genetic information for HD that has ever been assembled. Out of a total population of 12,000 people, 258 had the disease and all were direct decedents of a woman who lived in the 1800's. It is believed that this woman had the misfortune of having a spontaneous ('new') mutation, which was not present in her parents that caused HD in her and her family.

A large consortium called the Huntington's Disease Collaborative Research Group was begun, headed by Dr. James Gusella. After collecting samples in Venezuela and identifying individuals that were very likely to be homozygous for HD (the offspring of two afflicted individuals), consortium investigators in returned to their labs and began looking for RFLPs that were linked to the HD gene. Usually this process takes years, but these people got very lucky and almost immediately (in 1983) found a RFLP that was closely linked to the disease. This RFLP, containing a marker called *G8,* was always present in afflicted individuals and never present in wild-type individuals. In addition, because these investigators had obtained blood from HD homozygotes, they were able to determine

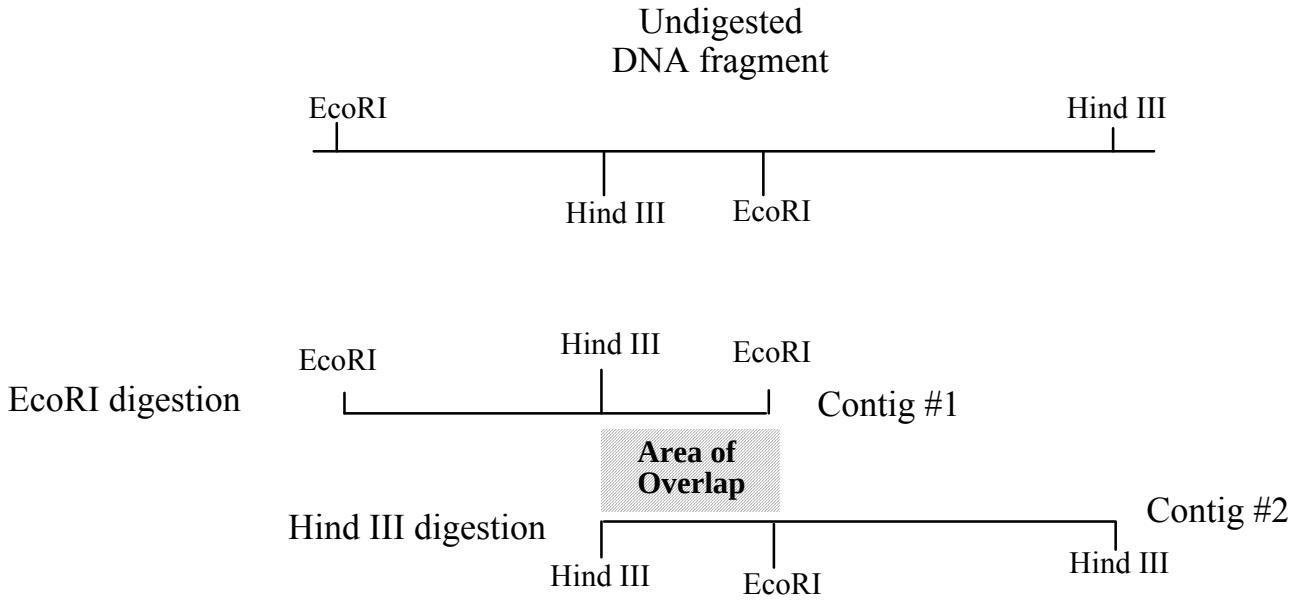
the RFLP fragment that contained the normal equivalent of the HD gene. It appeared that the quest for the HD gene was going to be short and sweet and everyone was very excited.

As was the case in CF (and every other genetic disease), as soon as a reliable RFLP is discovered, the disease can be diagnosed by looking for the normal and disease RFLP in a Southern blot. So very early on, a diagnostic test for Huntington's disease became available. The availability of this test forces people with an HD parent to make an agonizing decision. Should they have the test or not? If one parent had HD, they stand a 50% chance of having the disease themselves. Most of these people have watched the chronic deterioration of body and mind caused by this disease as their parent died or is dying. They now must make a choice about whether or not they want to know if that is the way they will die too. This example brings into sharp focus the impact of biotechnology on our lives. Because the test now exists, children of Huntington's victim must decide what they want to do. Even if they decide not to have the test, to let nature take its course, they have been forced to make a decision that, before the technology existed, was completely out of their hands. Increasingly, biotechnology forces us to decide—to withdraw a respirator, to conduct amniocentesis to detect fetal "abnormalities", to abort those fetuses we might consider undesirable, to register as a recipient or a donor of an organ transplant, to be tested to determine if we are genetically predisposed to cancer, or heart disease, or diabetes. Having to make decisions might be the most significant by-product of the biotechnology revolution.

But back to the quest for HD. As it turned out, all the euphoria about how quickly the HD gene would be discovered evaporated as it became apparent that the search would be long and arduous. One of the problems was that the HD gene was mapped to the very end of the short arm of chromosome #4; 4p. This area of chromosome #4 has been described as a "gene junkyard" and is peppered with many short segments of DNA that could encode short peptide sequences interspersed with intron sections. In addition to the difficulties posed by the "messiness" of 4p, the HD investigators did not have a clue about the protein or the cells involved in the development of HD. The CF investigators developed cDNA probes from respiratory or sweat gland cells. But HD investigators did not know which cells of the brain might be making the normal, or abnormal, version of the HD protein.

Well, investigators slowly narrowed their search on chromosome #4 by finding RFLPs that were linked more and more closely to the presence of HD (that means through a study of the recombination rates, the RFLPs segregated from the HD locus less and less frequently). This narrowed the DNA segment of the search to a 500 kb region (3 times smaller than the CF region). They couldn't narrow their search any more by linkage analysis because they had arrived at a point where the flanking RFLPs were so closely linked to the gene that a recombination event was never detected between them (an effective chromosomal distance of 0.0 map units.) So, it was time to walk down the chromosome and create a restriction map of this 500 kb region.

But, the HD investigators had to take a slightly different approach to this problem than the CF investigators because they would never be able to probe their mapped segment with cDNA. So now what? The HD investigators began by creating a set of overlapping fragments (called contiguous fragments of DNA), which were mapped, as was done for CF.



By digesting this segment of DNA with two different restriction enzymes, two contiguous and overlapping segments were created. Sixteen such contiguous segments were created from the 500kb region known to contain the wild-type allele of the HD gene. Each overlapping segment, or "contig", was put into a vector. Each vector was inserted into a strain of yeast that replicated the human DNA along with its own DNA. Thus, as long as these yeast kept dividing (and they do that as long as you feed them), they kept making copies of these 16 contiguous segments of human DNA which contain the HD gene. (By the way, this shows how simple DNA cloning is, but don't tell anybody because everyone will want to major in biology!)

Then they did something quite interesting called **exon amplification**. They transfected a monkey cell line so that 16 different petri plates of cells each received a different contiguous segment of DNA. Researchers then looked for an mRNA that resulted from transcription of any genes contained within that segment of DNA. If a segment contains a "real" exon from the HD gene—one that contains a coding sequence and is flanked by the proper recognition sites for mRNA splicing, then a new mRNA product will be created in the monkey cells. If the segment does not contain a real exon but only introns or "junk" DNA that does not encode a peptide, no new mRNA would be created.

By doing this and mapping the exons to their exact location based on which segments they were on, they identified a functional gene, isolated, cloned, and sequenced it. The wild-type gene is 210 kb long (ouch!) and creates a processed mRNA transcript that is 10,366 bases long (that's huge! and 120,000 bases of the gene are in introns!). This transcript is translated into a protein that is 3450 amino acids in length and has a molecular weight of 348 kD (a large protein).

Because huge computer libraries of amino acid sequences are available (and you can access them via WWW), you can enter the sequence of their newly found protein and ask the computer to compare the sequence of this protein to the sequence of all known proteins. When the HD investigators did this, the computer told them that their protein was not similar to any other known protein. It would have been nice to have a known homolog—it would have provided a clue about what this protein does, as was the case of the CFTR protein looking like other ion channels. But, alas, not so lucky. As it stands, we have no clues about the function of this protein, even though we know the sequence of the entire gene.

The HD investigators did notice something quite unusual about this gene however. At the 5' end of the coding area, the codon "CAG" repeats itself many times; CAG is the codon for the amino acid glutamine. This type of nucleotide pattern is called a **trinucleotide repeat.** In the normal HD gene from non-afflicted individuals, "CAG" is always repeated between 11 and 34 times in this region. In itself, this is not so unusual. Many functional genes contain trinucleotide repeats. However, the HD gene from afflicted individuals always contains from 38 to over 100 copies of "CAG" in this region. This increase in the number of codons is a kind of mutation called a **trinucleotide repeat expansion**. This mutation accounts for the difference between the HD gene and its wild-type allele—the number of times "CAG" is repeated at the 5' end of what appears to be the coding area of the gene. (We will see something like this during the last two weeks of lab).

While this doesn't give us much help in understanding the protein defect in the HD gene, this type of mutational change is also found in at least three other, less well-known genetic diseases: myotonic dystrophy, fragile X syndrome, and spinal bulbar muscular atrophy (see p 339 and fig 18.9). Therefore, this trinucleotide repeat expansion, a form of insertion mutation, is a type that has been shown to produce at least three other genetic diseases. This greatly strengthens the evidence that investigators have actually found the HD gene.

In November of 1995, researchers made a startling discovery. When the protein encoded by the HD gene (now this protein is called **huntingtin**) has 38 or more glutamines in a row, it has very different binding properties (form meets function again). Huntingtin normally binds to another recently discovered protein called **huntingtin-associated protein number 1** (**HAP-1**). When there are 38 or more glutamines in huntingtin, it appears to bind more tightly to HAP-1. The increased binding causes a change in the level of activity of a dimer of huntingtin and HAP-1 such that only one mutant allele of the huntingtin gene is sufficient to cause a dominant disease. In wild-type individuals, huntingtin and HAP-1 probably have the same function (still unknown as of July 2001) as in affected individuals, but this function is properly regulated in healthy people. Too much of this activity leads to neuronal cell death. Researchers have used immunofluorescence to determine that huntingtin is present in every cell of the entire body. Why only neurons are affected is unclear but this phenotype may have to do with which cells express HAP-1.

Locating and characterizing the function of huntingtin/HAP-1 will occupy investigators for a long time. The brain is one of the most complicated chemical systems in the body, and the most mysterious. But in the search for huntingtin, investigators will undoubtedly learn much about the biochemical function of the normal brain as well as coming to a better understanding and possibly cure of the biochemical defect that causes Huntington's Disease.

> NEWS ITEM: We know that the mutant form of huntingtin contains a poly-glutamine region that is toxic. There are fourteen known neurological diseases that have this sort of repeat or a trinucleotide repeat region that causes the protein to not be translated at all. Recently, a group from Canada has shown that PKR, a double-stranded RNA-binding protein, PKR, preferentially binds mutant huntingtin. Previously PKR had been linked to been linked to a form of virally-induced and stress-mediated cell death called apoptosis. Could it be that the neurological defects in Huntington's and other trinucleotide repeat-based diseases are due to mutant RNA transcripts interacting with PKR and initiating something like apoptosis? PKR is found in the 'right' tissues' at the 'right' time so researchers will be looking in this direction in the future. A. L. Peel et al. Hum. Molec Gen 2001, Vol. 10, 1531-1538 July 2001.

Study Questions:

1. In attempting to locate and characterize the genetic defect causing a disease, explain why it is helpful to have an individual in your sample who is homozygous for the disease trait.

2. The cells that express the HD protein are unknown to investigators. How did this fact change the approach taken by HD investigators when compared to that taken by CF investigators?
3. What is contiguous DNA and how was it used to clone the HD gene?

4. What is exon amplification and how was it used to pinpoint the location of the HD gene?

5. HD investigators determined that the normal version of the HD gene is not similar to any other known protein in structure. How do they know this?

6. What is the actual genetic defect in HD? What is this type of mutation called? Why does the presence of this type of mutation in the HD gene strengthen the evidence that investigators have located the gene that actually causes Huntington's disease?

7. Look at the figure called HD pedigree showing anticipation on the Bio111 Home Page: What do you think caused the patients to get HD at younger ages with each generation?

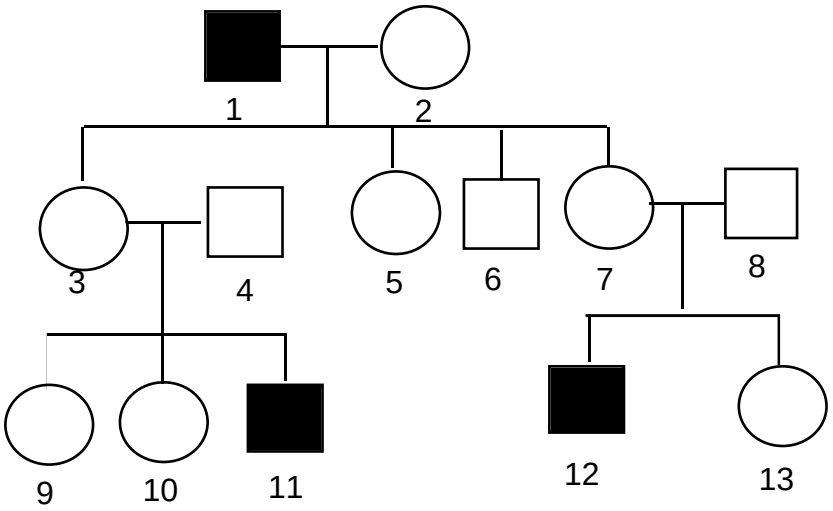
NEWS ITEMS: In January 1998, another international team has isolated a disease version of a potassium channel that appears to cause some cases of schizophrenia. The interesting cause of the defective ion channel is that there is a trinucleotide repeat which causes too many glutamines to occur in a row in this disease allele. It appears that this allele over stimulates some neurons in the CNS and causes other proteins to change their function which leads to the mental state we call schizophrenia. (see January 1998 issue of *Molecular Psychiatry*)

----------STOP----------

And finally, a note about sex-linked genetic disorders such as blue-green color blindness, hemophilia A, and Duchenne's muscular dystrophy (defined p 377 "structural proteins").

Focused Reading: p 195-6 "Genes on sex chromosomes…" to end of chapter
p 195-6 Fig. 10.25 & 10.26

Here is a pedigree for a family with hemophilia A:



Males are **hemizygous** (analogous to being haploid) for sex chromosomes because they are XY and the Y chromosome is greatly reduced in length compared to the X chromosome. Thus, the genes on the X chromosome have no corresponding alleles on the Y. Sex-linked genetic diseases all map to the X chromosome and most are expressed in a dominant fashion in males and in a recessive fashion in females. Because females can have a wild-type allele to counter balance the defective one on the X homologue, they frequently escape the effects of sex-linked genetic diseases. However, because males are haploid at the sex chromosome, if they inherit a single diseased copy, they have the disease.

In the pedigree above, individual #1, a male, has the disease. The disease is carried on his X chromosome. Therefore, he cannot pass the disease on to his sons because they must receive his Y chromosome in order to become male. However, all of his daughters will inherit his X chromosome (that is what makes them girls, they must inherit an X from both parents). Individual #12 inherited his disease-bearing X chromosome from his mother who inherited it from her father. Therefore, all mothers of hemophiliacs must live with the knowledge that they are the genetic source of their sons' disease. This is good news for researchers because a genetic disease that is sex-linked is easier to identify and isolate since the researchers start out knowing to which chromosome the gene maps.

Study Questions:

1. What are the genotypes of all of the individuals in the hemophilia pedigree, assuming individual #2 is homozygous normal? Assuming individual #2 is a hemophilia carrier?

2. How did individual #11 get hemophilia?

3. Given the genotypes of individuals bearing sex-linked traits, be able to predict the genotypes and phenotypes of the offspring. (e.g. Male with no disease crossed to a female carrier, etc.)

4. Test your understanding of the overall concepts in this Unit by thoroughly explaining this newspaper article to a classmate. How do you think these investigators approached this problem? Upon what classic genetic principles was their work based? What aspects of modern biotechnology made this discovery possible? Based on information in this article, would you classify Alzheimer's disease as a Mendelian genetic disease? Why or why not?

From the Minneapolis Star Tribune, August 13, 1993:

Most common form of Alzheimer's linked to cholesterol-processing gene

From News Services
Washington, D.C.

Researchers have linked the most common form of Alzheimer's to a gene that helps process cholesterol, enabling them to identify some patients who are virtually certain to develop the mind-destroying disease in their elderly years. The discovery could account for half of all patients with the common neurological disorder, they said, and it points the way toward devising treatments to block or at least delay the ultimately fatal symptoms of the incurable illness.

About 4 million Americans suffer from Alzheimer's and the number is expected to increase sharply as the population ages. In research on 42 families where late-onset Alzheimer's is common, Duke University scientists found a 90 percent risk of the disease by the age of 80 among people with two copies of a gene variant called apolipoprotein-E, type 4, or APOE-4. Copies of the APOE-4 gene also was linked to people developing Alzheimer's at an early age, said Dr. Allen D Roses of Duke. "What this shows is that APOE-4 increases the risk and lowers the age at which you get the disease." he said. "It looks like virtually all will develop it (the disease) by the age of 80 if they have two

A report on the study appears in today's issue of the journal *Science*. Dr. Zaven Kachaturian, director of Alzheimer's Research at the National Institute of Aging, one of the National Institutes of Health, said the research has caused "a great deal of excitement" among Alzheimer's researchers because it links the most common form of the disease with a specific gene factor, APOE-4, that can be measured. "It could become a diagnostic tool" said Kachaturian. "We may be able to screen for this and be able to make judgments about whether a person's likelihood of getting the disease is high or low, or early or late. It has that potential."

The Duke researchers cautioned that their conclusions now can be applied only to families where members have late-onset Alzheimer's, the most common form of the disease. Additional studies to verify the finding will be required before the conclusions can be applied to the general population, said Kachaturian.

In the latest finding, the researchers studied a gene that allows the body to manufacture apolipoprotein E, or ApoE, an essential protein that shepherds

Scientists have known for years that the gene comes in three varieties, called E2, E3, and E4, and they have known that patients with the E4 version of the gene have a small but notably elevated risk of cardiovascular disease. The new work demonstrates that possession of the E4 variant is an even greater risk factor for Alzheimer's disease than it is for heart disease.

Studying 234 people from 42 families afflicted with late-onset Alzheimer's, the researchers found that those patients with two copies of the E4 gene had eight times the risk of having the neurodegenerative disease that people had when their two copies of the apolipoprotein gene were some combination of either E2 or E3 varieties. (All genes of the body come in two copies, one donated by the mother, the other by the father.)

Even inheriting one copy of the E4 gene turns out to be bad news, doubling or tripling the risk of Alzheimer's over that of people having no E4 genes at all. Researchers do not yet know how the gene predisposes people to the disease, but, although it has not yet been proven, Roses say he believes there is a direct cause and

copies.”	cholesterol through the bloodstream.	effect.

Unit III: Bioenergetics

Brief Overview Reading: Chapters 2-6

Certainly one of the primary differences between biological creatures and inanimate objects is their structural organization. Organisms are built from cells and, while cell structure varies dramatically from one organism to another, all cells share many common features (e.g. plasma membranes, genetic material, enzyme systems, receptors, membrane transport systems, etc.). In trying to define what we mean by "life", this structural difference serves us well. If you look through the microscope and see cells, you are certainly looking at a biological creature.

But is it alive? How do you distinguish living cells from dead cells? Living creatures from dead creatures? Well, okay, dead creatures don't move, or vocalize, or breathe, or eat -- i.e. they can't do anything. Doing something requires the contraction of muscles, the beating of cilia or flagella, or the secretion of products. In addition, as far as we can tell, dead creatures don't sense anything. They don't see, hear, feel, or taste. That is, they have no "sensory function." And, again, as far as we living types know, dead creatures don't think about anything or have any emotions -- they don't remember, plan, enjoy, problem-solve, love, hate or do homework.

If you look for the common denominator in all these activities -- moving, sensing, thinking and feeling -- you find that they all require ENERGY. Only living creatures can use energy to accomplish these things, these characteristics of life. You can have the structure of a biological creature, but without energy, you cannot be alive.

Energy is an odd and sometimes hard thing to study. It's much less tangible than matter, which you can see, weigh, and measure directly. The effects of energy are manifested in movement (e.g. actin and myosin filaments sliding past one another, ions traveling up their concentration gradient) or in increases in temperature. We also have sense organs that can sense the presence of certain kinds of energy. For instance, our eyes can detect the presence of electromagnetic radiation with wavelengths between 380 and 750 nm (visible light). Our ears can detect the vibration of the air at certain frequencies (sound waves). However, we have no sense organs for many forms of energy including radio waves (your radio receiver can detect these, but you can't), radioactivity (a Geiger counter can detect these, but you can't) or neutrinos (they are passing through you right now, but you can't sense them.)

So, what is energy anyway?

Focused Reading: p 28 Most biological…" stop at "Acids, Bases…"
p 25-6 " Chemical Reactions…" stop at "Water:…"
p 95-99 "Energy conversions…" stop at "Chemical reactions…"

Energy is the capacity to do work. In order for this definition to make sense, you must think of work in the very broadest sense of the word -- work is anything that changes the position or state of matter. Matter at absolute zero (no energy) is absolutely still and immutable -- no movement or change of any kind. Any movement or change in the structure of matter requires the input of energy. And that is how energy is defined. It's circular reasoning, but reasoning all the same. That which moves or changes matter is energy. And ENERGY IS MEASURED BY THE AMOUNT OF

MOVEMENT OR CHANGE IN MATTER THAT IS PRODUCED. Big change or big movement equals big energy. Little change or little movement equals little energy.

In many ways, this is just common sense. Does it take energy to move a barge up river? Yes. Does it take more energy to move a large barge than to move a small sailboat? Yes. Does it take more energy to move a barge up river than down river? Yes. Energy and matter functioning on the molecular level are NO DIFFERENT. Does it take energy to move a molecule across a cell membrane? Yes. Does it take more energy to move a big molecule than to move a small one? Yes. Does it take more energy to move a molecule up its concentration gradient than down its concentration gradient? Yes.

Concepts, concepts, concepts -- there are only a few but they apply in many, many situations.

Study Questions:

1. What is energy? Give one of the classic definitions, and then define it in your own terms.

2. How is energy measured? How do you know that a lot of energy is being expended versus a small amount of energy?

3. How do kinetic and potential energies differ? Give some examples, not found in lecture or your textbook, of the two forms of energy.

4. What are the two laws of thermodynamics? Define them in everyday terminology.

This unit is about how biological creatures harvest energy from their environment and use it to live. The sun provides the energy we need to live, but in order to convert this energy into a usable form, biological creatures have had to develop elaborate systems for energy harvesting, storage, and use. This system is called **metabolism** and its study is the field of **bioenergetics**.

In this Unit, we will look at four examples of cells that harvest, store and use energy in different ways. We will find out why the United States government spayed paraquat on Mexican marijuana, why cyanide is used by terrorists to poison consumer goods, why vegetarians eat tofu, and how a rusty nail can kill you.

QUESTION #1: THE US GOVERNMENT V. MEXICAN MARIJUANA FARMERS

Rolling Stone, April 6, 1978

Whatever Happened to Mary Jane? by Michael Roger

The case of the poisoned Mexican marijuana started late in 1975, when the United States, faced with an abrupt increase in the amount of heroin entering from Mexico, began to assist that government with an elaborate program of spraying poppy fields with powerful herbicides. From the beginning, however, that aerial attack was equally aimed at marijuana fields.

The program has been a success; the Mexican heroin supply in this country has declined dramatically. But is has also meant that approximately twenty percent of the Mexican marijuana entering this country is contaminated with a dangerous herbicide, an estimate based on government analysis of marijuana samples confiscated recently in the Southwest.

The herbicide in question is **paraquat**, an exceedingly toxic chemical that, less than a month ago, was placed on the Environmental Protection Agency's restricted list - meaning that only licensed applicators may purchase it - and which some observers feel may be banned altogether in this country. Paraquat remains in the body even longer than DDT and has no know antidote, thus figuring occasionally in fail-safe suicides. At present, the maximum paraquat contamination that the EPA allows in foodstuff is 0.05 parts per million. Confiscated marijuana samples analyzed last November contained an average of 177 parts per million, with a high of 655. (One recently tested sample reportedly contained 2200 parts per million.)

It is not yet clear what paraquat will do when burned and inhaled, although the National Institute on Drug Abuse is doing its best to find out. The current and hopeful guess, of course, is that the compound is rendered harmless during combustion.

Even assuming that to be the case, however, what about oral ingestion - brownies, *majoun*, and the like? The Drug Enforcement Administration estimates that 2700 tons of marijuana enter this country each year from Mexico. Assuming that only one percent of that produce is eaten, and that only twenty percent of that has been contaminated with paraquat, that still means that almost fourteen tons of poisoned marijuana have been eaten in this country since the spraying program began.

It's not clear what sublethal doses of paraquat can do because most cases reported have involved lethal doses due to the ingestion of pure material. Evidence suggests that damage would occur first in the lungs, liver, and kidneys.

Rolling Stone, May 4, 1978

Poison Pot In the weeks since Michael Rogers's Alternating Currents column (RS 262) described the possible health hazards of ingesting herbicide-contaminated Mexican marijuana, the situation has changed for the worse. Scientific studies have revealed that the herbicide involved - paraquat - can survive the burning

process and be inhaled directly into the lungs. Paraquat is so exceedingly toxic that on March 12th, HEW Secretary Joseph Califano issued a warning that heavy use of contaminated marijuana could lead to irreversible lung damage.

At present, the only drug-analysis laboratory with an effective paraquat test is in California. Consumers may send a one-half gram sample (one joint) of suspected Mexican marijuana wrapped in foil to PharmChem Research Foundation, 1844 Bay Road, Palo Alto, CA, 94303. Enclose five dollars for lab costs plus any five-digit number. After ten days, the result of the analysis may be learned by calling (415) 322-9941 and giving the identification number. PharmChem also requests the following information about the sample: city and state where purchased, street price paid, and what it was sold to you as (Colombian, Mexican, Hawaiian, etc.)

_____

Brief Overview Reading: Chapters 7 & 8

Why would our government want to spend tax dollars to spray paraquat on Mexican marijuana? To answer this question, we need to know that paraquat is an herbicide. It kills almost all plants (except a few plants that are resistant to it). How does paraquat kill plants? Why might it be dangerous to humans? Before we can answer these questions, we need to understand how plants do what they do best - harvest energy from the sun and turn that energy into sugars, which are then used to support all life.

A plant is nothing less than a miracle. Plants are able to harvest the energy of the sun and use it to convert CO_2 into food. And in the process of doing this, plants produce a waste product called oxygen. All animal life depends on plants to harvest energy, make food from a gas in the air, and produce the oxygen we breathe. If there were no plants there could be no animals, whereas without animals, many plants would be just fine. We need them more than they need us. Remember that the next time you walk on the grass, or forget to water your houseplant, or lean against a tree, or read about the rapid loss of the planet's rain forests.

So, how do plants turn sunlight into sugars? How do they harvest energy, use that energy to create food from CO_2, and excrete oxygen? Energy is harvested and oxygen is produced in a process called the **light reactions of photosynthesis**. The creation of food (sugars, proteins, nucleic acids and lipids) from CO_2 occurs in a process called the **dark reactions of photosynthesis** or the **Calvin-Benson cycle.** Both processes occur in the leaves of plants. A typical plant leaf is illustrated in fig 8.19 (p149) of your text. The leaf is covered by a skin, or **epidermis,** which secretes a waxy coat, called the **cuticle**. The epidermis protects the plant and the cuticle prevents water loss on exposed surfaces. Under the epidermis lies the **mesophyll**, a tissue that contains the **photosynthetic** cells of the plant.

First, let's look at the light reactions of photosynthesis. Remember, during this process, the marijuana plant will harvest the energy of sunlight and give off oxygen. What do we mean by the term "harvest the energy of sunlight?" How would you harvest sunlight energy if you were asked to do so? The word "harvest" implies that the energy is gathered and stored in a form that can be used at a later time -- the harvest contains potential energy. Going out and eating a field full of corn would

not be considered "harvesting" the crop. So, using sunlight energy to do something (e.g. illuminate a room, warm your skin, dry your clothes) is not harvesting energy because you have already "used" it -- none of the energy is stored for use at a later date. (Of course, you can't destroy energy, and in the process of using it, the energy has simply been converted to another form, namely to heat energy which is eventually radiated into space.)

Have you thought of a way to harvest sunlight? One high tech example of harvesting sunlight energy is the solar cell. The cell collects sunlight and uses it to separate charge (create voltage). This voltage can produce current to run electrical devices. A low-tech example would be hanging a blanket out in the sun to warm it and then using the blanket to warm yourself. The radiant energy of the sun increases the kinetic energy of the blanket, which can be used to warm you as it is released from the blanket.

Study Questions:

1. What major events happen during the process of photosynthesis?

2. In general, what happens during the light reactions of photosynthesis? During the dark reactions?

3. We say we use energy to perform tasks. However, the first law of thermodynamics instructs us that energy cannot be created or destroyed. What happens to the energy we "use" to live our lives?

The task of the green plant is to collect the energy of the sun and store it in a form that can be used later to do work. In order to understand this, we have to know a little more about sunlight, a form of **radiant energy**, and **chemical energy**, the kind of energy organisms use to run their lives.

Focused reading: p 138-9 "Properties of Light and Pigment" stop at end of page)

Radiant energy comes in various forms including radio waves, microwaves, gamma rays, X-rays, visible light, and infrared. Each of these waves has a characteristic wavelength. The wavelengths of visible light are between 380 and 750 nm. Because we are primates and can see color, our eyes can distinguish the various **wavelengths** from one another, and we experience these different wavelengths as differences in color. For instance, when light at 400 nm hits our eyes, we experiences this as violet, while light at 600 nm will give us a yellow sensation. The relationship between colors and wavelengths is illustrated on page 139, Figure 8.5.

Visible light has some of the properties of waves and therefore has a **wavelength**. However, light also has some of the properties of particles. These particles are called **photons**. They can be thought of as packets of energy. Each photon has a certain quantity of energy (a **quantum** -- plural **quanta**). The energy level of photons is inversely related to the wavelength. Thus, a photon of red light (wavelength 750 nm) has about half the energy of a photon of violet light (380 nm); short wave length = high energy.

So during the day, you and the *Cannabis* plants are being bombarded by these photons of light (acting like waves and particles at the same time). Zillions of photons per millisecond hit us, each

one having a particular energy level, wavelength, and color. [There is no such thing as a white photon -- the color white is caused by photons of all the different energy levels or wavelengths (colors) striking your retina simultaneously. White sunlight contains blue photons and red photons and violet photons and yellow photons, etc., all mixed up together.

IT IS THE ENERGY OF PHOTONS THAT THE GREEN PLANT HARVESTS. But how? The first thing the plant has to do is **absorb** the energy of these photons. Most of the world around you absorbs photons. In fact, anything with any color or **pigment** is absorbing photons. The grand mixture of photons in white light hits an object; some of the photons are reflected back to your eye while some are absorbed by the object. If the object is colored, it contains a type of molecule, a **pigment molecule**, that is chemically structured in such a way that it can absorb some photons' energy. Each type of pigment molecule will absorb photons based on their energy levels. Some pigments only absorb blue photons, some only absorb red, some absorb yellow and blue, etc.

If no pigment molecules are present, then all the photons are reflected and the object appears white. If all the photons are absorbed, the object reflects no light and appears black. If only red photons are absorbed, the rest of the photons are reflected back, minus red photons, and the color will be a mixture of violet, blue, green, yellow and orange -- no red. If red, orange and yellow pigments are absorbed, the remaining colors (violet, blue and green) will be reflected back and the object will appear to be some shade of blue. [If a tree reflects green photons in the forest but no one is there to see it, is it really green?]

Because photons are a form of energy, when colored objects absorb photons, they are absorbing energy and become warmer (due to an increase in the kinetic energy of the molecules in the colored object). Thus, black clothing absorbs all photons and heats up while white clothing reflects all photons and remains cool.

Study Questions

1. Describe the components of white light. Which components have the highest energy? The lowest? What is the range of wavelengths spanned by visible light?

2. Chemically and physically, what makes something appear to have color?

3. Visible light is an example of electromagnetic radiation. What are some other examples of this type of energy?

Focused Reading: p 140-2 "Light absorption and…" stop at "Electron flows…"
WWW Reading: Relative Sizes: from glucose to cells and larger

Chlorophyll a and **b** are green pigments and **carotenoids** are shades of yellow and orange (as in carrots and fall leaves). Since chlorophyll a is the dominant pigment in most plant leaves, most plants appear green. But if you look at plant leaves closely, you'll note that each plant, and each leaf on each plant, is a slightly different shade of green. This is due to a shift in the proportions of the various pigment molecules in the cells.

Let's focus on chlorophyll a. It appears green. So, it reflects green light. So, it does not absorb green photons. Because colors are so complex, however, it's really hard to say what colors are absorbed. This has to be measured using a **spectrophotometer**. The **absorption spectrum** for chlorophyll a and b are illustrated in Figure 8.7 on page 140. You have used a spectrophotometer in lab and should understand how this piece of equipment works. You have also constructed an absorption spectrum so the figure should be easy to interpret. If it doesn't look familiar, refer to your laboratory manual (IDH labs).

This absorption spectrum shows that chlorophyll a absorbs maximally at about 450 nm (it prefers to absorb high-energy blue/violet photons) and also at 670 nm (orange/red photons). It does not absorb blue-green, yellow or true red photons so they are reflected back to the eye of the observer. The ultimate color produced by this absorption pattern is green. It is the energy of the blue/violet and orange/red photons that is harvested by chlorophyll a in the green plant. In addition to chlorophyll a, most plants have accessory or secondary pigments (e.g. chlorophyll b and the carotenoids) which absorb photons at other wavelengths. Therefore, plants can frequently harvest photons across the entire spectrum of white light.

Study Questions

1. Explain how a spectrophotometer works.

2. What is an "absorption spectrum"? How is it obtained?

3. What wavelengths and colors of light are absorbed by the chlorophyll a? Explain how this results in its green appearance.

4. Why does the absorption spectrum of chlorophyll a differ from the spectrum of an entire chloroplast?

So, the leaves of green plants are full of these pigment molecules that absorb photons. Where are these pigment molecules? Floating free in the cytoplasm? Attached to a membrane? (These are generally the two options in cell biology.) Well, photosynthesis is a COMPLICATED PROCESS. It involves dozens of enzymes performing dozens of tasks in a set order. This is very much like assembling an automobile; you cannot put in the stereo before you have assembled the dashboard; you have to do things in order. So instead of having the molecules involved in the process floating around haphazardly in the cytoplasm, they are attached to membranes in **macromolecular complexes**. These complexes are organized so that the molecules involved in each reaction are kept near the next molecule in the sequence. The membrane serves as a scaffolding, or frame, that holds these molecules in position and carrier molecules travel between those positions. The membranes that hold chlorophyll, and all the other molecules associated with photosynthesis, are found in the **chloroplast.**

Focused Reading: p 68-9 "Plastids photosynthesize or store materials" stop at "Other types…"

Look at the picture of chlorophyll **a** on page 141 (Figure 8.9). You'll remember that cell membranes have hydrophobic, lipid cores. Therefore, the non-polar **hydrocarbon tail** of chlorophyll

a dissolves with great stability in the lipid membrane of the thylakoid (in the **thylakoid membrane**.) The highly polar **porphyrin ring** containing the Mg atom is the portion of chlorophyll that interacts with light. Thus, part of chlorophyll is designed to anchor it to the membrane, maintaining its orderly relationship to the rest of the molecules of photosynthesis, and the other part is designed to harvest light energy.

Study Questions:

1. Describe, in general terms, the chemical structure of chlorophyll a. Focus on the structural characteristics of the molecule that are significant for its function.

2. Explain why it is advantageous to embed macromolecular complexes in cell membranes.

3. Describe the structure of the chloroplast including the structure and location of thylakoids, grana, and stroma. Describe the location of chlorophyll a in the chloroplast, and explain how the molecule is anchored into the membrane.

Focused Reading: p 139-40 "Absorption of a photon…" stop at "Light Absorption…"
p 140 Fig 8.6

Each pigment has a particular photon energy level that "fits" it perfectly such that photons of that energy level can be absorbed, while photons of other "misfitted" energy levels cannot be absorbed. This absorption results in a transfer of photon energy to an electron in the pigment molecule. This electron, normally at **ground state**, or in its normal non-excited position in an orbital around the nucleus, is boosted to a higher orbital (an **excited state**) by the absorbed energy. In regular pigment molecules such as the ones in your clothes, when the electron is in its excited state, it quickly returns to ground state and gives off the absorbed energy as light or heat. However, in the chloroplast, chlorophyll a is anchored in the thylakoid membrane in a macromolecular complex. One of the associated molecules is the **primary electron acceptor**. When the electrons of chlorophyll are boosted to an excited state, before they have a chance to fall back to ground state, this primary electron acceptor takes them away from the excited chlorophyll.

Because biologists and chemists have to give names to everything, this process -- where an electron is transferred from one molecule to another (or when an electron moves closer or farther away from a molecule without actually being transferred to another atom) -- is called **oxidation-reduction** or a **redox reaction**.

Focused Reading: p 11506 "Redox Reactions..." stop at "The coenzyme NAD…"
p 116 Fig 7.2

WWWeb Reading A quick reference for reaction equations can be found at http://www.whfreeman.com/purves6e choose the Math for Life icon and the 'Biochemical reactions' in the reference table section

Study Questions:

1. What is oxidation? Reduction? Give examples. What is a reducing agent? An oxidizing agent? Give examples.

2. What is electronegativity? Electropositivity? In which way will an electron naturally tend to flow -- from electropositive to electronegative or *vice versa*?

3. What would you suspect is true of the primary electron acceptor in the thylakoid membrane; it is relatively electropositive, relatively electronegative, or about in the middle? Explain your answer.

----------STOP----------

So this is how chlorophyll harvests light energy. Well, it's a long way from this little excited electron to using paraquat to kill marijuana. In order to get there, we have to talk about the other kind of energy involved in photosynthesis, **chemical energy**.

Potential chemical energy (food and fuels of all kinds) is said to be stored in the bonds of molecules. Covalent bonds, as you know, are shared electrons. These electrons are being shared because each element in the bond "needs" the stability that sharing electrons brings. (The elements are more stable or at a lower energy level if they are sharing electrons with one another.) Each atom is trying to fill an electron shell with the correct number of electrons and covalent bonds help the molecule do this. [Refresher? Read pages 20-24]

Molecules contain POTENTIAL CHEMICAL ENERGY. We say that the potential chemical energy is "in the molecule's bonds", but this is in many ways misleading. Potential chemical energy can be thought of as the capacity to produce molecular change (to do chemical work). Thus, if a molecule is fairly UNSTABLE, it is likely to change to a shape or configuration that is more stable. The energy that **is released** when this change towards a stable configuration occurs is the **heat of the reaction** (**ΔH, H** is referred to in the book as **enthalpy**) and it is a measure of how much potential energy was stored in that unstable molecule. [Actually, it is a measure of the difference in the potential energy stored in the reactant and the potential energy stored in the product, since the product could go on to react and become even more stable and release even more energy.]

By convention, when a reaction **gives off energy** (this energy is usually given off as heat, but it might also be light, electrical current or movement), the ΔH of the reaction is designated as negative. Thus, a reaction that **gives off energy** (e.g. burning fuel) has a **-ΔH** and is said to be **exothermic**.

Conversely, reactions that proceed only when energy is added (usually in the form of heat, but it might also be light, electrical current, or movement) the ΔH of the reaction is positive. Thus, a reaction that **requires the input of energy** (draws energy from the environment) has a **+ΔH** and is said to be **endothermic.**

Focused Reading: p 96-100 "Energy Conversions:…" stop at "ATP: Transferring…"
WWW Reading: Animation of Photosynthesis (from Virtual Cell)

When a chemical reaction gives off energy (e.g. when gasoline is burned in a car engine), most of the energy given off by the reaction is given off as heat or car movement. Heat and movement represent work. [Car movement = propelling the car; and Heat = increasing the movement (kinetic energy) of molecules.] However, some of the energy given off by the reaction is not represented in either heat or movement (is not represented by work.) Rather, this energy is represented by a change in the **entropy** of the gasoline molecules. The chemical reaction is:

Oxygen + Gasoline ---> Carbon Dioxide + Water

Gasoline (a long chain hydrocarbon) is **more organized** than CO_2 and H_2O. Therefore, gasoline has **less entropy** (or **randomness**) than CO_2 and H_2O. Some of the energy given off by this reaction has been used to **increase entropy** -- change low entropy (more organized) molecules into higher entropy (less organized) molecules. This change in entropy level (Δ**S**) is not available to do work (in this case, provide heat or movement.) Reactions (e.g. burning gas) in which entropy is increased have a **positive** Δ**S** while reactions (e.g. refining gasoline) in which entropy is decreased have a **negative** Δ**S**.

Thus, in all chemical reactions (the energy source for virtually all biological function), two kinds of changes occur -- changes in potential energy of the molecules (ΔH) and changes in entropy (ΔS). Usually, reactions that give off energy to do work (exothermic or -ΔH reactions) also involve an increase in entropy (have a +ΔS). Such reactions in biology include the burning of food for energy. Big, complex, organized molecules (proteins, carbohydrates, lipids and nucleic acids) are broken down to simple, small molecules of CO_2 and H_2O. Much energy is given off in the process (-ΔH), and the entropy of the molecules is dramatically increased (+ΔS).

Conversely, reactions that absorb energy (endothermic or +ΔH) usually involve a decrease in entropy (have a -ΔS); such reactions in biology include the building up of structures during growth. Simple, small molecules such as amino acids, nucleotides, and monosaccharides are linked together into large, organized molecules such as proteins, nucleic acids and polysaccharides. Much energy is required for this process (+ΔH) and the entropy of the molecules is dramatically decreased (-ΔS.) The production of sugars by plants is an example of an endothermic reaction that decreases the randomness in the world - a comforting thought.

Biologists are very interested in the ΔH of reactions because this determines when a given reaction will be able to supply energy for life and when a reaction will require the input of energy from the organism. However, another factor is also very important to biologists and that is the **free energy of the reaction (ΔG).** This factor determines whether a reaction will **proceed or not**. Reactions that proceed on their own without energy input from the cell (beyond activation energy) are called **spontaneous,** while reactions that will not proceed unless energy is added are called **non-spontaneous**. Spontaneous reactions are said to be **exergonic** and have a **-ΔG** while non-spontaneous reactions are said to be **endergonic** and have a **+ΔG**. Usually, exothermic reactions are exergonic and endothermic reactions are endergonic, but not always. If an endothermic reaction (takes heat from the environment, +ΔH) involves a large increase in entropy (+ΔS), then it may be spontaneous (have a -ΔG) even though it requires the input of energy. Melting ice is an example of a reaction that requires the input of energy (heat is removed from the environment -- thus the reaction

has a +ΔH) but results in a dramatic increase in entropy (+ΔS) as ice goes from an organized crystal to a disordered liquid form.

Cheat Sheet

-ΔS = product has less randomness	+ΔS = product has more randomness
-ΔH = rxn gives off energy (exothermic)	+ΔH = rxn takes in energy (endothermic)
-ΔG = rxn is spontaneous (exergonic)	+ΔG = rxn non-spontaneous (endergonic)

The friendly relationship between ΔH and ΔS (given on page 121 of your text) is:

$$\Delta G = \Delta H - T\Delta S \qquad \text{(T is the temperature in Kelvin units)}$$

Thus, to determine whether a reaction is spontaneous or not (and to determine how much of the reaction energy is actually available (or **free**) to do work, you must subtract any gain in entropy multiplied by the temperature (Kelvin degrees) from the total change in potential energy of the reaction.

We do not introduce all this to confuse you, although it must seem like it. Rather, we face a dilemma in teaching you. In chemistry, you are learning about ΔH, which is widely used by chemists to describe the "simple" thermodynamics of chemical reactions. Biologists, however, focus less on the chemical reactions themselves, and more on what the chemical reactions can do for biological creatures, (i.e. we are interested in that portion of the energy that is available to run biological creatures). Thus, we have to introduce ΔG, and we have to tell you how it relates to ΔH so you can integrate what you are learning in the two classes.

Study Questions:

1. How is energy stored in molecules?

2. Be able to explain these terms: exothermic, endothermic, -ΔH, +ΔH, endergonic, exergonic, +ΔG, -ΔG, spontaneous, non-spontaneous, entropy, +ΔS, -ΔS.

3. What determines whether or not a reaction will proceed without an input of energy from the cell? Given examples of the types of biological reactions that tend to be exergonic and examples of those that tend to be endergonic.

4. Explain the second law of thermodynamics as you would to a junior high school student in a science class. Give an example of how the second law of thermodynamics is important in the study of biological systems.

The second law of thermodynamics governs all chemical reactions (that means your entire life). Now this is just fine if you want to do something exergonic. If you provide an enzyme to lower the activation energy barrier, the reaction will proceed just fine. The problem is, most of what you really want to do (move, pump blood, breathe, think, see, hear, secrete, etc.) is decidedly ENDERGONIC. This is another way of saying that living is an energy-absorbing activity. Living is endergonic and endothermic. So, you have a problem because endergonic reactions don't occur spontaneously -- you have to add energy to the reaction to get it to proceed.

While ultimately, this energy you live on comes from the food you eat (which ultimately comes from plants which synthesize it using the energy harvested from the sun), the DIRECT SOURCE OF ENERGY FOR MOST ENDERGONIC REACTIONS IN LIVING THINGS IS ATP.

Focused Reading: p 100-102 "ATP: Transferring…" stop at "Enzymes: Biological catalysts"

Study Questions:

1. Describe and draw the reactions converting ATP to ADP, and *vice versa*. What is the ΔG of each reaction?

2. Be able to describe the process of energy coupling by phosphate transfer outlined in Figure 6.10 on p 102.

For the most part, as long as you maintain an adequate supply of ATP, you can live your life -- ATP will supply the energy required for your endergonic reactions. And the same is true of all other creatures on the planet, including marijuana. As long as we have enough ATP (or other high-energy nucleotides that function in the same manner as ATP), we can do all the endergonic reactions we must do to stay alive. However, you can see that ATP is converted to ADP during the process of providing energy for endergonic processes. Thus, living cells are constantly sapping ATP's energy so cells must have a way of constantly replacing this lost ATP. Marijuana plants can do this directly by using solar energy (photosynthesis) or indirectly by burning fuel molecules (cellular respiration). Animals can do this only by burning fuel molecules.

The *Cannabis* plant has two biological needs: 1) It must provide itself with enough ATP to stay alive and 2) it must provide enough nutrition in its seeds to nourish its offspring which will allow them to sprout, and in turn harvest energy on their own. In the process of harvesting energy and storing energy, marijuana leaves also happen to produce oxygen as a waste product, which animals gratefully inhale (we're referring to the oxygen, OK?!).

Harvesting Energy and Generating Oxygen: The Light Reactions of Photosynthesis

Focused Reading: p 136-8 " Identifying photosynthetic reactants…" stop at "Properties of light…"
p 100-2 "ATP: Transferring energy…" stop at "Enzymes…."

WWW Reading: p 142-6 " Excited chlorophyll…" stop at "Making sugar…"
http://www.whfreeman.com/purves6e
Ch8 Tutorial 81. Photophosphorylation
Diagram of $NADP^+$ conversion to NADPH
http://www.whfreeman.com/purves6e Math for Life"Biochemical reactions"

With the help of photosynthesis, CO_2 and H_2O are converted to sugars (e.g. glucose = $C_6H_{12}O_6$), lipids, amino acids (with the addition of nitrogen) and nucleotides (with the addition of nitrogen and phosphorus.) This process is HIGHLY ENDERGONIC and the energy to power this non-spontaneous process is provided by the sun.

Your text uses the example of the synthesis of glucose from CO_2 and H_2O and we will use this example too. However, you must remember that plants can make all the nutrient classes, not just glucose, *via* photosynthesis. The overall balanced reaction for glucose synthesis, then, is:

$$6CO_2 + 12H_2O + \text{light energy} \longrightarrow C_6H_{12}O_6 + 6O_2 + 6H_2O$$

This overall reaction is actually a **redox reaction**. The light reaction is as follows:

$$12H_2O + \text{light energy} \longrightarrow 6O_2 + 24e^- + 24H^+$$

At this point, we need to stop and talk a bit more about hydrogen and its propensity to fall apart into an electron and proton. Hydrogen is extremely **electropositive**, meaning that the nucleus of hydrogen (composed of only one proton and zero neutrons) does not have very much affinity for electrons -- it does not pull very hard, or attract very tightly, the electron in orbit around it. Thus, it is relatively easy for an **electronegative** molecule (one that has a high affinity for electrons and attracts them very strongly) to take hydrogen's electron away from it, rather than sharing the electron in a covalent bond. Thus, in the presence of electronegative molecules (such as $NADP^+$, NAD^+ and the cytochromes), hydrogen's electrons are more attracted to electronegative molecules than to the hydrogen nucleus -- they leave orbit and are added to the electronegative molecules (e.g. converting $NADP^+$ into NADP, etc.). The "naked" hydrogen nucleus, having lost its electron, becomes a proton, or hydrogen ion (H^+). These protons simply float around in the cytoplasm.

During the light reactions of photosynthesis, the 24 hydrogens on the 12 water molecules are removed, leaving 6 molecules of O_2, which the plant gives off to the environment. The 24 hydrogens are split into 24 protons (H^+) and 24 electrons (e-). The 24 electrons are added to "carrier" molecules called $NADP^+$. 12 carrier molecules pick up the 24 electrons (2 electrons per $NADP^+$----> $NADP^-$) and 12 protons (1 proton each, $NADP^-$ ----> NADPH). [To form reduced NADPH, $NADP^+$ (the oxidized form) picks up 2 electrons and one proton. One electron neutralizes the $NADP^+$ to NADP. The second electron plus the proton forms a hydrogen atom and is added to the molecule to form NADPH. The other 12 protons simply float free in the thylakoid space of the chloroplast, lowering its pH.]

By causing chlorophyll to lose an electron, solar energy converts chlorophyll into a powerful **oxidizing agent** (it will get reduced). Because chlorophyll "wants" to replace that electron very badly, it is able to take the hydrogens away from oxygen in molecular water. By taking water's

hydrogens, chlorophyll gains back the electrons it loses by photooxidation. Again, hydrogens are split into electrons, which enter the chlorophyll molecule, and protons, which float freely in the thylakoid of the chloroplast. Water is a very stable molecule, so removing its hydrogens is not easy -- oxidized chlorophyll is one of the most **electronegative** molecules known -- far more electronegative than oxygen. That's how it is able to take oxygen's hydrogens away in a water molecule.

Study Questions:

1. Explain why the addition of a hydrogen atom to a molecule is reduction. How does the electropositive nature of hydrogen allow it to function as a reducing agent?

2. Very specifically, how does solar energy cause the splitting of water into hydrogen and oxygen during the light reactions?

3. What happens to the oxygen released from the split water? What happens to the hydrogen released from the split water?

4. Describe how the carrier molecule $NADP^+$ works. Why is it called a carrier? What does it carry? To what molecule that you have studied in this course is it most closely related (besides NAD^+ and FAD)? Is $NADP^+$ a protein, lipid, carbohydrate, or nucleic acid?

5. Describe the processes of cyclic and non-cyclic photophosphorylation. What is being phosphorylated in these reactions? How do these processes different from one another? Which process evolves oxygen? Explain the mechanism through which this process evolves oxygen while the other process does not. Which of these processes produces NADPH? Explain the mechanism through which this process produces NADPH while the other process does not.

6. The ultimate products of the light reactions of photosynthesis are NADPH, ATP and O_2. Be able to describe how each of these products is formed.

7. What is a cytochrome? What is an electron transport system? Upon what basic concepts is this model based? (i.e. what attracts the electrons down the system?)

8. Describe the chemiosmotic theory and explain how it works to produce ATP in the chloroplast during the light reactions of photosynthesis.

9. Explain, in the simplest possible terms for a younger brother or sister, how green plants harvest sunlight energy.

----------STOP----------

Synthesizing Food: The Calvin-Benson Cycle

Focused Reading: p 148 Fig 8.17
p 137-8 “The two pathways…” stop at "Properties of light…"
p 146-8 "Making sugar…" stop at "Photorespiration…"

The light reactions of photosynthesis generate O_2, ATP, and NADPH. Oxygen is a waste product to the plant, but ATP and NADPH are required by the plant to make nutrients from CO_2. For the synthesis of nutrients, the plant requires an energy source (provided by the ATP generated during the light reactions), a source of carbon (CO_2 from the atmosphere), and a source of **"reducing power"**. Look at the second half of the photosynthesis equation -- the part that synthesizes sugar:

$$24e^- + 24H^+ + 6CO_2 + \text{energy (ATP)} \text{ -------> } \text{glucose } (C_6H_{12}O_6) + 6H_2O$$

In this reaction, CO_2 is reduced to glucose; the carbon atoms have 12 hydrogens added and 6 oxygens removed. This reduction requires a **reducing agent** and that reducing agent is NADPH, itself reduced in the light reactions with electrons and hydrogens from water. Ultimately, the hydrogens used to reduce CO_2 to glucose come from water.

Despite the fact that glucose was used in this example of photosynthesis, the molecule we should focus on is **glyceraldehyde 3-phosphate** (often abbreviated as **G3P).** The structure of this molecule is shown on p147. G3P is a three-carbon sugar and it is the starting molecule (**precursor**) for the synthesis of several sugars, not just glucose, and lipids. Amino acids used for protein synthesis can also be made using PGAL as a precursor, in addition to a source of reduced nitrogen.
The overall reaction of the Citric acid cycle is:

$$\begin{array}{ll} 3CO_2 \text{ (from the air or water)} & + 1 \text{ G3P} \\ + 9 \text{ ATP (from the light reactions)} \text{ ----------> } & + 9 \text{ ADP } + 8 \text{ } P_i \\ + 6 \text{ NADPH (from the light reactions)} & + 6 \text{ } NADP^+ \end{array}$$

The complete cycle is outlined in figure 8.17 on page 148.

Study Questions:

1. What role does "reducing power" play in photosynthesis? What molecules provide reducing power directly to the Calvin-Benson cycle? Where and how do these molecules obtain their reducing power?

2. Explain how the photooxidation of chlorophyll a is related to the reduction of CO_2 in photosynthesis. Trace the connection in general but accurate terms (i.e. you need not list every chemical in each pathway, but you need to list each pathway and discuss its significance.)

3. What is the product of the Calvin-Benson cycle? Why is this molecule of pivotal importance in the life of the plant?

4. **Rubisco** is one of the most important enzymes in the entire biological world. What does rubisco do that is so impressive? What is rubisco's full name? Explain this name in terms of the enzyme's function.

5. Explain the Calvin-Benson cycle in general terms. What is important about this cycle? What does it do? What are its products, what happens to them, and why are they important?

6. Based on what you know about the role of phosphorylation in chemical reactions, develop a hypothesis that explains why 3-phosphoglycerate is phosphorylated in the second step of the Calvin-Benson cycle? The phosphates come right off again in the next step. Why do you suppose the cycle doesn't simply convert 3-phosphoglycerate to glyceraldehyde phosphate in one step? This would save 6 ATP per cycle and would be of great adaptive advantage to the plant. Use an energy diagram to explain your hypothesis.

NEWS ITEM: Macromolecular complexes appear to be very common for proteins involved in photosynthesis. Darl-Heinz Süss has evidence that suggests that Rubisco is anchored to the thylakoid membrane via the ATP-synthase. If this is true, it demonstrates that many proteins may have a primary function (synthesis of ATP) and a secondary function (anchor rubisco). This is the kind of dual function allows a duplicated gene to give rise to two similar proteins with very different functions - the kind of variation that is critical to evolution. (See Süss. *Z. Naturforsch.* Vol. 45c: 633-637. 1990.)

Now the marijuana leaf has harvested sunlight energy and stored it in the nutrient G3P. As the first law of thermodynamics tells us, energy cannot be destroyed or consumed, it can only be converted to another form of energy. As a summary of the process of photosynthesis, let's briefly describe the harvesting of energy through the light reactions. Remember, you measure energy by the effect it has on matter, so to follow energy, we describe what gets "energized" during this process.

1. The electrons in chlorophyll get energized and jump to a higher orbital

2. These electrons pass across an electron transport system (ETS) and transfer their energy to the proton pumps, which use the energy to move protons up their concentration gradient. The energy of sunlight is now contained in the high concentration of protons in the thylakoid space.

3. The protons fall down their concentration gradient and transfer their energy to the ATP synthase, which energizes ADP by phosphorylating it to become ATP.

4. The high-energy electrons tumbling down the ETS in Photosystem I don't transfer all their energy to the proton pumps. Much of the energy remains in the electrons and is transferred to the $NADP^+$ as it becomes NADPH.

5. NADPH and ATP both contain much of the energy originally reaching the plant in sunlight. During the Calvin cycle, this energy is transferred to CO_2 (in the form of high-energy electrons and hydrogen ions) as it becomes G3P.

At every step in this process, the energy transfer is not 100% efficient; a percentage of the energy is not transferred to the next step but is given off as heat to the environment. This is the nature of energy transfers. They are never 100% efficient. However, the energy transfers of photosynthesis are among the most efficient.

Study Question:

1. Be able to explain the transfers of energy outlined in steps 1-5 above. Make sure you understand the nature of each energy transfer and the nature of energy transfers in general.

G3P has been synthesized and the marijuana plant uses G3P in the following ways:

1. The leaves send G3P to the mitochondria inside the mesophyll cells. G3P is oxidized in the mitochondria to CO_2 and H_2O. The energy released by this process is stored in ATP, which powers the living processes of these leaves.

2. Plants synthesize glucose, fructose, sucrose (a fructose-glucose disaccharide) and starch (polyglucose) in the chloroplast. The starch is a storage form of sugars that the plant can live on in times of darkness when photosynthesis cannot occur. The mono- and disaccharides are stored in the mesophyll as well, but are also transported to all the cells that do not photosynthesize (roots, stems, and flowers). These sugars are used:

   A. As an energy source - the sugars are burned for energy by these cells.

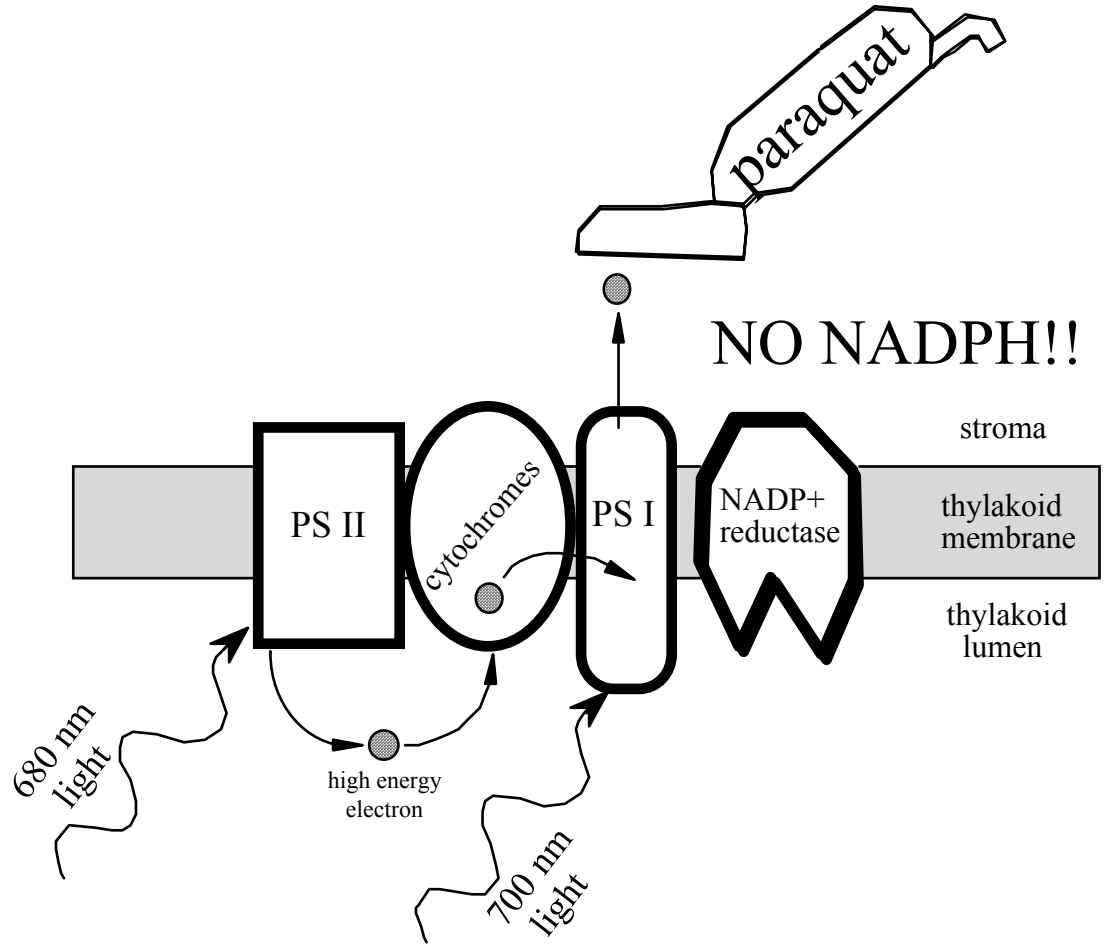
   B. As a source of glucose for the production of cellulose, the major structural component of cell walls.

   C. As a source of glucose for producing starch in non-photosynthetic plant cells.

   D. As a precursor for amino acids and nucleic acids made in the cells of the root.

See Fig 8.22 (p152) for a diagram of how the Calvin-Benson cycle fits in plant metabolism.

Back to our original question, “Why did the US government use paraquat on marijuana plants to kill them?" Paraquat is very, very electronegative and binds to a protein near photosystem I. When light hits chlorophyll at the reaction center, the electron is excited and sent to the primary electron acceptor. Rather than entering the electron transport pathway, it is sucked up by paraquat. As a result, no NADPH is produced. So.....? Later in this unit, we will see why it might be so harmful to humans.



Study Questions:

1. Why does paraquat kill marijuana? What affect would this have on marijuana's ability to make G3P?
2. If you wanted to design a weed killer what other steps in photosynthesis could be exploited?

The New York Times Magazine
November 19, 1978 (an excerpt from)

Poisonous Fallout From The War On Marijuana

by Jesse Kornbluth

Mexican marijuana growers had learned that parquat-drenched plants might still be sold as commercial-grade marijuana if they could be harvested before the herbicide turned the leaves brittle and the taste harsh. Because their illegal crop meant the difference between a subsistence income of $200 a year and a cultivator's income of as much as $5,000, the Mexicans unhesitatingly harvested the poisoned marijuana. And then they sold it to Americans.

The dangers of paraquat were no secret to the State Department. Swallowing as little as a half ounce is suicidal; paraquat gravitates to the lungs, where it causes such massive damage that death almost invariably occurs within two weeks. There is no known antidote. But whether paraquat that has been burned and then inhaled, produces those same deadly results was unknown. In 1975, when State started funding the Mexican program, there had been no inhalation studies. There would be none until 1977, when Senate investigators forced the issue.

This month, Secretary of Health, Education, and Welfare Joseph Califano announced the disturbing results of those tests: Heavy users of this tainted marijuana might develop fibrosis, an irreversible lung disease, and "clinically measurable damage" might befall less frequent smokers. In the furor that followed, the Administration explained that there was nothing it could do but warn smokers against Mexican marijuana - the Government of Mexico selected this herbicide independently, purchased it from a British company with its own funds, and sprayed marijuana mostly when opium-poppy fields, the true targets of the American-funded program, lay fallow.

Among the many accomplishments of the Mexican-American eradication program are these unforseen results:

• Contrary to the original, widely publicized White House announcement, this poisoned marijuana is generally indistinguishable from the ordinary Mexican product.

• Because of the distribution of patterns of Mexican marijuana, paraquat-sprayed marijuana is sold mostly on the West Coast to teen-agers, on the East Coast in ghettos, and across the nation to the estimated 200,000 Armed Forces enlisted personnel who smoke. These are the three groups least likely to have heard Secretary Califano's warning, or to believe it if they did.

• Conflicting statistics released by various Government agencies have caused widspread confusion. Secratery Califano's announcement indicated that one-fifth of the marijuana confiscated at the Mexican border had been contaminated by paraquat, some of it at concentrations 40,000 times greater thatn the Environmental Protection Agency allows for domestic use. In August, the Center for Disease Control tested paraquat-positive marijuana forwarded by PharmChem, the California laboratory which had received more contaminated samples than all other private labs combined. PharmChem's findings - that as much as 39 percent of its 10,000 samples were paraquat poisoned - had been widely publicized; when the CDC discovered that only 2 percent of this laboratory's "contaminated" samples were paraquat-positive, PharmChem rexeamined its testing procedures, found them to be inadequate, and suspended its operations. In the confusion which surrounded these developments, the CDC's warning against "paraquat test kits" - devices which might enable consumers to resolve their doubts at home - hurt sales of the one kit said to be reliable, a simple chemical test developed by University of Mississippi marijuana researcher Dr. Carlton Turner for Landis Labs of Horsham, PA. Last month, when the National Institute of Drug Abuse announced that paraquat was as prevalent and as potentally dangerous as Secretary Califano origianally indicated, this news went almost unreported.

----------STOP----------

Question #2 Why Do Vegetarians Eat Tofu?

Focused Reading: p 640-41 "nitrogen fixation…" stop at "Some plants
p 889 “Food provides…”, stop at "Animals need mineral elements…"
p 890 fig 50.5
WWW Reading: What are Tofu Pups?

Tofu is made from soybeans and soybeans are excellent sources of protein. It harvests the sun's energy, gives off oxygen for our consumption, AND stores nutrients in its seeds that contain an unusually high concentration of protein. As you know, protein gives us our structure, which allows us to function. Without protein, we can't produce any chemical reactions, pump any ions, phosphorylate any substrates, or send any electrical impulses (to name a few functions that depend on proteins). We animals need a constant source of protein in our diet. If you are a carnivore, you get much of your protein from meat (the muscle cells of other animals.) However, this is an inefficient source of protein. It takes at least 10 times more energy to create a gram of animal protein than it does to create a gram of plant protein. Therefore, with the human population explosion, and hunger and starvation a constant threat, it makes sense for humans to consume less meat and eat more plants in order to conserve the precious energy resources of the planet. Unfortunately, many plants are poor in protein, but the soybean is an exception.

A note here about "complete" and "incomplete" proteins. All 20 amino acids must be available to you on a daily basis in order for you to make the proteins you need to be healthy. You need to consume eight amino acids (the **essential amino acids)** in your diet every day (see figure 50.5, on p 890). From these, you can biosynthesize the other 12 (the **non-essential amino acids**), thus giving you all 20. [In case you are interested, the essential amino acids are isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, and valine.] Because most animals (especially vertebrates) are composed of the much same proteins you are, if you eat animal muscles, you will automatically take in the correct amino acids in about the right proportion for your dietary needs. [We’ll ignore fats and vitamins for now.] However, plants are quite different in their amino acid compositions and are present in proportions that vary significantly from humans. Therefore, if you eat only one kind of plant (say wheat, or corn, or rice), you usually get too much of some essential amino acids and not nearly enough of others. You will be eating "incomplete protein." Therefore, if you are vegetarian, you should eat legumes (peanuts, soybeans, garbanzo beans, navy beans, kidney beans, pinto beans, etc.) and grains (wheat, rice, oats, corn, etc.) These two types of plants provide a "complete" protein mixture by compensating for each other's missing essential amino acids.

Study Question:

1. When do you classify an amino acid as "essential?" In order to remain healthy, why must vegetarians eat meals containing both legumes and grains?

Focused Reading: p 635 “Autotrophs…”, stop at "How does…"

Sugars and lipids contain only carbon, hydrogen, and oxygen, like G3P. Therefore, G3P can be used as the precursor for the biosynthesis of carbohydrates and lipids without the addition of other elements. (As is true for all synthetic processes, these processes are ENDERGONIC and require an energy source in the form of ATP.) However, proteins, which are composed of amino acids, contain carbon, hydrogen, oxygen and nitrogen. Therefore, in order to synthesize amino acids, and therefore protein, the plant must have a source of nitrogen.

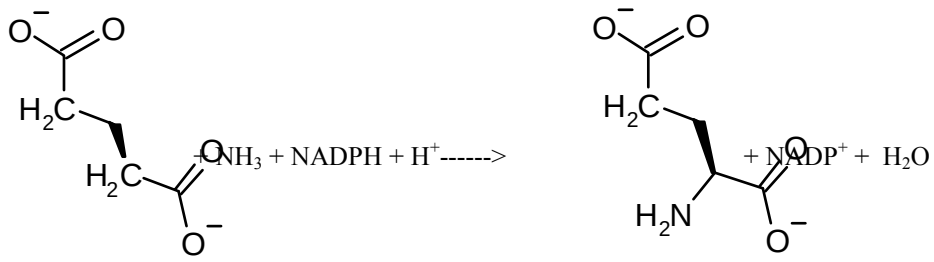
Focused Reading: p 465-66 "Nitrogen and Sulfur metabolism" stop at "Prokaryotes and their…"
p 641 "Some plants and bacteria…" stop at " Biological nitrogen fixation…"

The equation for nitrogen fixation is as follows:

$$N_2 + 8H^+ + 8e^- + 16ATP \text{ ----> } 2NH_3 + H_2 + 16ADP + 16P_i$$

As you can see, this is a redox reaction in which nitrogen is reduced; hydrogens are added to nitrogen. Thus, the reaction requires **reducing power**, which it gets from NADH produced during bacterial metabolism (see below). The reaction is also very endergonic requiring at least 16 ATP per reduced nitrogen molecule. (Some estimates of the overall energy requirements of nitrogen fixation place this figure at 25-35 ATP per nitrogen molecule.) Thus, the creation of amino acids, the raw materials of protein synthesis, is itself a costly endeavor for biological creatures. Nitrogen fixing bacteria contribute about 2 x 10^8 tons of ammonia (NH_3) to the soil each year for plant growth and produce many times more soil ammonia than is provided by agricultural fertilizers.

The soybean obtains nitrogen in the form of ammonium from the *Rhizobium* residing in its root nodules. Sucrose is transported to the plant roots where it is converted to alpha-ketoglutarate. Root cells synthesize the amino acid L-glutamate by combining alpha-ketoglutarate and ammonia as follows:



alpha-ketoglutarate L-glutamate

The amino acid L-glutamate can be used as a source of amino groups to make all of the other amino acids. The amino acids are transported all over the plant to meet its own protein synthesis needs. In the case of the soybean, amino acids are also supplied in large numbers to the developing soybeans. These soybeans, then, are a rich source of protein for humans and other animals when harvested.

And with this came the invention of soy-burgers and tofu pups - what some vegetarians eat on picnics and at ball games.

NEWS ITEM: The recommended daily allowance of vitamin E (10- 13.4IU) can be easily obtained in your daily diet. However higher levels of the vitamin (100-1000 IU) have been found to decrease your risk of heart disease and some cancers and improve the immune system. This level was almost impossible to reach without vitamin supplements--until now. Soybean oil is one of the main sources of dietary vitamin E and researchers in Nevada are working to bioengineer a better bean. A precursor of vitamin E is abundant in plant oils but most of it doesn't get converted into the vitamin. By cloning in a gene that overexpresses the necessary biosynthetic enzyme, scientists are able to shift the equilibrium of the reaction from precursor to product and increase the overall vitamin E content. (D. Shintani and D. DellaPenna *Science* 1998 vol 282 p 2098.)

Study Questions

1. Describe the various ways in which G3P is used by the green plant, in general terms.

2. In what form must nitrogen be supplied to plants in order for them to incorporate the nitrogen into amino acids? How is this form of nitrogen provided to non-legumes?

3. Describe the symbiotic relationship between legumes and *Rhizobium*. What does the legume gain from this relationship? What does the *Rhizobium* gain?

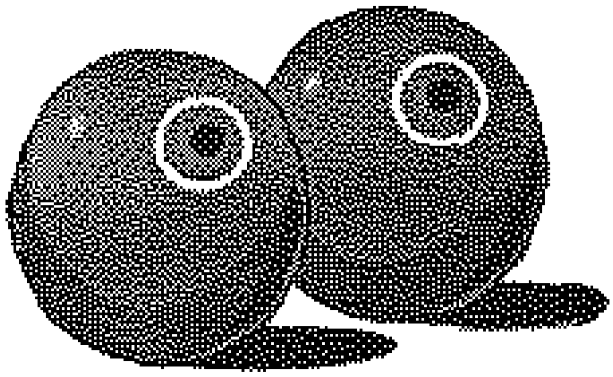
4. Describe the efforts of genetic engineers and selective breeding to increase the protein productivity of crops. Why is this work important? What is the problem with simply fertilizing crops to provide more ammonia and nitrates?

QUESTION #3: WHY IS CYANIDE THE TERRORIST'S POISON OF CHOICE?

The Cyanide Scare

A Tale of Two Grapes

by Bill Grigg and Vern Modeland (excerpts from *FDA Consumer)*



March 1989 marked the most intensive food safety investigation in Food and Drug Administration history. Millions of tons of fruit became suspect when a terrorist, 6,000 miles away, apparently made good on a phone call threatening to poison this nation's fresh fruit supply. Fruit in stores was returned or destroyed, and shipments coming into the country from Chile were halted.

In Chile, seasonable fruit and vegetable exports are second in importance only to copper to the national economy. In the United States, the cost of the terrorist's call might reach $50 million - the estimated value of 45 million crates of nectarines, plums, peaches, apples, pears, raspberries, strawberries, blueberries, and table grapes that faced destruction.

How did it happen?

****

Since it was his turn as duty officer, Dick Swanson wasn't surprised when the black box on his belt beeped at 7:20 p.m., Friday, March 3. Ever since the 1982 Tylenol tampering crisis, his wife only half counted on him on Fridays. A second beep sounded as he reached his door, so he headed straight to the telephone and called the number that had appeared on the beeper. A U.S. Customs official came on the line. He told Swanson that a cable from the U.S. Embassy in Santiago, Chile, had informed Customs:

ON MARCH 2 AT 1550 HOURS AN EMPLOYEE OF THE AGRICULTURE PUBLIC HEALTH INSPECTION SERVICE RECEIVED A CALL FROM A SPANISH SPEAKING MAN, WHO SOUNDED MIDDLE AGED AND WHO SPOKE WITH AN UNEDUCATED ACCENT. THE MAN STATED THAT FRUIT BEING EXPORTED TO BOTH THE UNITED STATES AND JAPAN WILL BE INJECTED WITH CYANIDE... IN ORDER TO FOCUS ATTENTION ON THE LIVING CONDITIONS OF THE LOWER CLASSES IN CHILE. HE FURTHER STATED THAT TOO MANY PEOPLE IN THE COUNTRYSIDE WERE STARVING DUE TO INCREASED LIVING COSTS AND WERE UNABLE TO BUY SUFFICIENT FOOD TO SURVIVE.

The caller said killing policemen and placing bombs had not solved the problem and he wanted to involve other countries. Although the Manuel Rodriguez Patriotic Front and the Leftist Revolutionary Front had been attacking policeman and placing bombs to bring about changes in the country and government of Augusto Pinochet, the caller did not say if he was involved with either group.

Saturday, FDA Commissioner Frank E. Young, M.D., Ph.D., and others met at FDA headquarters in Rockville, MD. They continued to confer on Sunday. But by Monday, the State Department had concluded the telephone call was “probably a hoax.” FDA then released news of the call and State’s view of it as a likely hoax. FDA said fruit had been temporarily held but was moving again. Few newspapers reported FDA’s announcement. The crisis appeared over.

****

The terrorist called the embassy in Santiago again on the eighth of March, and again on March 17, warning that the March 2 threat was no hoax.

FDA began to step up inspections, mostly at the Port of Philadelphia, where 80 percent of all Chilean fruit imported by the United States arrives.

First to be inspected would be the Almeria Star, which had sailed Feb 27 from Santiago with 364,000 boxes of fruit in her holds. On Sunday, March 12, investigators began examining a representative 12,000 boxes of fruit.

To examine the mountain of Chilean fruit, the FDA Philadelphia district office needed extra help. Among those assigned to the temporary duty was William Fidurski, from FDA’s North Brunswick NJ, resident inspection post. He was one of some 40 FDA people assigned to inspect fruit at the Tioga Fruit Terminal in Philadelphia.

“They were right on top of the box,” Fidurski recalls. The red seedless grapes were discolored. They had damaged skins. That’s about all he remembered about them, out of the 2 million grapes FDA investigators saw that day.

Being careful not to disturb anything in the box, Fidurski turned the crate over to his supervisor. It went, among others containing damaged or discolored fruit, to the FDA Philadelphia laboratory for closer examination. There, color photos were taken that showed rings of a crystalline substance surrounding what might be puncture sites. The grapes then were sliced carefully and placed in small glass flasks. In the flasks, the slices were squeezed with a glass rod to release the juice, and a solution of dilute sulfuric acid was added. Sulfuric acid will cause chemical changes to cyanide compounds, releasing hydrogen cyanide. This “cyantesmo test” would detect the presence of as little as 10-millionths of a gram of cyanide. Within minutes, it did. The analysts then did a Chloramine T test, which produces a pink-purple color in a reactive solution. The second test confirmed results of the first.

Those two red grapes contained cyanide in amounts far too small to cause death, or even illness, to anyone eating them. And, because crystalline potassium cyanide and sodium cyanide change to hydrogen cyanide gas in acid fruit which can then dissipate, FDA scientists couldn’t determine how much of the poison might have originally been injected into the grapes. But, cyanide *was* present.

FDA Commissioner Young said, “Very low levels. Very low... 0.03 mg vs. 20 mg to hurt an adult.” The newly confirmed Secretary of Health and Human Services Sullivan was briefed. The many political and financial ramification of a quarantine were discussed. They agreed that HHS and FDA weren’t charged with foreign policy considerations and commerce.

A news release was drafted, in case it was needed. Copies were passed around the table and quickly approved: "The FDA said today it has found and confirmed traces of cyanide in a small sample of seedless red grapes from Chile and as a result, is detaining all grapes and other fruit from that country...." The news was made public on the evening newscasts on March 13.

Why is cyanide poisonous? How does it kill people? To answer these questions, we need to learn how all organisms generate ATP from sugars like glucose.

Non-photosynthetic organisms are called **heterotrophs** (*troph* = to feed on; *hetero* = other; therefore, "one who feeds on others") as opposed to photosynthetic organisms which are called **autotrophs** ("ones who feed themselves.") Animals, many bacteria, most protists, and non-photosynthetic plant cells (roots, stems, flowers) must get ATP by non-photosynthetic means. These means are called **fermentation** and **cellular respiration**. Fermentation does not require the presence of oxygen (we'll cover this later), while, as the name implies, cellular respiration does. While all nutrient molecules can be burned to obtain energy, by far the molecule most frequently used for this purpose, under normal circumstances, is the monosaccharide **glucose**. Glucose is the predominant sugar in human blood. Homeostatic mechanisms maintain the plasma glucose concentration at about 80 mg per 100 mls of blood.

Focused Reading: p 108 "Metabolism and the…" stop at "Enzyme activity is subject…"
p 145 Fig 7.1--note relationship between autotrophs and heterotrophs
p 43-46 "Carbohydrates…" stop at "Derivative carbohydrates…"

The energy of sunlight has been harvested and stored in the glucose molecule through the process of photosynthesis. Humans eat sugars and complex carbohydrates, which are converted to glucose for consumption by the cells. The energy stored in glucose by the green plant is thus released to the cell (and converted to ATP) during the process of cellular respiration.

Complex carbohydrates come in three varieties -- **starch**, **cellulose** (made by plants), and **glycogen** (made by animals). All of these polysaccharides are polyglucose. Because they contain **alpha-glycosidic** linkages, people can break down glycogen and starch to glucose molecules that serve as fuel for the cells of the body. Because it contains **beta-glycosidic** linkages, cellulose cannot be broken down to glucose by us. You do not have the enzyme required to break the beta-glycosidic linkage, therefore lettuce, celery, carrots, broccoli, etc. actually contain thousands of calories but you can't get at these calories because you can't break down the primary bulk of the vegetables -- cellulose. So it simply passes through your body as "roughage."

The overall equation for cellular respiration is:

$$C_6H_{12}O_6 + 6O_2 + 6H_2O \text{ -----> } 6CO_2 + 12H_2O + \text{ energy (ATP + Heat)}$$

You will immediately recognize this as the reverse of photosynthesis. Photosynthesis is an endergonic reaction with a ΔG of +686 kcal/mole. Conversely, cellular respiration is an exergonic reaction with a ΔG of -686 kcal/mole. Thus, for every mole of glucose oxidized by the cell, 686 kcal of energy becomes available to do cellular work. However, as in the case of photosynthesis, each of

the many energy transfers in cellular respiration is inefficient. Thus, about 254 kcal of this total energy is given off as heat while only about 432 kcal is successfully stored in ATP. This heat is definitely used by humans to maintain their body temperature, and cannot be considered “wasted” energy. However, in warmer environments, much of the heat is "dumped" into the air by cooling mechanisms (most notably perspiration). Nonetheless, the transfer of energy from glucose to ATP is about 63% efficient (432 kcal stored out of 686 available). By the standards of other biological processes as well as those of human-built machines, this is an extraordinarily efficient process.

In addition to being highly exergonic, cellular respiration, like photosynthesis, is a **redox reaction**.

Focused Reading: p 115-16 "Obtaining energy and electrons…" stop at "An overview…"

Cellular respiration happens in two basic processes:

1. The oxidation of glucose and water -- glycolysis and **the citric acid cycle** (also called the **Krebs cycle** after the scientist who first described it)

2. The reduction of oxygen -- **oxidative phosphorylation**

While ATP is synthesized from ADP and P_i ($P_i = H_2PO_4$) throughout both halves of cellular respiration, the vast majority of ATP is synthesized during oxidative phosphorylation. During glycolysis and the citric acid cycle, the hydrogens (high-energy electrons plus protons) are removed from glucose, which becomes CO_2 in the process. This equation is:

$$C_6H_{12}O_6 + 6H_2O \text{ ------> } 24e^- + 24H^+ + 6CO_2$$

The 24 hydrogens that are removed from glucose and water come off as 24 high-energy electrons plus 24 protons. As in the case of photosynthesis, the 24 electrons and some of the protons are transferred to carrier molecules. In cellular respiration, this carrier molecule is NAD^+ (the same molecule as NADP but minus one phosphate) and, in one case, FAD. NAD^+ and FAD pick up the 24 electrons from glucose and water that are released during glycolysis and the citric acid cycle. As in the case of photosynthesis, the hydrogens picked up by NAD^+ and FAD are high-energy electrons plus protons. NAD^+ becomes NADH by picking up two electrons and one proton while FAD becomes $FADH_2$, by picking up two electrons and two protons. These high-energy electrons are carrying the energy that was originally carried to the earth as photons.

In the second part of cellular respiration, called oxidative phosphorylation, oxygen is reduced to water. The equation is as follows:

$$24e^- + 24H^+ + 6O_2 \text{ ----> } 12H_2O$$

The 24 electrons required to reduce oxygen are donated from the carrier molecules NADH and $FADH_2$, which picked up the hydrogens during glycolysis and the citric acid cycle. Some of the 24 protons come directly from NADH and $FADH_2$, while others come from the pool of H^+'s in the cytoplasm. This should all sound vaguely familiar. Same idea as photosynthesis -- shuttle high-energy electrons plus protons (hydrogens) from one molecule to another using a dinucleotide (e.g.

NAD) as an intermediate. Nature has a few good ideas and they provided a selective advantage over and over again.

One idea that should not escape you in all this is that electrons are not all equal in energy level. When an electron is bound into covalent bond between hydrogen and carbon (as in glucose, amino acids, lipids, etc), it has a relatively high energy level. When it is bound into a covalent bond with oxygen, as in water, it has a relatively low energy level. Thus, the transfer of hydrogens from a carbohydrate (forming CO_2) to oxygen (forming H_2O) constitutes an exergonic process in which the energy level of the electrons falls. This loss of energy is converted to ATP and heat which is released into the environment.

Study Questions:

1. Explain the process of homeostasis in relation to thermodynamics. What is it and why is it important?

2. In this unit on bioenergetics, you have now encountered the four biological processes that yield the ATP that living creatures use to power their lives. What are they?

3. Analyze the cellular respiration equation as a redox reaction. What is being reduced? What is being oxidized? During which processes do each of these reactions occur? Be able to do the same for the photosynthesis equation.

4. Explain how hydrogens are shuttled from one process to the other in cellular respiration. What molecules do the shuttling?

5. Explain the concept of high-energy electrons storing energy. How is this energy released? How is it stored in the first place?

----------STOP----------

So, how is this all accomplished? Cells in our body get glucose from the blood. The concentration of glucose is always very low in the cytoplasm for two reasons: 1) glucose is constantly being burned for energy and 2) as soon as glucose enters a cell, it is immediately converted to glucose-6-phosphate (whether it enters glycolysis or not.) Glucose-6-phosphate is not the same as glucose -- thus glucose is removed from the cytoplasm by phosphorylation and the cytoplasmic glucose concentration remains very low.

Being a hydrophilic organic molecule and not a simple ion, glucose must cross a cell's plasma membrane by being transported by a glucose transport protein and not a glucose channel. However, because the concentration of glucose is higher outside the membrane (in the blood) than in the cytoplasm, the process can be passive (not requiring ATP.)

Focused Reading: p 85-6 "Physical nature…" stop at "Osmosis…"
p87-8 "Diffusion..." stop at "Active transport…"

WWW Reading: Relative Sizes

The model for the glucose transport protein is very much like the active transport models you studied in Unit I. However, there is only one substrate binding site on the molecule -- a site highly specific for glucose. The affinity of this site does not change as the protein opens to the inside and then the outside of the membrane. Let's say the concentration of glucose is 100 fold higher outside a cell than inside. Thus, when the glucose transporter is open to the outside the cell, it is 100 times more likely that a glucose will hit its binding site on the transporter and stick before the transporter flips to the inside. After the transporter flips, because the bonds between glucose (a ligand) and its transporter are weak, glucose wiggles free through its own kinetic energy. Now it is 100 times less likely that a glucose molecule from the cytoplasm will bind to the site before it flips back to the outside. After it flips, it is 100 times more likely that glucose will bind.... Thus, for every glucose molecule that is transported outward, 100 are transported inward and the net transport is inward without the expenditure of energy in the form of ATP. One more thing, the glucose transporter does have one additional site on it for allosteric modulation. This causes the transporter to flip faster or slower, allowing the rate of transport to be increased or decreased.

Molecular oxygen is hydrophobic because the double bonds between the two oxygens are not polar; both oxygens have equal affinity for the electrons. Therefore, oxygen can enter a cell by passive diffusion across the phospholipid bilayer. Since oxygen is constantly being converted to water by cellular respiration, the oxygen concentration in the cytoplasm is lower than in the blood outside the cell. Therefore, oxygen enters down its concentration gradient. In fact, this is a self-regulating system since a rapid consumption of oxygen due to increased cellular respiration increases the concentration gradient across the plasma membrane and causes oxygen to enter a cell at a faster rate by passive diffusion.

Study Questions:

1. The transport of glucose into most mammalian cells is a passive process. Explain the conditions that make it possible for glucose to cross the membrane without the expenditure of energy in the form of ATP.

2. Explain the passive transport process of glucose. In what ways does it differ from active transport? How can the rate of passive transport of glucose be changed?

3. Explain how the delivery of oxygen to cells is a self-regulating process that adjusts as the rate of cellular respiration changes.

4. Along the lining of your intestines, there a different glucose transporter, a symporter, that uses the Na^+ gradient to power glucose uptake. (See Fig 5.13 p 90) Why is this necessary?

Once the glucose is inside the cytoplasm, it can be oxidized for energy.

Focused Reading: p 68 "Mitochondria are…" stop at "Plastids photosynthesize…"
p 31-2 "Functional groups…" stop at "Isomers have…"

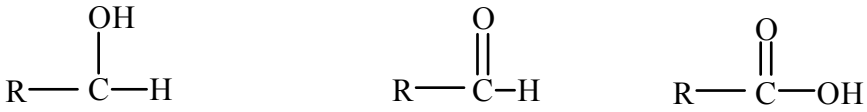
p 31 Table 2.20
p 116-25 “In the presence...” stop at “The respiratory…"
In lab we worked with IDH--Find the Krebs cycle step catalyzed by the mitochondrial form of IDH

WWW Reading: Glycolysis Summary

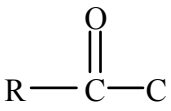
Look over all the steps in glycolysis and the Krebs cycle and try to understand each one. While you do not have to memorize these steps, you will understand the overall concepts a lot better if you have some understanding of the individual steps in the process.

Here are some chemical rules that will help you understand glycolysis and the Krebs cycle. In general, • alcohols end in **-ol** (e.g. ethanol, butanol, and estradiol)
- aldehydes end in **-aldehyde**; (e.g. formaldehyde)
- ketones end in **-one** (e.g. cortisone, acetone)
- acids end in **-ic acid** or **-ate** (e.g. carbonic acid/carbonate; phosphoric acid/phosphate.)

	oxidation		oxidation	
R-COH	-------->	R-CO	--------->	R-COOH
R-COH	<-------	R-CO	<--------	R-COOH
	reduction		reduction	



alcohol aldehyde acid



or ketone

When alcohol groups are oxidized, they become aldehydes or ketones. When aldehydes or ketones are oxidized, they become acids. Conversely, when acids are reduced, they become aldehydes or ketones, which become alcohols when they are reduced. For example:

	1		2	
CH_3-**C**H_2OH	<--------->	CH_3**C**HO	<-------->	CH_3**C**OOH
ethanol		acetaldehyde		acetic acid

The forward reaction is oxidation while the reverse is reduction. In reaction 1, the oxygen in the **hydroxyl group** breaks its bond with hydrogen and the carbon breaks its bond with one of its hydrogens, and carbon and oxygen form a double bond. This carbon=oxygen double bond is a **carbonyl group**. [If this occurs at the end of a molecule, it is an aldehyde group; if it occurs anywhere but the end, it is a ketone group]. This is the loss of hydrogen, or oxidation.

In reaction 2, the carbon breaks its attachment to the hydrogen and bonds with a hydroxyl group. When a carbonyl and hydroxyl are bonded to the same carbon, this is an **acid group**. During this process, the ketone or aldehyde gained an oxygen -- thus, this process is oxidation.

Sugars contain one carbonyl group and several hydroxyl groups. Therefore, they are not very highly oxidized or, in other words, they are highly reduced. Through the process of glycolysis and the Krebs cycle, more and more hydroxyl groups are converted to aldehyde, ketone and acid groups by the process of oxidation. Finally, the most highly oxidized form of carbon is produced -- carbon dioxide.

One more rule that may help, if a compound ends in -ate, it is the ionized form of an organic acid. For instance, because acetic acid is an acid, when you put it in water, it "donates" a proton as follows:

$$\underset{\text{acetic acid}}{CH_3COOH} \text{ <------> } \underset{\text{acetate}}{CH_3COO^-} + H^+$$

Thus, when you call a molecule glutamate, or pyruvate, or oxaloacetate, you are indicating that the molecules are acids that have ionized. In their non-ionized forms, they are glutamic acid, pyruvic acid, and oxaloacetic acid. Biologists frequently use the ionized and non-ionized names interchangeably, so don't be thrown off by this.

Study Questions:

1. What is the difference between substrate level phosphorylation and oxidative phosphorylation? What is being phosphorylated in each process?

2. For what purpose is ATP spent during the first few reactions of glycolysis? If glycolysis is supposed to yield energy, not cost energy, why is the cell doing this?

3. While glycolysis is considered to be a redox process, really only one step in the pathway is a redox reaction. What happens at this step?

4. Given just the names of compounds in reactions, be able to determine whether the reaction is an oxidation or a reduction. For example: formaldehyde to formate; phosphoglyceraldehyde to phosphoglycerate.

5. Explain the difference in the chemical structure of a molecule whose name ended in "-ate" as opposed to "-ic acid".

6. What is the overall reaction of glycolysis? What goes in and what comes out? What is the fate of all products?

7. What is the overall reaction of the Krebs cycle (including pyruvate oxidation)? What goes in and what comes out? What is the fate of all products?

8. What is cyclical about the Krebs cycle? Explain, in general, how carbons cycle through this pathway.

9. If you had to summarize the processes of glycolysis and the Krebs cycle in the simplest terms, how would you describe it? If your life depended on clearly conveying what happens in these processes in two or three sentences, what would you say?

10. How do our cells obtain glucose from the blood? Does this process require the expenditure of ATP? Explain.

11. While the overall reactions of glycolysis and the Krebs cycle yield energy, the process also costs some cellular energy in the form of ATP. What steps in the process require energy and why? By how much is the total ATP yield reduced by these endergonic steps?

NEWS ITEM: A group of researchers at Duke University have located a second protein that interacts with huntingtin, and it also interacts with HAP-1. The "new" protein is glyceraldehyde-3-phosphate dehydrogenase, the first enzyme in the "energy harvesting half" of glycolysis. This is the first protein in the HD story that has a known function. The scientists are entertaining the idea that HD and four other less common neurodegenerative diseases have reduced energy production due to a molecular interference with glyceraldehyde-3-phosphate dehydrogenase. (See related summary; Marcia Barinaga. *Science.* Vol. 271, pp1233-1234. 1 March, 1996.)
A group from Syracuse University has used antibodies to localize the enzymes involved in glycolysis in *Drosophila* flight muscles. Surprisingly, these "cytoplasmic" proteins were found spaced in regular intervals over the striations in the muscles. When mutations were made in these enzymes so that they were still functional but no longer located over the striations, the *Drosophila* was no loner able to fly. Therefore, glycolytic enzymes are necessary for energy production but it appears that this production must be located in specific areas inside some cells in order for the cells to function properly. (*Molecular Biology of the Cell* Vol. 8: 1665. September, 1997)
----------STOP----------

For every one glucose molecule and six molecules of water that enter glycolysis and the Krebs cycle, a cell makes six molecules of CO_2. This CO_2 is hydrophobic and it leaves the cell by passive diffusion across the lipid bilayer. As in the case of oxygen, increased levels of CO_2 in the cytoplasm (which would occur if cellular respiration rates increased) would increase the concentration gradient. This would in turn increase the rate at which CO_2 diffuses out of the cell. Thus CO_2 elimination is a self-regulating process, too.

To make this CO_2, cells transfer 24 hydrogens (24 high-energy electrons plus 24 protons) to carrier molecules, two at a time. You need 12 carriers to transfer 24 electrons: 10 NADH and two $FADH_2$. While these electrons have lost some of the energy they had when they were in glucose, they haven't lost very much, and they continue to be "high-energy."

In addition to the 24 hydrogens, we have a net synthesis of four ATP (two from glycolysis and two from the Krebs cycle) produced by **substrate level phosphorylation**. These four ATP are a net gain and can be used by the cell for anything it wishes. ATP made in glycolysis is in the cytoplasm ready to be used. The ATP generated in the Krebs cycle is in the mitochondria and can be used there or can be transported across the mitochondrial membrane into the cytoplasm for use there. Because ATP is in such high concentration inside the mitochondria, ATP can go down its concentration gradient on a transport protein into the cytoplasm by the process of passive transport.

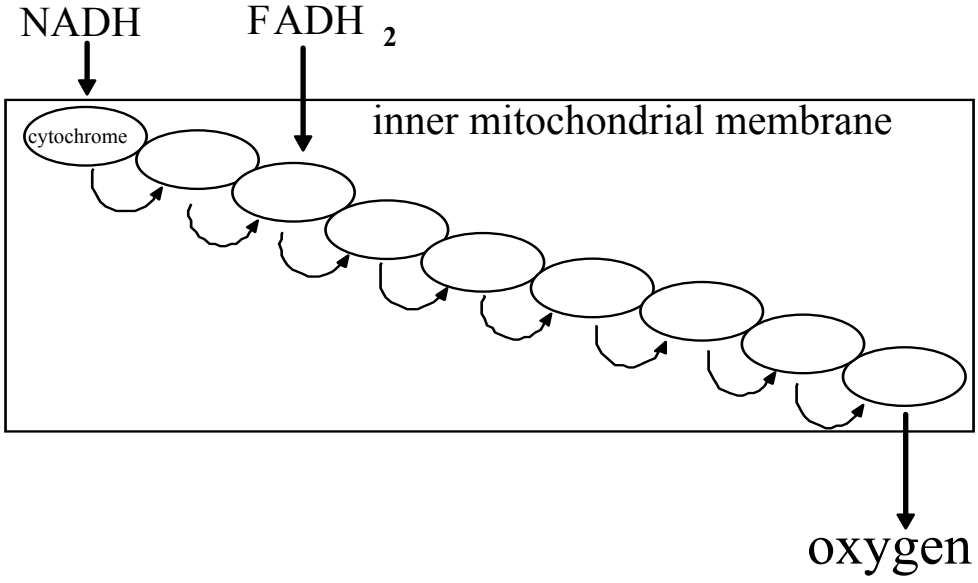
The majority of ATP is synthesized by the cell from the energy stored in the high-energy electrons found in NADH and $FADH_2$. The process of oxidative phosphorylation harvests this energy.

Focused Reading: p 125 "The respiratory chain:…" stop at "The respiratory chain transports…"
p 126-28 "Active proton…" stop at "Two experiments…"

WWW Reading: Animation of Photosynthesis

The process of ATP synthesis in photosynthesis and oxidative phosphorylation is essentially the same. This process must be very ancient, having evolved before plants and animals separated during evolution. In fact, because bacteria synthesize ATP this way too, it must be one of the most ancient "good ideas" in the biological world. Bacteria pump protons toward the outside across their plasma membranes. Protons then reenter the cell via an ATP synthase, and ATP is synthesized. Thus, in bacteria, the plasma membrane has a function that is equivalent to the inner mitochondrial and thylakoid membranes in eukaryotes.

NEWS ITEM: The study of apoptosis (programmed cell death) has become provided potential links to bioenergetics and to neurodegenerative disease. Some of the main players in apoptosis are proteases called caspases that help to kill the cell by degrading it from the inside out. Dr. Shimizu and colleagues have shown that caspases are activated when cytochrome C is released from the mitochondrion. When apoptosis is initiated, cytochrome C, which normally acts to help harvest energy, acts as a messenger of doom to activate the caspases so that they can destroy the cell. Caspases also appear to be activated in Huntington's disease. In this case, it seems that the presence of the CAG repeats activates the apoptosis cascade leading to neuronal death in both mice and humans. This finding leads to hope for treatment of HD since caspase inhibitors (injected into the cerebrospinal fluid) slowed progression of the disease. (See Nature 1999 vol299 p411 and 263.)



Study Questions:

1. Explain the process by which ATP is synthesized from ADP and P_i using the energy of the high-energy electrons from NADH and $FADH_2$. You need not memorize the name of the cytochromes, but you should understand the process and be able to explain it accurately.

2. What role does oxygen play in oxidative phosphorylation? Why is oxygen a good molecule to play this role (why not carbon, or neon, or hydrogen?)

3. Approximately how many ATP are synthesized in oxidative phosphorylation per glucose molecule?

Most cells in our body do not absolutely have to use glucose as a source of energy. They can oxidize lipids or amino acids to make ATP. [While most cells are like this, the cells of the brain must burn glucose -- no other fuel will do. Thus, if you suffer from low blood sugar (**hypoglycemia** (hypo = low; glyc = sugar; emia = in the blood), you may experience loss of ability to concentrate, to speak coherently, and even to stay conscious -- all signs of compromised brain function.]

If a source of lipids is available, a cell will burn lipids along with glucose for fuel. Fats are digested into glycerol and fatty acids--both of which can 'feed' into metabolism. Glycerol is converted into glyceraldehyde phosphate and used in glycolysis. This releases a little energy but most of the energy from fat is stored in the fatty acids. A process called **beta oxidation** breaks fatty acids into 2 carbon units that can enter cellular respiration as acetyl coA. If glucose and lipid levels are low, the cell will begin to burn amino acids for fuel. This can be detrimental to your health because the amino acids must be **deaminated** in order to be burned, and the brain and kidneys have a hard time dealing with the extra ammonia that is produced.

Focused Reading: p 130-32 "Metabolic pathways…" stop at "Regulating energy…"
p 49-51 "Lipids:..."; stop at "Phospholipids …"

Study Questions:

1. Fat stores more energy per gram than carbohydrates. What part of a fat molecule stores the most energy? How is that part broken apart so that components can enter metabolism?

2. How is glycerol burned for fuel? Where does it enter the cellular respiration pathway?

3. What must happen to amino acids before they can be burned as fuel? One common point of entry for amino acids into the Krebs cycle is at alpha-ketoglutarate. Refer back to the section on nitrogen fixation and the production of amino acids in this unit, and explain how an amino acid such as L-glutamate might enter the Krebs cycle at alpha-ketoglutarate.

Intestinal cells work very hard all the time and have a fairly constant metabolic rate. In contrast, cardiac myocytes have a fairly low metabolic rate when you are sleeping and a very high metabolic rate when you are exercising. Therefore, the rate at which glucose is burned must be regulated so that you don't waste energy (burning a lot of fuel when little energy is needed) or starve for energy (burn very little fuel when a lot of energy is required). All cells must be able to regulate the rate at which glucose is burned and ATP is created.

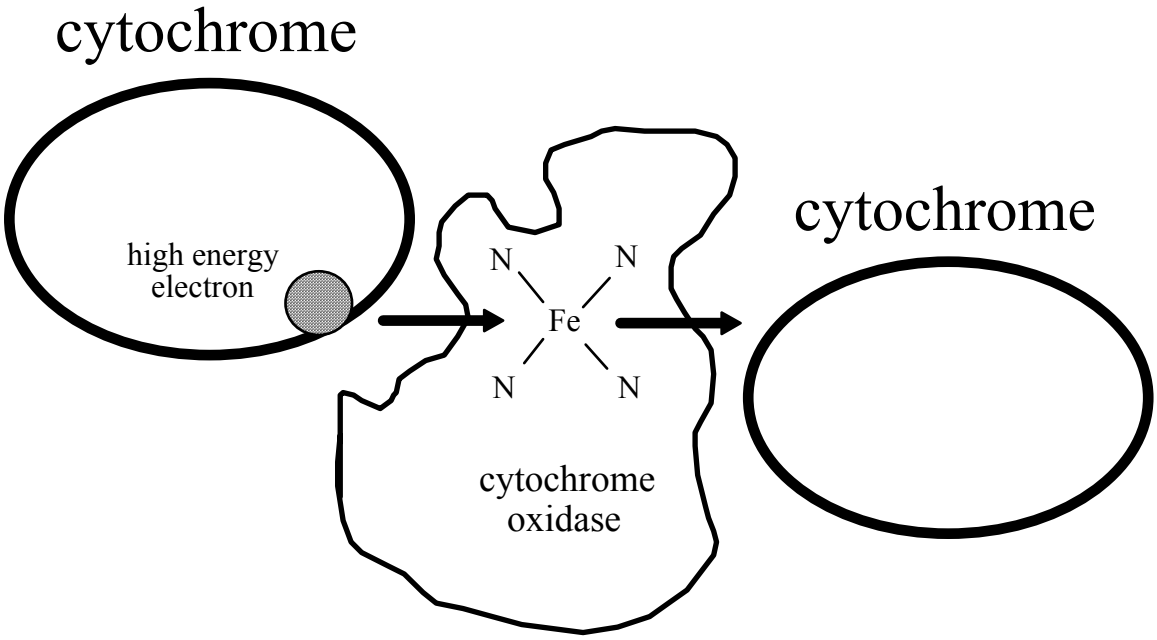
Focused Reading: p 132-4 "Regulating the energy pathways" to the end.

Study Questions:

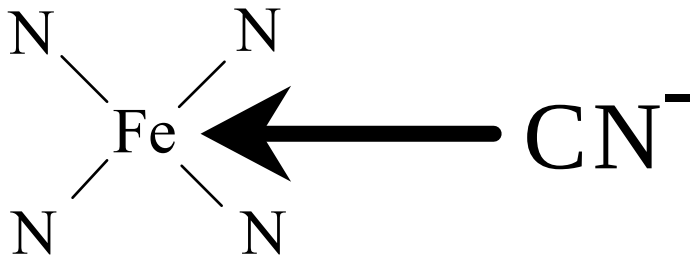
1. Explain in chemical terms how the rate of glucose oxidation is controlled by environmental conditions. Why is this evolutionarily adaptive?

2. Describe the structure of phosphofructokinase. How many sites binding does it need to perform its function? What molecules bind at each? Explain the name of the enzyme. What makes it a good enzyme to function as a rate regulator for cellular respiration?

Now that we understand how cells get energy from sugar, we can understand why cyanide is so lethal, and so popular with extortionists. Cyanide (its chemical formula is CN^-) has a negative charge, as the name indicates since it ends with the suffix "-ide". As you know from your basic chemistry, negative ions are attracted to positive ions (cations). Unfortunately, some of our vital enzymes use cations as a part of their structure. One class of enzymes that use iron ions is cytochrome oxidases. As the name tells you, they oxidize cytochromes by taking away an electron and these oxidases are located in the inner mitochondrial membrane. The high-energy electron temporarily binds to the iron in the cytochrome oxidase before the electron is passed onto the next cytochrome in the electron transport pathway, as seen in this diagram:

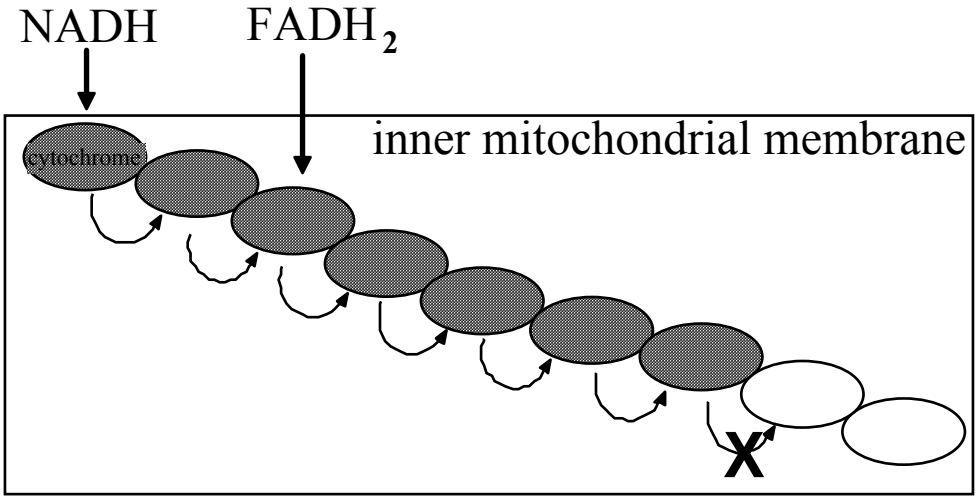


As a result, our enzymes, and our very lives, depend upon cytochrome oxidases being able to carry high-energy electrons temporarily. Cyanide has the unfortunate tendency to bind irreversibly to the iron ions in cytochrome oxidases. If the iron is occupied by CN^-, then it cannot accept another electron from a cytochrome that is carrying a high-energy electron.



But CN^- does not bind to all of the cytochrome oxidases, only the next to last one. So what's the problem??? As you know, most of the H+ ions are transported into the mitochondrial intermembrane space before this next to the last step of the electron transport pathway. How could missing out on the last two steps kill you?

Think of yourself in a bucket brigade where each person passes one bucket of water onto the next, and receives another bucket of water from the person "upstream". You are the next to last person passing on buckets in a long line of bucket passers. All of a sudden, the person you normally give your bucket to has stopped - he has been given an ice cold glass of sweet tea, and cannot be bothered to accept your bucket. What are the repercussions for every one else in this long line of bucket passers? Since you cannot get rid of your bucket, the person who normally passes a bucket to you cannot unload her bucket.... and a domino effect rushes backwards until every person in the long line is left holding a bucket of water with no one to accept it. This is what kills you with cyanide poisoning. The next to last cytochrome oxidase is gummed up with CN^-, therefore it cannot relieve a cytochrome of its high-energy electron and this clogs up the entire electron transport pathway.



Therefore, no H^+ ions are transported into the intermembrane space, no H^+ gradient is created, which results in an absence of the chemiosmotic generation of ATP. You die by a deprivation of ATP -- you run out of energy even though you have already generated lots of NADH and $FADH_2$.

Study Questions:

1. Given what you know about electron transport, why might paraquat be harmful to humans?

2. How does cyanide kill?

3. What do photosynthesis and cellular respiration have in common? How do they differ?

NEWS ITEMS: A single base pair substitution has been identified in patients who suffer from severe infantile lactate acidosis and encephalomyopathy. These symptoms were due to a genetic disease but surprisingly, the gene is not located in the nucleus. The base pair substitution occurred in a mitochondrial gene, which encodes one subunit of the mitochondrial ATP-synthase. The mothers of each patient contained 1:1 mixtures of wild-type and mutant mitochondrial DNA suggesting that each mother had inherited mutant DNA from their mothers too. (See *Houstek et al., Biochemica et Biophysica Acta*. Vol 1271: 349-357. 1995.)
In the 29 April, 1997 issue of *Proc. Nat. Acad. Sci.*, a team from UVa has identified a gene that is mutated in many patients with Alzheimer's Disease. This gene is located in mitochondrial DNA and encodes for one of the 13 proteins that make up cytochrome oxidase. These data are consistent with two points. 1) It has been suspected for many years that AD patients may suffer from poor energy metabolism. This would suggest that oxygen free radicals (O-) cause neuronal damage. 2) Children of mothers with AD are more likely to get AD than children whose fathers have AD. We inherit almost all of our mitochondrial DNA from our mothers. (*Science* Vol. 276: 682. May 1997)

4. Mitochondrial genes encode several components of the ATP synthase complex. One family has been identified that has a missense mutation in subunit a of the synthase and this gene is a
   mitochondrial gene (patients suffer from neurogenic muscle weakness). Draw a pedigree for this family.

5. Hypothesize why paraquat is probably toxic to animals.

Sour Grapes Land US in the Dock

by Dan Charles
excerpted from *New Scientist*, 16 March, 1991

The US Food and Drug Administration may have botched tests that appeared to detect cyanide in grapes from Chile two years ago. ON the basis of the tests, the US banned imports of all fruit from Chile for five days. Last month, Chilean fruit growers filed a legal claim against the US government, arguing that the FDA's mistakes in analytical chemistry cost them more than $400 million.

Manuel Lagunas-Solar, a radiochemist at the University of California, Davis, has spent the last two years injecting grapes with cyanide and trying to duplicate the FDA's results. From his research, which was paid for by Chile's fruit growers, one thing seems clear: the grapes were not contaminated with cyanide when they left Chile. Lagunas-Solar suspects that the grapes were never contaminated at all.

The central problem with the FDA’s results is that they found too much cyanide, says Lagunas-Solar. His tests show that the chemistry of grapes breaks down and detoxifies cyanide with remarkable speed.

The FDA detected 6.2 micrograms of cyanide in the pulp of the two grapes. According to Lagunas-Solar, this would mean large amounts must have been injected into the grapes just a few hours before the tests. But the grapes were on the docks in Philadelphia or in the custody of the FDA for longer than that before the tests were carried out. Working backwards, Lagunas-Solar estimates that a terrorist in Chile would have had to inject a minimum of 4000 micrograms of cyanide into the grapes in order to produce this result. It is more likely that ten times this much would be necessary, he says. But the larger of these quantities cannot physically be injected into grapes, and even the smaller amounts would have damaged the grapes and contaminated other grapes in the package.

The grapes the FDA analyzed were in good physical shape, and they did not find any other contaminated grapes, even in the same bunch. “We were able to rule out with confidence the hypothesis that cyanide tampering could have occurred in Chile, “ says Lagunas-Solar.

Bill Grigg, a spokesman for the FDA rejects Lagunas-Solar’s conclusions. The FDA’s own studies confirm that cyanide does disappear rapidly from grapes and other kinds of fruit. But in one FDA study, two grapes did retain large amounts of cyanide for between 3 and 6 days without having much effect on the look of the grapes. No one has been able to explain this result.

A further puzzle in the saga is that the FDA was also unable to find any traces of cyanide on the other grapes from the same bunch, even using their most sensitive techniques. Lagunas-Solar’s experiments show that traces of cyanide from contaminated grapes will show up throughout an entire crate of grapes.

----------STOP----------

QUESTION #4: Why would authorities ask you to update vaccinations after a flood ?

Bacteria are stunningly diverse and comprise an entire domain -- Eubacteria. It is impossible to cover the bioenergetics of this entire domain in any meaningful detail. However, brief introduction to the metabolic diversity of bacteria will help broaden your understanding of the variety of ways organisms can acquire energy.

Focused Reading: p 466-7 "A small minority…" stop at "Prokaryote…"
p 464-5 " Prokaryotes have…" stop at "Nitrogen and sulfur…"
p 469 Fig 26.12

WWWeb reading: http://www.cdc.gov/nceh/emergency/flood.html Choose 'Immunizations'

We will focus on one bacterium, *Clostridium tetanii*, the organism that causes tetanus. During the summers of 1993 and 1995, the Mississippi River and its tributaries flooded and caused billions of dollars in property damage and catastrophic losses for thousands of people in the Midwest. The Red Cross responded to this natural disaster by providing shelter, food, clothing, and tetanus vaccine. Why, in the midst of chaos and misery, did the Red Cross spend time and money delivering this vaccine (and what does this have to do with studying bioenergetics)? Well, herein lies the tale.

Tetanus causes all of the skeletal muscles of the body to contract into rigid paralysis. If untreated, the disease is fatal -- the diaphragm (the skeletal muscle that facilitates breathing) contracts into a rigid paralysis along with all the other skeletal muscles. Because it cannot relax, the victim cannot exhale and subsequently suffocates.

Tetanus is caused by a protein **toxin** (poison) released by the bacteria *Clostridium tetanii*. Because this toxin is released by the bacteria as a soluble molecule, it is called an **exotoxin**. Other bacteria (called **gram-negative** because they do not stain with a gram stain) such as *Salmonella* contain a toxic molecule in their outer membrane called **lipopolysaccharide** (or **LPS**). Because LPS remains bound to the bacterial membrane and is not released as a soluble product, it is called an **endotoxin**. Exotoxins are very dangerous and often lethal (e.g. tetanus, botulism, diphtheria, cholera, whooping cough), while endotoxins have lower levels of toxicity and are rarely fatal.

The exotoxin produced by *Clostridium tetanii* is called a **neurotoxin** because it attacks the nervous system. If the tetanus bacteria is growing somewhere in the body and is releasing this toxin, the toxin is carried throughout the body by the blood. When the toxin reaches the brain, it binds to and inactivates components within the membranes of neurons. The inactivation of these components inhibits the nerve impulse and blocks contraction of the muscles on the other side of synapse. While it is apparent that your brain causes muscles to contract, we often forget that your brain must also inhibit contraction (or cause relaxation.) For instance, in order for you to flex your arm, the muscles that extend your arm must relax. Otherwise, both sets of muscles would contract into a tug of war and your arm would be rigidly paralyzed. Tetanus toxin prevents the victim's muscles from relaxing and all movement is halted in rigid paralysis. (NOTE: Rigid paralysis can be contrasted with flaccid paralysis, a condition in which muscles cannot contract at all -- the body cannot move because it is limp or flaccid.)

Study Questions:

1. What is the difference between an exotoxin and an endotoxin?

2. What are the symptoms of tetanus? What happens at the cellular level to cause these symptoms?

The genus *Clostridium* also contains other pathogens (i.e. disease-causing agents) such as the organism that causes botulism (*Clostridium botulinum*), a form of severe and often fatal food poisoning as well as the organism that causes gas gangrene (*Clostridium perfringens*). Other *Clostridia* are non-pathogenic and are used to produce valuable fermentation products such as various alcohols and organic acids, or to fix atmospheric nitrogen. All bacteria in the genus *Clostridium* are soil bacteria and all are obligate anaerobes. Anaerobes harvest energy in the absence of oxygen. The metabolic pathways we have discussed so far need oxygen (that's why they are called 'cellular respiration'). So how do these bacteria generate ATP? They rely on a bioenergetic pathway that looks very familiar but has a different ending. They rely on **fermentation**, a metabolic pathway that oxidizes glucose to pyruvate using the reactions of glycolysis, producing NADH and ATP in the process (no oxygen required). Then, instead of having further energy harvested from pyruvate (via respiration), these organisms use NADH to reduce pyruvate to lactic acid or to ethanol and CO_2. The energy yield is less than aerobic cellular respiration but some ATP is harvested (2 ATP per glucose) and NAD is regenerated so it doesn't build up in the cell. (And a large industry has been built around the production of ethyl alcohol).

Focused Reading: p 129-130 "Fermentation..." stop at "Connections with other pathways"

Study Questions:

1. What is the difference between a facultative anaerobe and an obligate anaerobe? If you were to do a protein analysis of a facultative anaerobe and an obligate anaerobe, what differences would you find? In other words, what enzymes would you expect to find in the facultative anaerobe that would be missing from the obligate anaerobe and *vice versa*?

2. Explain the process of fermentation. The absence of oxygen is a requirement for the fermentation process. Explain why this is the case?

3. Compare and contrast the production of ATP through aerobic and anaerobic metabolic processes. How is ATP made in each process? Which process yields more usable energy for the cell? By how many fold? Explain. What are the end products of each process? Explain how these end products are produced.

4. Facultative anaerobes need a control mechanism that responds to presence or absence of oxygen. Based on what you know about molecular control systems, develop a reasonable hypothesis that describes such a functional control system for facultative anaerobes.

Because the *Clostridia* are obligate anaerobes, they are killed by oxygen. Thus, they must live in an environment in which oxygen levels are extremely low. While it is not clear how oxygen kills these microbes, the dominate hypothesis is that they are unable to detoxify (eliminate) the toxic by-products of oxygen reduction (hydrogen peroxide (H_2O_2), superoxide (O_2^-) and hydroxide radicals (OH^-). These by-products are toxic to all cells, but facultative anaerobes and aerobes contain enzymes that immediately destroy these substances as soon as they are formed, while obligate anaerobes lack these enzymes.

Getting away from oxygen on this planet is no small task. Thanks to the phototrophs, air is 20% oxygen, a lethal level for anaerobes. They must, therefore, live in places that are deprived of oxygen such as deep soil, sediments of rivers and lakes, bogs and marshes, canned foods, intestinal tracts of animals, sewage-treatment systems, or injured tissue which has had its blood supply interrupted. Because *Clostridium tetanii* lives in soil and the intestinal tract of animals, wounds that come in contact with dirt or animal feces are particularly susceptible to the development of tetanus. This is why people who work with animals professionally or as a hobby should be sure they have their tetanus vaccinations up to date. The common practice of bleeding a wound, especially a deep puncture wound, is a good one since the bacterium enters the body through a wound and blood carries oxygen to the area, which can kill the tetanus bacterium.

So why give tetanus vaccinations to flood victims? Because the floodwaters would wash the tetanus bacteria from soils, water treatment plants, and animal feces. If an open wound came into contact with this floodwater, it might become infected with the tetanus. Losing your possessions to a flood is one thing--losing your life is something else!

When *Clostridium tetanii* enters a wound, if the oxygen level is very low, it will begin to divide and produce a colony. This bacterial colony does not invade the body, but excretes the toxin that is carried from the wound into the body and eventually into the central nervous system. The exotoxin enters neurons by endocytosis (review on p 111) and travels retroaxonally (in reverse direction to nerve impulses) to reach the spinal inhibitory interneurons. Tetanus toxin is a protease (it degrades proteins) but it is a very selective one. Its substrate is VAMP - the integral membrane protein that facilitates neurotransmitter release. (Remember way back when we talked about neurotransmitter release? Why don't you go back and take a quick peak at Unit I. (For a good summary, see Montecucco and Schiavo *TIBS*. 18: 324-329. 1993.)) By blocking the release of inhibitory neurotransmitters, no muscles get the message to relax and get stuck in a contracted state. Tetanus vaccinations are aimed at this toxin, rather than the bacteria itself. The tetanus vaccine contains purified tetanus toxin, which has been denatured with formaldehyde. Because protein function is dependent on its 3-D structure, denaturation makes the toxin inactive. In this form, the toxin is called a **toxoid**. The immune system, however, will react to the toxoid in the same manner that it would a toxin. Thus, the body produces an immune response (antibodies) against tetanus toxoid which neutralizes the real toxin, should it ever be encountered, before it reaches the brain.

> NEWS ITEM: During the summer of 1997, millions of waterbirds died to a mysterious illness. The common symptom prior to death was flaccid paralysis and the disease was initially called "limberneck". The cause was eventually identified - an outbreak of botulism caused by *Clostridium botulinum*. (*Science* Vol. 278: 1019. November 1997)

Study Questions:

1. What kind of paralysis is caused by tetanus toxin and how does it cause paralysis?

2. Are we capable of anaerobic metabolism? If so, when and where?

3. What kind of toxin would cause "limberneck" and why would this be fatal?

As you know we are capable of burning glucose anaerobically for short periods of time. It has been determined that our muscles contain about 5 millimoles of ATP per kilogram. This ATP supply is depleted in a few seconds when we begin to exercise. After 10 seconds, we use ATP that has been

generated by an enzyme called phosphocreatin kinase that rips a phosphate from phosphocreatin and adds it to ADP. After one or two minutes of hard breathing, you will be using anaerobic metabolism (glycolysis) to generate ATP and lactic acid, which is why your muscles burn with extensive exercise. Eventually, this oxygen dept must be repaid so your muscles can return to aerobic metabolism, which is why we are obligate aerobes.

TWO RESEARCH QUESTIONS AND APPROACHES IN BIOENERGETICS

Focused Reading: p 331-2 stop at "Hemoglobin"
p 340-1 "Screening for abnormal…" stop at "There are several…"

Genetic Defects in Metabolism: Sometimes a mutation will occur in an organism that interferes with its ability to metabolize fuels properly. Because mutations alter genes, and each enzyme in the metabolic pathways is produced by one gene, each mutation in a single gene should alter or destroy only one metabolic enzyme. By studying the results of these mutations, some of the complexities of the metabolic pathways can be unraveled. For instance, 1 in 15,000 human infants is born with the inability to metabolize (oxidize) the amino acid phenylalanine. This condition is called **phenylketonuria** or PKU. Normally excess phenylalanine is converted to either fumarate or acetyl Co-A and is oxidized in the Krebs cycle for energy. However, this normal oxidation pathway cannot occur in people with PKU because ability to make the first enzyme in the pathway, phenylalanine 4-monooxygenase, is destroyed by a mutation in both alleles of this locus. Thus, excess phenylalanine must be eliminated by an alternative pathway in which it is converted to phenylpyruvate. Phenylpyruvate accumulates in the blood, is excreted in the urine, and causes irreversible brain damage and mental retardation. People with PKU can escape the effects of this disease by eliminating phenylalanine from their diet (including Nutrasweet).

Focused Reading: p 218-20 "One gene…" stop at "DNA, RNA, …"
p 219 Figure 12.1

Metabolic mutations (naturally occurring or induced by agents such as x-rays) in lower organisms, (e.g. yeast and bacteria) can be especially useful in helping us characterize metabolic pathways. For instance, in 1941, GW Beadle and EL Tatum induced a series of mutations in the mold *Neurospora.* (See Figure 12.1 on page 219 for a diagram of this study.) Normal or **wild type** *Neurospora* can synthesize all 20 amino acids. Therefore, it will grow on a simple medium containing only glucose and a source of nitrogen in the form of ammonia. Mutants, however, cannot grow on this simple medium but rather each mutant needed one nutritional supplement in order to survive. For instance, one mutant could only live if the amino acid arginine was added to the medium. This meant that this particular mutant was defective in one of the enzymes that synthesize arginine. Without arginine, the mold could make no proteins, and died. In looking further at the arginine mutants, Beadle and Tatum discovered that there were several different types of mutants. Class III mutants required arginine in the medium or it would die. Class II mutants could live on arginine or citrulline. Class I mutants could live on arginine, citrulline or ornithine. Thus, Beadle and Tatum were able to arrange the three terminal enzymes in the proper sequence of arginine synthesis. Since these experiments were performed, investigators have used metabolic mutations in other pathways to characterize many of the enzymatic steps in metabolism.

Spectroscopic Analysis of Cytochromes: The oxidation and reduction rates of the cytochromes can be measured using a spectrophotometer because the cytochromes change color as they become oxidized and reduced. This is very much like hemoglobin, the molecule that carries oxygen in

mammalian blood. When oxidized (by the gain of oxygen), hemoglobin is bright red. When reduced (by the loss of oxygen) the color shifts to a more bluish-red or reddish purple. Because spectrophotometers can analyze the wavelengths of light being absorbed by a substance and since the wavelength of absorbance changes as the color changes, the oxidized and reduced forms of the cytochromes produce different absorption spectra. When the cytochromes are spectroscopically analyzed across time, their rate of oxidation and reduction can be measured. The affect of various factors on the rate of electron transport can then be tested.

Study Questions:

1. If you are given data similar to that in figure 12.1 (p219), be able to determine the order of genes/proteins in a metabolic pathway.

----------STOP----------

Unit IV: Other Interesting Topics

In this Unit, we will cover:

I. Cancer
II. AIDS
III. Genetic Engineering

I. Cancer

WWW Reading: Cancer Statistics (20-year trends)

One person in five in the developed world will die of some form of cancer. For 25 years, cancer research has been among the top priorities of the biomedical research community in the United States. We have learned much about this disease and, in the process, about the function of normal cells.

Focused Reading: p 342-3 "Cancer Cells..." stop at "Some cancers…"

Cancer is defined as the presence of a **malignant tumor** in the body. A **neoplasm** or **tumor** is a relentlessly growing mass of abnormal cells that are dividing in defiance of normal restraints on growth. However, most tumors are **benign**, that is, all of the cells of the tumor remain in the tumor mass and do not invade other tissues. Benign tumors are not cancerous, but they can be life-threatening if they occur in places in the body from which they cannot be removed without causing serious damage (e.g. some places in the brain. Such tumors are said to be "inoperable.") However, most benign tumors are not life threatening, and can be easily treated by surgical removal.

A tumor becomes **malignant** or **cancerous** when its cells invade the other tissue(s). **Invasiveness** usually implies that the cells of the tumor can break loose, travel to a new site in the body through the blood or lymph, and establish secondary tumors. Such a tumor has **metastasized**. The process of spreading is called **metastasis** and it is the hallmark of cancerous tumors. While benign and pre-metastatic cancerous tumors are relatively easy to cure by surgery or localized radiation or chemotherapy, metastasized tumors are very difficult to treat. Because cancer arises from a single cell that is growing out of control, in order to cure the metastasized cancer, every single cancerous cell in the body must be destroyed. This is virtually impossible in widely disseminated cancers. The type of cell that becomes cancerous defines the name of the cancer.

Focused Reading: p 695 "Tissues…" stop at "Nervous tissue…"
p 695 Fig 40.2

Animal tissues come in four varieties: 1) Epithelia lines the inside and outside surfaces of the body (e.g. skin, lungs, blood vessels, etc), and provides the bulk of functional cells in internal organs (e.g. endocrine glands, liver, pancreas, kidney). Cancers of epithelial cells are called **carcinomas**. 2) Connective tissue is a very broad category of tissue, which includes blood, bone, cartilage, fat, tendons, ligaments, and the strong protein fibers that hold all the organs together. Cancers of the connective tissue cells are called **sarcomas**. Cancer of the white blood cells (the **leukocytes**) is called **leukemia**. 3) Muscle forms the mass of the skeletal muscles, creates the walls of blood vessels and internal organs (smooth muscle) and forms the wall of the heart (cardiac muscle).

Cancers of muscle cells are also called **sarcomas**. 4) Nervous tissue forms the brain, spinal cord and nerves in the body. Cancers of the nervous system are called **neuromas** if they involve actual neurons, and **gliomas** if the involve the supporting cells of the nervous system.

While these terms define broad categories of cancer, each type of cancer has its own distinguishing name. For instance, basal cell carcinoma is a kind of skin cancer caused by the cancerous growth of a basal cell in the skin (an epithelial cell). Melanoma is a different form of skin cancer caused by the cancerous growth of a melanocyte, the pigment producing cells of the skin. Both types of cancers are carcinomas, but they have very different characteristics, the former being very easily treated and almost never fatal while the latter is much more life threatening. Most cancers (85%) are carcinomas, and, in fact, an agent that causes any type of cancer is said to be a **carcinogen** or to be **carcinogenic**.

Below you will find a list of some of the most prevalent cancers in the United States. Lung cancer has the distinction of having the highest incidence of any single cancer at over 150,000 new cases per year and the highest death rate at 87%. (Actually, this method of calculating the death rate understates the threat of this disease. Over 90% of lung cancer victims will die within one year of diagnosis.) The other feature that distinguishes lung cancer is that, of all the cancers listed below it is by far the most preventable. The vast majority of lung cancer victims smoked cigarettes or lived with a heavy smoker. As in the case of AIDS, this sadly preventable disease continues to claim lives needlessly. Every year, lung cancer kills three times as many people as died in the Vietnam War.

Cancer Statistics, US

Site of Cancer	New Cases in 2000	Deaths in 2000	# Male Deaths, 2000	# Female Deaths, 2000
Colon	93,800	47,700	23,100	24,600
Lung and bronchus	164,100	156,900	89,300	67,600
Pancreas	28,300	28,200	13,700	14,500
Skin	56,900	9,600	6,000	3,600
Breast	184,200	41,200	400	40,800
Cervix	12,800	4,600	xxx	4,600
Prostate	180,400	31,900	31,900	xxx
Brain and other nervous system	16,500	13,000	7,100	5,900
Lymphoma	62,300	27,500	14,400	13,100
Leukemia	30,800	21,700	21,100	9,600

Adapted from: Greenlee, R.T., Murray, T., Bolden, S., Wingo, P.A. CA Cancer J Clin 2000;50:7-33

To begin looking at what might cause cancer, here are a few things we know from simple observation:

1) Cancers tend to run in families. However, very few cancers exhibit Mendelian inheritance ratios indicating that most are heavily influenced by non-genetic factors.

2) Exposure to certain agents in the environment (certain chemicals and irradiation) is associated with the development of cancer. We call these agents carcinogens.

3) If we perform the Ames test for mutagenicity, we find that all carcinogens are mutagens.

(However, not all mutagens contribute to the development of cancer.)

4) Malignant cancer cells have at least two things wrong with them:
   1) They can't stop dividing; and
   2) They leave their normal tissue beds and take up residence in areas of the body that are completely foreign to them.

So, at a minimum, based on this information, we should be able to hypothesize that:

1) Cancer is caused, or enhanced, by changes in the DNA which may be
   a. inherited mutations (since predisposition for cancer runs in families); and/or
   b. new mutations (since carcinogens cause mutations in the DNA)

2) Cancerous cells have a defect in the molecule(s) that control communication about:
   a. when to stop dividing; and
   b. in which tissue the cell should exist

The genetic changes could be in the molecules that control communication. The mutations could be in genes that encode transcription factors needed to transcribe the genes that encode communication molecules. Keeping this in mind, let's look at what we know about the normal signals that control cell division in normal cells.

Focused Reading: p 157-9 "Interphase…" stop at "Eukaryotic…"
p 158-9 fig 9.4 and 9.5

Cell division is almost always studied by placing cells in tissue culture, an experimental approach that grows mammalian cells in a petri dish. Many types of animal (and plant) cells can be removed from an organism and, if provided with the right combination of nutrients, the right gas mixture, and the right kind of substrate to sit on they will not only live, they will also continue to divide. Lots of picky details insetting up the system but a great way to get at how a cell really works without dealing with an entire pesky organism! The tissue culture cells can then be treated in such a way that their cell cycles are **synchronized**. Normally, cells divide on their own inherent timetables, regardless of what their neighbors are doing. Having a culture of cells that are all at different stages in the cell cycle is not very helpful. In order to study the changes that occur in the cell as it moves from stage to stage, it is easier to look at a large population of cells in one stage (in one dish) and compare them to a large population of cells in a different stage (in a different dish). Certain drugs are used which arrest cell division at a given stage. As each cell enters this stage, it gets stuck there. Since the cell cycle is just that--a 'circle'-- no matter what stage a cell was in when you added the drug, sooner or later it will come around to the drug-blocked step and get stuck. Given enough time every cell in the petri dish will be ready and waiting, stuck at the drug block. By removing the drug, all the cells resume dividing, but now they are all starting at the same point and will be in synchrony.

Why both getting synchronized cells? Here is an example of a kind of experiment you can do with them. Scientists hypothesized that a soluble factor in its cytoplasm stimulated cell to go past the G_1-S boundary. This boundary is a step that commits the cell irreversibly to DNA synthesis and mitosis. (This point is also known as the **restriction point** or **G_0**--get it a cell must pass 'go'). To test

the hypothesis, they synchronized one dish of cells in G_1 and another in S. They then mixed the cells together and caused them to fuse so that they ended up with giant “double cells”. This "double cell" has two nuclei with DNA in different stages of the cell cycle, but all of the cytoplasmic molecules have mixed together. Thus, after fusion, these two sets of chromosomes receive the same cytoplasmic signals. When cells in the S phase were fused with cells in G_1, the 'S' DNA stayed the same but the G_1 DNA began to replicate. Thus, there was some soluble signal molecule in the S phase cells that caused the G_1 cells to enter the S phase.

Investigators wanted to know if this factor was made in S phase and then stayed as soluble factor in the cytosol for the rest of the cycle or was the factor destroyed after the S phase. So they fused G_2 cells with G_1 cells. This fusion did not result in the replication of G_1 phase chromosomes. Thus, they hypothesized that this soluble factor was no longer present in the cell after the S phase was complete. This soluble factor was called the **S-phase activator**. A rise in the concentration of this molecule in the cell triggers or facilitates the transition of the cell from G_1 to S.

Normally, a cell that enters the S phase has passed the restriction point and will undergo mitosis. However, another control molecule must signal that the S phase is complete before the cell will enter mitosis (M). If S phase has begun but DNA synthesis is artificially blocked so that it cannot be completed, the cell will not enter mitosis until the block has been removed. Also, if a G_2 phase cell is fused with an S phase cell the G_2 phase chromosomes will wait for the S phase chromosomes to complete their duplication before they enter mitosis. Therefore, investigators hypothesized that there is a "delay" molecule, which prevents mitosis from beginning until the S-phase is complete.

After this "delay" molecule has been inactivated, the cell needs yet another signal to progress into mitosis, the **M-phase promoting factor (MPF).** If M-phase cells are fused with cells in any other phase, the “double cell” will immediately enter mitosis, even though the division will be unsuccessful for any chromosomes that have not replicated their DNA. Thus, MPF can override the "delay" factor, and therefore must not be present in the cell during S phase. Otherwise, the "delay" signal would be overridden and the cell would enter mitosis prematurely. The MPF is described further in your text.

The investigation of the cell cycle in wildtype and mutant yeast has proven to be one of the most powerful tools used to investigate the cell cycle. (In fact, many cancer researchers actually study yeast!) This is because a number of mutant yeast strains exist which are deficient in different proteins required at different stages of the cell cycle. These yeast strains are called **cell-division cycle mutants** (**cdc mutants**). Therefore, by determining which protein a given strain is missing, and correlating it with the stage of cell division that is eliminated or dysfunctional in that strain, the role of various proteins in the process of cell division can be determined.

So far, over 50 genes have been identified that act to control some phase of the cell cycle. In some cases, these genes are well-known biochemical entities in the cell. For instance, one cdc mutant strain that cannot go through the S phase has a defective gene for DNA ligase, while another such mutant cannot synthesize nucleotides from nucleosides. However, other genes encode true control molecules such as MPF, S-phase initiation factor, mitosis inhibition factor, etc.

In addition to identifying the intracellular proteins that control entry into the various stages of cell division, investigators have also recently identified a number of **growth factors** which exist outside the cell and whose stimulation is necessary for cell division in many cases. The following is a list of some of the major growth factors and the types of cells that responds to each.

Factor	Effect
Platelet-derived growth factor (PDGF)	Stimulates connective tissue cells and supporting cells of the brain
Epidermal growth factor (EGF)	Stimulates many cell types
Insulin-like growth factor I and II	Collaborates with PDGF sand EGF; simulates proliferation of fat cells and connective tissue cells
Transforming growth factor ß (TGF-ß)	ıcreases the sensitivity of most cells to other growth factors; regulates differentiation in come cells
Fibroblast growth factor (FGF)	Stimulates cell division in many cell types including cells of connective tissue, blood vessels, and muscles.
Interleukin-2	Stimulates cell division in T lymphocytes
Nerve growth factor (NGF)	Promotes growth in size of neurons
Many blood cell growth factors	Promote growth and development of all the cell types in the blood.

Study Questions:

1. Understand the meaning of the terms that are used to describe tumors and cancers.

2. List the phases of the cell cycle, including the phases of mitosis, and explain the significant events that happen in each phase.

3. Understand the mechanisms used by the cell to produce two genetically identical daughter cells during cell division. While these cells are genetically identical, they may not be identical in other ways. Explain.

4. Describe the factors that have been shown to play a role in controlling (triggering or inhibiting) cell division (e.g. nutrients, cell size, growth factors, etc).

5. What is the restriction point? When does it occur and what is its significance?

6. Discuss the structure and function of MPF. What is structure of this molecule? Through what mechanism does this molecule's concentration rise and fall in the cell? What is the role of this molecule in cell division? What specific function(s) does this molecule perform?

7. Discuss the following methods and their application to the study of cell division. Give one example for each method illustrating the type of information that can be obtained using this approach.
   A. Cell synchronization in culture
   B. Cell fusion
   C. Yeast cdc mutants

8. Be able to interpret results from a cell fusion experiment in which cells of different phases of cell division are fused. For instance, if you learned that, when G_1 cells and S cells are fused, the G_1 phase chromosomes replicated their DNA, what would you conclude?

----------STOP----------

What do we know now about cell division that will help us figure out what causes cancer?

1) Cell division is carefully synchronized and controlled by proteins (which must be encoded by genes).

2) Cells respond to signals from their environment to "decide" whether or not to divide. Each of these signals must be "received" by the cell and responded to through a receptor system and second messenger system mediated by proteins that are ultimately controlled by genes. External signals include:

    1. the presence of adequate nutrients
    2. in some cases, the presence of specific growth factors.
    3. the degree of contact with neighboring cells
    4. the degree of attachment to a substrate (Note: In this case, substrate, or substratum, means a layer of protein fibers that underlie cells and anchor them in position)

A defect in any of these processes may cause a cancerous transformation. And because cancers arising in different tissues or organs have very different characteristics, different cancers may have very different causes.

What approaches can we use to figure out what exactly is wrong in a cancerous cell? Well, one approach that has been extraordinary helpful in cancer research has been the experimental use of viruses that are known to cause cancer in animal cells. Such viruses are called **tumor viruses**. The first tumor virus to be discovered, the **Rous sarcoma virus**, (discovered by Dr. Rous) causes connective tissue tumors in chickens. Since this discovery, several other tumor viruses have been identified and characterized:

Virus	Species	Tumor
Rous Sarcoma Virus	Chicken	Connective Tissue
FBJ osteosarcoma virus	Mouse	Bone
Simian sarcoma virus	Monkey	Connective Tissue
Abelson murine leukemia virus	Mouse	Leukemia
Avian erythroblastosis virus	Chicken	Bone Marrow
Harvey murine sarcoma virus	Mouse	Connective Tissue
Avian MC29 myelocytomatosis virus	Chicken	Bone Marrow

Humans:	Tumor
Papillomavirus (H{V)	Uterine Cervical Carcinoma

Hepatitis-B	Liver Carcinoma
Epstein-Barr virus (EBV)	Burkitt's lymphoma (B cell cancer) Nasopharyngeal Carcinoma
Human T-Cell Leukemia Virus-I (HTLV-I)	Adult T-cell Leukemia/ Lymphoma
Herpes Simplex virus variant AIDs-related opportunistic infection	Kaposi's Sarcoma

Note that the only way you can be sure that a virus causes cancer is to inject the virus into an organism and watch for the development of the tumor (with the aid of proper control injections). Because you cannot do this with human subjects, and because viruses are species specific, you have no way of definitively determining if the human viruses listed above actually cause or contribute to the development of cancer. However, we do know that people who are infected with these viruses have an increased likelihood of developing the tumors listed in the table. On the other hand, infection with these viruses does not guarantee the development of a tumor -- it only increases the likelihood.

It is probably the case that human tumor viruses contribute only minimally to the overall incidence of cancer in humans. However, tumor viruses have been of exceptional importance in the study of cancer. When a known tumor virus is placed in culture with its target cell, the cell will become cancerous; a process called **cellular transformation**. (If a biologist tells you a cell population has been "**transformed**", you need not ask, "into what?" By definition it is a tumor cell.) By studying the differences between a cell population before and after transformation, scientists can gain an understanding of the changes that occur during the development of cancer.

What happens to these cells in the process of transformation? Well, it depends on the cell and the virus, but here is a summary of some changes that occur when cells are transformed:

I. Plasma membrane related changes
- A. Enhanced transport of nutrients
- B. Excessive blebbing of plasma membrane (small areas where the membrane balloons out, like a weak spot in a garden hose.)
- C. Increased mobility of the plasma membrane proteins

II. Adherence abnormalities
- A. Diminished adhesion to surfaces
- B. Disorganization of the cytoskeleton
- C. High production of protease causing increased extracellular protein degradation.

III. Growth and division abnormalities
- A. Growth to an unusually high cell density
- B. Lowered requirement for growth factors
- C. Less "anchorage dependence" (Can divide even without attachment to a solid surface. This is highly unusual in normal cells.)
- D. Can continue to divide indefinitely -- immortality in tissue culture.
- E. Can cause tumors when injected into animals.

The actual growth of tumor cells in culture is amazing to see! Depending on the cell type, they can be large, misshapen cells with little interest in attachment to the culture dish. They divide while

they float in the medium, draining the culture medium of nutrients in a very short time. If they are "fed", that is, given fresh growth media, they will continue to divide indefinitely. We would have absolutely no trouble filling Dana Science Building with the offspring of one, well-fed tumor cell in a surprisingly short period of time.

This week's "puzzlah" *a la* Car Talk (for fun only):

Average cell volume = 125 pL;
(1 picoliter = 0.001 nanoliter, and 1 nanoliter = 0.001 microliter);
Dana Science Building (50,000 m^3);
Average cell cycle = 1 division every 12 hours.
How long would it take to fill Dana Science Building?

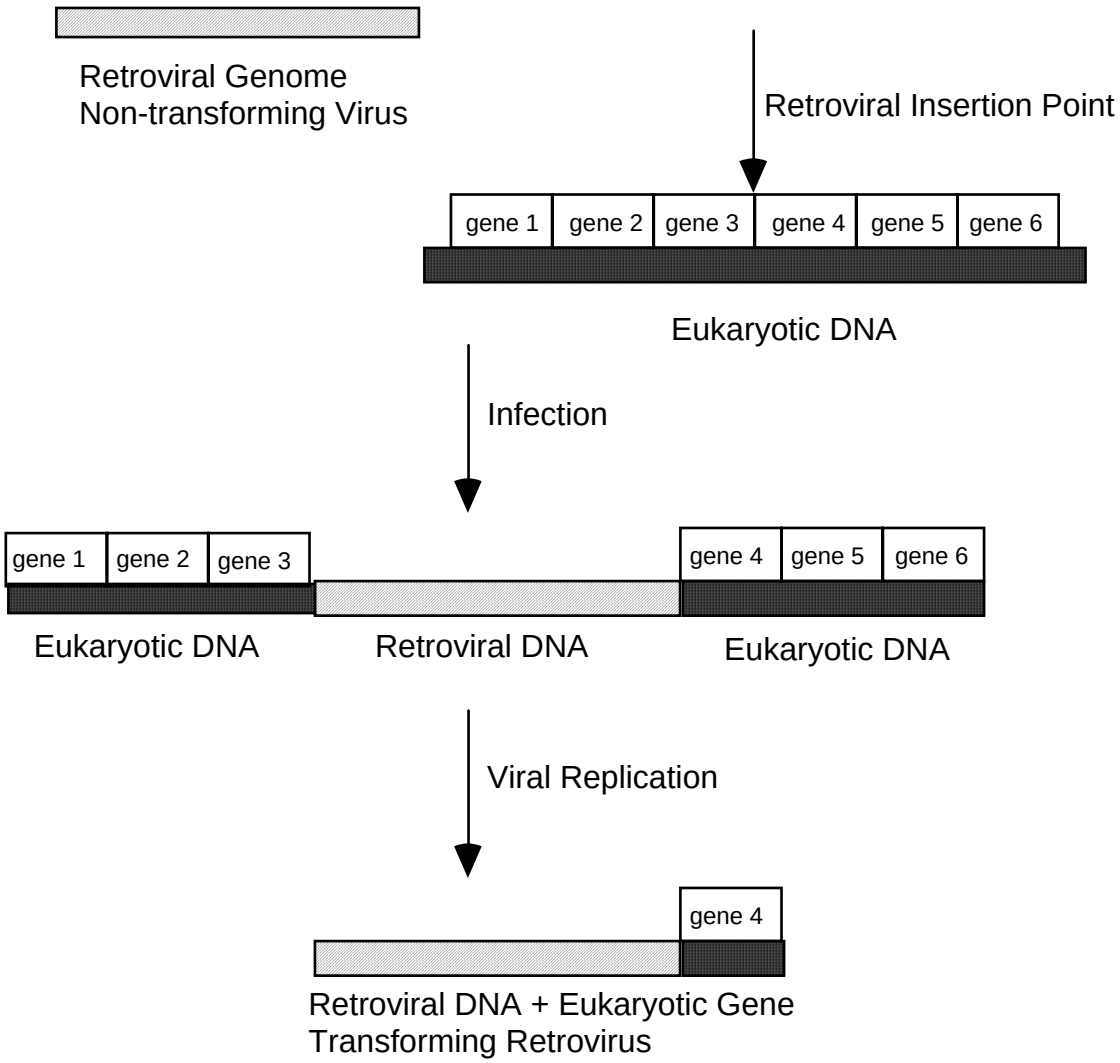
Focused Reading: p 240-1 "Viruses reproduce…" stop at "Bacteriophages..."
p 342 "Some cancers are caused…" stop at "most cancers…"
p 244 fig 13.5 (as example of a retrovirus)

The key to understanding cellular transformation is to look at the genetic changes that occur when the tumor virus infects the cell. To study this, investigators have focused on **tumor retroviruses** since, as you know, these viruses actually insert genes into the cell's genome that are then passed to the next generation of cell. Thus these genes become a genetic characteristic of the tumor. The first such tumor RNA retrovirus studied was the Rous sarcoma virus (RSV). RSV inserts its entire genome into the host cell during the transformation event, so it would be difficult to determine which of these viral genes is responsible for the cancerous transformation. However, as is the case with all viruses, RSV mutates at a rapid rate, and investigators were able to find a viral strain that seemed like a perfectly healthy virus (it was able to infect cells, insert its DNA, and make new virus), but it did not transform the cells. When investigators looked for the difference between this non-transforming RSV and the non-transforming strain was missing one gene. Investigators named this gene the *src* gene (pronounced "sark"). [By convention, the names of genes are italicized while the names of their protein products are not.] Investigators called this *src* gene an **oncogene** because it causes cancer. ("Onco-" is from the Greek *onkos* meaning tumor. The study and treatment of cancer is the field of **oncology**.)

What does *src* do? What does it encode that causes this dramatic change in the behavior of cells? As a next step in answering this, investigators created a radioactive DNA probe that was complementary to the *src* gene and probed the DNA of normal cells (using a Southern blot) to see what they could find. Surprisingly, they found a version of *src* in the genome of perfectly normal cells!! While these normal genes were not absolutely identical in structure to *src*, they had a lot of homology. They were so similar that they had to be alleles of one another -- versions of genes that encode the same trait. Investigators called this normal gene a **proto-oncogene**. [Michael Bishop and Harold Varmus (two recent Nobel winners) discovered Proto-oncogenes]. Also, because they had found very similar genes in both a virus and its eukaryotic target (in this case, chicken connective tissue cells), they needed a way to distinguish the viral gene from the eukaryotic gene. Thus, they called the viral version of the gene ***v-src*** ("v" stands for "viral") and the eukaryotic cellular version of the gene ***c-src*** ("c" stands for "cellular"). Since the discovery of *src*, over 20 oncogenes and their proto-oncogene versions have been discovered through their presence in retroviral genomes, and

over 50 oncogenes have been identified overall. It is worth mentioning that src is a kinase that often phosphorylates growth factor receptors. The viral form is about 20 times more active than the proto-oncogene cellular form, which helps explain why some viruses can lead to cancers.

As an aside, you might be wondering why a virus would contain a gene that causes cancer. These viral oncogenes don't appear to confer any survival value whatsoever to the virus. In the case of a retrovirus, the virus' direct ancestor probably picked up the gene from a host when it became incorporated into that host's DNA. Because retroviruses actually become part of the genome, pieces of host DNA can be included in the viral genome fairly easily. If the viral genome is transcribed from viral DNA plus some of the flanking human DNA, the viral genome will contain a copy of the host's gene. It is assumed that this is the way human genes get into viruses and, when the virus infects the next cell, it carries this human gene along with it and incorporates it into its new host's DNA.



In the case illustrated above, if gene 4 is a proto-oncogene, when the retrovirus picked it up in the process of replication, it would become a retrovirus carrying an oncogene -- the definition of a tumor virus.

While tumor viruses provide an invaluable approach to the study of cancer, we should not get too carried away at this point and give the impression that cancer is caused by little bits of human DNA attached to retroviruses. In fact, this occurs only in a few cases of animal cancer. **However, these viral oncogenes have led us to their normal counterparts, the cellular proto-oncogenes, and it is assumed that most cancers are caused when these normal proto-oncogenes become mutated, thus becoming oncogenes** (related to the development of cancer)**.**

Investigators assume that cancer-causing mutations are caused by the same mechanisms that cause other mutations, such as:

1) chemical agents that alter the structure of DNA;
2) irradiation (e.g. UV light) that breaks DNA or forms inappropriate covalent bonds
3) retroviruses that insert themselves in or near a gene, thus changing its proper regulation
4) normal mistakes ("typos") made when the DNA is replicated during cell division.

While all cancers are caused by mutations, not all mutations cause cancer. (Some, for instance, cause cystic fibrosis, or color-blindness, or a predisposition to heart disease.) It is assumed that what distinguishes cancer-causing mutations from other mutations is that cancer-causing mutations occur in proto-oncogenes. Thus, investigators have focused intensively on trying to find out what proto-oncogenes do.

Study Questions:

1. Know the name of at least one human virus that is thought to be associated with the development of cancer. Explain why it is difficult to prove that viruses cause cancer in humans.

2. What is a tumor virus? What is cellular transformation? Tumor viruses do not cause most human tumors. Given this, explain why tumor viruses have been the focus of such intensive research efforts. What types of information have we gained about cancer through the use of these viruses?

3. What are the characteristics of cells that have been transformed in tissue culture by a tumor virus? If you were looking through a microscope at a cell culture, what would you look for to determine whether or not you were looking at transformed cells?

4. How do retroviruses come to carry human genes?

5. Carcinogenic mutations are probably caused by the same agents as other mutations. What are these agents? How do carcinogenic mutations differ from other mutations? Why do these changes cause cancer while other types of mutations do not?

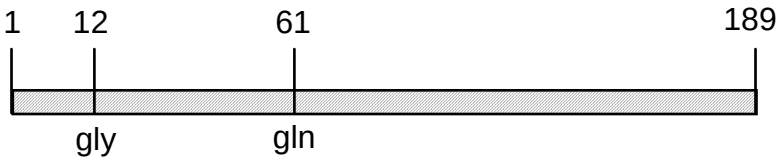
6. Understand the terms used to identify oncogenes. What does it mean when it is preceded by a "v?" By a "c?"

----------STOP----------

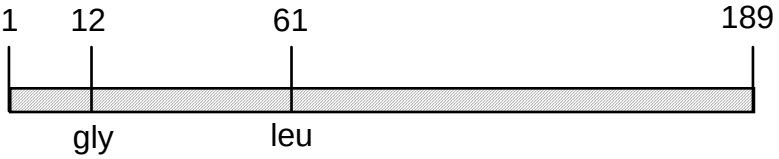
In order to illustrate some of the normal functions of proto-oncogenes; let's look at some specific examples of proto-oncogenes that have been fairly well characterized. First, let's look at the *ras* proto-oncogene (first identified in a rat sarcoma.) The proto-oncogene *ras* encodes a G-protein! Remember G-proteins? It's the protein that forms the link between a membrane receptor (for a hormone or, in this case probably a growth factor) and adenylate cyclase or phospholipase C.

Focused Review: p 282 “Receptors …" stop at end of page
p 286-7 “Protein kinase cascades…” stop at “Cyclic AMP”
WWW Reading: G-protein similar to *ras* with GTP bound - RasMol Image

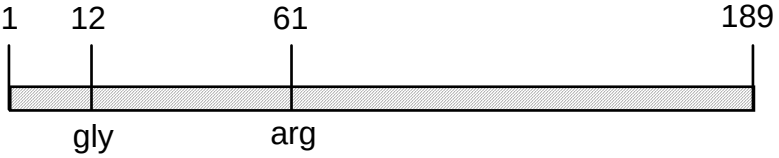
The normal *ras* proto-oncogene encodes a G-protein that contains 189 amino acids. So far, three oncogenic versions of this gene have been isolated from cancerous tissue. These oncogenes differ from the proto-oncogene at only one amino acid at position 12 or 61. The mutations are diagrammed below:



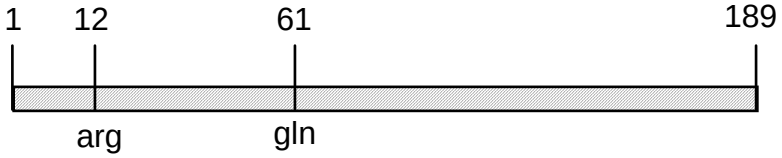
Normal ras protein



Protein encoded by H-ras oncogene



Protein encoded by N-ras oncogene



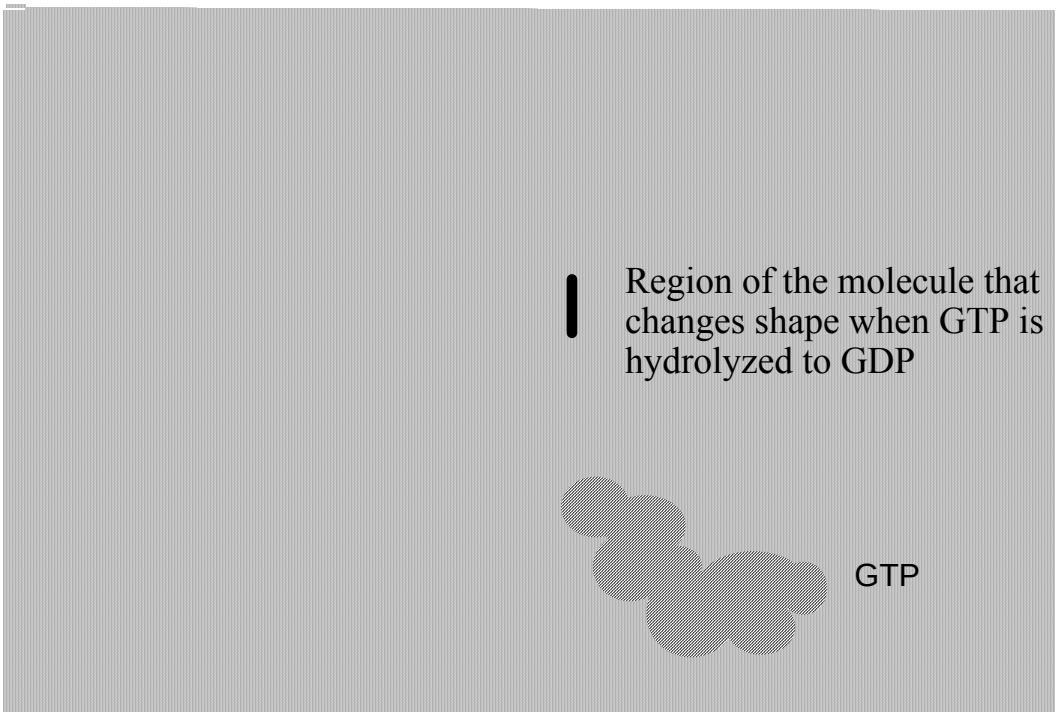
Protein encoded by K-ras oncogene

Turn to page 224 and compare the amino acids in the different versions of the ras protein. H-ras has a leucine instead of a glutamine at position 61. Leucine is non-polar while glutamine is polar. This mutation could change the folding pattern of the molecule significantly. The other two mutations (glutamine to arginine in N-ras: and glycine to arginine in K-ras) also change the characteristics of the amino acid significantly.

Biochemical studies show that the mutant *ras* oncogenes encode proteins that cannot hydrolyze GTP to GDP + P_i. As you will recall, the hydrolysis of GTP is the step that inactivates the G protein, making it unable to stimulate its enzyme target any longer. The ras G-proteins, therefore, are "stuck" in the "on" position. Once they become stimulated by the binding of a growth factor to a receptor and the subsequent binding of GTP to their active site, they are permanently on, and keep stimulating their target enzyme, which keeps making second messenger, which keeps signaling the cell to divide.

The drawing below is an illustration of a computer-generated structure for the normal *ras* G-protein. The GTP binding site is on the lower part of the molecule (bound to GTP). The areas of the molecule that change shape when GTP is hydrolyzed are drawn indicated by a black line. These

changes in shape represent the "on" and "off" conformations of the molecule. Oncogenic versions of the ras protein are stuck in the "on" conformation.



Further evidence of the linkage between the ras proteins, growth factors, and control of cell division comes from intracellular antibody binding studies. When anti-ras antibodies (against the normal version of the protein) are injected into the cytoplasm of normal cells, these cells are unable to divide in response to growth factors. Thus, the ras protein forms a link between the growth factor signal and the cell division response.

Study Questions:

1. What protein does the *ras* proto-oncogene encode? What is the normal function of this protein?

2. In general (you need not remember the exact changes), how are the *ras* oncogenes different from the *ras* proto-oncogene? How do these changes alter the protein's function? How do these changes lead to the development of cancer?

3. Describe the intracellular antibody binding studies that link the ras protein to the response of the cell to growth factors.

To illustrate the types of growth-related proteins that can be altered in cancerous changes, here is a list of some of the known proto-oncogenes and the normal proteins they encode. As you can see, proto-oncogenes come in four varieties: growth factors, growth factor receptors, signal transducers, and nuclear proteins involved in gene expression.

Proto-oncogene	Type of Protein Product	Protein Product
sis	Growth Factor	Platelet-Derived Growth Factor
fms	GF Receptor	Colony-Stimulating Factor-1 Receptor
erbB	GF Receptor	Epidermal Growth Factor Receptor
neu	GF Receptor	Protein with similar structure to Epidermal Growth Factor Receptor
erbA	GF Receptor	Thyroid Hormone Receptor
src	Signal Transducer	Tyrosine kinase, required for entry into G_2 of cell cycle
abl	Signal Transducer	Tyrosine kinase
H-ras	Signal Transducer	G-protein
N-ras	Signal Transducer	G-protein
K-ras	Signal Transducer	G-protein
jun	Nuclear Proteins	Transcription Factor AP1
fos	Nuclear Proteins	Transcription Factor AP1
myc	Nuclear Proteins	DNA-binding protein (transcription regulator)

In several cases, carcinogenesis is associated with **gene amplification**. In this situation, the gene is frequently normal in base sequence and may be located on the correct chromosome. However, hybridization studies show that the gene has been duplicated, sometimes hundreds of time, and is repeated over and over again in tandem sequences. Each gene is active, and therefore, the protein product of such gene amplification is over expressed and therefore over stimulates cell division. The oncogenes that cause some types of leukemia and lung, skin, colon and breast cancers are in this group.

Please note that oncogenic mutations can be inherited or can arise in the afflicted individual. In some cases, people get cancer because they inherited an oncogene from their parents. These types of cancer tend to run strongly in families (e.g. breast and colon cancers.) (For reasons we will discuss below, however, the inheritance of these oncogenes does not guarantee the development of cancer.) The majority of cancers, however, are probably associated with the development of new mutations in a proto-oncogene in one cell of the afflicted individual. This cell becomes cancerous and gives rise to the disease. Thus, in the case of many of the mutational changes associated with cancer, there is no way to test for the presence of the mutated gene since it has not been inherited, but rather is present only in the tumor cells and their descendants.

Study Questions:

1. Discuss the differences between oncogenes that are inherited and those that arise in the afflicted individual. In which case can a test be developed for the presence of the gene? Explain.

2. Describe the four types of protein products that proto-oncogenes are known to encode. Give an example of each.

Focused Reading: p 344-6 "Two kinds of genes…" stop at "Treating genetic…"

Oncogenes are usually expressed in cells as dominant traits, that is, only one copy of the oncogene is required for cancerous transformation. However, while the presence of an oncogene is required for the development of cancer, it is not sufficient. The cell has a number of **tumor suppressor genes** that function to prevent out-of-control cell division. If these tumor suppressor genes are functioning normally, one oncogene by itself will not produce a cancerous cell. Thus, at least two genetic changes are required for carcinogenesis: 1) changes that create an oncogene from a proto-oncogene and 2) changes that inactivate tumor suppressor genes. See figure 18-18 on page 346 for the steps required for colorectal cancer, the system we understand the best.

The good news about tumor suppressor genes is that usually both alleles at a tumor suppressor locus have to be destroyed before there is loss of growth control. A mutation that inactivates one allele will not have an effect (that is, will be recessive to the dominant suppressive effect of the other allele.) Often times, it is the inheritance of a defective tumor suppressor gene that predisposes us to cancers.

Study Questions:

1. Describe the relationship between tumor suppressor genes and oncogenes. What genetic changes must be present in these genes in order for cancer to arise?

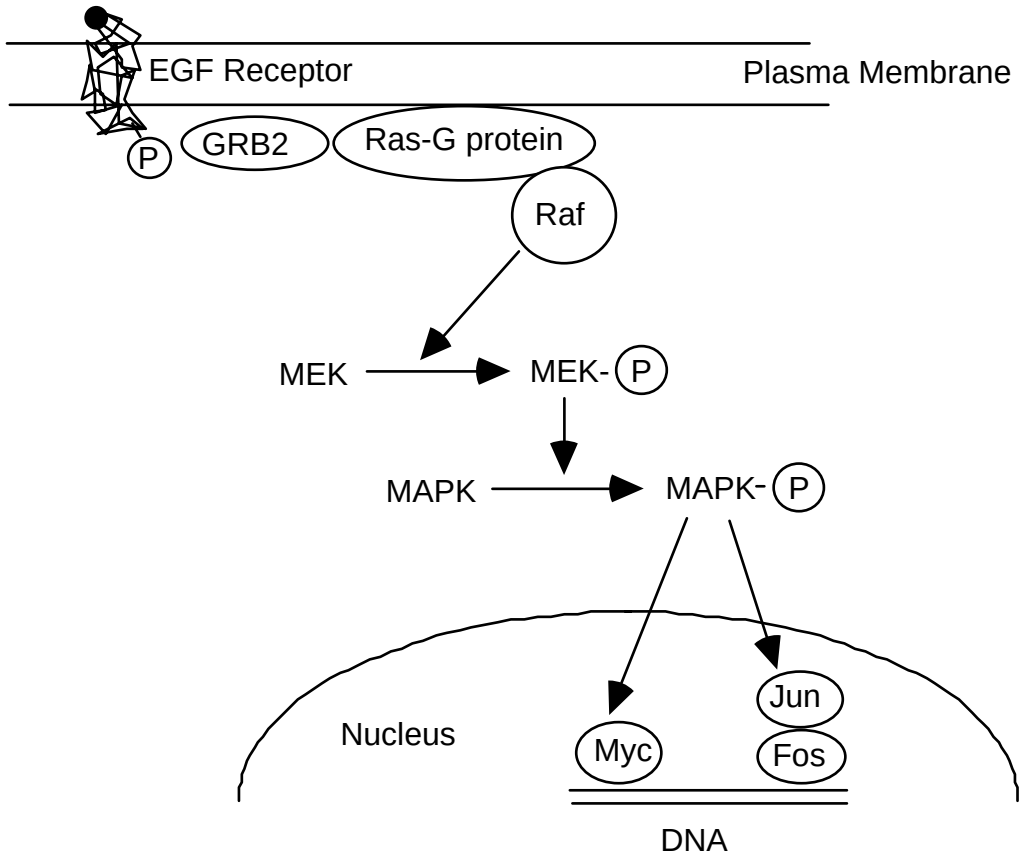
2. What is wrong with the phrase, "Some day we may find **the** cure for cancer"?

NEWS ITEM: A group at the Scripps Research Institute in La Jolla, CA have synthesized a bacterial compound called Epothilone A which can kill cancerous cells. Like the drug taxol, which is extracted from the bark of the yew tree in the northwest part of the US, this newly synthesized compound binds to microtubules and prevents chromosomes from separating during mitosis. However, there are two great advantages for Epothilone A. It can be manufactured in the lab and therefore is not dependent upon the slow-growing yew. Secondly, it is water-soluble and therefore it will be easier to administer to patients. (See summary by Robert Service. *Science*. Vol. 274: 2009. 1996)

WWW reading MAPK Signal Transduction

In a flurry of scientific papers recently, investigators have outlined the entire relationship between the *ras* protein and cell division. By examining this pathway, you should be able to get a clearer picture of the link between the development of cancer and changes in G-proteins, tyrosine kinases, growth factor receptors and nuclear transcription factors. Note: You do not have to memorize this pathway. It is presented here simply to help you see how oncogenes might cause uncontrolled cell division.

Some cells contain receptors for epidermal growth factor (EGF.) These receptors are membrane-bound tyrosine kinases. When EGF binds to its receptor, the receptor **autophosphorylates**, that is, it adds a phosphate group to its own tyrosine residues. This causes the receptor to change shape. This change in shape allows the receptor to bind to a cytoplasmic protein called **growth factor receptor binding protein** (**GRB-2**). This binding activates GRB-2, which then binds to the ras G-protein and activates it (in the classic manner, by causing it to bind GTP.) Activated ras activates a protein called Raf-1 (itself the product of a proto-oncogene.) Raf-1 is a kinase that phosphorylates and activates a protein called MEK. MEK is a kinase that phosphorylates MAPK. Phosphorylated MAPK travels to the nucleus where it activates transcription factors that are necessary for gene expression. These transcription factors are encoded by the proto-oncogenes *myc, jun and fos.* These transcription factors may allow the production of proteins (such as cyclin) that trigger cell division. Here is an illustration of this pathway:



Thus, the signal pathway initiated by EGF is filled with the products of proto-oncogenes. Any change in any of these proteins could cause abnormally high levels of expression of cell division signals, thus producing a cell dividing out of control.

Study Questions:

1) Given what you know about signal transduction and the cascades used in cellular communication, would you be more susceptible to cancer if you had a mutation that: A) left the EGF-receptor always activated, B) left the MAPK always activated, C) left the

transcription factors always activated, D) all of the above, E) none of the above. Be able to explain your answer.

NEWS ITEMS: In June 1996, a team of researchers found a species of voles that is resistant to mutations caused by radiation. When they analyzed their cells, they found that the voles had elevated levels of IDH, which they believe is protecting them from radiation-induced mutations. (See summary in *Science*. Vol 273. 19 July, 1996)

A new finding by Pascal Goldschmidt-Claremont from Ohio State Univ. suggests that *ras* also uses superoxides (an oxygen molecule with an extra electron) to communicate within the cell. They speculate that oncogenic alleles of *ras* may produce more superoxide than wild-type alleles do. You may remember from earlier News Items and our lab work using the Ames test, that oxidative damage to DNA can lead to mutations and thus cancer. Therefore, researchers are looking at antioxidants (yes, the same thing health-food stores have claimed will cure cancer) as potential drugs for treating cancer. This can be summarized in one phrase that you have heard all your life, "Eat your green vegetables" which are high in antioxidants. (See summary by Elizabeth Pennisi. *Science*. Vol 275: 1567-1568. 1997.)

2. Let's imagine that there is a wonder drug that is capable of protecting us from cancer-causing mutations and it works by stimulating the transcription of the cytoplasmic form of IDH. Explain to your grandparents how this drug might prevent cancer.

NEWS ITEMS**:** In 1998 National Cancer Institute launched a new web site as a part of the Cancer Genome Anatomy Project (CGAP). This is a spin off of the human genome project and its goal is to sequence all the cDNAs from healthy and cancerous tissues to compare what genes are expressed in each situation. Interestingly, a Davidson biology major alumnus ('97) was one of the first technicians to work on CGAP. The URL is <http://www.ncbi.nlm.nih.gov/ncicgap/>

Many people think that the hottest area in cancer research in on the enzyme called telomerase. Telomerase is the DNA polymerase that replicates the telomeres of our chromosomes and keeps them from "unraveling". It appears that normal cells do not have much, if any telomerase, while cancerous cells have a lot. Interestingly, telomerase has a lot of similarity to reverse transcriptase so there is some hope that drugs similar to AZT might be effective treatments for cancer.

A new type of cancer-causing mutation was found recently. A group at Johns Hopkins found that many people carry a particular allele for a proto-oncogene involved in the formation of colon cancers. For years, this sequence variation in the DNA was ignored because it was a silent mutation, causing no changes in the resulting protein. However, they recently learned that this particular mutation made the surrounding DNA susceptible to errors in replication. These subsequent mutations resulted in oncogenic mutations. Now labs around the world are going back over old data to see if any of these unstable mutations were overlooked. (see *Science* Vol. 277: 1201. August, 1997)

In focusing on mutations in the genes that control cell division, we shouldn't forget about the second criteria for malignancy, the ability to metastasize. In order to spread, cancer cells must be able to break free from the tissue bed they are in, enter a blood or lymphatic vessel, leave the blood or lymph and invade a different tissue bed. Leukocytes are the only cells of the body that can normally do this. Metastasis requires changes in cellular motility (most cells don't move they just sit there) which requires changes in the cytoskeleton, and changes in the secretory products of the cell since they have to digest their way across barriers. All these changes are caused by mutations (inherited or new) in the genes that control the cytoskeleton, secrete degradative enzymes, and receptors that interact with neighboring cells. Thus, even if a cell acquires an oncogene mutation and loses some tumor suppressor genes, this is still not sufficient to cause cancer. In addition, the cell must acquire mutations that allow it to metastasize, which is makes this story even longer.

NEWS ITEMS: A specific protease was identified that enabled breast cells to migrate out of the breast tissue. This protease cleaves a protein in the extracellular matrix called laminin-5. This may be a target to block metastasis. (*Science* Vol. 277: 225. July 1997)

We know that tumors become much harder, if not impossible, to treat when they metastasize. We also know that cancer is the result of inappropriate signaling. Dr S Wiley (Univ of Utah) has shown that a majority of cancers remain sensitive to signaling by EGFR (a growth factor receptor) and that blocking ligand release from this receptor can be enough to block metastasis. Drugs that block EGFR-ligand release are being tested on tissue culture cells and may provide a way contain cancer.

----------STOP----------

II. AIDS

Brief Overview Reading: Chapters 12, 13, 17, 18. & 19
Focused Reading: p 240-1 “Viruses reproduce…” stop at “Bacteriophages…”
p 221 "RNA viruses…" stop at "Transcription…"
p 242-3 "Animal viruses… " stop at "Many plant viruses…"
Pay special attention to figures 13.4 and 13.5

Some Definitions

Since its identification in the 1970's, Acquired Immune Deficiency Syndrome (AIDS), a preventable sexually transmitted disease (STD), has claimed the lives of over 2 million people worldwide. By 1983, the cause of this syndrome had been identified as the Human Immunodeficiency Virus (or HIV). AIDS is the clinical syndrome associated with chronic infection by HIV. Just as the flu (the disease) is caused by *Influenza* (the virus), AIDS (the disease) is caused by HIV (the virus). Unlike most viral infections, HIV infected (or **HIV**$^{+}$) individuals may be infected for months or years before they become sick with AIDS. This is called the **latency period** of the virus and it is one reason that HIV is such a dangerous organism. People can carry, and spread, the virus for many years without having any symptoms of the disease to tell them they have been infected. Because blood tests for HIV are not mandatory, we have no way of knowing exactly how many people in this country are HIV-infected.

Who can get infected with HIV?

In the United States, as of 1995, over 500,000 AIDS cases had been reported to the CDC and about 300,000 of these people have died from the disease. According to available information, it appears that 56% of AIDS victims are gay or bisexual men, 24% are IV drug users or their heterosexual partners, and 6% are heterosexuals. Two percent of AIDS victims (over 3,800) are children. Thirteen percent of children with AIDS contracted the disease from blood transfusions while the remainder contracted it at birth from their HIV-infected mothers. In 1997, the United Nations published the following list of estimated AIDS patients: 440,000 in East Asia; 1.6 million in Latin America and the Caribbean; 6 million in South America and South East Asia; 720,000 in Europe, North Africa and the Middle East; and 20.8 million in Sub-Saharan Africa. HIV continues to spread worldwide; the estimate of HIV-infected people has risen from 14 million in 1995 to 30.6 million in 1997.

In the United States, the epidemic has spread through African American and Hispanic populations at a faster rate than it has through the Caucasian population. For example, whites make up 80% of national population but represent only 38% of AIDS cases. Conversely, African Americans comprise 12.6% of the population but represent 40% AIDS cases. Hispanics make up 6% of the population, but represent 19% of AIDs cases.

HIV is spread when bodily fluids containing the virus contact the blood of an uninfected individual. The body fluids that contain the highest levels of virus are blood and semen. Entry can be gained through any breach in the skin or lining of an organ (eg. mouth, rectum, and vagina). The breach can be microscopic -- well below the size one would detect normally. High-risk behaviors include sharing needles during IV drug use, and participating in anal, vaginal or oral sex. Because semen contains the virus, if semen comes in contact with a small cut or tear, HIV can be transmitted. Because stretching and tearing of the anus and rectum can accompany anal intercourse, this practice

is a high-risk behavior. Vaginal intercourse is also presents high risks because the uterus and cervical area tend to be rich in blood vessels naturally, and abrasion during vaginal intercourse may cause areas of access for the virus. Also, if an open sore or cut exists in the mouth, nasal cavity, or esophagus, the recipient partner during oral sex is at risk. While it is possible for the virus to be transmitted from the recipient partner to the penetrating partner during any type of sex, the transmission rate is much lower in this direction. However, the virus can certainly be spread through any tiny cut or opening in the skin of the penis, fingers or hands. Finally, anyone who comes in contact with blood as part of his/her work (physicians, dentists, emergency medical technicians, etc.) or on an occasional or accidental basis (e.g. helping at the scene of a lab injury where blood is present) is at risk.

Because HIV spends most of its life inside human cells, the body fluids that do not contain cells also do not contain the virus or do not contain it in a transmissible form. These fluids include saliva, tears, sweat and urine. Touching, hugging, kissing, and sharing food or drink does not transmit HIV.

Worldwide, the vast majority of HIV transmissions are through heterosexual vaginal intercourse. In the United States, the disease has spread most rapidly (and continues to be predominantly present) in the gay male population. This selectivity is due to two unrelated factors. First, it appears that HIV has mutated into different strains, some of which adhere to vaginal cells better than other strains do. Therefore, heterosexual intercourse is the primary means of transmission where this strain predominates (currently in Asia and Africa). The second factor is because the disease appears to have entered the US via the homosexual population and HIV was spread through gay sexual encounters. If the disease had arrived from Asia and entered the country through the heterosexual population (still a possibility), the impact of the disease would be significantly different. It should be noted and understood that the disease is not a "gay" disease and makes absolutely no distinctions based on sexual preference. Heterosexual women make up one of the fastest growing HIV^+ subgroups. In 1992 at the Johns Hopkins Hospital emergency room, 30% of women delivering babies were HIV^+ (no statistics on their babies).

The disease is not spread by <u>being</u> in a high-risk group (e.g. gay, IV drug user), but by <u>doing</u> risky things (unprotected sex of any kind, sharing needles).

AIDS is a <u>preventable disease</u>. While we do not yet have effective drugs or vaccines to kill HIV, we do know precisely how the virus is spread. HIV infection can be avoided by avoiding contact with another person's blood or semen. Because HIV can enter the body through cuts or tears too small to detect, it is not enough to simply make sure that the blood or semen of another does not come in contact with an open wound. Rather, only <u>complete protection from contact</u> with the blood or semen of another person will guarantee safety. Protection includes abstinence from sexual practices, avoidance of situations in which you may come in contact with someone else's blood, and protection of yourself with latex barriers (condoms, rubber gloves, rubber dental dams, etc.) Condoms are normally used with contraceptive foam and they may be better at preventing the spread of HIV if the contraceptive foam contains Nonoxyl-9.

It can be difficult to approach the subject of protection with a partner, especially if the sexual encounter is of a more casual nature. It is easy to simply let it go, to tell yourself that the chances of contracting the disease are small and that it is too much effort, too embarrassing, too alienating, too unromantic, too nerdy, or too awkward to say anything. In heterosexual encounters, it is the woman who is at a far greater risk of contracting the disease than the man, and traditionally women are taught that being feminine includes being less assertive about sexual matters: not being pushy or

demanding about the use of protection. It is also tempting to tell yourself that everything will be okay because your partner looks healthy, is not in a "high risk group", says he/she has had a limited number of sexual encounters before you, or says he/she has just had an HIV test that came back negative. Even if your partner is telling you the truth, none of these is a guarantee that you will be safe. When you have unprotected sex, or come in contact with someone else's blood, you are at risk of contracting a disease that will kill you. You are gambling with your life. Even if the risks are low, the stakes are as high as they can be.

Table 1: People Living with AIDS (US, 2000)

	Year						
Race/ethnicity	1993	1994	1995	1996	1997	1998	1999
White, not Hispanic	80,480	86,703	91,756	98,615	107,273	114,895	122,860
Black, not Hispanic	60,678	71,863	81,287	92,274	105,306	117,426	129,943
Hispanic	31,245	36,524	41,072	46,194	52,121	57,443	62,995
Asian/Pacific Islander	1,295	1,460	1,617	1,859	2,094	2,318	2,609
American Indian/Alaska Native	569	662	718	803	888	969	1,085
Total[2]	174,475	197,471	216,796	240,184	268,242	293,702	320,282

Adapted from Center for Disease Control Surveillance reports (http://www.cdc.gov)

Based on random sampling data, the incidence of HIV infection on college campuses has been estimated to be **ten times higher than the general heterosexual population**. Based on these statistics, we would expect 2-3 students at Davidson (student population ≈1,700) to be infected with the virus. There is an especially disturbing trend in the rise in AIDS among those born since 1959 (figure 1). Notice that cases for those born before 1960 has leveled off, while those born after 1959 have continued to rise. Therefore, the rise in cases for people under 38 is responsible for the continued rise in reported AIDS cases for the entire country.

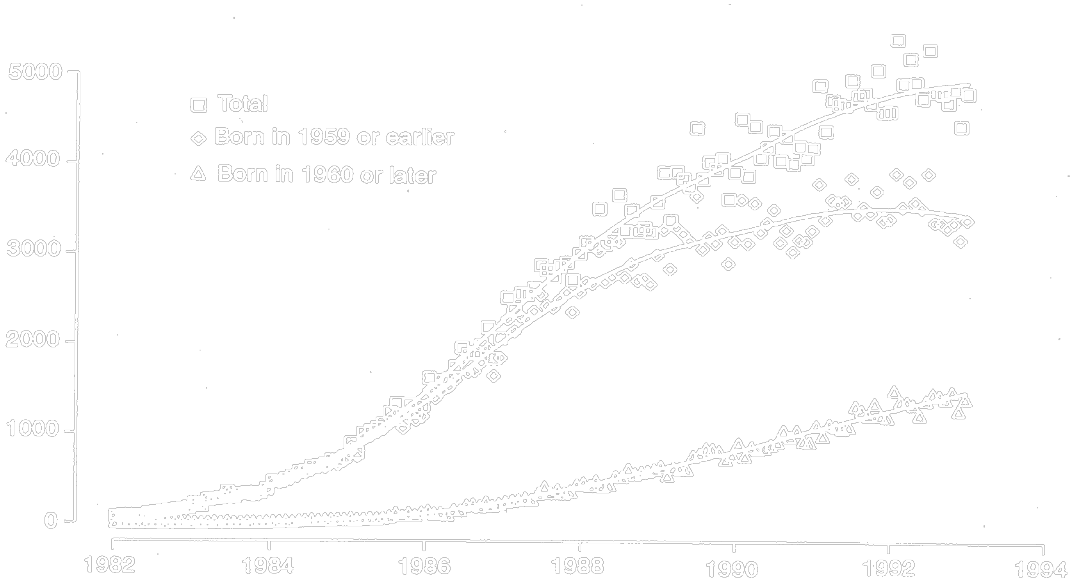


Figure 1. AIDS incidence (cases reported per month). This is the reported incidence of AIDS for the total population (squares), for individuals born in 1959 or earlier (diamonds), and in 1960 or later (triangles)

Table 2: New HIV infection reported July 1999 through June 2000 (in 34 areas with confidential reporting)

Age	Male	Female	Total
>13 years old	111	114	225
13-19 years old	327	536	863
20-24 years old	1,433	6,053	7486
25+ years old	12,820	385	13,205

Examining Table 2 you can see that the reported new HIV infection is highest in males above 25 years old and females between 20-24. This is a very high rate of new (reported) infection for the female category if you consider that it only includes a range of 4 years (and the high male # is for those 25 and up). A recent newscast reported that new infection in teenage girls is up 136%. The direct cause of this jump is not known but it certainly shows that HIV infection is not only not going away—in some age groups (like the one you are currently in) it's rising!

There is also a symptomatic difference between the sexes. A study from Johns Hopkins School of Public Health revealed that women who have 5000 copies of HIV/mL blood show similar AIDS symptoms to men carrying twice the viral load (10,000copies HIV/mL blood). While the development of full-blown AID appears to take the same amount of time in men and women, this

finding suggests that treatments that are started based on viral load 'requirements' may need to take symptoms and gender into account.

Some History of HIV/AIDS

In this Unit, we will look at what we know about HIV and AIDS. As is the case in the study of all diseases, we learn an enormous amount of basic biology as we learn about the disease. By studying HIV, we now know much more about all viruses and we certainly know a lot more about the human immune system (the target of HIV). One difference between this disease and others we have encountered is that AIDS was 'discovered' recently. Your professors remember the news about the first cases and the drama that surrounded identifying HIV (and we aren't all that ancient, honest).

In the early 1980s, investigators at the Center for Disease Control (CDC) in Atlanta noted that there was a dramatic increase in the number of adult males dying of a mysterious disease that appeared to compromise the immune system severely. The immune systems in these men were so weak that they could not fight off infections that are usually no match for a healthy immune system, most notably a kind of pneumonia that was often the cause of their deaths. In attempting to determine the cause of these deaths, the CDC tried to find out what all these men had in common. Four characteristics emerged which were called the "Four Hs" -- being a male Homosexual, Haitian, IV drug user (Heroine), or Hemophiliac. Very quickly, the investigators deduced that, at least in three of these cases, the underlying similarity is the increased likelihood of coming in contact with the blood of another person. Gay men, IV drug users, and hemophiliacs were known to be at increased risk for hepatitis B, spread by blood-to-blood contact. (It was later determined that the gay male population first infected with the disease vacationed extensively in Haiti, where some of the native population became infected. Being Haitian, in itself, has nothing to do with the disease.)

The disease spread exponentially in these three populations (gay men, IV drug users and hemophiliacs), reaching epidemic levels very quickly. Investigators in the United States and France began a frantic race to be the first to discover the presumably blood-borne agent that caused this disease. This race did result in the rather rapid characterization of the viral agent that causes AIDS, but it was fraught with fierce competition and accusation of foul play. Luc Montagnier, from the Pasteur Institute in Paris, and Robert Gallo of the National Institutes of Health (USA) share credit for the discovery of HIV as the causative agent for AIDS. (If you are interested, there are countless articles covering the legal and scientific battles between these two men and their associates.)

Study Questions:

1. What is the difference between HIV and AIDS? What is the difference between being HIV^+ and having AIDS?

2. How does the long latency period of this disease contribute to its spread?

3. How is HIV spread? What are "high risk behaviors" for contracting HIV?

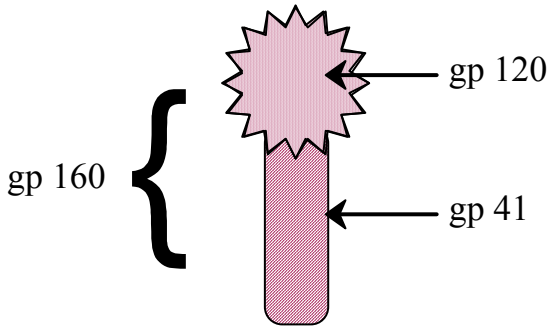
4. Some people believe that the AIDS epidemic has been wrought as a punishment by God against homosexuals. Based on the facts of transmission, how would you respond to this argument? Why is the disease so prevalent among gay men in the United States?

5. How can the spread of AIDS be prevented?

Focused Reading: p 240-244 "Viruses:…" stop at "Viroids:…"
p 373-5 "AIDs…" stop at end of chapter
p 244 Figure 13.5
WWW Reading: Life Cycle of HIV- Attachment

Structure of HIV

From your focused reading, you can see we know a great deal about what the virus looks like (structure) but we still have a lot to learn about how it works (function). Figure 13.7 is the best illustration of what HIV looks like, but there are a few special features we need to note.



The HIV genome is surrounded by a protein **capsid**, which is surrounded by a phospholipid membrane containing large glycoproteins. The lipid bilayer with embedded glycoproteins is called the **viral envelope** (remember it also contains human integral membrane proteins from the infected cell). The glycoproteins in the HIV envelope are called **gp160** (for "glycoprotein 160" because its molecular weight is 160 kilodaltons). gp160 is composed of two smaller subunits: gp120 (large star shape) and **gp41** (the stalk). gp120 is the protein that specifically binds CD4 allowing attachment and infection. The genome consists of two identical strands of ssRNA, which contain nine genes. Each ssRNA strand is bound to a molecule of **reverse transcriptase (RT)**, the enzyme required to transcribe ssRNA into ssDNA. Because human cells never do this, they do not contain RT, and therefore the virus must bring RT along with every virus particle. The two identical copies ssRNA are **reverse transcribed** (using reverse transcriptase, like what is used to make cDNA) into DNA and then inserted into the host genome where it can remain dormant, or latent, for months or years. (Human retroviruses were discovered by Robert Gallo, well before HIV.)

One more important feature of HIV biology is that when its nine genes are transcribed and translated, the encoded proteins are not made individually, but are made as a few multi-protein structures all stuck together like-this-compound-word. In order for these individual protein components to perform their functions, they must be cut free from each other. One of HIV's genes encodes for a **protease** that acts like molecular scissors to cut the multi-protein structures into their proper and functional subsections. The critical functions provided by the protease and RT have been the subject of a lot of research and pharmacological treatment of AIDS patients (see below).

Study Questions:

1. In general, describe the structure of a typical virus.

2. In general, how do viruses reproduce? What molecules must they encode in their own genome? Which molecules does the host cell provide?

3. Unlike bacteria that will grow on nutrient agar, viruses will not. What must you supply to support the replication and growth of viruses?

4. What special structures do animal viruses contain that allow them to enter and leave animal cells without having to cause the entire cell to rupture? Describe this process.

How HIV Infects Cells

We will begin looking at how your cells become infected with HIV by looking at the target of HIV, the immune system.

Brief Overview Reading:	Chapter 19
Focused Reading:	p 358-64 "Specific defenses..." stop at "Hybridoma..."
	Figures 19.6 and 19.11
	p 364-8 "T cells:..." stop at "MHC molecules are responsible..."
	Especially Figures 19.18 (animated version available at Purve6e site)
	p 372-5 "An inappropriately active…" to end of chapter
WWW Reading:	Cytotoxic T-cell Killing Its Target

The interactions of the immune system are extraordinarily complex and the subject of one of the frontier disciplines of biology, immunology. It is well beyond the scope of this Unit to delve deeply into the workings of this system. However, if you are to understand how HIV produces such a deadly effect in the body, you do need to understand a few things about how the immune system works.

As we discussed in Unit III, microbes are constantly invading your body, despite your best efforts to keep them out. You wash them away with mucus secretions in the lungs, you wash them away by sloughing off the outer layer of cells in the intestine and skin, you try to kill them with acid (skin, stomach, vagina), with enzymes (in tears, sweat, saliva), with antibodies (in all the secretions of the body) and still, they get in. Those resourceful creatures that make it through all these hostile defenses are met by an internal surveillance system so precise and deadly that all but the most virulent microbes are completely destroyed. Without this system of surveillance and destruction (the **immune system**) microbes would overrun your body and kill you -- fast -- by this time tomorrow.

The immune system functions by recognizing and attacking foreign molecular shapes (usually due to amino acid sequences that are not "self", that is, not part of any of your own personal proteins.) The cells of the immune system that do this are called **lymphocytes**. Lymphocytes have specific receptors in their membranes for foreign shapes.

Lymphocytes come in two varieties -- **T cells** (mature in the thymus) and **B cells** (mature in the bone marrow.) B cells make **antibodies**, the same specific proteins you have encountered in looking at the method of immunocytochemistry or immunohistochemistry. These proteins can bind

specifically to the foreign substance and trigger a number of responses that destroy it. T cells do not make antibodies, and they come in two varieties: **T helper cells** (**T_h**) and **cytotoxic T lymphocytes** (**T_c**). T_c s kill other cells directly by making membrane-to-membrane contact with them and inserting proteins in the cell's membrane that produce large holes. T_c s effectively punches holes in the membranes of other cells. This makes it impossible for the host cell that contains the pathogen to maintain any ion gradients across its plasma membrane and it dies. T_c s kill virally infected cells, cancer cells, and transplanted organs, a process called the **cell-mediated immune response**.

As their name implies, T_h cells help other cells perform their functions. They help B cells make antibodies; a process called the **humoral immune response** (The fluids of the body are called **humors** and antibodies were initially discovered in body fluids (blood plasma). In general, the humoral immune response neutralizes foreign proteins (e.g. bacterial toxins) and bacteria. T_hs also help T_c s become capable of killing. T_hs perform both helping functions by secreting various **lymphokines** that provide activation signals. Lymphokines function as local signaling molecules, binding to specific receptors and triggering cell functions through second messenger systems. Because both B cells and Tc s require its help, the T_h plays a pivotal role in all immune responses. Unfortunately, **the T_h is the primary cell that is targeted by HIV**. Thus, by interfering with the function of T_h, HIV cripples the entire immune capacity of the individual.

Viruses target certain cells based on specific binding between proteins in the virus' envelop and proteins in the cell's membrane. For example, the influenza virus binds specifically to proteins on the surface of the respiratory tract, the chicken pox virus binds to target proteins in the skin, the herpes virus binds to target proteins in the lips or genitals, etc. These virus-cell interactions are specific, just as are the interactions of enzymes and substrates, receptors and hormones, antibodies and antigens, transport proteins and transported substances, etc. Thus, viral targeting, attachment, and infection, just like virtually everything else in biology, relies on the interactions between molecules with specific three-dimensional structure.

The protein molecule on the surface of the T_h cell to which HIV binds is called **CD4**. (Immunologists have complicated ways of naming things, so this name doesn't stand for anything very meaningful.) HIV will bind to any cell that bears CD4 in its membrane. This includes T_h, macrophages and some supporting cells in the brain. However, the story is more complicated than this. CD4 is necessary for HIV binding, but not sufficient. For example, if the gene for human CD4 is transfected into monkey COS cells, HIV will not infect these COS cells. During the summer of 1996, several research teams (read lots of people working cooperatively in the labs) made significant progress in understanding HIV infection (*Science* 272: 809, May 1996; *Science* 272:1740, June, 1996; Science 274: 502. October, 1996). There are at least two types of molecules (coreceptors) that are also required for HIV infection: **CXCR4** and **CCR5** (such catchy names; see figure 2). As shown in the figure below, HIV requires cells to have CD4 and either CXCR4 or CCR5 in their plasma membranes. CXCR4 had been cloned previously and though its function was unknown, the cDNA sequence suggested that CXCR4 would turn out to be a G protein-coupled receptor (sound familiar?!) for an unknown ligand. CCR5 is a receptor for the chemokine RANTES. (**Chemokines, cytokines** and **lymphokines** are chemical messengers secreted by cells to alert the immune system; the significance of RANTES will be discussed later.) We now know that CXCR4 is a chemokine receptor too. What is especially interesting is that there are different strains of HIV that infect different types of $CD4^+$ cells at different times during a person's HIV infection. One strain infects macrophages during the first phase of infection, and another strain prefers T_h cells later after the disease progresses. As it turns out, macrophages express CCR5 and T_h cells express CXCR4. It has been known for years that when a person is first infected with HIV, macrophages get infected

first. A plausible explanation is that the strain of HIV that is responsible for initial infection requires CCR5 as a coreceptor and not CXCR4. As the infection spreads within a person, HIV is able to infect T_h cells, which means it requires CXCR4 as the coreceptor. These discoveries are very recent so their impact is uncertain, but they do help explain much about HIV infection.

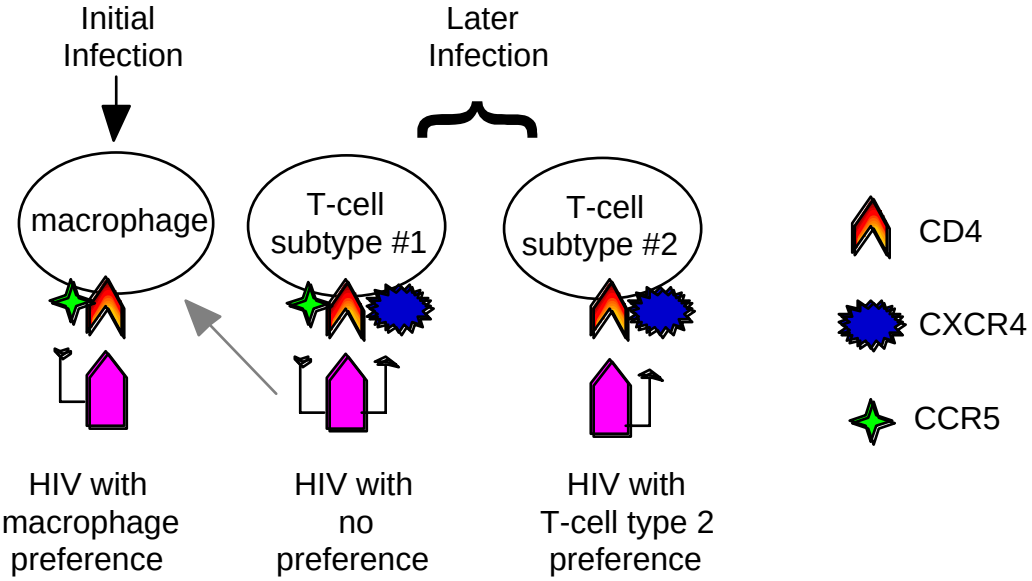


Figure 2. This diagram outlines what we know about HIV infection via its two coreceptors. HIV must bind to CD4, but also requires either CCR5 during the initial stages of HIV infection or CXCR4 during later stages of infection.

NEWS ITEM: The coreceptors CCR5 and CXCR4 were identified in 1996 and allowed the 'simple' model described above. As of now there are at least 13 known coreceptors for HIV and SIV (simian immune virus). Many of the co-receptors have unknown ligands and are expressed by different cells within the body. CCR5 and CXCR4 appear to be central to infection but the jury is still out. (See *Science* 1998. Vol. 280 p 825)

As is the always the case, these membrane proteins that bind viruses are <u>not</u> in the membrane for that purpose (this certainly would not be adaptive.) Rather, they are there for some other purpose, and the virus exploits their presence to gain entry into the cell. CD4 is one of the molecules that allows T_h to bind to antigen in order to become activated. Figure 19.18 illustrates the pivotal role of T_h. CD4 is an integral membrane protein on the surface of the helper T cell and interacts with the Class II MHC, T-cell receptor, and antigen. It stabilizes the interaction of these three molecules. Chemokines are secreted by a wide range of cells and they alert immune cells (T_h cells and macrophages) that there is need for immune cells to come to the area of chemokine secretion.

Study Questions:

1. What does the immune system do and, in general, how does it do it?

2. Which arm of the immune system is most effective against protein and bacterial antigens? Which arm is most effective against viruses and tumors?

3. How do viruses target specific cells? From an evolutionary perspective, explain why a cell would have a viral target in its membrane if this molecule allows the cell to be infected and killed.

4. What is CD4 and what does it do? How is this molecule related to HIV?

5. What are the other two coreceptors and where are they found?

6. Which cells of the immune system are primarily targeted by HIV and when? Why are these cells so important in immune function? What roles do they play in the immune system?
7. Describe the life cycle of HIV in detail. Understand what happens in each of the steps shown in the WWW reading.

8. What is a retrovirus? How does it differ from other viruses?

9. What is gp160? What does its name stand for? What are the names of the subunits comprising this molecule? Which of the subunits is involved in the attachment phase of the viral life cycle? How is it involved in this stage?

NEWS ITEMS: For many years it has been known that some people are exposed to HIV but never develop AIDS. This led some to hypothesize that HIV is not the cause of AIDS. New data have shed light on why a person can be HIV+ and not get AIDS. A group at the National Cancer Institute examined the amino acid sequence of CCR5 in 1,995 people. They found that there are a variety of CCR5 alleles in the population (genetic variation) and everyone they found who was homozygous for a "mutant" allele of CCR5 was not infected with HIV. This mutant allele has a 23 base pair deletion (note that it is not multiple of 3) which caused a non-sense mutation and the mutant protein never leaves the ER. A second study has been conducted with slightly different numbers, but both found that the HIV-resistant allele was more common in Caucasians (17% and 11% for the two studies) than in African Americans (1.7%) or Africans (0%). This may help explain part of why African Americans are over represented in the AIDS population. Some have speculated that one reason for this might be that HIV is not a new virus and that Europeans have been exposed to this before. Since genetic variation can lead to selective advantage (survival of an HIV epidemic), that would explain why Caucasians have a higher frequency of the resistant allele. (See summary by Jon Cohen. *Science*. Vol. 273: 1797-1798. 27 September, 1996; Samson *et al.*, *Nature*. 382: 722. 1996; or Liu *et al.*, *Cell*. 86: 367. 1996)

Recent results have made the mechanism of HIV infection even more complex and troubling. A French group has discovered that another molecule (called US28) can act as a coreceptor for HIV. Unfortunately, US28 is not a human protein but a viral one. The virus that contains the US28 gene is called cytomegalovirus (CMV) which is very common. It is estimated that 80% of the population has been infected with CMV. As it turns out, the molecular structure of US28 resembles CCR5. When the researchers put the US28 gene into cells that lacked either CCR5 or CXCR4, these cells that used to be resistant to HIV infection are now capable of being infected with HIV. So now the question is whether CMV has an active role in destroying the immune system in AIDS patients. For example, CMV might be able to infect cells that lack CCR5 or CXCR4 and thus provide a new host cell for HIV. (See summary by Michael Balter. *Science*. Vol. 276: 1794. 20 June, 1997.)

----------STOP----------

Treatments for HIV and AIDS

So, how can HIV's life cycle be inhibited in a way that harms the virus but leaves the HIV-infected individual unharmed? The major problem in finding effective anti-viral agents is that viruses use so many of <u>our</u> proteins in replication (e.g. DNA polymerase, RNA polymerase,

glycosylation enzymes, ribosomal proteins, spliceosomes, etc.). HIV contains only nine genes encoding nine proteins. (The simplest retroviruses contain only three genes.) All of the other proteins required for the viral life cycle come from our cells. For this reason, it is very difficult to inhibit a virus without inhibiting our own cells at the same time. Bacteria, on the other hand, are free-living organisms with their own enzymes. They have been separated from us by evolution for so many years that their enzyme systems are usually quite different from our own. Thus, we can treat bacterial infections with **antibiotics** which function by inhibiting the action of proteins or enzymes that are peculiar to bacteria and not shared by humans. Thus, you can fairly easily inhibit the growth of bacteria without harming yourself.

While researchers have had a hard time devising such an agent, our immune systems specialize in making such fine distinctions. Thus, when we become infected with the flu, mumps, measles, chicken pox, etc., our immune systems can usually eliminate the invading virus without harming us in the process. However, in the case of HIV, **the virus attacks the very cells that are responsible for its elimination**. Thus, it knocks out our defenses leaving us unable to kill the virus or, as the disease progresses, any other microbe. Defenseless against microbial attack, AIDS victims are ultimately killed by microorganisms growing out of control in the body.

Study Questions:

1. Why do strategies for producing anti-viral agents differ dramatically from those used to develop anti-bacterial agents?

2. In general, what are antibiotics and how do they work? Why don't antibiotics work against viral infections?

3. Why isn't HIV eliminated from the body in the same way that the viruses that cause colds, flu, chicken pox and measles are eliminated?

Due to the rush of recent research results, many new therapies are under development and at various phases of clinical trials. Here are some approaches that are being tested to cure AIDS.

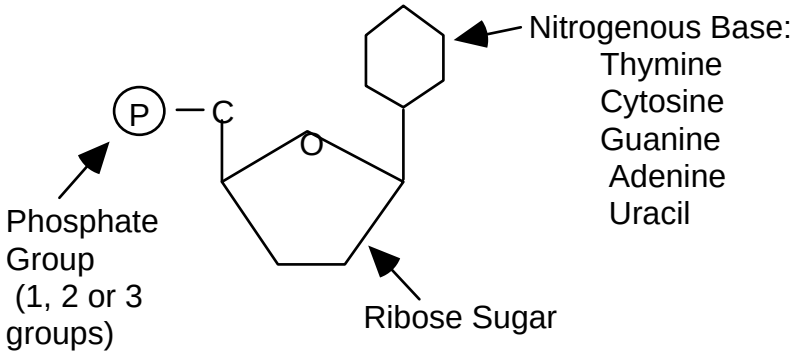
Focused reading: p 375 figure 19.25

Example #1 It has been known that T_c cells (also called CD8+ cells) are capable of secreting a "factor" that is capable of stopping the spread of HIV. At a meeting in December of 1995 (*Science* 270: 1560.), several research teams (including one headed by Robert Gallo who is helped discover HIV in the first place) announced that they had discovered this mysterious and elusive "factor". With hindsight, it's easy to see why identifying this factor was so difficult - it is actually three factors that work as a group. The factor is comprised of three chemokines **RANTES**, **MIP1-α**, and **MIP1-β**. (The names are acronyms that stand for **R**egulated-upon-**A**ctivation, **N**ormal **T** **E**xpresses and **S**ecreted; Macrophage Inflammatory Protein #1- a and 1-b.) Although the mechanism for inhibiting HIV replication is not known, the more recent discovery that CCR5 is a coreceptor is very exciting because it is known that RANTES binds to CCR5! For the first time in years there is a lot of optimism for discovering a way to treat and/or prevent AIDS. The most obvious explanation is that these three factors bind to CXCR4, CCR5 resulting in the inability of gp160 to bind to CD4+ cells. There are at least 14 pharmaceutical companies that are developing drugs that will interfere with

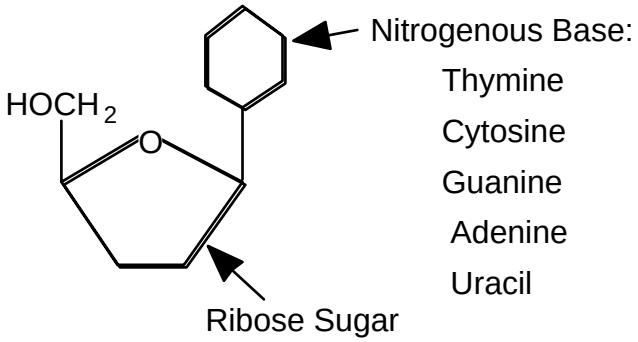
HIV's ability to bind to CCR5 and/or CXCR4. (For more information, see the excellent summary: *Science* 275: 1261-1264. 28 February, 1997.)

WWW Reading: Life Cycle of HIV - Reverse Transcriptase

Example #2: Interfering with reverse transcription of viral RNA: In a second therapeutic approach which interferes with the viral life cycle, about a dozen drugs have been developed which interfere with the process of reverse transcription. As of April 1997, the FDA (Food and Drug Administration) has approved 7 out of 11 of these drugs by for treating HIV infection and AIDS. The most popular drug treatment for AIDS is **AZT**. This drug's chemical name is 3'-**az**ido-2', 3'-deoxy**t**hymidine. AZT and six other drugs (only four of the six have FDA approval) are **nucleoside analogs**. One might ask, "What is a nucleoside?" Well, you know what a nucleotide is because you've encountered them over and over in looking at how DNA and RNA are synthesized and in looking at the energy molecule ATP (a triphosphonucleotide). The basic structure of a nucleotide is this:

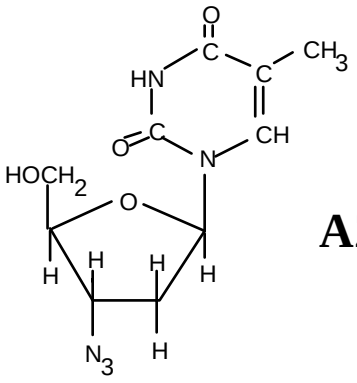


Nucleotides have three components: nitrogenous base, a ribose sugar and one, two or three phosphate groups. ATP, GTP, CTP, TTP, ADP and AMP are all nucleotides. A nucleoside is simply a nucleotide without any phosphate groups. It looks like this:



In making the nucleotides it needs to make DNA, RNA and the energy molecules, the cell takes nucleosides and phosphorylates them. Thus, nucleosides are the starting material for the manufacture of nucleotides.

A nucleoside analog is a molecule that looks so much like a naturally occurring nucleoside that the cell mistakes it for the real thing, makes it into a nucleotide and then incorporates it into DNA or RNA in the place of the naturally occurring molecule. For instance, AZT looks very much like the nucleoside precursor of thymidine. Below is the structure of AZT. (Compare it with the structure of thymidine).



AZT

You will notice that the nitrogenous base component (thymine) of both compounds is identical. The ribose of AZT does not have an oxygen on carbon #**2'** making it this sugar deoxyribose. The only difference in the molecular structure between normal deoxyribose (p 84) and this deoxyribose is the N_3 group (the azido group is N_3, the same compound we used in the Ames test - sodium azide) on carbon #**3'** in AZT. If you look at the chemical name of the compound, it actually 3'-azido-2',□3'-deoxythymidine. It tells you that the molecule is thymidine (has a normal thymine base in it), that it is deoxythymidine (meaning that it contains deoxyribose (missing an oxygen on carbons), that it also is missing an oxygen on carbon #3' and that it has an azide group there instead. Chemical names are exquisitely meaningful if you know how to interpret them. They tell you the actual structure of the molecule (take organic chemistry to understand biology fully).

Because the thymine part of the molecule is identical in thymidine and AZT, the cell mistakes AZT for thymidine. Thus, AZT functions as a **thymidine analog** in the cell. While you could certainly make nucleoside analogs for cytosine, adenosine, and guanosine, if you are trying to interfere with DNA replication, you are much better off using a thymidine analog because RNA does not use thymidine (it uses uracil instead) and therefore the normal process of protein synthesis so important to cellular life will not be affected.

When reverse transcriptase incorporates AZT into the growing DNA strand instead of thymidine, no further elongation of the DNA strand can occur. In other words, AZT stops replication. Normally, in DNA replication, the next nucleotide is added by dehydration synthesis to the OH group of the 3' carbon of the previous nucleotide. However, in AZT, this OH group has been replaced by an azide group and, thus, the next nucleotide cannot be added (no hydrogens and oxygens to 'dehydrate' into water). You have encountered this concept before in looking at DNA sequencing technology where dideoxynucleotides prevented strand elongation. In fact, the other two FDA approved drugs in this class are dideoxynucleotides that stop DNA synthesis in the same manner as AZT.

Because they inhibit DNA synthesis, AZT and other nucleoside analogs inhibit the ability of reverse transcriptase to make a cDNA copy of itself. This step is crucial to the viral life cycle. If it is inhibited, viral replication will be blocked and the virus will "die."

The principle limitations of AZT therapy are: 1) it is not a cure for the disease; 2) the half-life of the drug is fairly short, requiring that patients take tablets approximately every 4 hours; 3) its ability to extend the life of the AIDS victim diminishes with time (drug "tolerance" develops); 4) the drug does not appear to delay the onset of AIDS in asymptomatic HIV^+ individuals; 5) the drug is expensive, costing approximately $7,000 per year; and 6) AZT has a number of toxic side effects including nausea, rash, insomnia, vomiting, malaise, headache and severe anemia. Only 60% of AIDS patients can tolerate AZT therapy for more than one year.

A problem with the widespread use of AZT is the development of AZT-resistant strains of HIV. The use of any anti-microbial drug will act as selection pressure on the microbial population (evolutionary selection at a microscopic level). If a mutation occurs that that allows the microbe to live in the presence of the drug, the widespread use of the drug will give this mutant a competitive advantage over non-mutated microbes that were killed by the drug. Thus, the widespread use of AZT is undoubtedly favoring the development of an AZT-resistant strain of HIV. Such a strain, or strains, certainly exist and may be responsible for some of the cases in which AZT has lost its effectiveness in certain individuals.

WWW Reading: Life Cycle of HIV - Viral Protease

Example #3 is what made helped make David Ho (an AIDS researcher) *Time* magazine's Man of the Year for 1996 and *Science*'s Breakthrough of the year 1996. This approach is the second generation based off of the previous example and utilizes a cocktail of three drugs simultaneously - AZT, a non-nucleoside analog that inhibits RT (two of the four available have FDA approval), and a protease inhibitor (there are four FDA approved inhibitors available). As you will remember, HIV must have its multi-protein complexes cut into functional pieces in order to survive. Once the three dimensional structure of the protease was determined, many companies and researchers rushed to develop drugs to block its action. This has lead to a significant reduction in the number of HIV particles in the blood of AIDS patients. Unfortunately, there are still some negative side effect and now the cost of treatment has risen to about $12,000 per person per year.

> NEWS ITEM: There is an ethical dilemma when it comes to testing drugs. As you know from your laboratory work, every experiment must have a control. When new drugs are being tested, you must administer a placebo to a subset of the people in order to see how well they do without any treatment. The triple drug cocktail has been so successful, that the experiments have been canceled before they were completed because the group getting the treatment was doing so much better than the control group. But the fact remains that the experiment was not carried out completely. If allowed to continue, would the control group have appeared more similar to the experimental group? No one knows for sure and when testing a live saving drug, it is difficult to watch the control group get worse knowing that you might be able to prolong their lives if they were unluckily placed in the control group. (See summary by Jon Cohen. *Science*. Vol. 276: 520-523. 25 April, 1997)
>
> A new compound called calanolide-A has been isolated from trees in Malaysia that blocks reverse transcriptase in the lab. This is being developed for clinical trials. But an even cleverer trick has been designed. The problem has been to deliver a deadly compound to lymphocytes that are infected with HIV but not healthy lymphocytes. Researchers at the National Institutes of Health have created a virus that expresses CD4 and coreceptors for HIV. These engineered viruses should bind to cells expressing gp160 and infect them. If the virus delivered a deadly cargo, then only infected lymphocytes would be destroyed. (Science 277: 1606. September, 1997)

Study Questions:

1. Explain the mechanism AZT uses to produce its anti-HIV effects.

2. What is a nucleoside? How does it differ from a nucleotide?

3. If you are given the structure of 2'-deoxythymidine, be able to change the structure into AZT.

4. Explain how the widespread use of an anti-microbial drug actually stimulates the development of a drug-resistant microbial strain.

5. What is a protease inhibitor and how does it fight AIDS/HIV?

6. What drugs are in the triple cocktail drug treatment for AIDS?

Vaccines for HIV?

On May 18, 1997, as a part of a commencement address at Morgan State Univ. in Baltimore, President Clinton called for the production of an AIDS vaccine within the next 10 years to be "a new national goal for science in the age of biology." Earlier, the National Institutes of Health (NIH) named Dr. David Baltimore (a Nobel laureate) to head a new AIDS Vaccine Research Committee.

NEWS ITEM: (An example of politics and science) Dr. Baltimore acknowledged in an interview that he was hesitant to accept the position until after the November 1996 elections. Had the Democrats retaken control of the House of Representatives, Rep. John Dingell (D-MI) would have become chaired the subcommittee that oversees scientific misconduct. Dingell had aggressively accused Baltimore of being a knowing coauthor on a research paper that contained falsified results - Baltimore was later proven innocent. "I certainly did feel that if the House became Democratic, I had to come to some understanding with [Dingell] before I could take the job." (See *Science*. Vol 274: 2005. 20 December, 1996.)

In order to understand how vaccines are developed, we need to return to the immune system and see how vaccinations protect against disease.

Focused <u>Review</u>: p 360 "Immunological memory ..." stop at "Animals distinguish..."

Before vaccines were developed, the only way for a person to become immunized to a disease was to get the disease and survive it. Given the nastiness of most infectious diseases, this was a pretty grim prospect, and most individuals died in their youth of infectious disease. If an individual contacts and survives a disease, he/she is **immune** to that disease, at least for a while. Thus, if you survived the bubonic plague, you could safely care for other victims and be protected from contracting the disease again. This immunity to disease is due to a feature of the immune system called **immunological memory**. When lymphocytes encounter an infectious organism for the first time, they are not prepared to fight off the infection and you become sick. Slowly, through expansion of the anti-microbial lymphocyte population and genetic changes in the lymphocytes themselves, you acquire memory for the infectious organism. If you survive the first round of illness, this memory remains in place and the next time you encounter that same microbe, you

"remember" it and can fight off the infection before the microbe makes you sick. Immunity is specific for a given microbe. Thus, immunity to influenza will not protect you from tetanus. Because lymphocytes interact specifically with foreign antigens, they develop specific memory.

Because the immune system functions by recognizing foreign molecular shapes, it will respond the same way regardless of whether or not an antigen is harmful. This immune system characteristic is exploited in the development of vaccines. A vaccine is a harmless version of a pathogen that has the same shape as the pathogen but has been altered in some way to make it unable to cause disease. Vaccines are impostors -- they look like dangerous microbes to the body, but they are not. The body raises an immune response (including a memory response) against that particular foreign shape, and the next time you encounter that shape (this time in the form of the real pathogen), your immune system will "remember" the previous encounter and destroy the pathogen before it can make you sick. Thus, you get the immunity without having to contract the disease.

In the developed world, childhood immunizations for many viral and bacterial diseases can be routine. We can vaccinate against the viral diseases measles, mumps, rubella, polio, rabies, yellow fever, small pox and hepatitis B; and against the bacterial diseases tetanus, diphtheria, whooping cough, pertussis, cholera, plague, tuberculosis, *hemophilus influenza* type b, meningitis, and pneumococcal pneumonia.

The very first vaccines were **surrogate pathogens**. Surrogate pathogens are microbes that naturally look like the real thing, but are not pathogenic. The best example of this is the very first vaccine ever developed the vaccine against small pox. Small pox was a virulent and deadly scourge that, along with the bubonic plague, has threatened most of the known world since the beginning of recorded history. Edward Jenner, an English physician in the 18th century, noticed that milkmaids very infrequently contracted small pox, even when the disease swept through their villages, afflicting almost everyone else. Jenner noted that cows sometimes contracted a very mild disease that had some of the symptoms of small pox (most notably open skin lesions.) The cow version of the disease was called "cow pox." Suspecting that milkmaids were is some way protected through their contact with cowpox, Jenner, who must have been a very gutsy guy, scraped some of the tissue from one of these open sores from an infected cow, and injected the material into a young boy. He then exposed the boy to small pox (from an open sore of a small pox victim.) The boy did not become sick from small pox. (Biomedical ethics committees would have you locked up for doing such a thing today!) Thus, Jenner discovered a way to protect against small pox. He called this potion a "vaccine" (after "vacca", Latin for "cow.") [This is also an example of how important it is to keep your eyes open and study many different organisms -- prevention of a lethal human disease can be aided by knowing cow diseases!]

NEWS ITEM: Small pox has been completely eradicated from the human population. The small pox virus is present in only two places on earth -- in a vial at the Center for Disease Control in Atlanta, and in a vial in a comparable institution in Moscow. Scientists are currently debating whether these vials should be destroyed, thus causing the extinction of the small pox virus. Some argue that the risk of escape is too great and the human population is now largely unprotected. Since the 1970s, people have not been immunized against small pox because the disease is no longer a threat. An escape of the virus now would cause heavy casualties. (Anyone up for writing an international terrorist novel around this theme?) On the other hand, scientists argue that it is a mistake to cause the extinction of this important virus before we do not really know what questions we would like to ask about its structure and function. We know so little about viruses that we don't even know what questions to ask, and if we destroy the virus now, we will

lose the answers to those questions forever; answers that may save lives if applied to the control of other viruses. What do you think?

We have come a long way since Jenner scraped cow sores and injected them into people. Today, we have a dazzling array of genetic engineering techniques at our disposal in the development of hi-tech vaccines. Since 1986, more than 15 AIDS vaccines have been engineered and tested in humans. Here are two examples:

1. Live, attenuated virus. These vaccines are living viruses that have been altered in some way to make them non-pathogenic even though they remain alive (like removing the fangs of a snake). Microbes can be attenuated by treating them in various low tech ways (e.g. adding certain chemicals to their media) or high tech ways (e.g. removing a gene that is necessary for infectivity, but not necessary for life). Live, attenuated vaccines give the most vigorous immunity because they behave like the real thing in the body -- they go to the same tissues, actually invading the body as a pathogen would, and are seen by the immune system in the same way as the pathogen. However, in the case of HIV, investigators have been reluctant to use this approach. Because the disease is virtually 100% fatal and because the attenuation process may not be 100% successful, the chance of a pathogenic virus being included in the vaccine is too great. Also, investigators have felt that, with all the other recombinant DNA technology available, they should be able to develop a safe, effective vaccine without resorting to the use of live, attenuated organisms. However, so far, alternative methods have failed to produce a vaccine, and, in December 1992, a group of investigators reported that they could prevent infection by Simian Immunodeficiency Virus (SIV) using a live, attenuated SIV virus with one gene removed. These are the best results to date, and may cause the AIDS research community to rethink their resistance to the use of live, attenuated vaccines. In December of 1995, HIV+ individuals who have never contracted AIDS were studied. In one study, all of the individuals had HIV strains that lacked the *nef* gene, which is necessary for a vigorous infection. There is increasing interest in this approach for vaccination.

NEWS ITEM: Dr. Baltimore's group has recently shown that the nef protein can actually make HIV undetectable to our immune system. It appears that when a cell makes *nef*, it also makes less MHC I molecules, the same molecules that help T_c identify which cells are virally infected. Maybe this explains why the *nef*- strains of HIV are not as potent as their wild-type relatives. (See *Science*. Vol 276: 1196-1197. 23 May, 1997.)

2. Cloned Envelope Glycoproteins (also called Subunit Vaccines because they contain only a subunit of the virus, not the entire organism.) These are the safest vaccines because there is no virus present to cause an infection. By applying genetic engineering techniques (all of which you have encountered already in this course) investigators have cloned gp160 and gp120, placed the cloned genes in expression vectors, and made large amounts of the glycoproteins. The idea, of course, is that gp160 and 120 are foreign to humans and should elicit an immune response. This immune response should then be able to see the natural gp160 or 120 on the surface of a real HIV, and target it for destruction (thus destroying the virus).

   These vaccines have been shown to produce an antibody response that reacts with HIV. However, they are not especially effective at preventing infection by HIV, though antibodies can protect us from other viral infections. However, these glycoproteins are not being presented to the immune system in the same manner that they would be if they were

embedded in the envelope of a virus. Thus, the immune system may respond with the wrong kind of immunity. When foreign soluble proteins (such as recombinant gp160 and 120) are injected into humans, an antibody or humoral response predominates. Antibodies are effective against soluble antigens because they can bind up and neutralize soluble protein. However, when membrane-bound molecules are presented to the immune system, they tend to stimulate a cell-mediated immune response aimed at killing the cell bearing the antigen. This is the arm of immunity that is responsible for eliminating viral infections. For that reason, research is underway to attempt to bind gp160 and 120 into more natural, membrane-bound configurations (e.g. binding the glycoproteins into liposomes or into large lipid-protein complexes) in an attempt to stimulate the correct type of immune response to protect against viral infection. As a result, many researchers are looking for vaccines that will stimulate a T_c response.

NEWS ITEM: In June, 1994, two gp120 vaccines were considered by the NIH for large scale trials but the NIH determined that the data were not strong enough to justify the expense so neither vaccine was examined further. The companies took their vaccines to Thailand, where there is a severe HIV epidemic and convinced members of the Thai government to conduct a large-scale trial. Now other members the Thai Ministry of Health want the trial canceled since the data analyzed by the NIH are not strong. The manufacturer of one vaccine, Genentech, has raised $24 million dollars from investors for the Thai trial and calls the attacks "myopic". They expect the trial to begin in 1998. (See *Science*. Vol. 276: 1197. 23 May, 1997.)

While progress toward an HIV vaccine has been slow, this degree of difficulty is typical in the development of viral vaccines. The vaccine for hepatitis B took 17 years to develop. However, HIV presents some unique problems to investigators who are trying to develop effective vaccines.

1. HIV has an extraordinarily high mutation rate in the genes for its membrane glycoproteins. The membrane glycoproteins are really the only part of the virus that immune system will be able to "see" since immune cells can only see the outside of structures. These glycoproteins mutate at a very high rate. Thus, a glycoprotein vaccine developed against one strain of HIV may be entirely useless against another as the virus continually changes the shape of its surface glycoproteins. In the case of influenza, a new vaccine must be developed by the Center for Disease Control every year because the changes in the surface protein shape caused by viral mutations make last year's vaccine unusable. HIV mutates 65 times faster than influenza. [Retroviruses tend to mutate at high rates, possibly because reverse transcriptase has poor editing abilities. Thus, the mistakes that are usually fixed by DNA polymerase during DNA replication are not fixed by reverse transcriptase. These mutations get incorporated into the viral genome and are passed on to the next generation of viruses.

2. HIV is a retrovirus and, after it has integrated into the host genome as a provirus, it can lie dormant for many years. During this period, it produces no protein products so it cannot be detected by the immune system. Thus, the immune system is powerless to eliminate the virus when it is in its latent stage.

3. The lack of a suitable animal model for the disease. Because the disease is species specific, no animal model can be used to test vaccines in a faster, more efficient manner than are allowed by the ethics of human trials. Chimpanzees (our closest relatives) do become infected with HIV, but they do not develop AIDS, and their use as test animals poses an increasing threat to the already dwindling chimpanzee population. While the pharmaceutical

industry is pushing the World Health Organization to relax restrictions on the importation of chimpanzees from Africa, scientists warn that this could have a devastating effect on wild chimpanzee populations, threatening their extinction. Some degree of relief to the primate population has come with the bioengineering of a mouse that contains a human immune system (called the SCID/hu mouse.) This mouse normally has a severe genetic immunodeficiency disorder that destroys its own immune system. A human immune system can then be seeded into the animals at birth. While HIV does not infect these animals in exactly the same way it does humans, some limited experiments are possible using this model.

NEWS ITEM: With the identification of the coreceptors for HIV, many research teams are racing to develop animal models for HIV. They can introduce human CD4, CCR5, and CXCR4 genes into animals in hopes that they will be able to be infected with HIV and develop AIDS. Unfortunately, not even this is a simple as you might think. It turns out that mouse cells grown in culture do not support the growth of HIV as well as human cells do. However, rabbit cells appear to be better hosts, so some teams are trying to engineer rabbits instead of mice. To give you an idea how specific HIV is for CCR5, the mouse CCR5 cDNA has been sequenced at it is 82% identical to the human protein and yet HIV cannot bind to the mouse CCR5. (See Atchison *et al.*, *Science*. Vol 274: 1924-1926. 20 December, 1996)

Study Questions:

1. How is immunity developed? What is immunological memory?

2. How do vaccinations work? What features of the immune system make vaccination a viable approach to the prevention of microbial disease?

3. Discuss the aspects of HIV infection and AIDS that make it especially difficult to develop a vaccine.

----------STOP----------

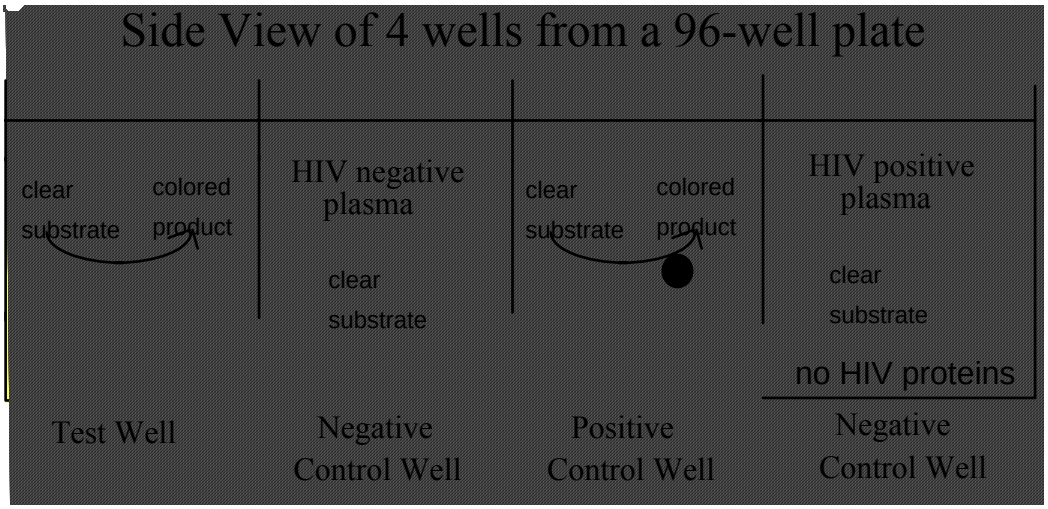
Diagnosis of HIV$^+$ Individuals

WWW Reading: ELISA for HIV

A blood test for HIV infection has been available for over 10 years. This test does not actually detect the virus in the blood, but rather it detects the presence of anti-HIV antibodies in the blood. If you are infected with HIV, you will make antibodies against the virus, thus allowing the detection of the virus through this indirect route. Antibodies are found in the **serum** (the fluid part of the blood minus the proteins that cause blood clotting) and, therefore, if the test shows that you have antibodies against HIV, you are said to be **seropositive**. If you do not have antibodies against HIV, you are said to be **seronegative**. If you were seronegative, but are now seropositive, you are said to have **seroconverted**. Because it takes from six weeks to six months for the level of anti-HIV antibody to rise to detectable levels, you can be HIV$^+$, but seronegative. If you think you may be infected with HIV, get a blood test. If it comes up negative, get another blood test six months later. Ninety-five percent of HIV$^+$ individuals will seroconvert within six months of infection. However, some investigators have reported that seroconversion may not occur for up to 36 months in rare instances.

The screening test for HIV is called an **ELISA** (**Enzyme-Linked ImmunoSorbant Assay**; invented by Eva Engvall of Sweden). This assay is based on the same principles as immunocytochemistry. In one version of the assay, the HIV virus glycoproteins are purified and stuck onto the bottom of the wells in a 96-well plate. Blood is drawn from the individual being tested. The blood cells are removed by centrifugation leaving the fluid component, called **plasma**. The individual's plasma is diluted and placed in a well containing HIV protein. As in all good experiments (especially ones that determine if someone has a lethal disease) control wells are included in the test. Negative control wells are filled with plasma from a person known to be HIV negative and plasma from the person being tested is put in a well that does not contain any HIV antigen. Positive control wells are filled with plasma from a person known to have high concentrations of anti-HIV antibody in his/her plasma.

The next steps should seem familiar. After an incubation period, the excess plasma is washed off, and a secondary antibody is added, usually a **mouse anti-human immunoglobulin** that has horseradish peroxidase conjugated to it (similar to turnip peroxidase). Antibodies are immunoglobulins, so everywhere human anti-HIV antibody has bound to the HIV glycoproteins lining the well, the secondary mouse antibody will bind, bringing along the enzyme peroxidase. If no antibody against HIV is present in the serum, nothing will bind to the HIV glycoproteins and the secondary antibody will also have nothing to bind to, so it will be washed away along with its peroxidase. In the final step, a peroxidase substrate is added to every well. This substrate is colorless when added but peroxidase will turn it into a colored product. Thus, a change in color in a well indicates a positive result. The ELISA is diagrammed below.



The change in color is measured by a **plate reader** (just like we did in lab) and the results are expressed in optical density units (OD units). A low OD indicates a negative well with no colored product, while a high OD indicates the presence of antibody against HIV, or a positive test result.

The ELISA assay is the most inexpensive assay for the presence of HIV antibodies. However, it is not the most reliable assay available. The American Red Cross estimates that the ELISA is accurate 99.8% of the time. In two times out of 1000, however, it will give a **false negative** or **false positive** reading. A false negative is a test that fails to detect the presence of anti-HIV antibody when it is present in the plasma. A false positive is a test that detects the presence of anti-HIV antibody when it is not present in the plasma. In the case of HIV, both types of errors can be devastating. Therefore, if a blood sample scores a positive result in the ELISA, a second test is performed. This second test is

called a **Western blot** and it is more reliable than the ELISA is, although considerably more expensive due to the time involved.

You have already encountered the **Southern blot** in Unit II. In this technique, restriction fragments of DNA are electrophoresed and then transferred to a piece of nitrocellulose where the DNA is hybridized with a probe. Two other types of blots are based on the same idea. In the **Northern blot**, RNA is electrophoresed and then blotted and probed. In the **Western blot**, protein is electrophoresed and then blotted and probed with an antibody rather than DNA. [A scientist named Dr. Southern developed the Southern blot. In naming the Northern and Western blots, the developers took advantage of the fortunate coincidence that Dr. Southern's name has three directional alternatives. No Eastern blot exists but a scientist with Asian heritage and a sense of humor developed a Far Eastern blot (detects protein binding)]

In the Western blot for HIV, the virus is highly purified and taken apart into its individual protein molecules. These molecules are electrophoresed and separated by molecular weights and blotted to nitrocellulose. As was done in the ELISA, these Western blots are incubated with plasma from the individual being tested, washed, and a secondary antibody conjugated to peroxidase is added. The blot is washed and soaked in a clear substrate that precipitates and turns dark when acted upon by peroxidase. Thus, all bands to which anti-HIV antibody is bound will turn dark when the substrate is added. Dark bands indicate a positive test, and, because the individual HIV proteins are separated by this technique, the test will also show against which HIV proteins the individual's antibodies are directed. If the Western blot results come back positive, the individual is considered HIV^+ and is notified of that fact.

Both screening and confirmatory tests for seropositivity test only the presence of antibody to the virus. While it is not used to screen the general public because of it expense, there is a test available that detects the presence of the virus inside T helper cells. The test is based on a general technique used to make many copies of a specific piece of DNA called the **polymerase chain reaction** (**PCR**; the same method we use in our last two labs). This test is used in experimental situations where it is absolutely essential to know whether or not someone is HIV^+.

Focused Reading: p 214 "The polymerase chain reaction…" to end of chapter
p 215 fig 11.21
WWW Reading: Cartoon of PCR Method

When PCR is used to clone DNA, one can start with a single copy of the human genome. In three to four hours, over one billion clonal copies of the DNA of interest can be made. Because the DNA primers are specific for the HIV gene you wish to amplify, in many cases you need not purify the DNA before you begin. In using this technique to detect the HIV virus, DNA is extracted from the white blood cells (which include T helper cells) of the individual being tested. This DNA is incubated in the presence of a pair of DNA oligonucleotides to act as DNA polymerase primers (of about 20 bases) which are complementary to a base sequence present only in the HIV viral genome and not humans. Thus, these primers will begin the process of amplification only if the viral DNA has been incorporated into the white blood cells of the individual. The resulting PCR product is electrophoresed to see if the band of the expected size is present. The PCR technique is so sensitive that it needs only one copy of the viral DNA in order to amplify it and allow its detection. Conversely, it only takes one stray cell to contaminate the sample.

Study Questions:

1. What is seroconversion? Why is it called this? What is the difference between being seropositive for HIV and being HIV^+?

2. Describe the ELISA as it is used as a test for HIV.

3. What is a false negative result? A false positive?

4. What is a Western blot? A Northern blot? A Southern blot? What do all these blots have in common? How are they different?

5. Describe the Western blot as it is used as a test for HIV? Why is this test used as a confirmation of a positive ELISA result?

Note: Another detection method hat relies on Western blot technology is the home pregnancy test. These tests are so reliable that gynecologists now tell women to use them rather than ordering tests from an outside lab. http://www.whfreeman/purves6e tutorial 19.2 goes over how they work and how they have been designed to include the all important 'control'

6. Describe the polymerase chain reaction. What reagents are required? What does this procedure do? In general, what are the steps in this procedure?

7. Describe the use of PCR to detect the presence of HIV. Why is this test far more accurate than the Western blot? Why is it not used as the routine screening test for HIV?

Future Directions

The major unanswered question is how the virus actually suppresses the immune system. T_h cells play a pivotal role in the function of the immune system. Because HIV infects T_h cells, it has been assumed that HIV spreads from T_h to T_h, killing the cells as it goes, until so few T_h cells remain that normal levels of immunity cannot be maintained.

It is certainly the case that T_h cells are destroyed during the progression to AIDS. Normal T_h cell levels are about 1,000 cells per ml of blood. By the time of the onset of AIDS, these levels have usually fallen to 200 cells/ml, and may fall to zero by the time of death. When the T_h cell level falls below 500 cells/ml, **opportunistic infections** begin to occur, and by the time the cell count falls to 200 cells/ml, these infections begin to occur regularly. AIDS used to be diagnosed at the onset of opportunistic infections. However, because the disease progresses differently in different individuals and the diagnosis of AIDS brings government-sponsored medical benefits to the individual, a more uniform guideline for AIDS diagnosis was required. Since April 1992, AIDS has been diagnosed when the T_h cell count falls below 200 cells/ml (an 80% reduction). This new definition increased the official number of AIDS cases in the United States by 55%.

While we know that AIDS victims have very low T_h cell counts and suffer and die from infections that are caused by the absence of a functional immune system, we do not know how HIV produces this crippling state. There are currently three competing theories for how HIV destroys the immune system:

1. HIV kills T_h cells directly
2. HIV stimulates other components of the body (T_cs?) to kill T_h cells
3. HIV causes T_h cells to commit suicide

For years, theory #1 was assumed to be true. However, several years ago it was found that, at the time in disease progression when the patient is losing T_h cells at the fastest rate, very little virus was present in the blood. This caused several investigators to wonder how HIV could be directly responsible for T cell death. However, there was considerable resistance among AIDS investigators to the idea that the direct killing hypothesis may not completely explain the disease. This hypothesis was vindicated to some degree by PCR analysis of lymph node cells from AIDS patients which showed that virus infects T cells in the lymph nodes and spreads in these organs throughout the course of the disease. Thus, the "latent" period of HIV infection may not be classical latency at all, but rather a period of incubation in the patient's lymph nodes.

Despite these findings, some troubling contradictions remain unanswered by the direct killing hypothesis. For example, investigators have known for some time that some strains of HIV are not able to kill T_h cells in culture (***in vitro)*** while others are. Yet, in experiments using mice with human immune systems, investigators found that the non-cytotoxic strains were able to deplete T_h cells in the animal (***in vivo)*** at a faster rate than the cytotoxic strains. It may be the case that the virus makes the T helper cell a target for destruction by T_cs or some other immune system cell. Thus, according to this theory, the virus simply marks the T helper for destruction, but does not destroy the cell itself.

In support of the third hypothesis, T helper cell suicide, investigators have shown that, if you take HIV^+ T helper cells from the body and stimulate them with antigen, they will commit suicide, a process called **apoptosis** or **programmed cell death**. Normal cells will begin to divide and differentiate, but HIV^+ cells will die. Thus, according to this hypothesis the virus does not directly kill the T helper cells, but rather it programs it in some way to kill itself at a later time. Spooky, eh? Of course, these three theories are not mutually exclusive, and all three processes may be acting to destroy T helper cells.

NEWS ITEM: It is believed that macrophages (or other immune cells with analogous functions - e.g. dendritic cells throughout the body, astrocytes and microglia in the CNS) are the other central player in HIV infection that needs further study. Many believe that the macrophage is the reservoir for HIV. Think about this - where do all the viruses come from if T_h cells are mostly dead? Secondly, many HIV proteins are neurotoxins and an HIV+ macrophage kill neurons and lead to the development of neurological symptoms that up to one third of all AIDS patients develop. Another factor is how HIV can cross the blood-brain barrier. It is reported that astrocytes can be infected but produce few viruses. Finally, the reason T_h cells die has never been explained, but many feel that infected macrophages may induce apoptosis in astrocytes and maybe T_h cells as well. (See summary by Michael Balter. *Science*. Vol. 274: 1464-1465. 29 November, 1996.)

Finally, everyone wants to know how HIV can evade cytotoxic T cells so well. Activation of the T_c requires the interaction of the T cell receptor on the T_c with a MHC Class I molecule that is displaying a viral peptide. In November 1995, it was shown that viral peptides in MHC I molecules that vary only slightly from the T_c recognizable peptide can inactivate (or **anergize**) the T_c (remember the News Item describing the effects of *nef* on MHC?).

We raise the issue of how HIV causes AIDS to allow you to see that the "obvious" answer is not always the right one, and it is extremely important to keep an open mind about things, even when a

dominant theory makes perfect sense. For every natural process there are many, many explanations that make perfect sense, though most are false. Truth in science does not depend on the quality of a rationale. Rather, it depends on the quality of evidence, gathered through work at the laboratory bench.

Major Stories to Follow:

1) Dr. Mary Klotman at Mount Sinai School of Medicine has isolated a different factor (a very small protein) that appears to suppress HIV. She calls this factor CD8+ Antiviral Factor (CAF). *Science*. Vol. 276: 1197. 23 May, 1997.

2) The late Dr. Angeline Douvas at UCLA has found that people infected with a harmless goat virus (caprine arthritis encephalitis virus - a distant relative of HIV) produce antibodies that bind to HIV. She will pursue this line of research by conducting epidemiological study to see if people infected with the goat virus have a lower incidence of HIV infection. *Science*. Vol. 276: 1197. 23 May, 1997.

3) Dr. Miles Cloyd at UT-Galveston has found that HIV was unable to reproduce in about 15% of the T_h cells he isolated from randomly chosen individuals. The significance of this story is that HIV is able to infect the cells, but once inside, it fails to reproduce. If this story holds up, it would suggest a different mechanism is available from simply blocking the entry of HIV. *Science*. Vol. 275: 1258. 28 Feb., 1997.

4) Antisense therapies are making a comeback in many areas and AIDS is no exception. Antisense technology is fairly simple in theory, but has many practical obstacles. Since all proteins are derived from mRNA, if there were a way to insert a molecular sponge to soak up all the HIV mRNA, then you would have killed HIV. To do this, you synthesize a short piece of RNA or DNA that has the complementary sequence to your target mRNA. When these two sequences get together (base pair), the mRNA cannot be translated and it is destroyed by the cell (see p 364 and fig 16.13 for more) The two big tricks are; a) which sequence do you choose that will bind to only HIV mRNA and b) how do you get these antisense molecules inside cells? There as been a great deal of improvement in part b, and only trial and error will solve part a.

----------STOP----------

III. Genetic Engineering

There are two major areas of genetic engineering - cloning and creating transgenic organisms. We will look briefly at cloning and then focus on transgenics.

Focused Reading: p 296-99 last paragraph, stop at "Genes are…"
WWW Reading: How to Clone Your Own Dolly

Cloning Organisms

Plants are very easy to clone, in fact many do it naturally. When a plant sends out a runner and establishes a new individual without reproducing sexually, that is cloning. Cloning means the production of genetically identical individuals. When you take a clipping from one plant, put it in some water until it has roots and then plant it, you have cloned an organism. Scientists have learned how to clone more plants by starting with single cells and growing them in tissue culture. But the big news in 1997 was cloning a mammal.

Animals in general and mammals in particular have been more difficult to clone. Amphibians have been cloned before but it was not until Dolly stunned the world (see the cover of March 10, 1997 issue of Newsweek). Previously, mammals had been cloned naturally and in the lab by separating embryonic cells and allowing each cell to grow into a different individual (thus identical twins). But Dolly was the product of a mature nucleus and an undeveloped cytoplasm (see WWW reference). But there are some problems with this type of cloning. Your chromosomes are like batteries - they are designed to keep going for a set length of time but eventually, they do expire. (Even the Energizer Bunny will die at some point.) The telomeres of chromosomes are the limiting factor and every time your chromosomes replicate, a little bit of the life span of your telomeres is lost. So what will happen to Dolly? Will she die at an early age? If so, with a sample size of one sheep, can you make any conclusions?

> NEWS ITEMS: As of August 1998, sheep, cows, and mice have been cloned from somatic cells. Some wonder if this technology could lead to way to save some endangered species from extinction. This leads to a hot debate on where conservation money should be spent, but cloners do have a unique argument. In species such as cheetahs where the gene pool is too small for long-term survival, there is a need to introduce new alleles into the breeding population. Years ago, researchers isolated and froze cells from adults. Now it might be possible to use the nuclei from these frozen cells to produce new animals with different alleles to be introduced into the population via normal matings. (*Science* Vol. 276: 1329. May, 1997)

To date, cloning is primarily used in plants and certain animals, so it is more a novelty at this point compared to the bigger technology of producing transgenic organisms.

Transgenic Organisms

Focused Reading: p 325-8 "DNA manipulation…" stop at "DNA fingerprinting…"

WWW Reading: A Portable Gene Gun
Making Trangenic Livestock

Throughout Units II and IV, we have introduced the idea that genes can be moved from one organism to another where they can be expressed as the protein product. The transfer of genes to

expression vectors is an example. However, this technology can also be used to move genes into more complex, multicellular creatures such as laboratory animals, livestock, and plants. Such transplanted genes are called **transgenes** and the organisms that bear these genes are said to be **transgenic** organisms.

In the case of unicellular organisms or cells in culture, you simply have to put the DNA in with the cells, create conditions that enhance DNA uptake, and wait for the cells to take up the DNA. However, if you want to create and entire multicellular organism that contains the transgene in every cell of its body, you have to put the gene in the embryo of the organism (for animals at least, see below for plants). In that way, the transgene will be replicated along with all the other genes of the organism, and passed on to every daughter cell. This type of genetic engineering is called **germ line** engineering because, once the gene is incorporated into the embryonic cells, it is present in all of the cells of the resulting adult, including its sperm or eggs. Thus, the gene is passed on to the next generation of organisms. Once you get one male and one female transgenic animal, you can have a **transgenic strain** simply by breeding them to one another.

To create a transgenic animal, you give a female animal fertility drugs which cause her to "superovulate" -- that is, make many, many eggs. You then harvest the eggs just before they burst from the surface of the ovary and place them in a dish with sperm collected from the male of the species. (By the way, this procedure is also done in humans and is called **in vitro fertilization**. In humans, so far anyway, the purpose is to enhance fertility, not to manipulate genes.) The egg and sperm join and form a zygote. At this stage, the transgenes (constructed with a promoter that will turn the genes on at the appropriate time or in the appropriate cell) are **microinjected** into the zygotes. The zygotes are allowed to grow in the tissue culture dish to the 2-8 cell stage and are then implanted in the uterus of a pseudo-pregnant female (having the hormones of pregnancy without actually being pregnant. The investigator administers the hormones.) When the offspring come to term, they are tested to see which of them carry the transgene by either a Southern blot or by PCR. Given all the steps in this procedure at which something could go wrong, the chances of producing a transgenic offspring are about 1 in 10 births, and much lower odds if you count every implanted embryo.

Plants are a bit easier to work with than animals because in many species the entire plant can be regenerated in tissue culture from a single adult cell. Thus, you do not have to manipulate the plant embryo. You simply have to insert the transgene into an adult cell and then grow the cell under the correct conditions in plant tissue culture. A new plant will grow, and every cell of the new plant will contain the transgene. Plant cells can be infected with a plasmid, or virus, bearing the transgene or virus.

The US Department of Agriculture regulates the field trials of transgenic crops and livestock. More than 370 permits have been issued in 35 states for field tests of transgenic crop plants. These plants include:

Plant	Transgene
Alfalfa	Herbicide tolerance, virus resistance
Apple	Insect resistance
Oilseed rape	Herbicide tolerance, insect resistance, modification of seed oils
Cantaloupe	Virus resistance
Corn	Herbicide tolerance, insect resistance, virus resistance, wheat germ agglutinin
Cotton	Herbicide tolerance, insect resistance
Cucumber	Virus resistance

Melon	Virus resistance
Papaya	Virus resistance
Potato	Herbicide tolerance, virus resistance, insect resistance, starch increase, and modifications to make a variety of non-potato products such as chicken lysozyme.
Rice	Insect resistance, modified seed protein storage
Soybean	Herbicide resistance, modified seed protein storage
Squash	Virus resistance
Strawberry	Insect resistance
Sunflower	Modified see protein storage
Tobacco	Herbicide tolerance, insect resistance, virus resistance
Tomato	Virus resistance, herbicide tolerance, insect resistance, modified ripening, thermal hysteresis (frost resistance.)
Walnut	Insect resistance

From Kareiva, *Nature*, Vol 363, pg. 580, June 17, 1993

While, as you can see, most of the first generation of transgenes tested so far confer resistance to viruses, insects and herbicides, the long-term hope is also to be able to engineer the plant product such that it provide more nutrition, e.g. higher levels of protein. Scientists are working to develop plants that can fix their own nitrogen, thus eliminating the need for nitrogen-based fertilizers.

NEWS ITEMS: Dr. Rafael Palocios and his colleagues from the University of Mexico in Cuernavaca have produced a better nitrogen fixing bacterium. They started with wild-type *Rhizobium* and then added additional *Rhizobium* DNA to these cells along with an antibiotic resistance gene. They then used evolution to find the best transgenic bacteria. They gradually increased the antibiotic concentration and selected bacteria that could still survive. Then they search for improved nitrogen fixation among the survivors. They do not know which genes have provided the improved fixation, but this method can be used to other bacteria to produce nitrogen fixation in non-legume plants. (Source: Daily InScight (http://www.europe.apnet.com/inscight/) sponsored by Academic Press and *Science* - research published in May issue of *Nature Genetics*, 1997.)

A transgenic mouse has been created that develops sickle cell. This mouse will be useful for testing treatments and potential cures. See Transgenic Mice from the Bio111 web page for more information.

The tobacco industry is under fire (no pun intended) and the farmers of North Carolina need to look for alternative crops. Their future may be linked to transgenic tobacco grown on **ph**arms. It has been shown that tobacco plants can produce functional human antibodies if they are given the correct DNA. Likewise, they can produce other **ph**armaceutical products like growth hormone, blood clotting factors, and insulin (look over pages 324-5 for how and some examples). Instead of a few dollars per bushel, these plants may well be worth their weight in gold, if not more!

The primary animal that has been used with transgenics has been the mouse. Researchers have been altering the genes of mice for many years and there is a large number of transgenic mice being studied (see WWW site for a small sampling). One mouse has been made to have a human immune system so we can better understand our immune system. Another mouse has been made that has twice the normal amount of skeletal muscles. This could be used to understand and perhaps treat muscle diseases like muscular dystrophy. In addition, now that we know how to make a "mighty mouse", we could make mighty cattle and produce twice the beef. A very popular transgenic approach is called the "**knockout mouse**" which means that both alleles at a particular locus have been deleted. This allows us to understand the role that the encoded protein plays in a living organism by determining the phenotype of a knockout mouse. Pharmaceutical companies

are making many of these mice in order to develop new therapies. In fact, Merck has hired (for $8 million) Lexicon Genetics of TX to make 150 new knockout mice. (see p320-1 and fig 17.10 for overview of the knockout technique)

The existence of transgenic plants and animals is, of course, troubling to many. There is legitimate concern that these genetically engineered species are not tested by natural selection and, if they escape and breed with natural populations, may confer a defect to the species and threaten its extinction. It is also of legitimate concern that the bioengineered species will exert selection pressure on viruses, weeds and insects to evolve into forms that can overcome the genetic trait of the transgenic organism. As in the case of AZT, by giving one species an artificial advantage, you always change the selection pressure on competing species, thus changing the current niche and altering their evolution.

A large and complex area of patent law has arisen along with transgenic technology. In 1988, the first transgenic mouse was patented. Of course, if companies go to all the trouble to produce a transgenic mouse strain, they want the proprietary rights to the animal. Normally, if you make a product and want exclusive rights to its sale, you get a patent. But, no one had ever tried to patent a living creature before. This raises all kinds of problems. For instance, what if I buy a transgenic mouse (or hog or goat) from someone that holds the patent. Then I want to breed this animal and produce my own line of transgenic animals. Can I do this? Or does the original patent owner own the exclusive rights to breed? Here's another problem. What if a transgenic organism is patented and then someone comes along and changes one base pair in the transgene and creates a second transgenic organism that makes an identical protein product? Slightly different transgene, but identical product. Does the original patent cover this transgene? If you are interested in biology and law, this might be the career for you, since it is a good bet that this controversy will be raging for years to come

Possibly most troubling, however, is the capability that these transgenic organisms represent. We know from almost a century of biomedical research that our biology is not essentially different from that of other mammals. If you can bioengineer the germ-line of a mouse or a goat or a hog, you can bioengineer the germ-line of a human being. In fact, as mentioned above, we already do one of the hardest steps of this process-- harvesting eggs and fertilizing them *in vitro*. The Human Genome Project coupled with transgenic technology will mean that we might be able to bioengineer virtually any genetic trait into the germ-line, as an inheritable feature. While this could be a great benefit to families with inherited genetic diseases, this technology raises unprecedented ethical questions. What will be bioengineered? Cures for diseases? IQ? Skin color? Classical beauty? What do we mean by "normal"? What pressures will parents be under to ensure that their offspring are genetically "normal"? If you don't bioengineer your offspring, will they be able to sue you for negligence? What will your family and community think of you if you choose to "go natural" and conceive your child the old fashioned way? Will bioengineering coupled with genetic testing create whole new categories of discrimination? People predisposed to cancer (would you hire them? What about health care costs?), people predisposed to violence (would you want them teaching in our schools?), people predisposed to forgetfulness (would you want them fixing the airplanes you ride in?), etc. What if only the wealthy can afford to bioengineer their children, but everyone is genetically tested? Right now these questions are the plots of novels but soon, who knows. (Remember the entire field of recombinant DNA manipulation didn't even exist 30 years ago.) If this area is interesting to you, you should take some of the medical ethics courses.

As was the case with nuclear energy, the revolution in biotechnology provides immense power to those who control it. Power that can be used for the tremendous benefit of society or in the service of evil. We humans don't have the best track record in using power wisely and for the good our fellow humans. While we cannot predict what the future holds, we can predict that the biotechnological revolution will dramatically change our lives and the lives of our descendants.

Study Questions:

1) Define what a transgenic organism is and how this differs from the way gene therapy is used to treat cystic fibrosis.

2) Give one benefit and one disadvantage inherent in creating either a transgenic animal or plant.

3) Describe the techniques used to introduce a transgene into the potential host cell, either plant or animal.

4) What is a knockout mouse?

5) What is antisense technology and how does it work?

6) Be able to cite examples of transgenic organisms and the product they are designed to produce.

7) How might gene therapy be used to generate a T cell-mediated vaccine for AIDs?

8) In an attempt to treat people with high blood cholesterol levels, I have decided to create a transgenic cow that will produce human apolipoprotein C2 (APOC2) in her milk. APOC2 binds to cholesterol in the blood and so it might be useful as a treatment for people with high cholesterol. I would like to employ you as my biotechnology consultant so you could advise me on how to design the transgene. What advice would you give me with regards to the best promoter to use and correct targeting of the APOC2 protein? In other words, how could you get this new protein to be expressed only in the milk and nowhere else?

9) Describe how the famous sheep Dolly was cloned?

----------STOP----------

Study Questions for Comprehensive Aspect of the Final

1) Describe the common themes found in cellular communication (e.g. the roles of calcium, ion gradients, phosphorylation, ligand binding to receptors, etc.). You do not have to know each enzyme in every pathway that we studied though specific examples used correctly will enhance your answer.

2) Explain the concepts of:

a) signal transduction
b) receptor-ligand interactions
c) amplification of the message
d) second messenger
e) turning off the signal

3) Be able to interpret a pedigree: predict the outcome of a Mendelian cross (e.g. 1:2:1); predict the probability of a certain genotype if you are given the phenotypes of the parents; know the major steps of mitosis and meiosis and how the two types of nuclear division differ, on the macro scale.

4) Understand the big picture of gene expression. Do **not** focus on every component (e.g. single-strand binding protein) but be able to explain the major events, especially in regards to cancer, AIDS, and transgenics.

5) Know what is consumed and produced in:

a) photosynthesis (light and dark reactions)
b) cellular respiration
c) fermentation
d) chemiosmosis

Do **not** worry about step-by-step details, but focus on the overall process of each. This is not the same as saying to memorize the overall equation - be able to follow the energy in a general sense.

6) Be able to interpret a Southern blot or a DNA gel. You will **not** need to deduce a restriction map.

7) Be very familiar with the Ames test and PCR experiments that you conducted in lab. This includes the theoretical aspects as well as the logistical ones (hint: be able to name the parts on the 3D model of DNA).

8) For example, here is a particular question that requires you to use what you have learned in different sections. In 1992, it was learned that human eggs secrete a protein that binds to a receptor located in the middle piece (see page 740, fig. 42.9) of the sperm tail. These receptors, which resemble the odor receptors in your nose, help the sperm “smell” the egg and swim towards it. Design a contraceptive that uses this aspect of a sperm’s ability to locate an egg.

----------Don’t STOP here… ----------