# *Unit I: Cellular Communication*

Living beings can be composed of a single cell (*e.g.,* bacteria, cyanobacteria, and protists such as *Paramecium* and *Chlamydomonas* (a single-celled, photosynthetic organism that you will meet in lab)) or many cells. Not surprisingly, organisms composed of many cells are called **multicellular organisms**. An adult human is a very organized collection of about 70 trillion cells. (If you counted these cells at a rate of one cell per second, it would take you over two million years to count every cell in your body.) With a few exceptions (*e.g.,* red blood cells), each individual cell in a multicellular organism is a living entity with a complete set of genes and life maintenance equipment. Each cell maintains its own existence in addition to making 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 when your brain sends a message to contract. 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 each other in order to coordinate their functions and maintain the organism. 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, unicellular organisms must swim toward nutrition or sunlight if they are photosynthetic and must be able to sense when conditions are right to reproduce.

In this unit, we will examine four examples of cellular communication:

1) how liver cells secrete glucose
2) how cardiac muscle increases force
3) how neurons tell muscles to contract
4) how an egg knows it is fertilized

Each system uses a slightly different communication system, and taken together, these four systems represent many of the cellular communication systems scientists understand thus far.

### 🕮 Overview Reading

***Note**: Yes, three chapters is a lot to read, but keep in mind this is **overview** reading and should be **skimmed** at this point. These chapters will be discussed throughout this unit, and we will go into more detail as indicated by "Focused Reading.' You do not need to remember every detail in this reading – just try to get the main concepts and a good idea of the topics we will encounter.*

- Chapter 2 • The Chemistry of Life
- Chapter 3 • Macromolecules & the Origin…
- Chapter 5 • The Dynamic Cell Membrane

## The Liver Produces Glucose in Response to Stress

Glucose ($C_6H_{12}0_6$) is the primary sugar that biological creatures use as fuel. Humans, like other creatures, burn (oxidize) this sugar into carbon dioxide ($C0_2$) and water ($H_20$), using the energy released by this oxidation process to perform life's many functions (discussed in detail in Unit III). To ensure that cells have enough glucose to burn (and, therefore, enough energy to perform essential functions), the body maintains a constant supply of it in the blood (about 1 mg glucose per 1 ml blood).

### 🕮 Focused Reading

***Note**: Whenever you see the heading "focused reading" you should read these short sections of your textbook carefully BEFORE continuing to read in this Study Guide.*

- p 40-42 "The structures…" to "3.1 Recap"
- p 49-53 "3.3 What are…" to "Chemically…"

### 💻 thelifewire.com Reading

- Tutorial 3.1 • Macromolecules
  (just consider the carbohydrate subsection for now)

**Web Reading**

- Carbohydrate section of CancerQuest
  http://www.cancerquest.org/index.cfm?page=32

However, our bodies respond to stressful stimuli by increasing the blood glucose level to ensure enough fuel to fight or flee from what is scaring us. This extra sugar comes from glucose stores in the liver. During meals, glucose enters the body, is transported in the blood, and is removed from the blood and stored for later use by the liver. To store glucose, the liver attaches many of the sugar molecules together (**polymerizes** them) to form a large storage molecule called **glycogen.** When glucose is needed (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 (**monomers**), which will be dumped into the blood to provide fuel for all of the cells of the body.

**Study Questions:**

***Note**: You will frequently encounter "study questions" throughout this* Study Guide. *Some answers to study questions can be found within in the preceding* Study Guide *text, but many answers will come from assigned text and/or web reading. You do not need to submit written answers to these study questions, but you should be prepared to discuss study questions in class and on exams.*

1. What is **glucose** used for in biological creatures? What is the function of **glycogen**? What is the relationship between glucose and glycogen? Why does your liver go to the trouble of converting glucose to glycogen then back to glucose? (Why not store glucose?)

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**? An **oligosaccharide**? Be able to recognize a monosaccharide and polysaccharide when you see one drawn.

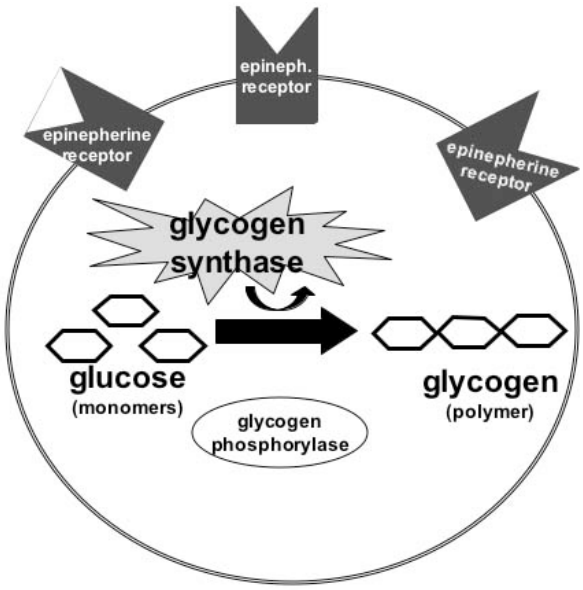
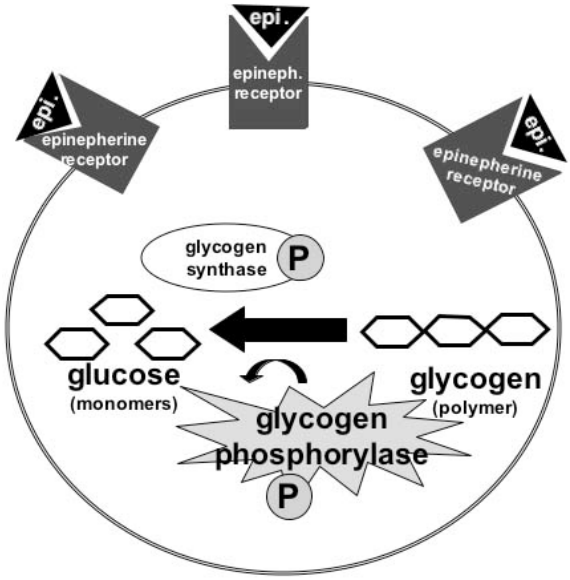
3. Glucose molecules are joined together to form glycogen by a process called **dehydration synthesis** (or **condensation synthesis**). Glycogen is broken down to form glucose by the process of **hydrolysis**. "Hydro-" means water and "lysis" means to break apart. What does water have to do with these 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.

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

***Note:** You will find interesting current research results included throughout the Study Guide as "News Items." You will not be tested on information in the News Items, but you can often enhance your understanding of important concepts and how modern biological research is conducted by reading about these recent advances in biology. For example:*

✎ NEWS ITEM: Cellulose microfibrils are strong fibers that encircle developing plant cells and provide structural support. In the summer of 2006 scientists discovered that microtubules (part of the cytoskeleton) tethered to the cytoplasmic side of the cell membrane actually help direct the deposition of very ordered layers of cellulose fibers via the enzyme cellulose synthase (which catalyzes the synthesis of cellulose). The researchers watched these molecules move in living cells by using fluorescent (glow-in-the-dark) versions of both the cellulose synthase enzyme and the microtubule protein tubulin in the model plant *Arabadopsis* (a model plant – sort of a fruit fly or lab rat for botanists). Their movies showed the cellulose synthase enzymes actually move along the microtubule tracks. [*Science* 312:1482]

After a meal, glucose molecules are joined together to form the polymer glycogen in the liver for storage. This process is called **glycogenesis** ("genesis" or creation of glycogen). An **enzyme** called **glycogen synthase** catalyzes the formation of each alpha glycosidic bond between glucose molecules. The following reading assignment describes enzyme function. Because enzymes are **proteins**, this reading assignment also describes proteins and protein structure. Further, to understand protein folding, you need to understand hydrophobic and hydrophilic groups.

| GLYCOGENESIS | GLYCOGENOLYSIS |
|---|---|
| A liver cell during **low stress** (no epinephrine) stores glucose by converting it into glycogen using the enzyme glycogen synthase. | A liver cell during **high stress** (+ epinephrine) breaks down glycogen by converting it into glucose using the enzyme glycogen phosphorylase. |
|  |  |





### 🕮 Focused Reading

- p 25-30 "2.2 How do atoms..." to "2.2 Recap"
- p 42-48 "3.2 What are:…" to "Environmental…"
- p 125-135 "6.3 What are enzymes?" to "6.5 Recap"

### 💻 Web Reading

- Protein section of CancerQuest
  http://www.cancerquest.org/index.cfm?page=34

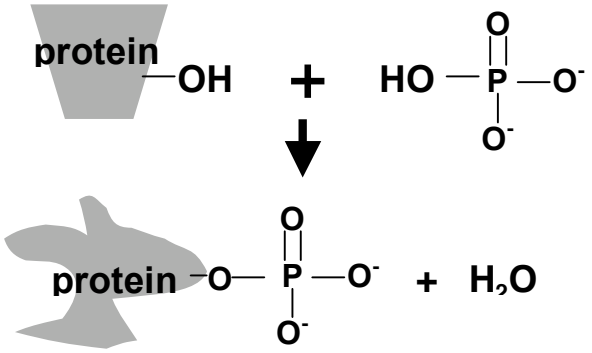
The enzyme glycogen synthase, then, lowers the activation energy barrier and allows glucose molecules to be linked together to form glycogen at a reasonable rate at normal body temperatures. Without glycogen synthase to catalyze the formation of alpha glycosidic linkages between glucose molecules, it would take a very long time to perform glycogenesis. **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 is increased when the cell increases the rate at which glycogen synthase catalyzes the reaction.

If the liver can make glycogen polymers, it must also be able to break down glycogen into glucose monomers. The enzyme that breaks down glycogen to glucose is **glycogen phosphorylase**. The breakdown of glycogen to glucose is called **glycogenolysis** ("**lysis**" for degradation). A liver cell can increase or decrease the rate at which glycogen is broken down simply by increasing or decreasing the catalytic rate of glycogen phosphorylase. (Note: glycogenolysis is different from the process of glycolysis, which we will discuss in detail in Unit III.)

The enzymes glycogen synthase and glycogen phosphorylase are turned on and off by the process of **covalent modulation**. This process is similar to **allosteric modulation** or **regulation,** except that covalent modulation depends on the process of **phosphorylation**. Phosphorylation is simply the covalent addition of a phosphate group

($PO_4^{3-}$) to an enzyme via dehydration synthesis. Phosphate groups can be added to molecules as simple as hydrogen atoms (H-$PO_4$) to enormous proteins.

Proteins are polymers of amino acids. Phosphate groups can only be added onto the side chains of certain amino acid residues by standard dehydration synthesis (onto an –OH group). Dehydration synthesis is also called a 'condensation reaction' --see page 41 fig. 3.4A). After a phosphate is covalently bound to an amino acid, the protein is said to be **phosphorylated.** Look at table 3.2 on page 43 in your text to figure out which three amino acids are the only amino




(not to scale--phosphate group is tiny compared to the whole protein)

acids that can be phosphorylated. (Hint: Look for 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 modification that acts as a kind of switch (or signal) to change the rate of an enzyme's activity. In order to 'turn off' this switch, a second enzymatic reaction is required. In contrast, allosteric modulation uses weak bonds (not covalent bonds) to regulate enzyme activity.

When you are scared, your liver slows the rate of glycogen synthesis and increases the rate of glycogen breakdown. Fear causes a series of reactions in the body that lead to 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 liver 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 (dephosphorylation occurs when you calm down), the rate of glycogen synthesis increases and the rate of glycogen 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, in chemical terms, 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 in chemical terms 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. How are enzymes turned on and off by **allosteric modulation**?

12. 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?

13. Give an example not covered in class 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.)

14. Explain the catalytic cycle (E + S ⇆ ES ⇆ P + E). Using this explanation as background, explain how each of the following events would increase the rate of an enzyme-catalyzed reaction: *(You will perform some of these manipulations in the IDH enzyme activity labs in coming weeks.)*
    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

15. 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.

---

We now know how the liver can liberate glucose from glycogen to increase blood glucose levels, but we are missing some very important elements of the system. That is, how does the liver "know" that the body is under stress? Your liver is sitting quietly in your abdomen -- it cannot see or hear stressful events. It has to be "told" that such an event is occurring. In multicellular creatures, the nervous system (brain and nerves) and endocrine system (hormones) tell the liver that something stressful is occurring.

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 collection of glands in the body that secrete hormones. Hormones are chemical messenger molecules that travel in the blood. A hormone travels throughout the blood system and affects **target organ(s)** (the liver in this case.) The hormone binds to receptors in or on the cells of the target organ and causes some change to occur in those target organ cells. Examples of hormones include insulin (lowers the blood sugar level among other things -- its absence causes diabetes mellitus), growth hormone (stimulates growth -- its absence causes dwarfism), and thyroid hormone (increases metabolic rate -- low levels cause coldness, weight gain, and lethargy.)

---

📖 **Overview Reading**
- Chapter 41 • Animal Hormones

📖 **Focused Reading**
- p 875-876 "41.1 What are…" to "Hormonal…
- p 879 Figure 41.4 (Epinephrine stimulates….)
- p 887-888 "The adrenal gland…" to "The adrenal…"

---

In the mobilization of glucose in response to stress, the hormone epinephrine (also called adrenaline) tells the liver that something stressful is occurring. Epinephrine is made (synthesized) and secreted by the **adrenal glands** (located just above the kidneys) 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 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, reducing physiological responses to stress is not very easy.]

The nerve impulse from the brain reaches the adrenal gland and causes these cells to secrete epinephrine into the blood. (The interaction between the nerve and the cells of the adrenal gland is an example of intercellular communication. We will discuss signaling by the nervous system later.) For now, understand that epinephrine enters the blood and goes everywhere -- to all the cells in the body.

Even though epinephrine goes everywhere, the hormone does not affect every cell of the body. Epinephrine *only* affects the cells that have **epinephrine receptors** on their surface. These receptors are proteins that are embedded in the cell membrane and can bind specifically to epinephrine in the same way enzymes bind to their substrates.

To summarize thus far, the adrenal gland secretes epinephrine when the brain "decides" that something stressful has happened. Epinephrine travels everywhere in the body via the bloodstream, but the hormone only binds to those cells that 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, the epinephrine bound to its 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**. Most cells transduce signals through a **second messenger system** that relays information from a cell surface 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.

### 🕮 Overview Reading

- Chapter 5 • The Dynamic Cellular Membrane

### 🕮 Focused Reading

- p 97-101 "5.1 What is…" to "Membrane carbohydrates"
- p 54-56 "Lipids…" to "Carotenoids"
- p 334 "A signal…" to "Responder"
- p 337-339 "G-protein…" to "cytoplasmic..."

### 💻 Web Reading

- Lipids section of CancerQuest
  http://www.cancerquest.org/index.cfm?page=36
- Crystal model of a lipid bilayer
  www.umass.edu/microbio/rasmol/cutctw.gif
- Fluid model of a lipid bilayer
  www.umass.edu/microbio/rasmol/cutftw.gif
- Combinations… section of CancerQuest
  http://www.cancerquest.org/index.cfm?page=40

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 smaller molecule that binds to a larger protein at a specific site. Hormones are ligands that bind to binding sites of hormone receptors.] A cell has many copies of a receptor that binds a given hormone in its plasma membrane. In addition, each cell has many different kinds of receptors -- one kind of receptor for every different extracellular signal molecule recognized by the cell. As an example, the liver interacts with epinephrine, growth hormone, thyroid hormone, insulin, glucagon, and many other hormones. Liver cell membranes therefore contain many copies of each of these different receptors, and each receptor type binds to specific hormones. The inclusion of many receptors within a cell membrane is part of the "mosaic" of the fluid mosaic model.

✎ NEWS ITEM: An international research team has found that a single ligand (the hormone estrogen) can bind to two different estrogen receptors, called alpha and beta. When the common ligand estrogen binds to the alpha receptor, the ligand-receptor interaction initiates gene activation. In contrast, when estrogen binds to the beta estrogen receptor, the interaction inhibits gene activation. Thus, the very same ligand can result in two very different outcomes (signals), depending on which receptor is present in the cell. [*Science* 277: 1508.]

❖ **Study Questions:**

1. What is a ligand? Give some examples.

2. Draw a diagram of a phospholipid that illustrates its distinguishing characteristics.

3. Explain, in chemical terms, why phospholipids are excellent molecular building blocks for cell 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 that floats freely in the cytoplasm. How would an integral membrane protein 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 chapter 5 figures for ideas). Be able to state the function and give a specific example for each type of protein.

7. Anti-estrogens are molecules used as drugs to treat and prevent breast cancer. Hypothesize a molecular mechanism to explain how the same ligand could give two different signals, such as described in the news item above.

We are now ready to put the elements of this story together by introducing the **cAMP** (pronounced "cyclic AMP" not "camp") **second messenger system** that links the epinephrine receptor to the phosphorylation of liver cell enzymes. This molecular 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 a category of molecules called nucleotides because cAMP (and ATP and GTP) are nucleotides.

📖 **Focused Reading**

- p 57-60 "3.5 What are…" to "DNA reveals..."
- p 338 fig. 15.7 (A G Protein-Linked Receptor)
- p 342 fig. 15.12 (The Formation of Cyclic AMP)
- p 340-341 "Second mess…" to "Second messen…"
- p 345 fig. 15.16 (Regulation of Signal Trans… )
- p 346 "Enzyme activities…" to end of page
- p 347 fig. 15.18 (A Cascade of Reactions…)
- p 893 "A Hormone can act…" to "41.4 Recap"

💻 **thelifewire.com Reading**

- Tutorial 15.1 • Signal Transduction Pathways

When you are in situations of low stress, your liver cells are busy synthesizing glucose monomers into glycogen polymers via the enzyme glycogen synthase. When you are in a high stress situation, your liver cells break down glycogen into glucose via the enzyme glycogen phosphorylase.

How does the liver cell get the message that something stressful has occurred? 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 text) to help you understand what is happening at the molecular level. Read the following sequence of signal transduction events sentence by sentence. Be sure you understand each event before you continue.

1) A stressful thing happens that the nervous system detects and then responds by sending a signal to the adrenal gland, which secretes the hormone epinephrine.
2) Epinephrine enters the blood and travels throughout the body.
3) When epinephrine encounters liver cells, it binds to epinephrine receptors in their membranes.
4) This binding of epinephrine to its receptor causes the receptor molecule to change its native conformation.

5) This change in shape reveals a binding site on the cytoplasmic (intracellular) side of the receptor.

6) The newly revealed cytoplasmic binding site is recognized by G-proteins in the cytoplasm that bind to the G-protein binding site on the epinephrine receptor. G-proteins are attached via lipid modification to the inner membrane.

7) This binding causes the G-protein to change shape (allosteric modulation).

8) This change in the G-protein's shape causes a GTP binding site in the G-protein to lose its affinity for the GDP nucleotide and gain affinity for a GTP instead.

9) The GDP leaves the G-protein and a GTP binds to the G-protein. (Note: a new/different GTP is bound. The 'old' GDP is not phosphorylated.)

10) This binding causes another alteration in G-protein shape that allows the G-protein to bind to the enzyme **adenylyl cyclase**, an integral membrane protein associated with the cytoplasmic side of the cell membrane.

11) When the G-protein binds to adenylyl cyclase, adenylyl cyclase changes shape, activating an enzymatic site on adenylyl cyclase

12) Activated adenylyl cyclase now binds the nucleotide **ATP** (substrate) and converts ATP into **cyclic AMP** (product).

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

14) PKA becomes activated when cAMP binds to it (non-covalent/allosteric modulation).

15) The activation of PKA causes this protein kinase to phosphorylate another enzyme inside the liver cell called **phosphorylase kinase**.

16) Activated PKA also phosphorylates another enzyme called **glycogen synthase.** [Valuable hint at no extra charge: All **kinases** phosphorylate a substrate. The word before "kinase" in the enzyme's name usually tells you which molecule the enzyme phosphorylates. For instance, hexokinase phosphorylates a hexose (a six carbon sugar). Phosphofructokinase phosphorylates phosphofructose (another six carbon sugar).]

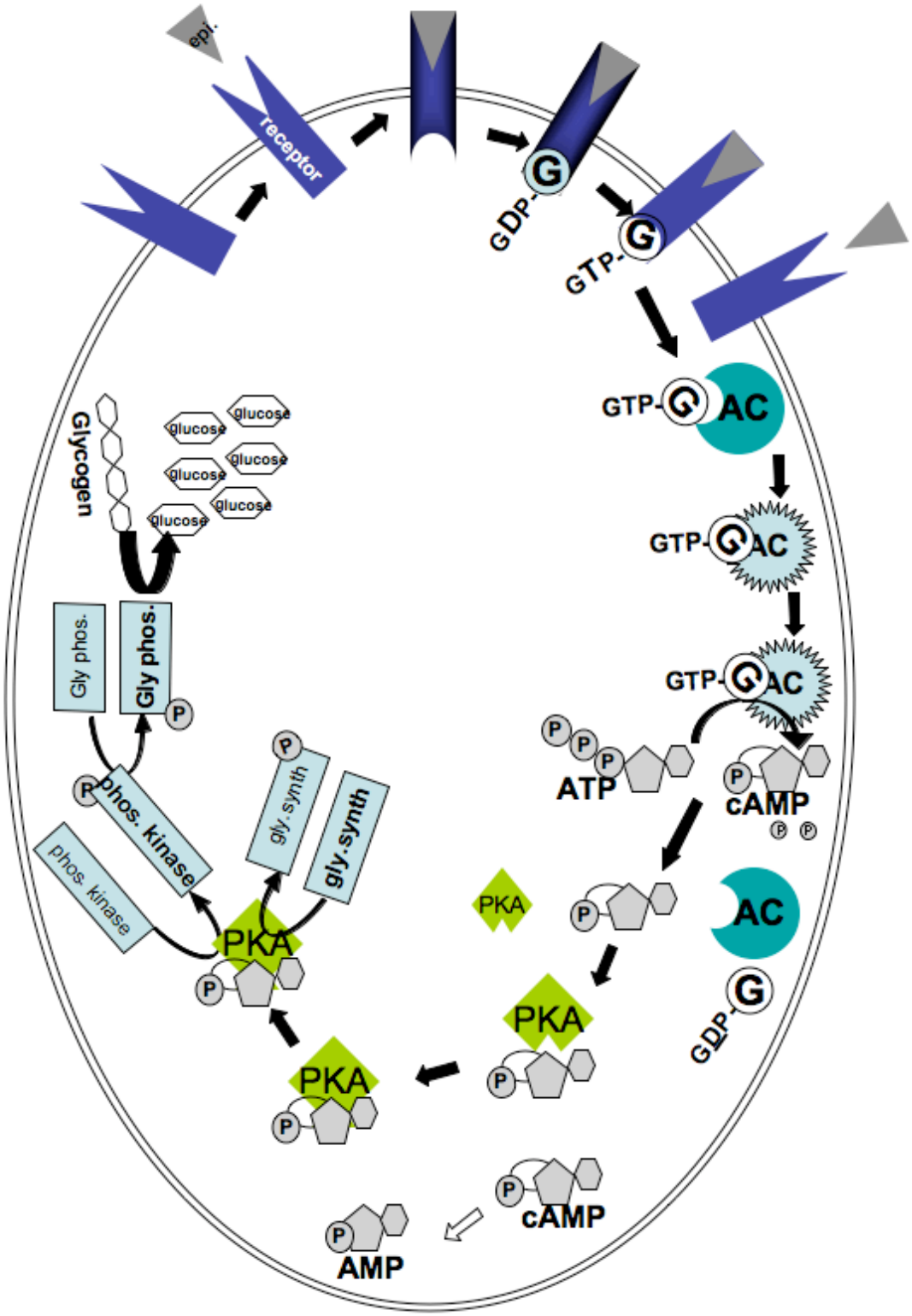
17) Phosphorylation by PKA activates phosphorylase kinase, which itself then goes on to phosphorylate another liver cell enzyme, **glycogen phosphorylase**.

18) The phosphorylation of **glycogen phosphorylase** activates this enzyme, thus allowing it to break down glycogen polymers into glucose monomers (**glycogenolysis**).

19) At the same time, **glycogen synthase** (mentioned as phosphorylated by PKA in step #16) is inhibited by phosphorylation. Therefore, **glycogenesis** (producing glycogen) is inhibited in the presence of stress, thus helping to keep glucose in its monomeric form.

20) Rapid glycogenolysis (and reduced glycogenesis) releases more glucose into the blood, and the blood levels of glucose rise, providing the organism with extra energy to react to the stressful situation.

This diagram summarizes the signal transduction steps in a liver cell when epinephrine triggers the hydrolysis of glycogen into glucose. The events in this diagram are also animated in Tutorial 15.1 (thelifewire.com) and are shown in fig.15.18 (page 347) of your textbook



❖ **Study Questions:**

1. What are the three major components of all **nucleotide** molecules? How are these components chemically 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 outlined above and describe how ligand binding affects the shape of each protein. What action is each protein 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 adding a phosphate group?

4. The cAMP second messenger system is an **enzyme cascade**. Why do you think this series of molecular interactions is called a cascade? What is adaptive (extra useful) about such a cascade? Why didn't the second messenger signaling 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. Do not stop until 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.

**Turning Off the cAMP Cascade:**

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 sugar high! We will now discuss three ways to turn off the "stress" signal delivered by epinephrine.

1) Epinephrine signal decreases
2) G-protein activation is inherently transient
3) Phosphodiesterase (PDE) converts cAMP into AMP

**1) Decreased Epinephrine**

Epinephrine (like 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. Thus, the epinephrine molecule eventually wiggles free from the receptor's ligand binding site simply because of constant motion due to kinetic energy. If the circulating epinephrine level is still high (*i.e.,* your nervous system is still stimulating epinephrine release by the adrenal gland), then another epinephrine *molecule* is probably in the neighborhood and will take the old ligand's place; therefore, the epinephrine receptor will remain activated. When the stress has ceased, the nervous system will no longer cause the adrenal gland to release epinephrine. Epinephrine is not a long-lasting molecule; the body quickly destroys free epinephrine. Consequently, epinephrine levels decrease when the stressful stimuli cease. Without the stress, when a molecule of epinephrine wiggles free of the epinephrine receptor, there will be very few molecules nearby to replace the ligand, and the hormone binding site on the receptor will remain unfilled or empty. If there is no epinephrine binding to its receptor, then the cAMP second messenger system will not be activated, and glucose will no longer be liberated from glycogen.

The "law of mass action" (*i.e.,* Le Chatelier's Principle) from chemistry also explains how the concentration of epinephrine in the blood is directly related to the amount of glycogen that is being broken down into glucose. According to this law, when the concentration of reactants increases, the rate of the forward reaction will

increase; this will use up the surplus reactants and favor formation of the products, thus reestablishing equilibrium. We can look at the free hormone and its binding site using the following chemical notation:

$$[H] + [B] \leftrightarrows [HB]$$

[H] = concentration of free hormone in the blood

[B] = concentration of free (empty) receptor binding sites

[HB] = concentration of binding sites containing hormone

As the concentration of free hormone ([H]) increases due to adrenal gland release, the rate of the forward reaction increases, and more hormones bind to receptors ([HB]). Likewise, as the hormone concentration decreases, the rate of the reverse reaction is increased and more hormones come free ([H]) of their receptor ([B]).

All hormones and signaling molecules have this direct relationship with their receptors. Therefore, the strength or degree of signaling depends on the hormone concentration. More hormone (ligand) molecules will cause more signaling. Less hormone (ligand) molecules will cause less signaling.

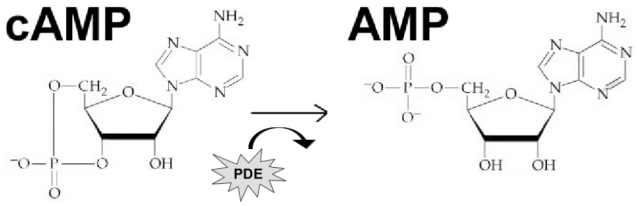
When the hormone concentration falls, the receptor has no hormone bound to the active site. This lack of a ligand causes the receptor to return to its original shape. In this original shape, the intracellular portion of the receptor cannot bind to or activate G-proteins. Thus, if there is no hormone bound, then there are no G-proteins activated.

### 2) Transient G-Protein Activation

G-proteins automatically turn themselves off (*i.e.,* they have transient activity). The G-protein very slowly (in about a minute) cleaves its GTP into a GDP by removing the terminal (last) phosphate from GTP. GTP thus keeps two phosphates and becomes GDP. With GDP bound instead of GTP, the G-protein goes back to its original shape and loses its ability to bind to the adenylyl cyclase enzyme. Therefore, no new adenylyl cyclases can be activated.


### 3) Conversion of cAMP to AMP by PDE

There's yet another molecular "off switch" inside liver cells—cAMP is degraded to AMP. The "cyclic" bond between the phosphate and the third carbon of the ribose is broken by the enzyme **cAMP phosphodiesterase (PDE).** A PDE breaks phosphodiester bonds—"diester" because it contains two oxygens (an ester linkage contains one oxygen) and "phospho" because it also contains a phosphate group. With or without stress, cAMP PDE constantly and quickly breaks down cAMP into AMP as soon as cAMP is formed. Consequently, cAMP concentrations are normally very low inside cells. During a stress response, PDE's cleaving cannot keep up with the amount of cAMP being made, so cytoplasmic cAMP concentrations rise and an important signal is transduced.




All three mechanisms that stop the cAMP cascade ensure that signal transduction is brief. In that way, if you need to continue to make extra glucose for the blood, your adrenal glands must continue to release epinephrine in response to input from your brain. Your brain, therefore, initiates 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.

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

4. Describe how enzymes are named. How can you determine what an enzyme does (even if you haven't encountered it before)? Here are some enzymes to practice on:
   pyruvate dehydrogenase
   ribulose bisphosphate carboxylase
   *(hint: look at carboxyl groups on text page 40)*
   tyrosine kinase
   DNA polymerase
   peptidyl transferase
   aminoacyl-tRNA synthase
   phospholipase

---

The cAMP second messenger system was the first signal transduction system to be characterized. Earl Sutherland received a Nobel Prize for this research in 1971. Since then, we have learned that many, many different types of cells use this cAMP signaling system to execute a very wide variety of cellular functions. Here are some of the many other examples of cell functions that use cAMP as a second messenger:

- 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* (single cell organisms that you will meet in lab)

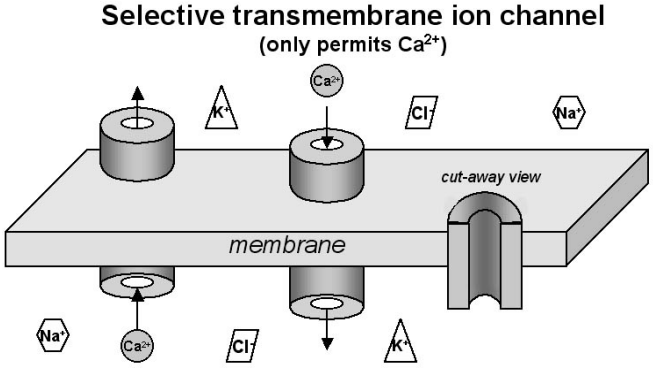
# How the Heart Pounds in Response to Stress

Everyone has experienced heart "pounding" during fear. The writer Edgar Allen Poe frequently mentioned this physiological response as a means to heighten the sense of terror in his readers. Fear is an emotion produced in response to things that are physically harmful, such as predators or dangerous situations. The heart pounding physiological response to fear prepares an organism to flee or fight the situation that is causing alarm. Both running away and fighting are physically demanding events that require more oxygen be delivered to muscle cells. (Active muscles need to burn more fuel for energy.) Blood carries oxygen to all tissues, but it can only carry so much. Thus, the only way to increase the oxygen supply is to increase the rate at which blood is delivered to the tissues. To increase blood flow, the heart beats faster and harder, making your heart contract more forcefully and feel as if your chest is pounding.)

Epinephrine released in response to fear (a type of stress) controls heart contraction force. As you know, the blood delivers epinephrine throughout the body to all tissues. Similar to liver cells, heart muscle cells bear epinephrine receptors in their membranes, called **beta-adrenergic receptors.** [Receptors that bind epinephrine (a.k.a. adrenaline) are called adrenergic receptors. If there's a beta ($\beta$), there's usually an alpha ($\alpha$), so it is not surprising that some other cells also express epinephrine receptors. FYI: $\alpha$-adrenergic receptors do not use a second messenger system.] This $\beta$-adrenergic epinephrine 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 (AC), and AC then converts ATP to cAMP. The cAMP then allosterically activates cAMP-dependent protein kinase (PKA). The similarities end here. The PKA in cardiac muscle cells phosphorylates two proteins that we will consider: (1) $Ca^{2+}$ channels in the plasma membrane and (2) myosin heads, part of a muscle cell's contractile apparatus.

In order to understand how the heart pounds more forcefully in response to fear, we need to look carefully at each of these two 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 have a hollow center (pore) filled with water that allows the passage of certain ions through the cell membrane. Each type of ion channel protein 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." Channels 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 that only allows calcium ions ($Ca^{2+}$) to pass through the pore:




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.

---

**Focused reading**

- p 27-28 "Ionic bonds…" to end of p 28
- p 105-106 "5.3 What are the..." to "Osmosis is…"
- p 108-109 "Diffusion…" to end of p 109
- p 108 fig. 5.10 (A Gated Channel Protein…)
- p 109 fig. 5.11 (The $K^+$ Channel)
- p 966 fig. 45.1 (Sensory Cell Membrane…)

**thelifewire.com reading**

- Tutorial 5.1 • Passive Transport

---

Although ion channel proteins allow ions to cross a cell's plasma membrane, the channels do not tell the ions which direction to flow (inside to outside or outside to inside). Ions can only flow through ion channels passively that is, **down their concentration gradient** from the region of higher concentration to the region of lower concentration. In most situations, the concentration of any given ion in the extracellular fluid (the fluid surrounding a cell) is very different from that ion's concentration in the cell's cytoplasm (cytoplasm can also be considered a fluid), so there is always a concentration gradient for a given ion. In the case of the $Ca^{2+}$ channel we are considering, $Ca^{2+}$ concentration is very high on the outside of the cell ($10^{-3}$ M) and very low on the inside ($10^{-7}$ M). Thus $Ca^{2+}$ moving through an open calcium channel in a cell membrane will follow its concentration gradient, flowing from the outside of the cell toward the inside of the cell.

Note that channel proteins cannot "pump" ions up their concentration gradient. If the concentration of $Ca^{2+}$ outside the cell (extracellular) is always higher than the cytoplasmic (intracellular) $Ca^{2+}$ concentration, then $Ca^{2+}$ will always flow from outside to inside and never from inside to outside. Ion channels can, however, be open or closed. When ion channels are closed, they do not allow any ions to move either direction across the cell membrane. When ion channels are open, they do allow ions to cross membranes. Any ion channel that can open and close is said to be **gated** because they have a 'gate' determining if ions can move through that 'path.' The cardiac $Ca^{2+}$ channel that we are discussing is gated. Different types of ion channels can be opened or closed in response to different types of stimuli (voltage, stretch, ligand binding, *etc.*).

The cardiac $Ca^{2+}$ channel is a "voltage-gated" channel; it opens and closes in response to a change in the **voltage** across the heart muscle cell's plasma membrane. These voltage changes occur rhythmically (about 80 times per minute), producing the normal heartbeat. 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, know that ion channels that open and close in response to changes in voltage are **voltage-gated channels**. Other types of channels open and close in response to ligand binding (**ligand-gated channels**) or to stretch (**stretch-mediated**

**channels**). We will consider ligand-gated channels later in this unit.

While the cardiac $Ca^{2+}$ channel we are discussing is voltage-gated, it can also be covalently modulated by being phosphorylated by PKA. When cardiac $Ca^{2+}$ channel is phosphorylated, it stays open longer than normal, thus allowing more $Ca^{2+}$ than normal to enter the heart muscle cell. The resulting higher concentration of intracellular $Ca^{2+}$ produces a more forceful contraction of the heart muscle cell.

---

❖ **Study Questions:**

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

2. Describe the chemical structure of an ion channel protein.

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

4. What does "gated" mean? What is a gated channel? What are the three types stimuli that might cause an ion channel to gate?

5. 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.

6. $Ca^{2+}$ 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?

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. The table below demonstrates why the magnitude of the difference between two concentrations (not the arithmetic difference of absolute values) regulate the direction and rate of ion flow.

While both concentration gradients have an arithmetic difference of 100 mM, 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.

| | Concentration Inside | Concentration Outside | Arithmetic Difference | Fold Difference |
|---|---|---|---|---|
| Concentration Gradient A | 1000 mM | 900 mM | 100 mM | 1.11 X |
| Concentration Gradient B | 200 mM | 100 mM | 100 mM | 2.0 X |

---

You know epinephrine activates the cAMP second messenger system in heart muscle cells (called **myocardial cells**) and that this increase in cAMP 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. In order to make logical connections that explain why these molecular events cause myocardial cells to contract more strongly, we need to address the following issues: Why is $Ca^{2+}$ always found at higher concentrations outside the cell? What creates this concentration gradient, and how is it maintained? How does more cytoplasmic $Ca^{2+}$ help the myocardial cell contract with greater force?

First, we will consider 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 (cytoplasmic) concentration of $Ca^{2+}$ very low when they are "at rest" - that is, when they are not receiving a signal. Then, if a signal (*e.g.,* a hormone) causes an increase in cytoplasmic $Ca^{2+}$ concentration, this rise in intracellular $Ca^{2+}$ provides an important signal to the cell, indicating that conditions have changed. Low cytoplasmic $Ca^{2+}$ levels tell the cell, "Do not secrete," "Do not contract," or "Do not pump ions" (whatever the cell does for a living -- do not do it). High $Ca^{2+}$ levels mean, "Secrete," "Contract," or "Pump ions" (*i.e.,* whatever the cell does for a living – when $Ca^{2+}$ is high inside it is the time to do your job).

This $Ca^{2+}$ signaling system must have two regulatory elements present in order for it to work correctly. First, the cell must have a way to keep the cytoplasmic $Ca^{2+}$ levels very low under normal, resting conditions. Secondly, the cell has to have a way to increase the cytoplasmic $Ca^{2+}$ concentration very quickly when a signal arrives. Because $Ca^{2+}$ plays such a critical signaling role in the cell, $Ca^{2+}$ is

frequently called a **second messenger** (or a third messenger, though most scientists do not distinguish between second and third. Do not take the numerical term "second" in "second messenger" too literally. In general, a second messenger is an intracellular messenger that is simply not the "first messenger.")

A rapid increase in cytoplasmic $Ca^{2+}$ concentration occurs when $Ca^{2+}$ ion channels open. Something happens (ligand binding, cell stretching, or a voltage change) that 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 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. Low intracellular $Ca^{2+}$ is maintained by an **active transport** system in the cell membrane and in the membrane of the endoplasmic reticulum that transports $Ca^{2+}$ out of the cytoplasm.

✎ NEWS ITEM: We sense heat because of ion channels in our skin. Each heat-sensitive ion channel works within an optimal range of temperatures. One channel in mice is able to sense heat above 33 °C and is produced in keratinocytes, cells which in the past were described as dead, flattened, and nearly functionless. Characterized in 2005, these same ion channels are also ligand-gated by camphor, the active ingredient in many commercial products producing a feeling of cooling similar to that of menthol. The effect of camphor is, in effect, sensitization to heat. These keratinocytes must contain a signal transduction cascade that eventually stimulates neurons to detect heat. [*Science* 307: 1468]

---

**📖 Focused Reading**

- p 111-113 "5.4 How do…" to "5.5 How do…"
- p 124-125 "ATP couples..." to "6.2 Recap"

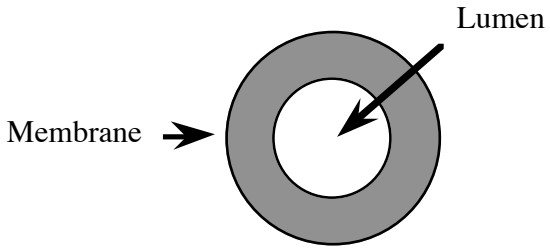
**💻 Web Reading**

- How the Calcium Pump Fills the SER
  www.bio.davidson.edu/misc/movies/pool.mov

---

Active transport is the movement of substances up their concentration gradient. Active transport appears to violate the second law of thermodynamics (that everything tends toward maximum randomness or entropy – page 119-120 of your text) and therefore active transport requires the cell to spend energy moving a molecule up its gradient. When the cell burns glucose to carbon dioxide and water, energy is given off. The cell harvests this energy and stores it within the phosphate bonds of ATP. When ATP is converted to ADP (by breaking off the terminal phosphate), the stored energy is released and cellular work can be performed using this energy. Very frequently, though not always, the terminal phosphate released during ATP breakdown is then covalently bonded to another molecule (*e.g.,* glucose, or a protein). The other molecule is thereby **phosphorylated**. You have encountered phosphorylation previously in its ability to activate or inactive enzymes by covalent modulation. Now you are encountering phosphorylation again. In this example, phosphorylation is used to provide the energy required to "pump" ions against their concentration gradient. It takes one ATP molecule to move two $Ca^{2+}$ ions against (up) the $Ca^{2+}$ gradient.

Myocardial cells have two sets of **$Ca^{2+}$ pumps**, or active transport systems, that remove $Ca^{2+}$ from the cytoplasm. Some pumps are proteins located in the plasma membrane that move $Ca^{2+}$ from the cytoplasm toward the outside of the cell. The other pumps are proteins located in the membrane of the endoplasmic reticulum (an organelle abbreviated ER) and they move $Ca^{2+}$ from the cytoplasm into the lumen of the ER. A **lumen** is the inside of a tube or hollow structure. 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.* Almost every cell has an ER, which is one of the many organelles in a cell. 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**.




Now that we're starting to discuss the SR, we need to refresh our understanding of all organelles in cells. Most, but not all, organelles are also tiny, specialized compartments in cells that are delineated by their own membranes.

### 🕮 Overview Reading

- p 10-13 "1.2 How is all…" to "1.2 Recap"
- Chapter 4 • Cells: The Working Units of Life
- p 70 fig. 4.1 • The Scale of Life

### 💻 thelifewire.com Reading

- Tutorial 4.1 • Eukaryotic Cell Tour

### 💻 Web Reading

- Cell Structure section of CancerQuest
  http://www.cancerquest.org/index.cfm?page=41
- Immunofluorescence Labeling of ER
  www.bio.davidson.edu/courses/Bio111/IMF.html
- Virtual Cell www.life.uiuc.edu/plantbio/cell/

*When 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 option will allow you to see more and more detail. You should also use the search function to select organelles.*

### ❖ Study Questions:

1. A major tenet in the biological sciences is that form follows function. Give an example that illustrates how a cell's form or composition allows that cell to perform a certain function. Be able to explain how your example illustrates this 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? What else is this size? How much smaller is a cell than a marble or a bowling ball or a typed period -- "."? What are the limiting factors in cell size? (*i.e.,* Why cannot cells be larger than they are? Why aren't they smaller?)

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

6. For each organelle and 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 <u>distinguishing features</u> 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.*) *Note: thelifewire.com includes 'flashcards' for each chapter that are useful for learning and practicing definitions.*

7. Today's reading includes two animations of the 'same' content. Compare Virtual Plant Cell and thelifewire.com, 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, and a nucleus.

✎ NEWS ITEM: The giant sulfur bacterium named *Thiomargarita namibiensis* is remarkable because it is a prokaryote with cells that grow to have diameters 750 μm (see "Relative Size' website for illustration of how big, or small, this is. For reference, 750 μm = 0.75 mm, making this bacterium visible to the naked eye). Scientists previously thought that a prokaryotic cell this large is not possible, but it indeed exists off the coast of Chile. How can this organism survive? You cannot change the laws of nature, so does it use mechanisms we haven't seen before? So far, scientists have shown that many big bacteria contain cytoplasmic inclusions (compartments or nearly empty bubbles) that reduce the amount of true cytoplasm, bringing the effective surface area to volume ratio back to a more predictable range. [*Science* 284: 493-95]

Now, back to the cardiac muscle cell. Both pumps <u>remove $Ca^{2+}$ from the cytoplasm</u>, either by pumping it outside the cell or into an organelle, the SER. The process of pumping $Ca^{2+}$ into the SER 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 below), which is outlined in the steps below:

1. The pump is a transmembrane protein with $Ca^{2+}$ binding and enzymatic abilities. We begin our study of the $Ca^{2+}$ pump cycle at an arbitrary point: the $Ca^{2+}$

pump is dephosphorylated and its $Ca^{2+}$ binding sites are vacant and facing the cytoplasm, possessing a very high affinity for calcium ions.

2. $Ca^{2+}$ floating in the cytoplasm binds to the $Ca^{2+}$ 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 ions that bump into the pump's binding sites will bind tightly and stay there. The conformational change resulting from $Ca^{2+}$ binding to the pump now causes the vacant ATP binding site to possess a very high affinity for ATP, and thus ATP also binds to the pump.

3. When ATP binds to the pump, ATP's terminal phosphate is transferred to the pump, **phosphorylating** the pump (and consequently dephosphorylating ATP into ADP).

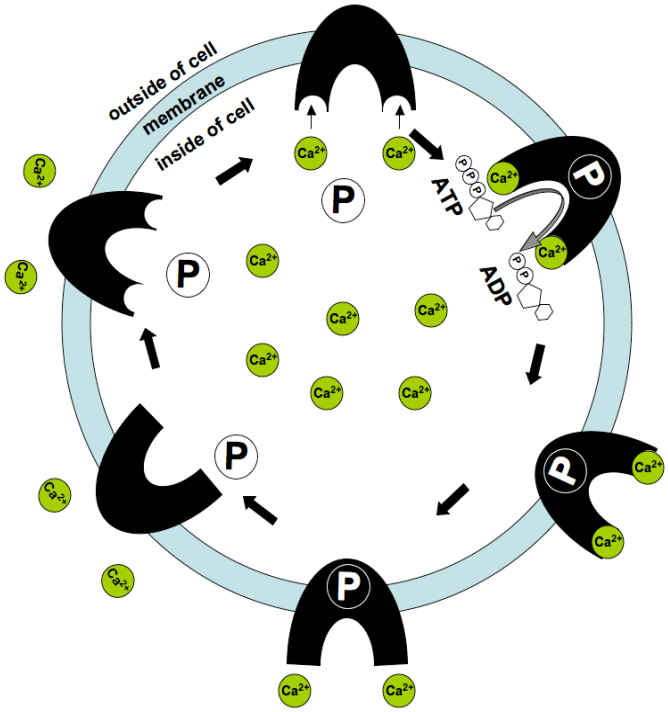
4. This phosphorylation causes the pump to change conformation and "flip," presenting the $Ca^{2+}$ binding sites to the other side of the membrane. For pumps in the plasma membrane, this flip delivers the $Ca^{2+}$ binding sites to the outside of the cell, but for pumps in the SER membrane, $Ca^{2+}$ is now delivered to the lumen of the SER. (If this concept seems confusing, review the thelifewire.com Tutorial 5.2 on active transport.)

5. Flipping its $Ca^{2+}$ binding sites to the other side of the membrane causes the $Ca^{2+}$ binding sites of the pump to decrease their affinity for $Ca^{2+}$. This decrease in affinity releases the $Ca^{2+}$ from the pump, allowing $Ca^{2+}$ to diffuse into the extracellular fluid (or lumen of SER).

6. The release of $Ca^{2+}$ from the pump's binding sites causes another 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.

The process of moving $Ca^{2+}$ against its gradient by the $Ca^{2+}$ pump is called **ATP-dependent $Ca^{2+}$ transport**. The pump is called an **ATP-dependent $Ca^{2+}$ transporter.** ATP plays an important role in $Ca^{2+}$ transport, providing 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 that allow phosphorylation or dephosphorylation. We spend a lot of energy pumping ions. It is estimated that we spend 10-20% of all the calories we consume just in the active transport of ions. Maintaining ionic gradients across cell membranes must be very essential to life if we expend so much energy on the process. Active transport accomplishes several other functions, but here we will focus on one of the main functions, signal transduction.




❖ **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.

---

OK, let's get back to the myocardial cell. To summarize so far, the resting myocardial cell had maintained an ionic gradient using membrane proteins that actively pumped $Ca^{2+}$ into the extracellular space and inside the SER lumen, spending ATP in the process. However, the brain has interpreted something in the environment as frightening, and it has sent nerve impulses to the adrenal gland to stimulate epinephrine secretion. 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. Epinephrine binding has triggered the cAMP second messenger system, which has activated PKA that has phosphorylated (using ATP as the phosphate source) voltage-gated $Ca^{2+}$ channels in the myocardial cell membrane. This phosphorylation has caused this ion channel to remain open longer than normal. $Ca^{2+}$ has moved down its concentration gradient into the cell through the open $Ca^{2+}$ channel. The phosphorylated $Ca^{2+}$ channel remains open longer than normal allowing more $Ca^{2+}$ than normal to enter the muscle cell.

How does this extra $Ca^{2+}$ in the heart muscle cause an increase in myocardial cell contraction strength? In order to address this question, we need to look at how muscle cells contract. All cells use their **cytoskeleton** to maintain their shape and to move when necessary. All cells have a cytoskeleton (cyto = cell), but only some cells move. Cells that are specialized for contraction have very specially organized cytoskeletal protein components. These components are specialized **microfilaments** (described in general on page 86-87) called **actin** and motor proteins called **myosin**.

---

📖 **Focused Reading**

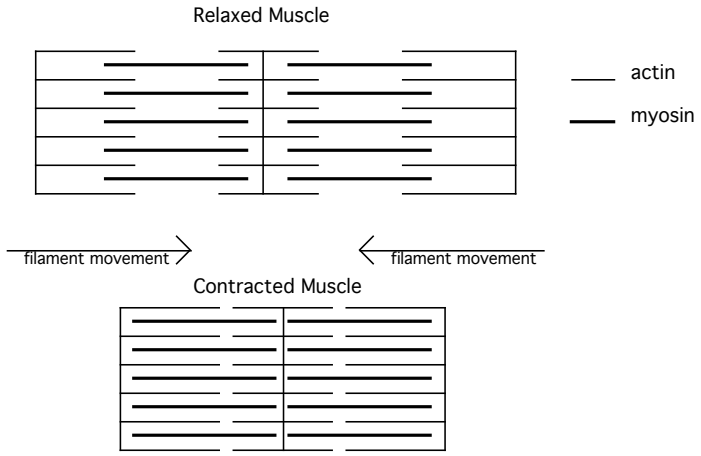
- p 1006-12 "Sliding…" to "Single skeletal..."
- p 1006 fig. 47.1 (The Structure of Skeletal Muscle)
- p 1007 fig. 47.3 (Actin and Myosin Filaments…)
- p 1008 fig. 47.5 (T Tubules in Action)
- p 1009 fig. 47.6 (The Release of $Ca^{++}$…)

💻 **thelifewire.com Reading**

- Tutorial 47.1 • Molecular Mechanism of Muscle Contr…

---

According to the **sliding filament theory**, muscles contract when actin and myosin filaments slide past one another as shown in the diagram below and in tutorial 47.1 at thelifewire.com




The actin and myosin protein fibers overlap one another within muscle cells. When they slide past one another the cell contracts. This sliding movement requires significant amounts of ATP to occur. Vigorous movement of parts of the myosin filament, called **myosin heads**, produces the sliding by forming **cross-bridges**. Myosin heads bind to

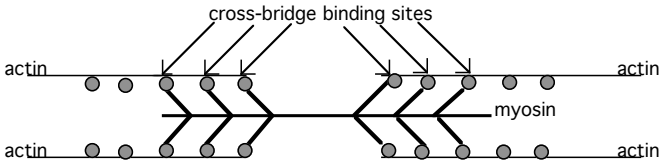
the actin and pull, then release and reset, then bind and pull, then release and reset. This process is very much like rowing a boat.

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

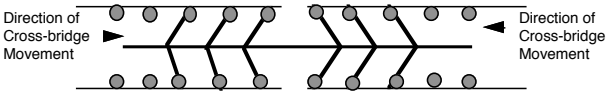
When cytosolic $Ca^{2+}$ levels are very low, $Ca^{2+}$ is not bound to troponin (for the same reason that ligands are not bound to proteins when their concentrations are low -- same concept, different example). When $Ca^{2+}$ is not bound to troponin, troponin has a particular shape that allows tropomyosin to cover the cross-bridge binding sites. The myosin heads cannot bind and contraction cannot occur. These protein configurations change when $Ca^{2+}$ levels are high. $Ca^{2+}$ binding causes troponin to change shape and this change in troponin shape pulls tropomyosin away from the cross-bridge binding sites. The myosin heads (always ready to bind to actin) 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 cytoplasmic $Ca^{2+}$ concentration occur $\geq$70 times per minute (on average) in the heart when you are not frightened; this oscillation in $Ca^{2+}$ concentration produces your regular heartbeat. However, epinephrine's effect on the plasma membrane $Ca^{2+}$ channel (via phosphorylation by PKA) allows the $Ca^{2+}$ channel to remain open longer, allowing more $Ca^{2+}$ than normal into the heart muscle cells. Higher $Ca^{2+}$ levels uncover more cross-bridge binding sites than normal. Uncovering more myosin-binding sites on the actin allows more cross-bridges to be formed -- more oars pulling in the water means more strength—thus, the force of the cardiac contraction is increased.

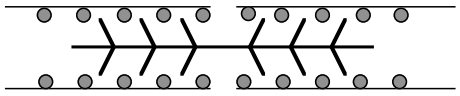
Besides phosphorylating the $Ca^{2+}$ channel, PKA has an additional action to stimulate the heart—it phosphorylates the myosin heads. These phosphorylated heads are capable of "rowing" at a faster rate. Therefore, the phosphorylated myosin heads can produce more strokes per millisecond. Because the 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.

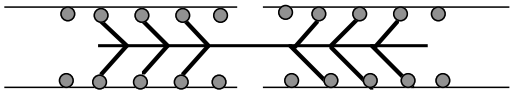


The myosin heads make cross-bridges when they bind to actin.



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.

---

❖ **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 sibling.
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 commonly used to lower blood pressure and to ease the strain on a weakened heart. Explain the mechanism by which beta-blockers reduce heart strain.

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

---

✎ NEWS ITEM: Psychophysiologists at the University of Southern California have observed a correlation between low heart rate at rest and aggressive and antisocial personality traits. (*Journal of the American Academy of Child & Adolescent Psychiatry*. 36: 1457-1464) A colleague commented that this was an interesting finding, but what do you do with this information? Is this a good hypothesis? Can you design an experiment to test this hypothesis?

✎ NEWS ITEM: You now know that a G protein-coupled receptor participates in the initial response to epinephrine in the liver and the heart. G protein-coupled receptors (GPCRs) are involved in many signalling events in many cell types, with the outcomes sometimes triggered by second messenger responses other than cAMP. Drugs that block GPCRs are currently on the market for many medical conditions such as high blood pressure, migraine headache, asthma, and psychosis. See the article by Terry Kenakin in the October 2005 issue of *Scientific American* for information on how similar drugs might eventually be marketed to combat HIV and obesity.

✎ NEWS ITEM: The drug bucindolol, prescribed for people with heart problems, is a "beta-blocker"--it binds beta adrenergic receptors like the epinephrine receptor and therefore inhibits the signal transduction pathway you just learned about. As a result, the heart is prevented from beating hard in response to stress, lessening the chance of a rupture in a weakened blood vessel. Recent research shows that not everyone responds in the same way to bucindolol! There are two types of the beta-1 adrenergic receptor, (to which the drug binds), varying at one amino acid site in the protein. A person can have two copies of the glycine variant, two copies of the arginine variant, or one copy of each, depending on what is inherited from one's parents. Bucindolol works only on people with two copies of the arginine variant. [*Science* 307: 1191]

# How Neurons Signal Muscles to Contract

---

📖 **Overview Reading**

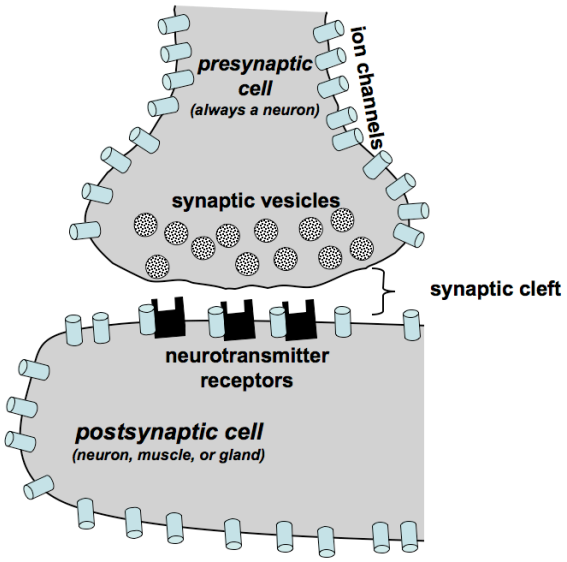
- Chapter 5 • The Dynamic Cell Membrane
- Chapter 44 • Neurons & The Nervous System

---

The previous two systems we studied involved the endocrine system releasing a hormone into the bloodstream to send a communicating molecule (first messenger). Now we are going to focus on communication mediated by the other major integrating system, the nervous system (brain, spinal cord, and nerves). The nervous and endocrine systems both communicate through **chemical messengers.** The endocrine system uses hormones, while the nervous system uses **neurotransmitters**. In both systems, the messenger molecules are secreted by one cell, travel to the target cell, and bind to specific receptors in the plasma membrane of the target cell. Furthermore, in both systems, this chemical binding triggers biochemical changes in the target cell. The endocrine system broadcasts hormones throughout the body by secreting hormones into the bloodstream. The nervous system sends neurotransmitters only a very short distance between two specific cells (the neuron secreting the neurotransmitter and the target cell that has neurotransmitter receptors to receive the message).

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 $10^{-9}$ meter, or a billionth of a meter, or a millionth of a millimeter.) This very small 0.1 nm gap where neurotransmitter diffuses from the neuron to the target cell 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 cell**. The postsynaptic cell can be another neuron, a muscle cell, or a gland cell. Below is a diagram of a synapse; you will also find a synapse diagram on page 956 (Fig. 44.13).




When a presynaptic neuron synapses on a postsynaptic muscle cell, this type of synapse is called a **neuromuscular junction (nmj)**. The chemical process of communication from the pre-synaptic neuron to the post-synaptic muscle cell is essentially the same as when two neurons communicate (or when a neuron synapses on a gland cell).

---

📖 **Focused Reading**

- p 944-946 "Neurons are…" to "Glial Cells…"
- p 946-947 "44.2 How Do…" to "Ion Pumps"

---

Just as there are many different hormones, there are also many different neurotransmitters. We will study the neurotransmitter **acetylcholine**, the chemical messenger between the nervous system and skeletal muscle cells. (Different types of neurons and synapses use different chemical neurotransmitters. A list of neurotransmitters can be found in table 44.1 on page 959 in your text.)

In looking at how the nervous system causes skeletal muscles to contract, we will start at the beginning of the process and look at the cellular and chemical events that produce a nerve impulse. (Note: skeletal muscles are the voluntary muscles in your body that are attached to your skeleton, such as your biceps, hamstrings, *etc.* The two other categories of muscle include cardiac (heart) muscle and smooth (involuntary) muscle that lines your digestive tract and arteries). Nerve impulses are electrical events, that is, they are caused by the flow of charged particles (in cells, ions are the moving charged particles that constitute the current, while the electrical current in power lines is provided by moving electrons.) Any flow of charged particles is a **current**. **Voltage (or electrical potential)** is the force that moves charged particles such as ions or electrons (causing them to flow (causing current)). Voltage is a separation of charge. Cells pump ions across their membranes to separate charges and create a **membrane potential**. According to the second law of thermodynamics, charged particles move in such a way that electrical neutrality (an equal distribution of positively and negatively charged particles) is produced. 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 is, the bigger the voltage will be (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,** which exists in all cells (including resting neurons that are not propagating an impulse (more on impulse propagation coming up).

This voltage (a.k.a. membrane potential or potential difference) allows the creation and propagation of a nervous impulse. Before we can understand how cells use voltage to transmit signals, we need 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 947-948"Ion pumps..." to end of p 948
- p 952 fig. 44.10 (The Course of an Action…)
- p 112 fig. 5.14 (Primary Active Transport…)

**thelifewire.com Reading**

- Tutorial 5.2 • Active Transport

---

The **ATP-dependent $Na^+/K^+$ pump** operates very much like the ATP-dependent $Ca^{2+}$ pump we have already considered. This pump is a bit more complicated because it transports two ions in opposite directions across the membrane. The $Na^+/K^+$ pump is, therefore, called an **antiporter** (two substances are pumped in 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 $Na^+$ because they are being transported in opposite directions across the membrane.

The unequal transport of potassium ions versus sodium ions 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). As a consequence, the outside of the cell is slightly positive and the inside is slightly negative. The magnitude of this charge difference is 60 mV (technically, -60 mV because 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" (rest means no signal is being sent, rest 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 of $Na^+$ 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 of $K^+$ 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. The unequal pumping of Na+ and K+ by the Na+/K+ pump produces this voltage.

❖ **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.

The neuron exists with its resting membrane potential around -60 mV and its concentration gradients for $Na^+$ and $K^+$ well established. Neurons that synapse on muscle cells and tell them to move are called **motor neurons** ("motor" because they cause 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 motor neurons in your spinal cord that control leg muscle contraction (you also have motor neurons in your spinal cord that control other muscles on your arms, torso, *etc.*). The neurons coming down from the brain synapse on the motor neurons in the spinal cord and secrete a neurotransmitter onto their membrane. Neurotransmitter secretion causes a change in the motor neurons that causes nerve impulses (called **action potentials**) to be transmitted across the motor neuron out to the muscles of the leg.

📖 **Focused Reading**

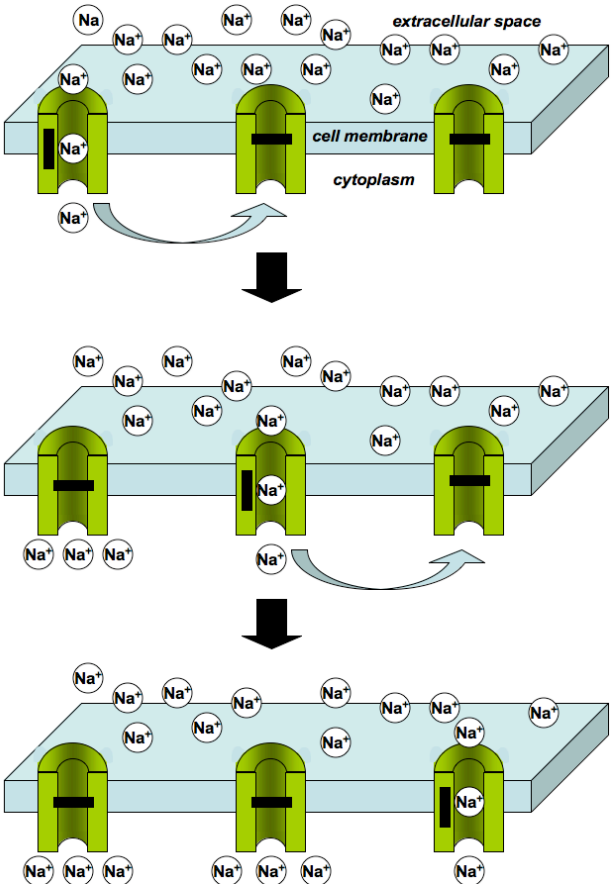
- p 950-53 "Gated ion..." to "Action potentials …"

💻 **thelifewire.com Reading**

- Tutorial 44.1 • The Resting Membrane Potential
- Tutorial 44.2 • The Action Potential

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. 44.14 on p 957 of your textbook). When the neurotransmitter binds, it causes a change in receptor shape (surprised?). This change in shape causes the $Na^+$ channel to open. This type of channel 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 down its concentration gradient. Because $Na^+$ concentration is higher outside the cell, $Na^+$ quickly moves into the cell. All that $Na^+$ moving into the cell causes the inside of the cell to become more positive. This movement of positive charge causes a change in the membrane voltage. At rest, the neuron's membrane potential is around -60 mV but as sodium enters the neuron, the membrane potential becomes -59, -58, -57, -56, *etc.*, until it reaches -50 mV. -50 mV is called the **threshold potential** because when the neuron's membrane potential is above this threshold value, another type of $Na^+$ channels open up. (Note: different neurons have very different resting and threshold potentials, but the resting potential is usually more negative than the threshold potential.) Neurons also have **voltage-gated $Na^+$ channels** near their ligand gated channels at the synapse. As their name implies, these voltage-gated $Na^+$ channels detect changes in voltage and consequently change shape when the voltage across the membrane reaches −50 mV. This change in shape causes these voltage-gated $Na^+$ channels to open, allowing <u>even more $Na^+$</u> to flood into the neuron. As a result, the neuron continues to become more positive (-30, -20, -10, *etc.*) until the inside of the cell actually becomes more positive than the outside. The neuron can become +50 mV, a pretty drastic change from its resting potential. Any change in membrane voltage away from the negative resting potential toward a more positive membrane voltage is called **membrane depolarization** (because the original poles (negative inside and positive outside) have been obliterated).




Ligand-gated $Na^+$ channels are located specifically at synapses (the only location where neurotransmitters will be released), but voltage-gated $Na^+$ channels are located all along motor neuron membranes. Motor neurons can have very long axons (the motor neurons in your spinal cord that innervate the muscles in your toes have axons that are several feet long). It is the voltage-gated channels that get the message from the spinal cord to the muscle. As each area of membrane reaches threshold (-50 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." Action potential propagation is analogous to doing "the wave" in a stadium. You cannot stand up until the person next to you stands up. After you stand up then the person on the other side of you will then be able to stand up. A voltage gated ion channel cannot open up until the channel next to it has opened up and allowed $Na^+$ in to depolarize the membrane potential. The action potential delivers the command signal from the spinal cord out to the muscle and causes the muscle to contract.

✎ NEWS ITEM: Think the analogy of action potentials and stadium waves is silly? Biophysicists in Hungary actually used computer models of action potential propagation in neurons and heart muscle cells to characterize the dynamics of the wave in German soccer stadiums and published their research in a very prestigious scientific journal. Spectators were modeled as "excitable units" that existed in one of three states: excitable (ready to participate), active (participating), or refractory (resting). As you know, ion channels can be put into these three states themselves. Like neurons, stadium waves also have thresholds – just in case you are wondering, it takes 25-40 people to initiate a successful soccer stadium wave. [*Nature* 419: 131-32.]

---

### ❖ 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 (other than "the wave").

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 channel.

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

---

✎ NEWS ITEMS: Chemicals that block sodium channels prevent action potentials and can kill. Some animals such as scorpions, frogs, octopus, pufferfish (blowfish), and newts produce sodium channel blockers to help them deter predators. The pufferfish (*Fugu*) makes tetrodotoxin (TTX) a potent Na+ channel blocker that can kill humans. Fugu is actually a prized sushi in Japan, but it must be prepared by a specially licensed sushi chef because a slip of a knife can contaminate the meat with TTX, causing a diner to have convulsions and gasp for air, in a culinary version of Russian roulette.

So how does the pufferfish survive itself if it makes such a potent toxin? In late 2005, scientists in Singapore compared the voltage-gated Na+ channels of the pufferfish and found that a single amino acid difference near the pore of the pufferfish Na+ channel as compared to the Na+ channels of other animals. This single amino acid difference changed the Na+ channel shape in pufferfish such that TTX cannot block the pore and thereby makes the pufferfish "immune" to it's own TTX. Moreover, when the scientists altered the amino acid in the pufferfish to the type of amino acid in the rat Na+ channel, that single substitution was enough to cause TTX to block the pufferfish channel. [*Curr Biol* 15:2069]

In another example, some garter snakes can safely dine on North American newts that make TTX in their skin. The snakes are slowed down a bit after eating the newts, but they are

resistant to TTX's toxic effects. Action potentials in snakes who can safely eat newts are far less affected by TTX, than action potentials in snakes who are not resistant to TTX. [*Science* 297:1289]

You might be surprised to learn that small Canadian company piloted low doses of TTX under the name Tectin as a way to ease pain in terminally ill cancer patients who have not had success with other pain medication. While they did determine that Tectin reduced chronic pain, the drug did not make it to market due to safety concerns. A related drug, Tetrodin, is being tested as a potential way to help patients undergoing heroin withdrawl. Another TTX-related drug, Tocudin, is being tested as a potential local, topical anesthesia. [www.wexpharma.com]

The cell membrane has to be returned to its resting state before it can send another signal through an action potential. (In the wave analogy, if everybody remained standing after they did the wave in a stadium, it would be impossible to propagate another wave.) When the membrane potential returns to the resting potential the process is called membrane **repolarization**. So what causes repolarization? Neurons also have voltage-gated $K^+$ channels located all along their membranes. The $K^+$ channels are also opened when the local membrane potential reaches threshold, but voltage-gated $K^+$ channels are just a bit slower to open than the voltage-gated $Na^+$ channels. When voltage-gated $K^+$ channels open, $K^+$ flows down its concentration gradient towards the outside of the cell (remember that there's more $K^+$ inside the neuron than outside). As the neuron loses positive $K^+$ ions to the extracellular fluid, the inside of the cell becomes more negative. Another factor that contributes to the repolarization of the neuron is the duration of the channel open time. Both $Na^+$ and $K^+$ channels can only stay open for only so long (in the stadium wave analogy, you eventually get tired of standing and sit down). So by shutting $Na^+$ channels and opening $K^+$ channels, the membrane repolarizes and the inside of the cell is back to its negative resting potential.

### 🕮 Focused Reading

- p 953-54 (Stop at "Action potentials can jump..")
- p 952 fig. 44.10 (The Course of an Action….)
- p 953 fig. 44.11 (Action Potentials Travel….)

*Note: If you can explain action potentials to your non-science friends, you understand them.*

### ❖ 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.
3. What would happen to a neuron that contained the defective $K^+$ channel described in the news item below?

✎ NEWS ITEM: A research team from Australia and Germany found a defect in a $K^+$ channel in the brain that causes a certain hereditary type of epilepsy, called benign familial neonatal convulsions. The convulsions start about three days after birth, but usually disappear within a few months. The channel protein is missing the last 300 amino acids. If the channel protein is incomplete, then it cannot work properly (structure – function relationship again). [*Science* 279: 403.]

Action potentials are very fast and the wave of depolarization rapidly reaches the end of the motor neuron axon. As you know, the motor neuron axon synapses on a skeletal muscle cell. Synapses between motor neurons and muscles cells have a special name, the **neuromuscular junction**. This neuromuscular junction looks very much like the brain neuron-motor neuron synapse you have already encountered in the spinal cord, except the postsynaptic cell is a muscle rather than a neuron. When the action potential reaches the motor neuron's presynaptic terminal, the membrane depolarizes (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 motor neuron's 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.)

### 🕮 Focused Reading

- p 113-114 "5.5 How Do…" to "5.6 What…"
- p 113 fig. 5.16B (Endocytosis & Exocytosis)
- p 956 fig. 44.13 (Synaptic Transmission Begins…)
- p 955 "44.3 How do…" to "The arrival…"

### 💻 Web Reading

- Movie of Calcium Influx into a Neuron
  *(warmer colors indicate higher $Ca^{++}$)*
  www.bio.davidson.edu/misc/movies/neuron.mpg

**thelifewire.com Reading**
• Tutorial 44.3 • Synaptic Transmission

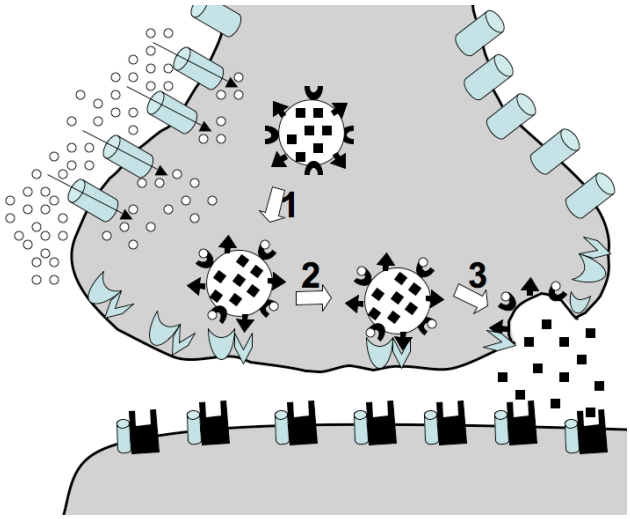
The synaptic terminal of the motor neuron contains secretory vesicles full of the neurotransmitter **acetylcholine**. When $Ca^{2+}$ enters the terminal, these vesicles fuse with the plasma membrane and release their contents into the synaptic cleft. The secretory process is an example of **exocytosis**.

Exocytosis and endocytosis are mirror image processes. 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, 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 requires no outside stimulus or trigger. Other cells, such as neurons, store their secretory product and wait for a signal to secrete; this is called **regulated secretion**. The signal to secrete is usually a rise in intracellular $Ca^{2+}$.

How does a rise in the level of cytoplasmic $Ca^{2+}$ trigger secretion in most cells and, specifically, in motor neurons? We do not know the complete answer, but a story has emerged that is very popular. According to the evidence scientists have thus far, proteins associated with the surface of synaptic vesicles and the presynaptic membrane first test the waters by tethering the vesicle to the plasma membrane. Once the tethering proteins have allowed the two membranes to remain close, a vesicle membrane protein called **VAMP** (vesicle associated membrane protein) binds to syntaxin proteins in the presynaptic plasma membrane. Both VAMP and syntaxin are **integral** membrane proteins (that means they go through a membrane, remember) and each acts as a receptor for the other. The bulk of the VAMP protein is on the cytoplasmic side of the vesicle membrane and the bulk of the syntaxin protein is on the cytoplasmic side of the plasma membrane. The cytoplasmic portions of VAMP and syntaxin act as 'snares' that allow the vesicle to come in contact with the plasma membrane. When VAMP and syntaxin bind to each another, the vesicle and plasma membranes get very close so that their lipid bilayers can fuse and the contents of the vesicle (in this case, acetylcholine) are secreted into the synaptic cleft. The fusion of synaptic vesicles with the presynaptic membrane is facilitated by additional cytoplasmic proteins (called NSF and SNAPs) that are involved in fusion between many types of membranes in the cell. Because VAMP specifically binds syntaxin (and *vice versa*), this interaction helps to assure that synaptic vesicles fuse only with presynaptic membranes and not with other membranes (such as nuclei, ER, or dendritic membranes).

While high cytoplasmic $Ca^{2+}$ levels clearly trigger vesicle binding to specific locations on the presynaptic membrane so that neurotransmitter is released into the synaptic cleft, how this rise in intracellular calcium triggers exocytosis is not completely understood. Another protein in the vesicle membrane, **synaptotagmin**, can bind $Ca^{2+}$. Investigators hypothesize that, when synaptotagmin binds $Ca^{2+}$, it changes shape. This change in shape causes a change in shape in the proteins keeping the vesicle 'docked' on the inside of the cell membrane, permitting exocytosis.

✎ NEWS ITEM: Scientists have known for more than 50 years that neurotransmitter release requires a rise in $Ca^{2+}$ at the nerve terminal. Within the past few years, scientists learned that the synaptotagmin protein bound $Ca^{2+}$ via aspartate residues in two of its cytoplasmic regions ($C_2A$ and $C_2B$) to trigger vesicle fusion. Researchers at Colorado State demonstrated that if they substituted asparagines for aspartates in the $C_2B$ region of the *Drosophila* (fruit fly) syanaptotagmin protein, that neurotransmitter release decreased by over 95%. Researchers at Harvard demonstrated that if they substituted asparagine for aspartate in the $C_2A$ region, that neurotransmitter release was normal. Now we know that even though syanaptotagmin binds $Ca^{2+}$ in two places, it is the $Ca^{2+}$ binding in only one region that actually triggers transmitter release. [*Nature* 418: 336-39; *Nature* 418: 440-43.]




$Ca^{2+}$
neurotransmitter
ion channel
SNAP-25
syntaxin
synaptotagmin
VAMP
neurotransmitter receptor
synaptic vesicle

✎ NEWS ITEM: Release of neurotransmitter has been visualized by using genetic engineering and the lightning 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. [*Proc. Natl. Acad. Sci. USA*. 94: 3402.]

✎ NEWS ITEM: Nicotine (found in cigarette smoke) binds to acetylcholine receptors and causes a rise in intracellular calcium at the nerve terminus. This rise in calcium leads to an increase secretion of other neurotransmitters (*Science* 269: 1692.). 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. [*J. Pharm. Exp. Ther.* 283: 675.]

In addition to mediating secretion (exocytosis), this process of tethering and then snaring vesicles and target membranes is the way substances are transported and sorted within the cell. For instance, all proteins are made on ribosomes. A secreted or transmembrane protein is made on a ribosome that is associated with the surface of the ER. As the protein is made, it is translocated ('moved across' the membrane) 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 receptor (as cargo) then bud off the ER and form the **cis face of the Golgi.** The protein product then transported through the Golgi until the **trans face** is reached. Then vesicles bud off the trans face and are targeted to the membrane, where they fuse. In this case, fusion acts to deliver new components to the plasma membrane, namely our friend the neurotransmitter receptor, and the other proteins in the lipid bilayer of the vesicle. Secreted proteins are released outside the cell.

Evidence is building that all of the budding, targeting, and fusing processes (called **vesicular transport**) are mediated by 'SNARE' protein complexes that function like the VAMP-syntaxin-cytoplasmic protein complex that mediates secretion. In fact, the current hypothesis describing membrane fusion is called the SNARE hypothesis. VAMP and syntaxin are examples of a family of vesicle and target membrane proteins collectively called SNARE proteins. Thus, if you asked how a vesicle that buds off the ER "knows" to fuse with the Golgi and <u>not</u> 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 <u>specific for</u> (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) 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 fusion of delivery vesicles from the endoplasmic reticulum to the Golgi. [*Science* 279: 696.]

✎ NEWS ITEM: 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. [*J Cell Biol* 143: 973-990.]

But we digress; let's get back to the neuromuscular junction. After synaptic vesicles fuse with the presynaptic membrane, acetylcholine is secreted into the synaptic cleft and 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 **ligand-gated $Na^+$ channels**. Thus, binding of acetylcholine to acetylcholine receptors on the *post*synaptic muscle membrane triggers an action potential that spreads across the muscle cell membrane in exactly the same way that the action potential spread along the motor neuron via voltage gated $K^+$ and $Na^+$ channels along the muscle membrane.

The action potential that spreads across the muscle cell membrane triggers a rise in (guess what?) intracellular $Ca^{2+}$ levels in the muscle cell. (Understanding how $Ca^{2+}$ works in cells is a very hot area in biological research.) 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 receptors will remain filled as long as acetylcholine is secreted by the presynaptic neuron.

Keeping the acetylcholine concentrations up requires effort because an enzyme (acetylcholinesterase) in the synaptic cleft 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 (or any other stress). At this point things get a little vague (and complicated), but know that neuroscientists are hard at work doing experiments to understand how higher neurological functions such as emotion, perception, and memory are controlled by chemical processes in the brain.

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 do not 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 $Ca^{2+}$ channels in the SR membrane. $Ca^{2+}$ then is free to flow down its concentration gradient into the cytoplasm. We do not know the exact mechanism whereby the membrane action potential communicates with $Ca^{2+}$ channels in the SR. However, recent research suggests that the protein **triadin,** which spans the gap between the plasma membrane and the $Ca^{2+}$ channels in the SR, is in the right place to do this job. Unfortunately, no one has any idea how this job could be accomplished. Whoever figures it out first will become very famous because the link between membrane depolarization and $Ca^{2+}$ release from the SR is a long-standing mystery in muscle cell physiology.

---

### ❖ 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 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 moving your arm. Tell this story using chemical and cellular language as you would for a traditional exam question (or explaining it to one of your Bio 111 classmates). Then tell it in simpler terms as you would to a younger sister or brother. Use as many good analogies as you can.

---

# How An Egg Learns It Has Been Fertilized

We have studied three specific and related cases of signal transduction thus far:

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
3) neurotransmitters bound to skeletal muscle receptors to tell the muscles to contract

Now we will examine another example of signal transduction, fertilization.

You have probably seen many film clips of sperm fertilizing an egg. The image is striking, one egg is surrounded by hundreds (or thousands) of sperm trying to penetrate the egg's plasma membrane. Then, 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 sperm might enter the egg at about the same time, a condition called **polyspermy**. If polyspermy occurred, the resulting zygote would be in trouble because it would have three (or more) haploid genomes (three copies of each chromosome), instead of the normal two copies (we will cover this issue of "ploidy" later in Unit II). Any diploid embryo that ends up with extra sets of chromosomes cannot develop properly and will very likely die shortly after fertilization.

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).

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 and nerves, egg cells have a resting potential of about -50 mV with a higher concentration of $Na^+$ ions outside the cell than inside. Fusion of egg and sperm membranes causes $Na^+$ channels in the egg's plasma membrane 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 -50 mV to about +30 mV. For unknown reasons, sperm cannot fuse with eggs that have positive membrane potentials. As you know from your studies of neurons and muscle cells, changes in membrane potential can occur very quickly, thus the depolarization of the egg induced by sperm fusion is called the fast block to polyspermy.

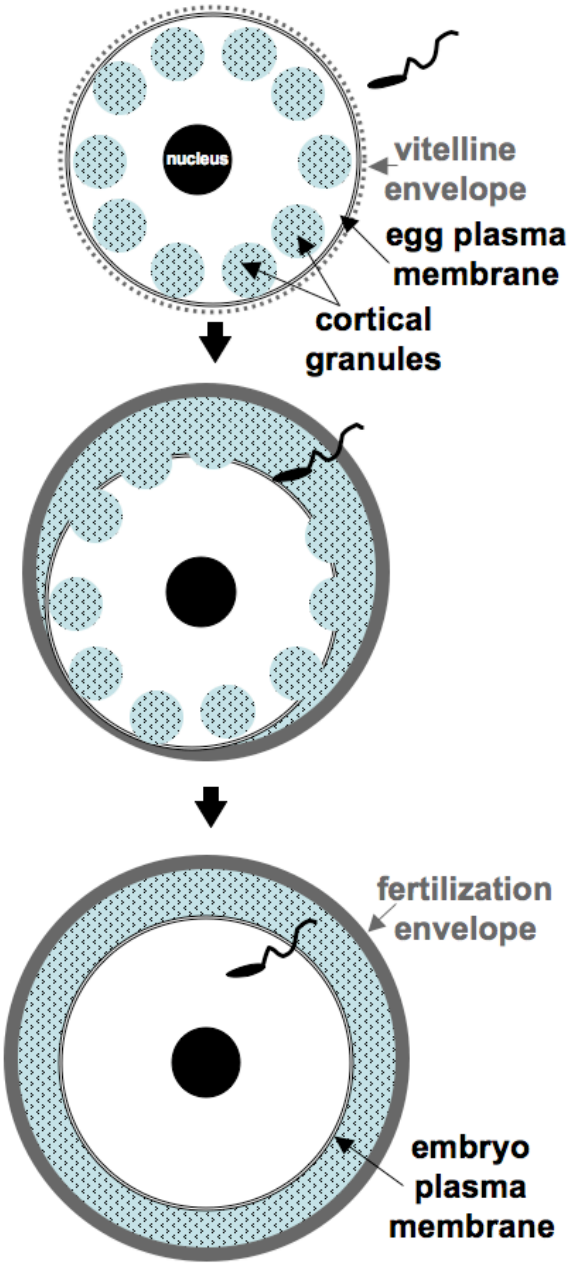
So why does an egg need a second, slower block to polyspermy? Shouldn't the fast block to polyspermy do the job? Think about what you know about ion channels. Once they are open, do they stay open? Recall that ion channels in muscle and nerve eventually close. Ion channels in the egg membrane are similar – they also close, thus the fast block is not a permanent block. The second, slower block to polyspermy creates a permanent physical barrier to sperm entry. How does the egg create a permanent physical barrier after it has been fertilized? If you examine an unfertilized egg in cross section, you see lots of small vesicles, **cortical granules**, just below the plasma membrane. Inside these cortical granules are **proteases** (protein-cleaving enzymes) and **mucopolysaccharides** (sugars). At fertilization, these cortical granules are exocytosed, they fuse with the egg's plasma membrane and release their contents into the extracellular space around the egg.

The exocytosis of cortical granule contents causes two significant events to happen in close succession:

1) the protease enzymes 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

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), that 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 that cortical granule exocytosis creates 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.




One thing is certain -- when nature develops a good system (for moving, for secreting, for transporting, for communicating, *etc.*) it keeps using it over and over again. Most animal eggs use the same communication system to signal the arrival of a sperm cell. This signal transduction system is called the **inositol triphosphate ($IP_3$) second messenger system**. The communication system used by an egg to sense fertilization is so ancient that it arose in a common ancestor that gave rise to

sea urchins (round, fist-sized marine invertebrates that resemble pin cushions), frogs, fish, *etc.* Odds are that evolution would not have produced the very same $IP_3$ communication system in so many very different animal species through random mutation and selection. Therefore, this method of transducing the fertilization signal probably evolved well before the evolutionary split between vertebrates and invertebrates. Thus many types of animal eggs kept this "good idea" as they diverged into different species.

---

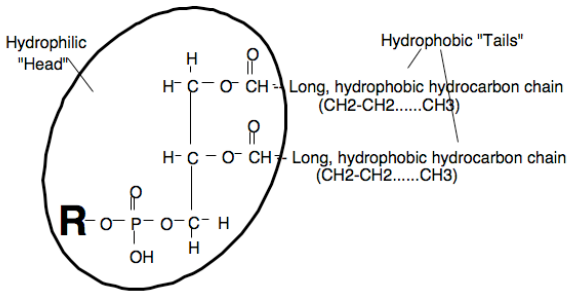
**📖 Focused Reading**

- p 341-43 "Second messengers.." to "Calcium ions…"
- p 343 fig. 15.13 ($IP_3$/DAG 2nd Messenger System)

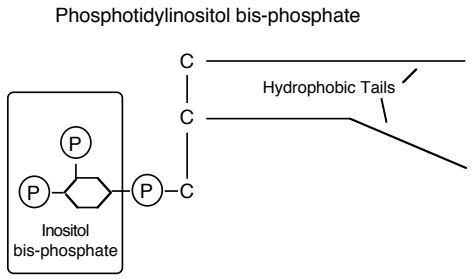
---

The inositol triphosphate ($IP_3$) second messenger system uses receptor kinases, molecules embedded within the plasma membrane, and a specific enzyme (sound familiar?). The enzyme **phospholipase C (PLC)** plays an important role in transducing the message that a sperm has begun to fertilize the egg. Like most enzymes, PLC's name tells you something about what it does. PLC cleaves (cuts up) a phospholipid. (PLC's name also implies that there are several types of phospholipases such as phopholipase A, phospholipase B, *etc.*, but we will only discuss PLC.)

Before we discuss the molecules that convey the fertilization message within an egg, we need to talk a little more about phospholipids. You should already know that phospholipids make up the cell membrane. Recall that phospholipids have hydrophilic "heads" that dissolve in the aqueous (watery) cytoplasm and in the watery extracellular fluid (or seawater) outside the cell (see fig. 5.2 on page 99 in your text if you need a quick reminder). Phospholipids also have hydrophobic tails that avoid water and dissolve in each other in the hydrophobic center of the lipid bilayer that makes cell membranes. There are many different kinds of hydrophilic molecules that can be added to the phosphate on the phospholipid. These various molecules are added at the "R" in the diagram at right. Regardless of what is added at the "R" site, phospholipid molecules are all highly polar, and many of them are charged, greatly enhancing the hydrophilic nature of the "head" of the molecule. Some examples of molecules that are added to the phosphate group at "R" include 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, the phospholipid would be called phosphotidylcholine (see fig. 3.20 on page 56 for the structure). And if inositol were added, the resulting phospholipid would be called **phosphotidylinositol**.

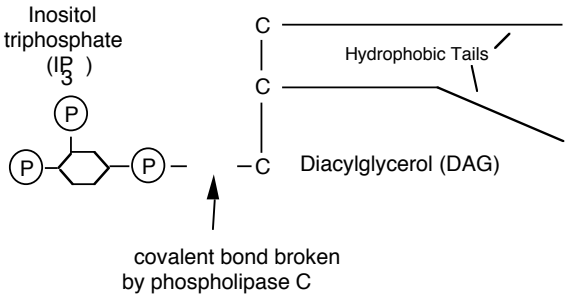



Some phospholipids have **inositol bis-phosphate** added to their phosphate group. "Bis" means "two." So, inositol bis-phosphate is simply inositol with two phosphate groups on it. When inositol bis-phosphate is added to a membrane phospholipid, the resulting molecule is called **phosphotidylinositol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




$PIP_2$ sits in the inner layer 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 cytoplasm. When PLC is activated it cuts inositol off of $PIP_2$ in such a way that all the

phosphates go with inositol and none remain on the lipid in the membrane. The products of this cleavage look like this diagram below:




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 second messenger. The remaining part of the molecule is called **diacylglycerol** (like triacylglycerol with two instead of three fatty acids chains), abbreviated **DAG**. DAG remains embedded in the membrane, but nonetheless also acts as a second messenger. So the cleavage of $PIP_2$ by phospholipase C results in two cleavage products: $IP_3$ and DAG, that can 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 (DAG), phosphotidylinositol bis-phosphate, and inositol triphosphate. (If you need more information about lipids and phospholipids, see pages 54-57 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?

Evolution is a process of natural selection; natural selection allows organisms with favorable traits or abilities to reproduce. When organisms with advantages are more successful at reproducing (and organisms with less advantages have more difficulty reproducing), advantageous traits or abilities are then maintained in the population and less advantageous traits may disappear (or become diluted in the population). Natural selection plays an important part in all levels of molecular and cellular biology; 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 survive to produce new individuals 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) produced an egg that permits only one sperm to fertilize it? Evolution is not a wasteful process; it recognizes 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 whole, new second messenger system to facilitate signal transduction, evolution thought to herself, "How can I tell the egg that a sperm has just arrived without inventing a totally new molecular 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 within newly fertilized eggs.

### 🕮 Focused Reading

- p 899-902 "42.2 How do animals…" to end of p 902
- p 902 fig. 42.5 (Fertilization of the Sea Urchin Egg)

### 💻 Web Reading

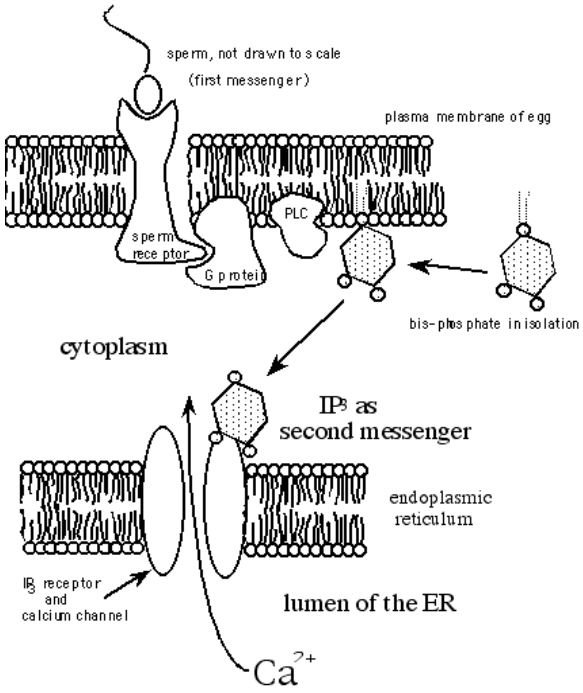
- Sea urchin sperm competing to fertilize
  www.bio.davidson.edu/misc/movies/SPERMRAC.MOV
- Mechanism for $IP_3$ production and $Ca^{2+}$ ion wave
  www.bio.davidson.edu/courses/Immunology/Flash/IP3.html
- Movie of Calcium During Fertilization
  *($[Ca^{++}]_i$ indicated by white in right panel)*
  www.bio.davidson.edu/misc/movies/PHASECAL.MOV
- Movie of Sea Urchin Fertilization
  www.bio.davidson.edu/courses/Bio111/images/urchinfert.MOV
- Movie of $IP_3$ injection
  www.bio.davidson.edu/courses/Bio111/images/IP3.mov

### 💻 thelifewire.com Reading (6th edition)

- Fertilization in the Sea Urchin Animation
  (43.1 in the 6th edition of Purves *et al.*)
  www.whfreeman.com/thelifewire6e/con_index.htm?43

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. Eggs are, however, often much larger than most other cells (think about the size of a chicken egg for example). The egg also has many unique features including the **vitelline envelope** (called the zona pellucida in some species) outside of the plasma membrane, an extracellular matrix (analogous to a plant cell wall). The vitelline envelope contains many copies of a sperm-binding receptor protein (called **ZP3** because it was the third protein identified in the zona pellucida). ZP3 interacts with **bindin** proteins on the surface of sperm cells and initiates the **acrosome 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, mouse sperm will bind to ZP3 on mouse eggs, but not to ZP3 on hamster eggs. Interaction between ZP3 in the vitelline envelope 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 time to fuse with the egg.

Now it is time to put $IP_3$ into the picture. The sperm plasma membrane proteins interact with the sperm receptor in a manner similar to the diagram (not drawn to scale).




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

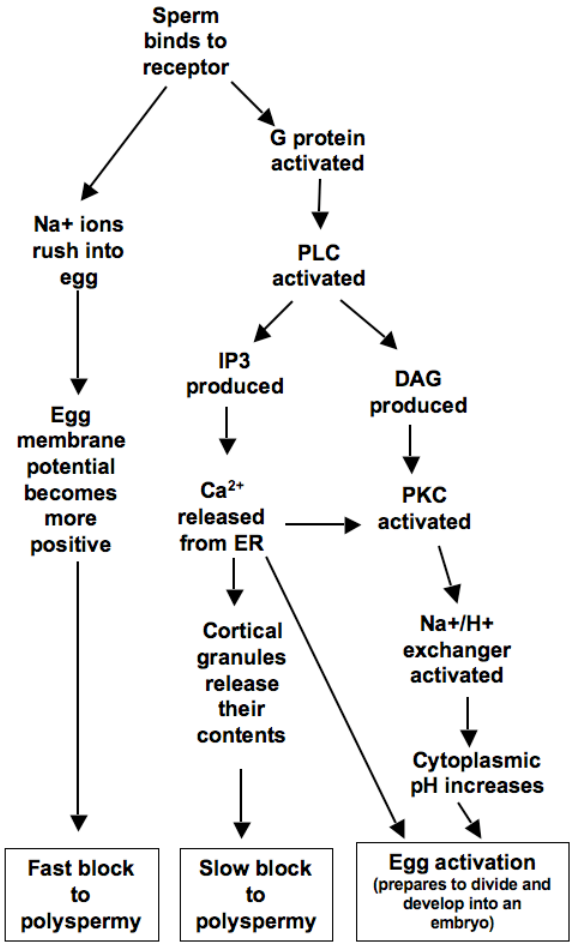
When the sperm receptor binds its ligand (ZP3), the receptor changes shape, activates the associated G-protein, which stimulates phospholipase C, which cleaves phosphotidylinositol bis-phosphate into two parts: $IP_3$ and DAG. $IP_3$ is a second messenger that diffuses throughout the cytoplasm where it eventually bumps into the **$IP_3$ receptor** located in the ER membrane. The $IP_3$ receptor is a **homotetramer** (composed of four 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 you 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. Thus, the cytoplasmic $Ca^{2+}$ concentration rises, which is the signal 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 a wave, from the point of sperm penetration throughout 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, a phenomenon called **calcium induced calcium release (CICR)** 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 wave of adjacent activation of CICR is analogous to the way an action potential is propagated in a neuron, using $Ca^{2+}$ instead of $Na^{+}$ dependant depolarization. Exactly how CICR works is an area of intense research. It is clear that in the slow block to polyspermy, $Ca^{2+}$ 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.

As with all second messengers, we need a way to turn off the signal. When calcium concentration reaches 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 binding initiates a chain reaction of enzymes (each can amplify the original single event) that 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: A team of Japanese scientists recently identified a protein on the surface of mouse sperm. They named the protein izumo, which refers to a Japanese shrine to marriage. Mice missing both copies of the izumo gene (knock-out mice) produced normal sperm that could bind to and penetrate the zona pellucida, but could not fuse with eggs. Scientists will be investigating ways to block izumo as possible new, hormone-free contraceptive methods. [*Nature* 434:234]

## ❖ 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. Explain how this second messenger is used in fertilization.
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?

✎ NEWS ITEM: People with high blood pressure-related cardiac hypertrophy have problems with cardiac muscle excitation-contraction cycles, and scientists have found that the "heart" of the matter is in the efficiency of CICR. [*Science* 276: 800.]

✎ NEWS ITEM: A group of collaborating scientists in Massachusetts, California, France, and Mexico have demonstrated 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 $Ca^{2+}$ channels. If ZP3 signaling continues, the pathway activates a sustained $Ca^{2+}$ influx mechanism and this sustained increase in $Ca^{2+}$ drives the acrosome reaction. [*Molec Biol Cell* 11: 1571-84.]

✎ NEWS ITEM: As you know, mammalian sperm must exhibit remarkable swimming ability to reach and penetrate an egg's zona pellucida. A $Ca^{2+}$ channel, CatSper, found only in the sperm tail membrane appears to play a crucial role in a sperm's ability to swim. Male knockout mice lacking the *CatSper* gene are healthy, but infertile, presumably because the sperm cannot penetrate the zona pellucida without the CatSper channel. Interestingly, sperm without this channel do fertilize eggs that have been stripped of their zona pellucidas. Consequently, researchers speculate that the CatSper channel might be involved in giving the sperm a "turbocharge" as it penetrates the zona pellucida. Contraceptives that block the CatSper channel would avoid the disadvantages of hormonal contraceptives. For example, CatSper-targeted contraceptives could be taken by a man OR a woman for a potentially short period of time. Some forms of male sterility might even be caused by CatSper defects. [*Nature* 413: 603-09.]

Many other cells use the inositol triphosphate ($IP_3$) second messenger system for a wide variety of functions. Below are a few examples:

- The secretion of 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 invasion

Note: You can learn more about some basic techniques described in this section in the *Study Guide* appendix on Experimental Techniques

Note: There are some good questions to help you study for the first review immediately following the next section on other communication systems.

# Other Cellular Communication Systems

It could be true that the majority of cells communicate through the four 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 biologists 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 that we now know something about:

### The cGMP 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) that 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 that 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, PDE5 converts cGMP into GMP. With

Viagra around, cGMP levels remain high and promote erection. Some not so well known Viagra facts are: 1) Viagra was originally developed to combat angina (pain that results from insufficient oxygen delivery to heart muscles), 2) Viagra is NOT an aphrodisiac, and 3) Viagra can cause distorted color vision. Why the vision changes? It is not because the users see the world through rose-colored glasses, but instead because the retina also uses cGMP as a second 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. [*Science* 273: 323-324]

### 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 into 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 cells. 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 the intracellular part of the molecule phosphorylates, it is called a kinase. And because it adds the phosphate to a tyrosine residue of itself (tyrosine is an amino acid), the receptor is called a **tyrosine kinase** (an enzyme than phosphorylates tyrosine.) When insulin binds to its receptor, binding causes a number of changes in the cell, including stimulating the transport of glucose into the cell, stimulating glycogenesis, and the synthesis of triacylglycerol. Interestingly, several genes associated with the development of cancer (called **oncogenes**) 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**, which makes them unusual. Prostaglandins, leukotrienes, and thromboxanes 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 (and out of) the cell easily. Steroid hormone receptors thus do not need to be located on the exterior of the cell and can be found in the cytoplasm and/or 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. We will discuss gene expression in Unit II.

✎ NEWS ITEM: A new family of about 100 genes has been discovered that function as human pheromone receptors. These receptors reside in a 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 pheromones. Pheromones are usually fatty acids or steroids and their receptors appear to span the membrane seven times and are linked to G-proteins. (*Science* 278: 79)

### 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 that 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 that 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. (*Science* Vol. 276: 912.)

### 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 play 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 (NO) 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* 276: 2034.]

---

### ❖ Study Questions:

*(These questions are good preparation for the first exam)*

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 your professor, to a Bio111 classmate, to a Davidson student who has taken no biology since high school, to your parents, and 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. Cancer researchers have studied second messenger systems extensively 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 does each of them do? In what ways are these processes similar? In what ways are they different?

5. Myasthenia gravis is a disease that 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 produces paralysis. Curare has been used by hunters on the tips of arrows to paralyze their prey.) 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 (commonly called "botox") 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. Why is this toxin used (carefully) by plastic surgeons to paralyze facial muscles?

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

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?

# *Unit II: Genetics*

**📖 Overview Reading**

- Chapter 3 • Macromolecules & the Origin of Life
- Chapter 4 • Cells: The Working Units of Life
- Chapter 9 • Chromosomes, Cell Cycle, & Cell Division
- Chapter 10 • Genetics: Mendel and Beyond
- Chapter 11 • DNA and Its Role in Heredity
- Chapter 12 •From DNA to Protein: Genotype to Phenotype
- Chapter 14 • The Eukaryotic Genome and Its Expression

*Note: you have reviewed much of this reading already*

The earth is teeming with living things. We can see some of the larger organisms—trees, grass, flowers, weeds, cats, fish, squirrels, dogs, insects, spiders, snails, mushrooms, and 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 differences between the code for a dog and the code for a geranium are the number and order of those letters in the code. The DNA from a human, rearranged slightly, would produce a mouse. If you took the human genome, doubled it, and rearranged the letters in the right way, you could produce a common toad. One seventh of the human genome, rearranged slightly differently would produce a poplar tree. 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 out the workings of this code. Many scientists over the past century have contributed to our understanding of DNA. James Watson and Francis Crick, using data from Maurice Wilkins and Rosalind Franklin, determined the 3D structure of DNA in 1953 and set the stage for later researchers to show that 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 all organisms 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: **Sickle Cell Disease (SC), Cystic Fibrosis (CF),** and **Huntington's Disease (HD)**. 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 genetic diseases include cystic fibrosis (CF), Huntington's disease (HD), phenylketonuria (PKU), Down's syndrome, Tay Sach's disease, sickle cell disease (SC), muscular dystrophy (MD), and hemophilia A. In other cases, such as heart disease, or cancer, one can inherit a genetic **predisposition** to a disease, but environmental factors also play a major role in determining which individuals develop the disease and which escape it. Most disease conditions are probably in this genetic predisposition category, which certainly includes diabetes, hypertension (high blood pressure), and many forms of cancer.

**📖 Focused Reading**

- p 376-77 "Altered membrane.." to "Altered structural…"

**💻 Optional Web reading**

- Cystic Fibrosis Web Site
  www.cff.org/AboutCF/
- Sickle Cell Disease Web Site
  www.sicklecelldisease.org/about_scd/index.phtml
- Huntington's Disease Web Site
  http://www.mayoclinic.com/health/huntingtons-disease/DS00401

The three diseases we will investigate in this Unit, sickle cell disease (SC), cystic fibrosis (CF), and Huntington's disease (HD), are caused exclusively by genetic defects. SC is the most common genetic disease among people of African heritage, 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 (RBCs) are biconcave in shape (shaped like tiny doughnuts with a membrane across the hole) in unaffected individuals, but in SC, some RBCs take on the shape of a crescent moon, or sickle, which causes several problems. 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.

CF is the most common genetic disease in Americans of European descent, occurring in 1 out of every 2500 births. One in 25 Caucasians is a healthy carrier of the CF defect. CF occurs with a frequency of 1 in 17,000 people of African descent and with less frequency in other races. In the US, approximately 1000 new cases are diagnosed each year. Cystic fibrosis patients 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 because 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 often the actual cause of death, not the thick mucus. Untreated children usually die by the age of four or five and the average life expectancy with medical care is 35-40 years.

✎ NEWS ITEM: For people with cystic fibrosis, one of the major causes of death is a chronic infection with *Pseudomonas aeruginosa* (PA), a bacterium that often is difficult to treat with conventional drugs. Researchers at Harvard Medical School made a discovery that may lead to the development of new, more effective drugs for PA. A protein associated with virulence in *Vibrio cholera*, Hcp1, is produced by PA and secreted from the bacterium via a novel protein secretion mechanism. New drugs targeting this protein or the protein secretion pathway may be useful in treating PA infections. [*Science* 312:1526]

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 generally do not begin to appear until after the age of 40, usually after an individual has already had children. Thus, until recently, people with HD in their families have had to reproduce without knowing whether they have the disease and run a 50% risk of transmitting it to their offspring.

The search for the causes and cures of these and other genetic disorders is a long-standing goal of biomedical research. 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 passed 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 182-84 "Eukaryotic cells…" to "9.1 Recap"
- p 187-92 "9.3 What happens…" to end of p 192
- p 187 fig. 9.7 (Chromosomes)
- p 190-191 fig. 9.10 (Mitosis)
- p 233-234 "11.1 What…" to "The transforming…"

### 💻 thelifewire.com Reading

- Tutorial 9.1 • Mitosis

### 💻 Web Reading

- Cartoon of Mitosis
  www.cellsalive.com/mitosis.htm
- Movie of Mitosis
  www.contexo.info/DNA_Basics/mitosis%20movie.htm

### 💻 Optional Web Listening (~4 min.)

- Reflections on Living with Sickle Cell Disease
  www.npr.org/templates/story/story.php?storyId=1238650

---

For any cell to survive and function properly, it must maintain the right number of chromosomes at all times. (Remember that the blocks to polyspermy prevent too many sperm from

producing a cell with too many chromosomes.) The necessity to maintain a constant number of chromosomes presents a problem for the average cell that is ready to divide. Let's say the cell has 23 pairs (it is **diploid**) of chromosomes and it wants to make two new cells. 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 into two cells. In its simplest form, **mitosis** is duplication of DNA that is then divided equally into two nuclei. The individual steps of mitosis are outlined in fig. 9.8. You should be familiar with the major steps of mitosis (technically mitosis does not include the cell cycle phase called interphase); 1) prophase; 2) metaphase; 3) anaphase; and 4) telophase (all four phases are reviewed in your text page 189-192). Two points to note: 1) the text includes a fifth phase called 'prometaphase' and 2) mitosis technically does not include cytokinesis (though mitosis and cytokinesis are two cellular events that are closely associated).

Now that you have a handle on mitosis, we need to see what gametes (sperm and eggs) do when they are formed. You know that to form a new individual by sexual reproduction, two gametes fuse to form a zygote. Because each gamete brings a set of chromosomes to **syngamy** (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 the process of **meiosis**.

### 🕮 Focused Reading

- p 900 fig. 42.3 (Gametogenesis)
- p 193-199 "Reproduction by.." to "Meiotic errors.."
- p 196-197 fig. 9.16 (Meiosis)
- p 198 fig. 9.18 (Crossing Over Forms…)

### 💻 thelifewire.com Reading

- Tutorial 9.2 • Meiosis

### 💻 Web Reading

- Movie of Meiosis
  www.bio.davidson.edu/misc/movies/MEIOSIS.MOV
- Side-by-Side Animations of Mitosis & Meiosis
  www.pbs.org/wgbh/nova/baby/divi_flash.html

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. Meiosis 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** (two 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** (fig. 9.18). Crossing over between chromosomes 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. What are the major steps in mitosis and meiosis?

2. What is the significance of meiosis in relation to creating variation in the next generation?

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 and the power to genetics is numbers. The more progeny available for study, the easier it is to discern the pattern of inheritance. Genetic experiments with humans are neither ethical nor practical, because 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 genetic diseases. Certainly one of the first things people noticed about CF, SC, and HD was that these diseases run in families. Now just because families usually live together and share a common

environment, you cannot always conclude that a disease or condition 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 (such as diet, stress levels, *etc.*) 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 constructing a pedigree, certain rules are observed:

• Squares represent males.

• Circles represent females.

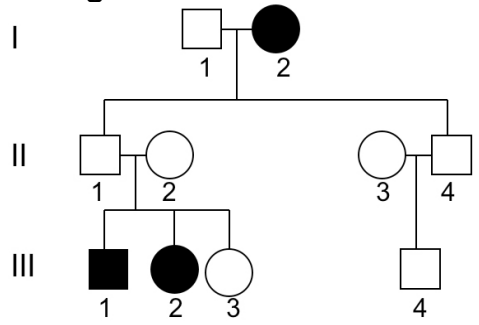
• Non-affected individuals are blank (or solid white) Affected individuals (people with the disease or condition) are colored or patterned in some obvious way (often solid black)

• Lines between a circle and square indicate a mating union and all offspring of a mating union (*e.g.,* siblings) are drawn on the same horizontal level.

• Generations are numbered with Roman numerals (I, II, III, *etc.*) and individuals within a generation are numbered with Arabic numerals (1, 2, 3, 4, *etc.*).

Below is an example—a pedigree for a family with cystic fibrosis. In this family, the woman (I.2) in the first generation (grandma) had CF and yet survived long enough to have two children. Neither of her sons (II.1 and II.4) had CF. Individuals II.1 and II.2 had three children, two of whom (III.1 & III.2) have CF. II.3 and II.4 produced son #III.4, a normal, unaffected, or **wild type** (wt) child.

**CF Pedigree:**




❖ **Study Questions:**

1. Given information about a family, be able to draw a family pedigree that complies with standard rules.

2. Be able to interpret a pedigree drawn by standard rules.

3. Draw a pedigree for the cross that is outlined in fig. 10.3 (page 211).

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 how genes are inherited, we have to go back 140 years to the Austro-Hungarian Empire and a Catholic monastery. Here a monk named **Gregor Mendel** conducted 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. Mendel defined laws that govern the simple inheritance of traits. Traits that are inherited in this straightforward manner are said to be **Mendelian traits** that obey the laws of **Mendelian genetics**.

📖 **Focused Reading**

• p 207-219 "10.1 What are…" to "10.2 Recap"

❖ **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. Look at the CF pedigree shown a few paragraphs ago. 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 generations 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 Mendelian 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 II.1 and II.2 above represents the F1 of a **monohybrid cross**. 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. In peas, yellow seed color is dominant to green. State the colors of the offspring of the following crosses:
   homozygous yellow x homozygous green
   heterozygous yellow x homozygous green
   heterozygous yellow x homozygous yellow
   heterozygous yellow x heterozygous yellow

7. If two animals heterozygous at a single locus mate and produce 200 offspring, about how many of the offspring would be expected to have the phenotype of the dominant allele?

8. 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?

9. 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?

10. 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.

(Questions #6-10 from *Biology* by Vilee *et al.)*

11. For more practice, try the questions at the end of chapter 10.

---

✎ NEWS ITEM: Collaboration between researchers at the Oregon State and the University of Bristol (in the UK) has cloned the gene associated with the dwarf pea plant 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 result 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 because their stems are shorter. Genetic engineers who want to produce wind-resistant food crops are obviously interested in understanding this gene and the enzyme it produces. (*Proc. Nat. Acad. Sci.* 94: 8907.)

---

### 🕮 Focused Reading

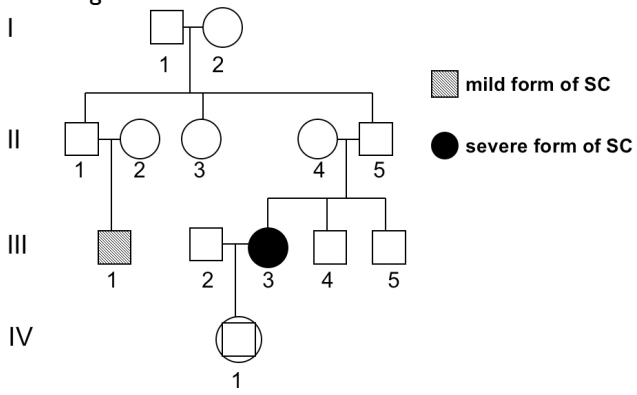
- p 220-222 "The environment" to "10.3 Recap"

---

When considering CF, an individual either expresses the phenotype (has the disease because s/he has two copies of the CF allele and is considered a **homozygote**) or does not express the phenotype (does not have the disease) as a **heterozygote** with one copy of the mutant allele OR as a **homozygous** wild type with no copies of the mutant allele. This phenotypic expression pattern is 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 (below). Here we see individuals that have 'mild' cases of the disease. How can this be? Doesn't one allele 'win' over the other? Well, not always. 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 observable in heterozygotes, so the mild form of the disease is said to be not fully penetrant. Environmental factors can affect 'observing' 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 hiking at a high altitude. (See fig. 10.12 on page 218 of your text for an example

of incomplete dominance in flower color.)

SC Pedigree:




### ❖ Study Questions:

*(These questions refer to sickle cell pedigree above)*

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

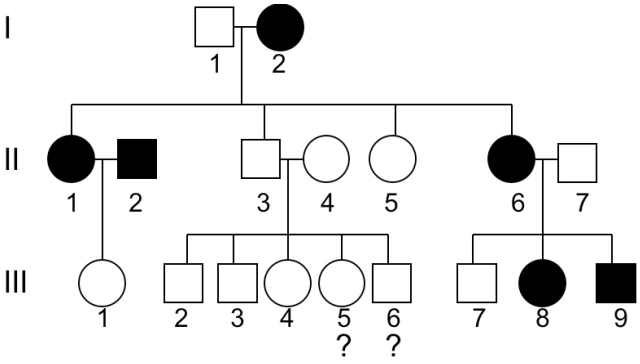
2. Label the generations in this SC pedigree using Mendelian terminology.

3. What are the genotypes and phenotypes for all individuals 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 mating union of II.4 and II.5 above represent (*e.g.,* Monohybrid cross, test cross)? Draw a Punnett square for this cross.

5. Individual IV.1 is still in the womb and consequently the baby's gender is unknown. For each of the following outcomes of this pregnancy indicate the genotypes of the parents III.2 and III.3 and the odds of these three outcomes:
   - IV.1 is homozygous wild-type
   - IV.1 is heterozygous
   - IV.1 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 people in the HD pedigree above? Individuals III.5 and III.6 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.

4. Individual III.1 did not inherit HD. What were her chances of inheriting HD?

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 384-386 "17.3 How does…" to "17.3 Recap"
- p 195 fig. 9.15 (The Human Karyotype)

❖ **Study Questions:**

1. Individuals II.1 and II.2 from the CF pedigree (found earlier in the genetics section of this *Study Guid*) are considering having another baby and come to you as a genetic counselor. They want to know the chances that this next baby would have CF. What will you tell them about this baby's chances of having CF? Here is the pedigree again:

2. Before III-4 was born, Individuals II.3 and II.4 from the CF pedigree wanted to know the chances their baby would have CF. Individual II-3 didn't know her family history, having been adopted. What would you have told them? (Assume that a person picked from the population at random has a 1 in 25 chance of being a carrier of a mutant CF allele, given the population frequency of carriers.)

3. Individuals II.1 and II.2 from the SC pedigree (found earlier in this section) want to know the chances their next baby will have the severe form of SC. (Assume that a person picked from the population at random has a 1 in 100 chance of being a carrier of a mutant allele.) Similarly, individuals III.2 and III.3 from the SC pedigree also want to know the chances their next baby will have the severe form of SC.

4. What would you tell individuals II.3 and II.4 from the HD pedigree about the chances of their child developing HD? Individuals II.6 and II.7?

5. A couple planning to have children 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 carry 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 all three will have severe SC?
   E. What is the probability that the first two children will have severe SC and the third is a carrier?
   F. What is the probability that all three of the three children will be homozygous wild type?
   G. What is the probability that all three 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) 1/25 x 1/4 = 1/100*
*3) 1/100 x 1/4 = 1/400*
*1/100 x 1/2 = 1/200*
*4) 0*
*1/2*
*5) A 1/4*
*B 3/4*
*C 1/4*
*D 1/4 x 1/4 x 1/4 = 1/64*
*E 1/4 x 1/4 x 1/2 = 1/32*
*F 1/4 x 1/4 x 1/4 = 1/64*
*G 1/2 x 1/2 x 1/2 = 1/8*

## Dominance and Recessiveness at the Molecular Level

So far, through pedigree analysis of the afflicted families, we know that CF is a recessive trait and that Huntington's disease is a dominant trait. 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, sugars, and fats) 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 that 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 CF, 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 (because mucus becomes thicker when water is removed), *etc.* In the first part of this discussion, we 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, this terminology 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. Codominance 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 wild type 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 wild type 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, makes a form of the protein that is hyperactive, or makes a form of the protein that is inappropriately subversive or destructive. 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 becomes apparent 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 can be classified as 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 other traits.)

---

### How Genes Encode Proteins

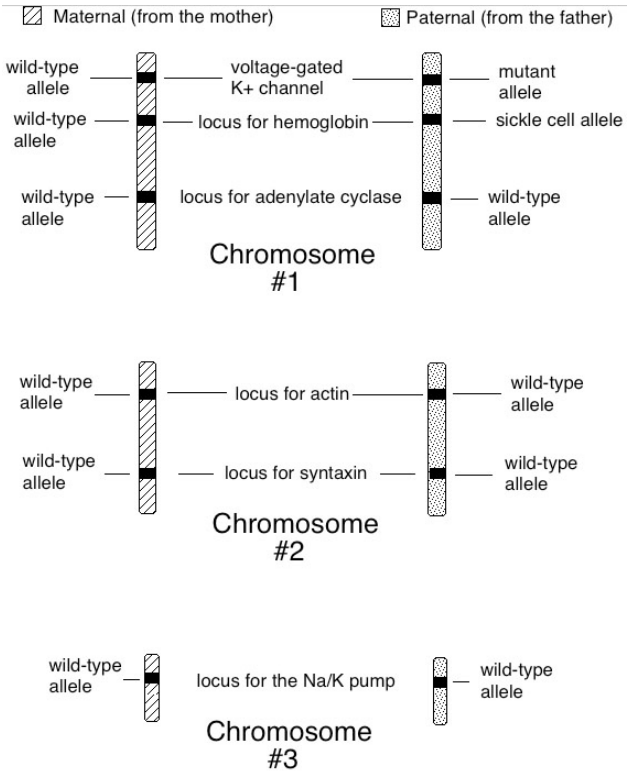
We need to stop and look at how wild-type genes produce wild-type proteins. Genes do not exist as individual strands of DNA, but rather, they

sit one after another in very long complexes called **chromosomes**. Chromosomes contain DNA as well as proteins that are involved in packaging chromosomes (to fit in 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. Most 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.

### 💻 Web Reading

- Cancer Biology Videos Genes I: Chromosome to DNA
  http://www.cancerquest.org/videoseries1_english.html

The order of loci is always the same on homologous chromosomes, with rare exceptions. 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 vary somewhat. 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.




### ❖ Study Questions:

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

2. What is a genetic locus? An allele?

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 comosome 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 at least <u>one</u> 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 HD is so rare, it is very unlikely that

two people with HD would mate and produce a homozygous HD offspring.)

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

---

**📖 Focused Reading**

- p 238-250 "11.2 What is the..." to "11.5 What are…"
- p 240 fig. 11.9 (Base Pairing in DNA Is…)
- p 242 fig. 11.10 (Three Models for DNA…)
- p 243 fig. 11.11 (The Meselson-Stahl Experiment)
- p 260-261 "RNA differs…" to "RNA viruses…"
- p 263 "The information…" to "12.3 Recap"

**💻 thelifewire.com Reading**

- Tutorial 11.1 • DNA Replication Part 1
- Tutorial 11.3 • DNA Replication Part 2

---

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 cytosine (C). 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, the base sequence of DNA determines the amino acid sequence of the protein. Because amino acid sequence determines native conformation and native conformation determines function, the nucleotide sequence controls most cellular 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 fig. 12.6 (page 264). 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 275-278 "Point mutations..." to "12.6 Recap"

---

You can see by studying the genetic code on page 264 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 and these are called silent mutations. 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 264 and see which mutations would make such a difference. For instance, the code for serine (Ser) is UCG (there are actually six 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 43 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 three-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 page 264. 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 43 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 because 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 because 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 some 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). Then of course, our cells can make mistakes in DNA replication that can lead to mutations too. 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 almost always destroy the gene's ability to produce a product. Explain why this is so.

4. Explain how a missense mutation may <u>increase</u> 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 how a five base pair insertion mutation could cause 300 amino acids to be deleted in the resulting protein.

---

📖 **Focused Reading**

- p 250-253 "What are some…" to "11.5 Recap"

---

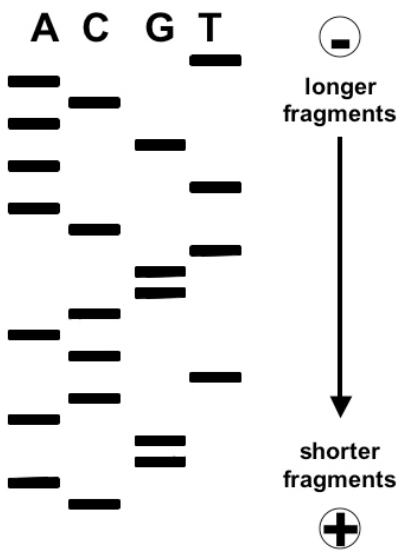
❖ **Study Questions:**

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 such as the one below to give the correct base sequence of a DNA segment (with the correct 5' to 3' orientation.)




---

When we talk about mutations, it is a common misconception that we are always talking about changes in the DNA that occur <u>in the individual bearing the trait</u>. This is <u>not the case</u>, 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 are thought to be enhanced by new mutations within individuals. <u>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.</u> 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.

---

📖 **Focused Reading**

- p 252 fig. 11.24 (Sequencing DNA)

💻 **Web Reading**

- SRY paper
  www.bio.davidson.edu/courses/Molbio/srypaper.html

---

At this web site, you will find a virtual reprint of an article that illustrates how important each and

every nucleotide is. A couple has had problems conceiving a child and both of them go to a fertility specialist for some advice. This woman has a point mutation with dramatic system-wide phenotypic consequences. She has a mutation in the SRY gene, a gene located on the Y chromosome. A functional copy of SRY is required for embryos to develop as males rather than females.

Note: If you want to learn more about SRY, Scott Gilbert's *Developmental Biology* text gives an overview of how the SRY gene was identified at www.devbio.com/chap17/link1702.shtml

---

❖ **Study Questions**:

1. What were the clinical symptoms of the woman described in this SRY 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 if she wanted to compete in the Olympics and was subjected to a karyotype analysis?

6. What is SRY, what is its function?

---

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 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. (The precise advantages of thinner hair remain a topic of debate.) The skin became more exposed to the harmful ultraviolet radiation in sunlight. These high-energy rays can mutate thymidine bases, increasing skin cancer risk. Early, thin-haired humans had to rely on the expression of genes that control the enzymes that make **melanin**, the dark pigment of skin. Individuals who produced more melanin didn't get skin cancer as often because their dark skin pigment blocked the penetration of UV light. Consequently, they were healthier and more able to reproduce and raise offspring to maturity. Dark skinned individuals became the wild type phenotype in the population. Pale individuals represented spontaneous mutations in genes that caused lower melanin production. Because the pale skinned individual was more susceptible to UV light damage and early death, dark skin came to be the dominant trait as the species evolved into *Homo sapiens*.

Mutations occur all the time (on average, one mutation per $10^{10\text{-}12}$ bases of DNA per cell division). While some early humans had mutations that increased melanin production, 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 persisted in the human genome because they did not enhance survival and reproduction.

Later, in humans that migrated north, melanin-reducing mutations were advantageous, allowing UV-induced vitamin D production in the skin in the weak northern sunlight. The need for vitamin D outweighed the skin cancer risk in these populations in this environment.

Your body contains some new mutations that 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 such mutations 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 humans, that take a lot of energy to produce, are virtually always diploid because diploidy provides enormous adaptive advantages.

The presence of a potentially lethal or harmful new mutation makes you a carrier of a defective gene. If you have a child 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. Consequently, genetic diseases are more frequent when close relatives mate. 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 two first cousins mated, they would have a dramatically increased probability of producing an offspring with two mutated alleles, a homozygous individual with serious or lethal genetic problems. Most cultures have laws or traditions discouraging such incestuous relationships.

If mutations must 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 because 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.

4. Most genetic diseases are recessive. Explain why. If maladaptive mutations are selected against, how do dominant and recessive inherited diseases remain in the population despite their detrimental effects 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

- p 263 fig. 12.5 (DNA is Transcribed to Form RNA)
- p 261-262 "12.3 How is the info…" to end of p 262

- p 313-315 "14.2 What are the…" to "Gene families…"
- p 316-319 "14.3 How are…" to "Contrasting eukary…"
- p 324-325 "14.5 How is…" to "Small RNAs…"

**thelifewire.com Reading**
- Tutorial 12.1 • Transcription

---

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 1)transcription (*e.g.,* RNA polymerase or transcription factors), 2) RNA processing (splicing, adding a cap or poly-A tail), or 3) transporting the mRNA from the nucleus to the cytoplasm. However, if possibility 1) is the case, then other proteins controlled by the same transcription factors would also be affected. If 2) or 3) were the case, the cell could make virtually no proteins because all proteins use the same polymerases, 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**. The promoter 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 be 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 not 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 hypothesis is viable because 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. The mutation 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
   B. The promoter
   C. The spliceosome
   D. snRNPs
   E. mRNA transport proteins (in nuclear pores)
   F. Introns and Exons
   G. Enhancer
   H. Transcription factors

---

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 265-271 "How is RNA…" to end of p 271
- p 326-327 "How is gene…" to "14.6 Recap"
- p 79-80 "The endomem..." to "Golgi..."

**thelifewire.com Reading**
- Tutorial 12.3 • Protein Synthesis

Optional Web Animation
A Protein Primer/Molecular Happening (circa 1970s)
- biology.kenyon.edu/slonc/Micro/protein_synth102105.mp4

---

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 die.

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 80-81 "Golgi…" to "Lysosomes"
- p 270-274 "Polysome…" to "12.5 Recap"

---

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' that instruct 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?

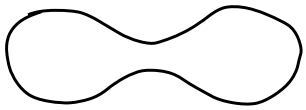
2. Explain how a mutation in a gene's signal sequence could produce a genetic disease.

---

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 impressive 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 ~25,000 genes in the right places at the right times.

**Sickle Cell**

Certainly, if you want to know how to treat 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 defective protein in SC is normally expressed in RBCs and is likely a protein that controls the RBC shape. RBCs are normally shaped like this:



side view

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 do not do something special, a cell will always assume a spherical shape. So, in order to maintain the RBC in this unusual biconcave shape, the cell has to distort and support the membrane with proteins. One such cytoskeletal 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 changes shape when it binds to oxygen is called **hemoglobin** (see page 47, fig. 3.9). Hemoglobin is a molecule much like chlorophyll (we will talk about chlorophyll more in Unit III) with a porphyrin ring structure containing an atom of iron (in hemoglobin) instead of magnesium (in chlorophyll—see page 165, fig. 8.7 to see what a metal-bearing porphyrin ring looks like.) Hemoglobin's iron atom 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 RBCs very easily. Osmotic pressure can burst open RBCs when they 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 of bursting RBCs 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 the Appendix of this *Study Guide*). 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 to **electrophoresis** (also described in the Techniques Appendix). If SC and normal hemoglobin 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 no change in mobility. This result indicated that wild type hemoglobin and SC hemoglobin were similarly sized proteins.

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 at the time. While you certainly do not have to understand the details of protein sequencing, the following brief discussion should give you an idea of some of the technical difficulties involved in this process.

### 🕮 Focused Reading

- p 1084 Table 50.3; *look at enzymes digesting proteins/peptides*

Traditional methods of amino acid sequencing rely on the use of analytic chemical techniques to identify amino acids released after digestion of the protein with enzymes that cleave peptide bonds at specific sites in the protein. These enzymes are called proteases or peptidases. For instance, if you subject a protein to **carboxypeptidase**, the enzyme will cleave off the very last amino acid in the chain, the amino acid at the carboxyl terminus of the protein. (Note that 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 44, fig. 3.5 for an illustration.) Another example of a protease is the enzyme trypsin, which cleaves a protein chain on the carboxyl side of a lysine or arginine. Similarly, the enzyme chymotrypsin will cleave on the carboxyl side of phenylalanine, tryptophan or tyrosine.

Various chemical processes can be used to tag or label the C- or N-terminus of a protein with a trackable chemical addition so that when that amino acid is released by a protease, it can be purified and identified by analytical procedures. For instance, in the common Edman degradation procedure, the compound phenylisothiocyanate reacts with the amino terminus of the protein. The tagged amino acid is released and identified with chromatography. The process is then repeated with the remaining portion of the protein to identify the second, third, etc., amino acids in the chain.

Nowadays, researchers more commonly use **mass spectrometry** to determine the amino acid sequence of a purified protein. You might use this same technique in chemistry class to determine various characteristics of smaller compounds. The technique can be extended to analyze proteins. In general, the researcher first fragments a purified sample of a protein into pieces with the protease trypsin; then the structure and composition of each piece is determined by analyzing the mass-to-charge ratio after ionization. For more information, see the following web site:
https://www.msu.edu/~gallego7/MassSpect/MSandPMM.htm .

### 🕮 Focused Reading

- p 44-48 "The primary…" to "Chaperonins…"
- p 274-275 "Mutations:…" to end of p 275
- p 1058 fig. 49.12 *(note RBCs must fit through capillaries)*

### 💻 Web Reading

- Sickle Cell Disease
  http://www.ygyh.org/sickle/whatisit.htm

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 (six 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 264, you see that the difference between the code for glutamic acid (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 six instead of glutamate. You can see on page 43 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 location in the protein. This subtle change in protein sequence and shape causes a critical change in protein function such that the mutant version of hemoglobin does not carry oxygen as efficiently and does not allow the RBCs to squeeze through capillaries as easily.

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. A protozoan that spends part of its life cycle in the RBC causes this disease. SC heterozygotes are resistant to this phase of the disease and are therefore somewhat more protected from malaria than are normal individuals. 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 information helps explain why SC is so prevalent in African Americans and also provides an example of why mutations are not inherently advantageous or harmful — it depends on the environment in which that organism exists.

✎ 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 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, it is a research direction that deserves 'air time.' [*Science* 284: 463-66].

---

### ❖ 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 sixth amino acid if it is encoded by the seventh codon?

✎ NEWS ITEM: In 1998 the FDA approved a sickle cell anemia drug that had been used to treat cancer for over 30 years. The drug, hydroxyurea, has the ability to activate the transcription of a gene that is normally silent in adults. This gene encodes a form of hemoglobin produced only in fetuses. Fetal hemoglobin works just as well as adult hemoglobin, and since it is encoded by a separate gene unaffected by mutation, it is still functional in sickle cell patients.

💻 Optional Web Listening

- Sickle Cell Drug (*hydroxyurea*) Reduces Death Risk (<4 min.)
  www.npr.org/templates/story/story.php?storyId=1216809

6. Is the treatment described in the news item above considered a cure? Will those being treated still be at risk of having children with SC?

---

✎ NEWS ITEM: In December 1999, the Associated Press reported the success of a new cure for sickle cell anemia. Stem cells from the umbilical cord of an unrelated infant were infused into a thirteen-year-old suffering from SC. (Stem cells are undifferentiated cells found in bone marrow that develop to produce red blood cells.) The transplant, performed Dec 11, 1998, was the first time unrelated cord blood as been used to treat sickle cell anemia. This treatment is much less painful than typical bone marrow transplants sometimes used for SC treatment. The patient acquired a self-renewing source of healthy

red blood cells. After one year the cord cells have taken hold in the boy's bone marrow and are making healthy blood cells, so doctors declared the child 'cured.' Do you consider this a cure? If he should have children, would they be at risk of having SC?

✎ NEWS ITEM: Researchers from seven universities or hospitals and one company report that they have cured SC in a mouse using gene therapy. The therapy inserts a globin gene variant they call *gemisch* into hematopoetic stem cells (cells that develop into blood cells) using an HIV-based vector. After ten months, 99% of the red blood cells show no sign of sickling. Moving this therapy from mice to humans will not be easy; the patients' own hematopoetic stem cells must be removed (killed), and researchers need to demonstrate that the HIV-based vector is safe. (*Science* 294: 2368.)

---

**Optional Web Listening**

• My So-Called Lungs
Laura Rothenberg, a college student with CF, chronicled her disease (including a lung transplant) in a powerful and award-winning 22-minute radio autobiography
www.npr.org/programs/atc/features/2002/aug/socalledlungs/index.html

• Cystic Fibrosis Update (~5 min long)
A follow up interview on the current status of CF treatment
discover.npr.org/features/feature.jhtml?wfId=1147900

• Remembering Laura Rothenberg (~4 min long)
www.npr.org/templates/story/story.php?storyId=1199420

• Genetic Testing, Part 2: Reading Genes for Disease (~9 min)
One Couple's Decision Against Testing for Cystic Fibrosis
www.npr.org/templates/story/story.php?storyId=1763554

---

## Cystic Fibrosis

The sickle cell disease puzzle was solved relatively early because the cellular defect was visible through the microscope and the protein affected by SC was an obvious suspect. Unfortunately, the overwhelming majority of inherited genetic diseases are much more difficult to investigate. In the case of cystic fibrosis (CF), investigators knew that the disease compromised the way the lungs and pancreas handle mucus. 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 didn't immediately point to a suspect protein. We have been referring to the CF gene as a "mucus gene," but that name doesn't explain all of CF's symptoms.

In 1984, a real breakthrough in CF research came from a lab investigating the differences between respiratory cells of CF patients versus wild-type individuals. These researchers tested the ability of these 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 (a.k.a. PKA) that, 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 research group (Sato and Sato) found that respiratory cells from CF patients were unable to pump $Cl^-$ in response to cAMP activation. They asked if this failure to pump $Cl^-$ might be because cAMP cannot activate cAMP-dependent protein kinase, and they found that the protein kinase does become activated, but it does not activate any $Cl^-$ pumping action. While scientists are usually wary of jumping to conclusions that are insufficiently supported by the data, this result was a very exciting finding because 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 not follow and the mucus normally found on these internal surfaces will remain thick and dry. Thick, dry lung mucus will harbor bacteria, causing pneumonia (a bacterial infection that can be fatal). Thick, dry pancreatic mucus will also block the passage of digestive enzymes from the pancreas into the intestine.

At this point it looked as though the CF allele 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. Sequencing a gene allows investigators to use the genetic code to determine the amino acid sequence of the protein

the gene encodes. 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 could 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, and hopefully develop an effective treatment. (You saw in lab how easy it is to isolate genomic DNA.)

The human genome (the sum of all of the DNA in all 23 pairs of human chromosomes) contains about $6 \times 10^9$ base pairs and about ~25,000 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 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. The ethics of this project are widely discussed. The knowledge of the entire base sequence of the human genome gives 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 terminate pregnancies when the fetus has **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, heart disease, homosexuality, baldness, obesity, *etc.*? Further, *in vitro* fertilization now allows the predetermination of genetic traits. Egg and sperm can be fused 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. The numerous ethical issues here are clear, but the answers are being discussed.

Obviously, this discussion 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 and we will lose the perspectives of differently abled people. 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 cellular defect.

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

3. Is the CF gene mutation present only in lung and pancreatic cells?

4. Does it matter what cell is used to isolate DNA for analysis?

---

✎ NEWS ITEM: The salt concentration in the surface airway liquid of CF patients is unusually high, and high salt concentrations hinder immune cells from combating bacterial infections there. A team from the University of Iowa found that the sugar xylitol could lower the airway salt concentration when delivered by an inhaler. Airway bacteria do not use this sugar for energy, so xylitol treatment could enhance the ability of the immune system to kill bacteria without promoting bacterial growth. Xylitol inhalers are being developed as a means to prevent bacterial

infections in CF patients and others prone to lung infections. (*Proc Natl Acad Sci USA*. 97: 11614-9.)

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 Review Reading**

- p 207-218 "10.1 What are…" to "Many genes..."
- p 222-224 "10.4 What is the…" to end of p 224
- p 224 fig. 10.21 (Steps toward a Genetic Map)
- p 225 fig. 10.22 (Map These Genes)

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 (trait that shows variation in the population), 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 are variant alleles of the same gene, we will use the capital letter I as the gene symbol, with the A, B, and o as superscripts for the different alleles, thus $I^A$, $I^B$, and $I^o$. The possible phenotypes and their corresponding genotypes are listed below:

| Phenotypes | Genotypes |
|---|---|
| A | $I^A I^A$ or $I^A I^o$ |
| B | $I^B I^B$ or $I^B I^o$ |
| AB | $I^A I^B$ |
| O | $I^o I^o$ |

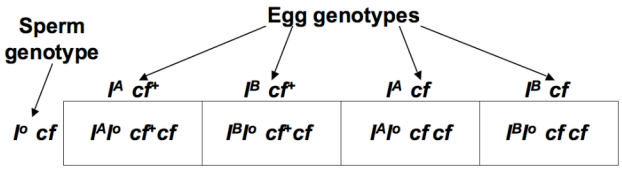
Now, if CF and blood groups are on different chromosomes, these traits will be independently assorted when they are passed down to the next generation. Here is an example. Let's say Maria has blood type AB and is a carrier for CF. Maria's genotype is $I^A I^B cf^+ cf$. Louis has blood type O has CF. Louis' genotype is $I^o I^o cf\ cf$

According to one of Mendel's postulates, the alleles at one locus segregate independently of the alleles at another locus when gametes are formed. (We know now that this applies only SOME of the time--when the loci are far apart from each other in the genome.) 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: $I^A cf^+$
Egg type 2: $I^B cf^+$
Egg type 3: $I^A cf$
Egg type 4 $I^B cf$

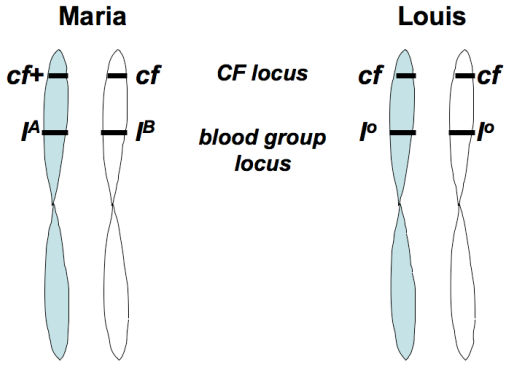
Louis's alleles also segregate independently during meiosis, but because he is homozygous at both loci, all of his sperm would get one $o^+$ and one

*cf*. If Maria and Louis should produce offspring (and this is fairly unlikely in this case because 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:

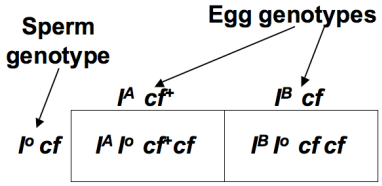



This is a classic Mendelian test cross in which a dihybrid is mated with a homozygous recessive individual. If blood group and CF are on different chromosomes, there are four possibilities for the children: carriers of CF with blood type A or B and afflicted individuals with blood type A or B. All possibilities 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.)

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




Because $I^A$ is linked to $cf^+$, the two alleles go together (assort together) into the gametes. Likewise, because $I^B$ 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 the diagram below. In this case, there are only two alternatives for the offspring. They are either 1) blood type A and a carrier or 2) blood type B and afflicted with the disease.




As a geneticist trying to determine if CF is linked to the blood group locus, you gather data from many families with CF-afflicted members, and you determine the blood type of each individual (afflicted or not). By analyzing this information you can see if the traits follow either of the two patterns presented above (either with four equally-represented combinations, or with two combinations such that, for example, a child that has blood type B always has the disease). If blood type B is often inherited with cf, then the two loci are 'linked' on the same chromosome. If you know which chromosome carried the blood type 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 independent assortment. What exactly does it mean if two loci show independent assortment?

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. 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?

5. 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?

6. What kinds of diploid matings result in the following phenotypic ratios? 3:1, 1:1, 9:3:3:1, 1:1:1:1

7. 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.

8. 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 deduce whether or not the traits are linked.

---

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

---

### 🕮 Overview reading

• Chapter 16 • Recombinant DNA & Biotech.

### 🕮 Focused Reading

• p 353-354 "Restriction…" to end of p 354
• p 355 fig. 16.2 (Separating Fragments of DNA)
• p 381-382 "Genetic markers…" to "Disease-…"
• p 381 fig. 17.8 (RFLP Mapping)

### 💻 Web Reading

• Southern Blot Animation (must use browser with Shockwave plugin)
http://www.dnalc.org/vshockwave/southblot.dcr

• Real Southern Blot
www.bio.davidson.edu/courses/Bio111/RealSouthernblot.html

---

Restriction fragment length polymorphisms (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 loci 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 cannot do genetic analysis. The problem is, as mentioned earlier, many human traits are not polymorphic. For many genes, 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.

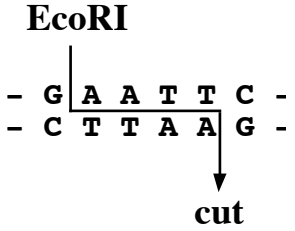
In order to establish the presence of a RFLP on a chromosome, or segment of chromosome, you have to have a way of labeling certain DNA or RNA sequences so that they can be seen with the naked

eye. You do this with a **probe** that is tagged with radioactivity (because it contains radioactive phosphorus in its phosphate groups) and complementary in nucleotide sequence to some chosen sequence of bases. (Currently, most researchers use probes tagged with non-radioactive molecules because such tags are cheaper, more sensitive, and safer, but the principles are the same – you have a DNA probe with a tag on it so you can detect the probe.) Some 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. Because the mouse and the human versions of glycogen synthase have a highly conserved structure and function, we assume that the nucleotide sequences for the two genes would also be conserved.

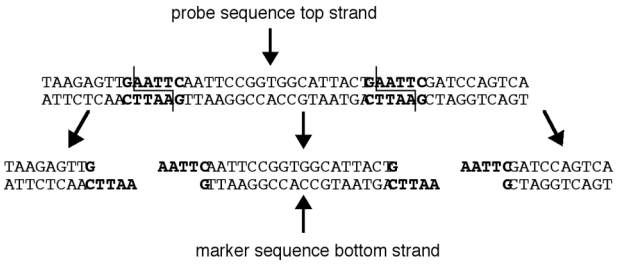
Another type of probe is called an **oligonucleotide** (oligo- means a polymer of unspecified length; nucleotides are what get polymerized). Oligonucleotides are short stretches of single-stranded DNA that are synthesized by a machine called a nucleic acid synthesizer. On this instrument is a four-letter 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. (FYI: the primers you use in the PCR lab were produced in this way.) The machine is loaded with dNTPs (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' that becomes our DNA **marker**. It is 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 because we are looking for a polymorphism or genetic variability between individuals. 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 because sampling merely requires drawing blood. You then incubate the DNA from the chromosomes with a **restriction enzyme** that will cut the DNA. 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). EcoRI 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 cut illustrated below:




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

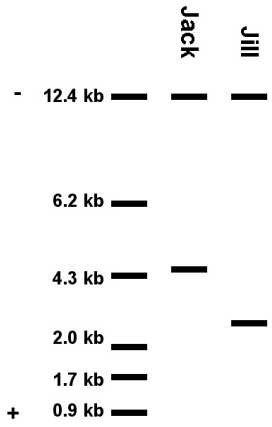



Thus, a restriction enzyme cuts the large, chromosomal DNA into small fragments (called **restriction fragments** because they are

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 if we stained the gel with a general DNA stain such as ethidium bromide we would see a continuous smear of DNA all the way up and down the gel. This smeary gel 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 or nylon) 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 create a **Southern** blot by transferring DNA from an electrophoretic gel onto nitrocellulose. The process is very simple. Basically, if you can make Jello™ and handle paper towels, you can perform a Southern blot.

The DNA is **denatured** after being transferred to the filter paper. Denaturing DNA is different from denaturing protein. When you denature DNA, you unzip the double helix and convert the molecule into two single strands. (Remember from our PCR experiment that high heat will denature double stranded DNA.) You then apply your radioactive probe and allow the probe to bind, or **hybridize**, to its complementary sequence. You wash the blot to remove unbound probe. Finally, to see where the radioactive probe has bound, you place an X-ray-sensitive film (also sensitive to the radiation from your probe) 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 type of film exposure is called **autoradiography**. This whole process is diagrammed and explained in more detail in your web reading.

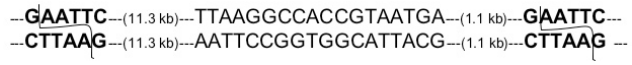



Above is a Southern blot from Jack and Jill after it was hybridized with the radioactive probe and the resulting blot exposed to X-ray film. DNA was loaded at the top (-) and ran toward the bottom (+). When restriction fragments are electrophoresed, molecular weight, or fragment length, markers (the "ladder") were electrophoresed at the same time. These markers are DNA fragments of known length and are often referred to as a DNA ladder or DNA standards. 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. (Note: you saw a DNA ladder in the PCR lab.)

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 Jack1 (~12.4 kb) and Jack2 (~4.3 kb). In Jill's case, EcoRI created many, many fragments, two of which contained the same marker sequence. We'll call these bands Jill1 (~12.4 kb) and Jill2 (~2.5 kb).

If we focus only on the restriction fragments that bear the marker (the only ones we can see in a Southern blot), Jack1 and Jill1 are the same length (about 12.4 kb). For one of their two chromosomes, the DNA carried by Jack and Jill is 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

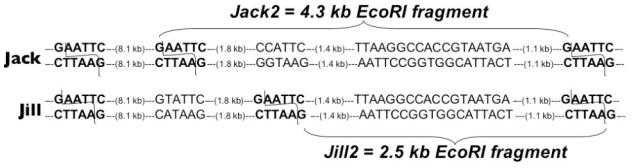
in the parentheses being hypothetical distances between the given sequences for Jack1 or Jill1:




These fragments are flanked by two restriction target sites for EcoRI and contain the marker sequence. While we cannot say that Jack1 and Jill1 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 two bands indicating that the probe hybridized with two different size restriction fragments of DNA. For this to have happened Jack's two copies of this chromosome must not be identical, the copy of the chromosome containing the region we call Jack2 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 2.5 kb piece of DNA.

Comparing Jack2 and Jill2 we see that these two bands are not the same size. Jack2 Is 4.3 kb, while Jill2 is 2.5 kb. Remember, both of these fragments must 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.




In this case, Jack has a 4.3 kb fragment bearing the marker sequence and flanked by two EcoRI sites (in bold). 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 12.4 kb 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 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 inherit 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 the following 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'AATTCCGGTGGATTACT3' probe. This type of RFLP analysis can be used to produce a "**DNA fingerprint**" that can be used as a very accurate form of identification in forensics. (You performed a different kind of "DNA fingerprinting" in the PCR lab.) 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 analyze several different RFLPs (*i.e.,* different combinations of restriction enzymes and probes).

---

### ❖ 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 DC28035, 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?

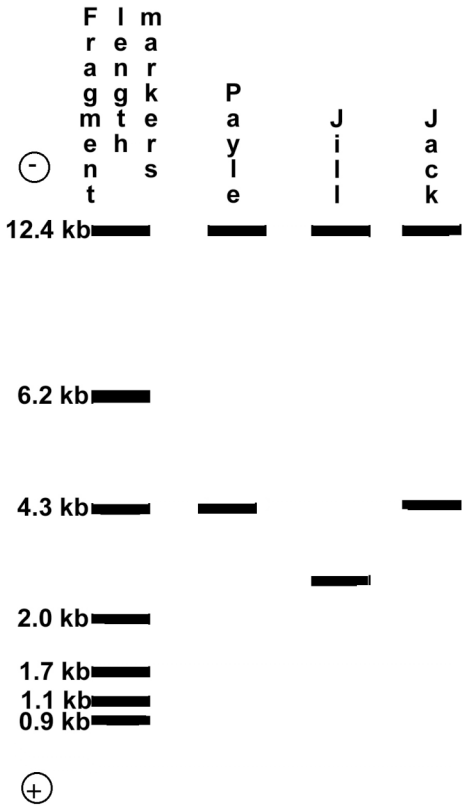
10. How are RFLPs related to the process of DNA fingerprinting?

---

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, what would you tell a jury?

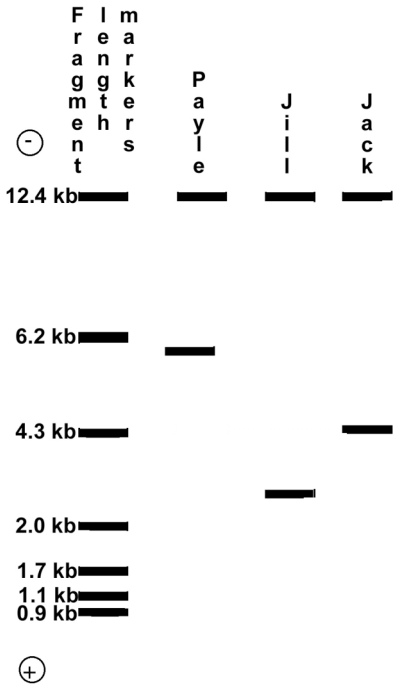
In addition to identifying individuals, RFLPs are passed on to children just as alleles are passed on. To illustrate the power of this multigenerational analysis of RFLPs, let's say that Jack and Jill have a child together. They name the child Payle. Let's say we did the same genetic analysis of Payle's DNA that we did to Jack and Jill, and nearby is a diagram of the resulting Southern blot. In analyzing these gels, remember that the marker sequence (3'TTAAGGCCACCGTAATGA5') cannot simply disappear (except through a new mutation and we will assume here that new mutations have not happened). Jack has two copies of the marker sequence (Jack1 and Jack2) and Jill has two (Jill1 and Jill2). Payle inherited two copies of the marker sequence too. It appears that he inherited Jack2 and Jill1 and he did not inherit Jill2 or Jack1.

How can we say that Payle inherited Jill1 but not Jack1? Don't parents have to pass their genes on to their offspring? And how is it that Payle didn't inherit Jill2? 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 <u>half</u> of their chromosomes to their offspring. Because Payle had to fetch something from his mother and he did not inherit Jill2, he had to inherit Jill1. Likewise, because Payle inherited Jack2, he could not also inherit Jack1 because Payle can only get one copy from each parent.




Above is one example of the many things you can determine by analyzing family RFLPs. In addition, you can tell whether an offspring is actually the child of a particular couple. Let's say that the RFLP analysis appeared as in the blot below. It is of little concern that Payle did not inherit Jack2. Payle could have inherited Jack1. 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 (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 oversimplified 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. [*Science* 285: 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 page 355-356 "DNA Fingerprinting…" where VNTRs are discussed. What is a VNTR5? How is it similar to a RFLP? How is it different? (NOTE: you determined your own VNTR pattern for the D1S80 locus in lab.)

---

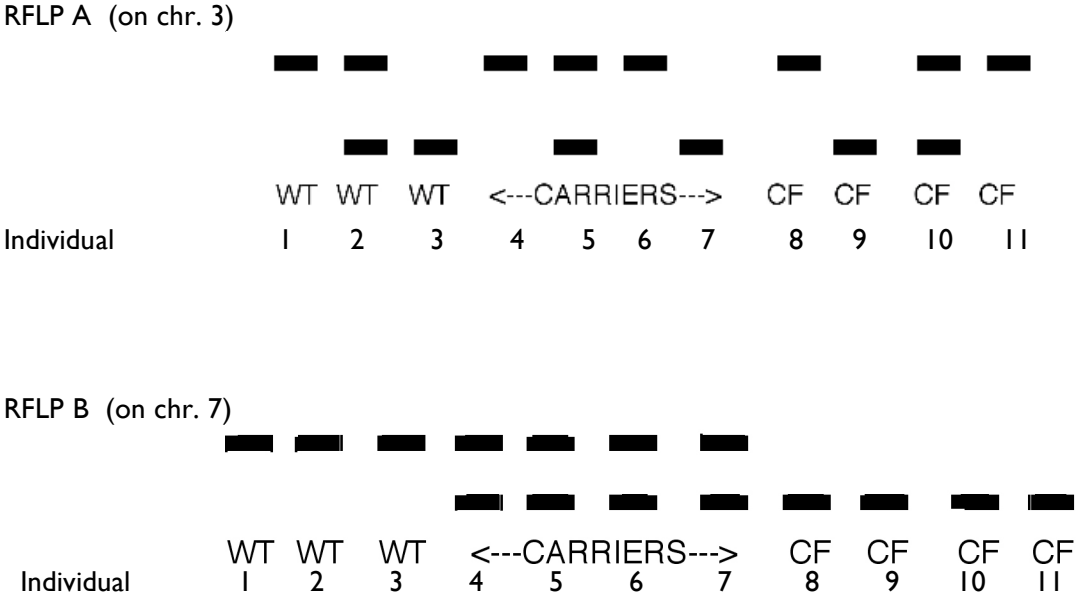
Lap-Chee Tsui, John R. Riordan, and Francis Collins determined that the CF gene was on Chromosome #7 by finding that it was linked to a RFLP that was located on that chromosome. To do this, they gathered DNA from hundreds of families—families without any CF history as well as families with the disease. 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 markers on Chromosome #7. Their results showed no linkage to markers on other chromosomes. Below is a simplified version of their Southern blot data.

It should be noted here that many different restriction enzymes and probes, each in a separate parallel experiment, are used to do RFLP analysis. Within a given experiment, you must use the same restriction enzyme and the same probe to look for that particular RFLP in everyone.

Individuals 1-3 on the CF Southern blots below are from families without CF; individuals 4-7 are carriers (parents of a CF patient without the disease themselves), and individuals 8-11 are CF patients. For RFLP A (known to be on chromosome 3), the different RFLP alleles do not correlate with disease genotype, indicating non-linkage. For RFLP B

(known to be on chromosome 7) 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, which is the only one present in wild-type individuals (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. These results indicate that the disease gene is very close to RFLP B on chromosome 7.

**Southern Blot RFLP Analysis for CF**




The chromosome 7 restriction fragment with the marker sequence may actually contain the CF gene, but it may also just be very close by. Such a tighly linked RFLP can allow for a Southern blot test to **diagnose** the disease state. For example, if the Southern blot for RFLP B were performed on a person from the same family with unknown disease status, and if 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 ability to determine genotype has been a 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 different 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 to find

a n unknown gene, you have to know something about a process that occurs naturally called **recombination** (or crossing over).

### 🕮 Focused Reading

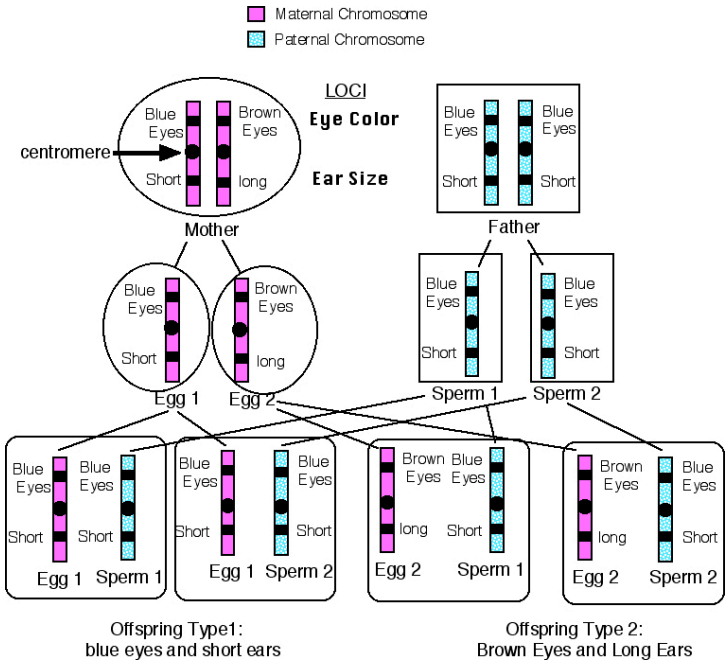
- p 211-214 "Mendel's first…" to "Punnett squares…"

### Meiosis Review Reading

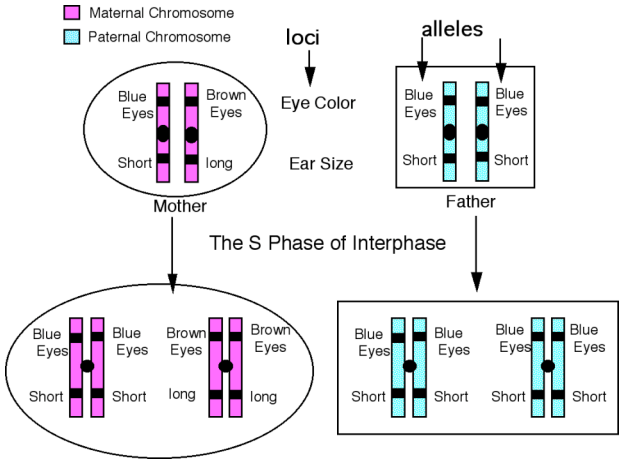
- p 195-199 "9.5 What…" to "Meiotic errors:…"
- p 222-224 "10.4 What is…" to end of p 224
- p 225 fig. 10.22 (Map These Genes)
- p 384-386 "17.3 How does…" to "17.3 Recap"
- p 386 fig. 17.13 (DNA Testing by Allele…)

In this example, the long ears gene stays with brown eye gene and short ears gene stays with blue eye gene. 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. (We'll modify that statement later.)

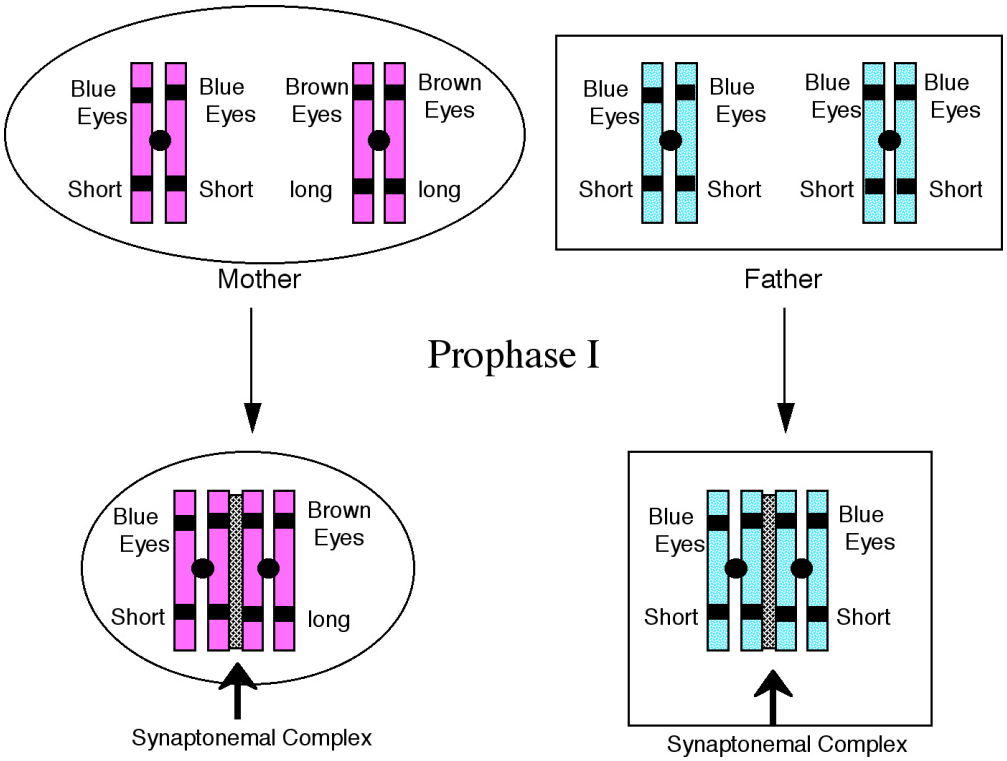
The homologous chromosomes segregate during meiosis and are independently assorted into the gametes. Thus, at your own fertilization, you received chromosome 1-23 from your mother and chromosome 1-23 from your father. Thus, you have two of each chromosome—homologous pairs. You inherit your genes in these 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 diagram illustrates the inheritance of only one chromosome. Remember that recombination occurs in all 23 pairs of chromosomes.




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 in the diagram below:



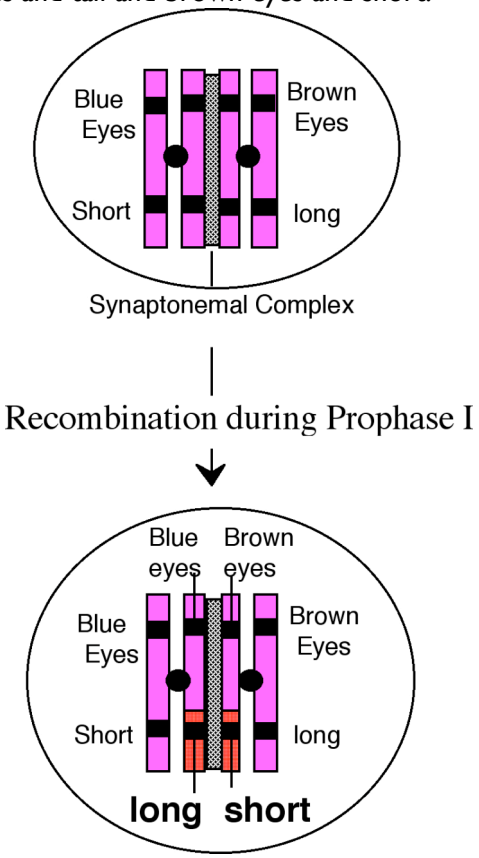

Each chromosome makes an exact copy of itself. The copies are attached to one another by the **centromere.** Each half of this double chromosome is now called a chromatid (remember fig. 9.7 page 187). The homologous pairs, (which have been ignoring one another in the cell up to this point), find each other and join together through a protein association called the **synaptonemal complex** as follows in the diagram below:




This process, in which 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 of recombination, the mother's chromatids would look like the diagram. The two outer chromatids are the original ones or

the **parental chromatids**. Because of the recombination event, the two inner chromatids are now different from any combination of genes on the mother's chromosomes. They are called **recombinant chromatids**. As meiosis dispenses 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.



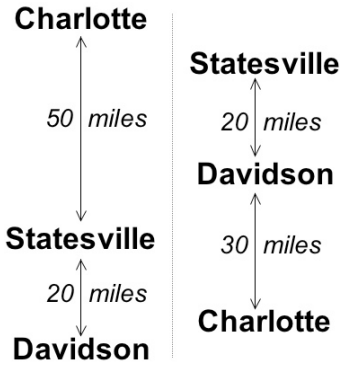

This feature of meiosis and inheritance (discovered by Alfred Sturtevant, an undergraduate student working with the geneticist Thomas Hunt Morgan) is used by nature as a way to increase the diversity in a population, thus giving natural selection a greater variety of organisms to work on.

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 made by Sturtevant, Morgan, and others about recombination: 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 <u>not</u> to have their alleles separated on different but homologous chromatids.

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 four chromatids, not just the inner two chromatids as shown in this simplified diagram.

But how does your understanding of recombination 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. Recombination frequencies and map units cannot tell you precisely how may kilobases apart two genes are, but it does give you an approximate distance to use as a valuable starting point in your hunt for an unknown gene.

You can tell how far apart three 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 we asked you to draw a map of these cities with this information, you would have two possible maps:




In order to choose between these two possibilities, you have to know the distance between Davidson and Charlotte. If it is 70 miles, then the map on the left is correct. If it is 30 miles, then the map on the right is correct.

This mapping strategy is exactly how you map genes on a chromosome. You need three points, three loci, and you find out how far apart each of the pairs of loci are 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. 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?)

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

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

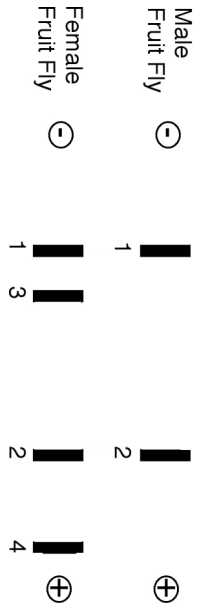
4. Given genetic data, be able to construct a genetic linkage map.

📖 **Review Reading**

- p 224 "Geneticists can…" to end of p 224
- p 224 fig. 10.20 (Recombination Frequencies)
- p 224 fig. 10.21 (Steps toward a…)
- p 225 fig. 10.22 (Map These Genes)

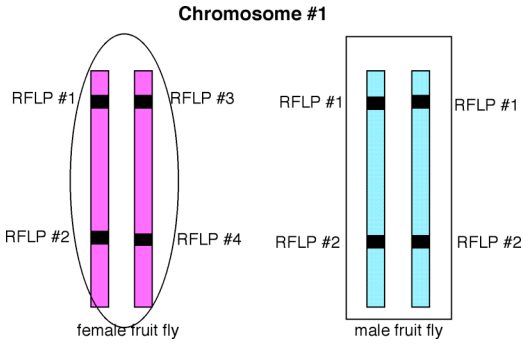
How do you determine recombination frequency? You must be able to detect the alleles and follow them as they are inherited. In the example above, determining recombination frequency was fairly easy—you can see eye color and height thus you can follow the alleles. Even though it is a bit more technical, 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 an 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 X a homozygous recessive). Below is an example of such a Southern blot:




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 Southern blot above. 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 female is a heterozygote.

While we do not know exactly which RFLPs correspond to which loci, chromosome #1 in these flies might look something like the diagram below:

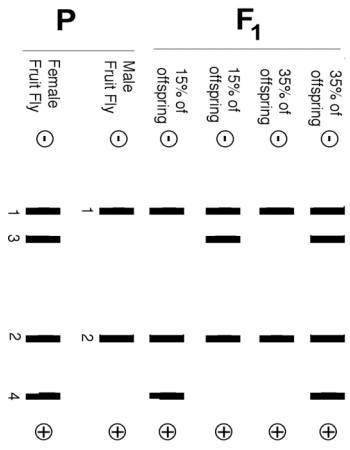



As it is drawn, band 1 and 3 on the female fly's Southern blot are alleles of the same locus, and bands 2 and 4 are alleles of a second locus. Thus, this female fly is a heterozygote at both loci. The male has identical alleles at the first and second loci; thus he is a homozygous at both loci. Now when this female fly creates her eggs, she can 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 knowledge that the frequency at 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 perform a Southern blot on the offspring and obtain the data below:

**Southern Blot for RFLPs on Chromosome #1**




Because the male fly is a homozygote, he always passes on RFLP 1 and 2. Thus, all of the F1 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 result 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.

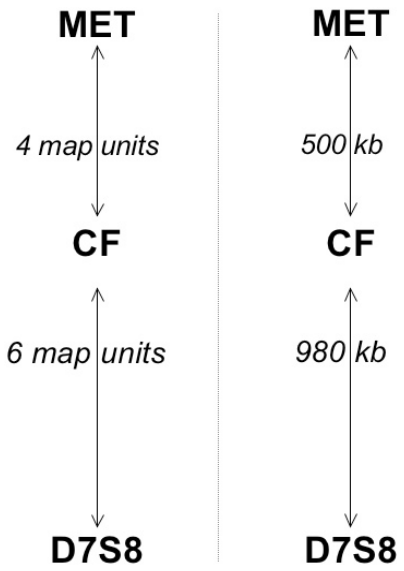
15% of the offspring inherited RFLP 1 and 2 from their father and RFLP 3 and 2 from their mother. A chromatid bearing RFLP 3 and 2 is a recombinant chromatid, and thus, these offspring are the products of recombination. Likewise, 15% of the offspring inherited RFLP 1 and 2 from their father and 1 and 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 (three 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 cannot accurately calculate recombination frequencies in a single family. Rather, you have to look at an entire population to determine recombination frequencies. Thus, you have to gather many, many samples and run many, many Southern blots. If you remember, in the case of CF, the disease is linked to two markers on chromosome #7 -- they are called D7S8 and MET (a.k.a. SWSS842). The lab that characterized the marker chooses the names and the names can mean almost anything, so do not try to look for a scheme to these names—there is none. You can think of them as human names. You name your kid D7S8 or MET (or SWSS842) 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:

| Loci Analyzed | Recombination Frequencies |
|---|---|
| *MET* (SWSS842) & *D7S8* | 10% |
| *MET* (SWSS842) & CF | 4% |
| *D7S8* & CF | 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 shown in the diagram below.




Thus, investigators were able to determine that *MET* (SWSS842) and *D7S8* flanked the cystic fibrosis gene. Knowing that *MET* (SWSS842) and *D7S8* are very often inherited with CF 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* (SWSS842) and *D7S8* and both of these markers are identifiable by the presence of restriction target sites and marker sequences. Recombination frequency analysis determined the order of and distance between these RFLPs to be as diagrammed below. The total distance between *MET* (SWSS842) 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. RFLP analysis can be 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 include? Whose DNA would you sample?

5. What is a map unit?

6. Be able to map a DNA segment given the outcome of dihybrid test crosses for a certain gene.

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

### 🕮 Focused Reading

- p 381-382 "Genetic markers…" to "Disease-causing…"
- p 393-395 "What have we…" to "The use of…"

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 once 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 that encodes insulin or the gene that encodes glycogen synthase.

How long does a probe have to be to meet this criterion? If there are $6 \times 10^9$ base pairs in diploid

human genome, a base sequence should be long enough to have a probability of existing at the frequency of one in six billion. How long is that? What are the chances that a given base sequence starts with "A"? The answer: one chance in four because 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, one in four. But the chance 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. With six 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 (six 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 one in six billion? The answer is 14; that is $(1/4)^{14}$ = about $3 \times 10^{-9}$.

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.

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 flanked by two RFLPs, MET (now known as SWSS842) and D7S8. So, what now? Linkage analysis was pointless because the distances between loci in this region are so small that recombination doesn't occur often enough to be detected. Researchers had no probes for CF itself, since they had no idea what the gene sequence was. So investigators had to continue their studies with a classic technique called **chromosome walking.** At the time, the human genome had not been sequenced, and the database of genetic markers was tiny compared to what is currently available. The investigators used chromosomal walking to clone adjacent pieces of chromosome 7 gradually until they found a piece of DNA that encoded CF. Every time they cloned an adjacent piece of DNA, they tested it to see if it encoded a gene expressed in lungs and sweat glands. Eventually they did locate, clone and sequence the CF gene.

Chromosome walking is rarely used now that whole **genomes** have been sequenced. A genome is the total genetic information of an organism. The human genome was completely sequenced in an effort that concluded in April, 2003, exactly 50 years after Watson and Crick published the structure of DNA.

### Uncovering the CF Gene

Now we are going to use some of the genomic tools that are freely available online. These tools were developed as a part of the human genome project and access to these tools is one legacy of the human genome project. As you may know, "the" human genome was not deduced from a single person, nor does it represent an "average" sequence of human DNA. It is more of a reference genome sequence the same way the tree closest to your home is an example for its species, but it should not be considered "the" tree or an "average" for all trees. The databases we are about to explore were produced by labs all over the world and funded by government and private organizations. Using those databases, let's go find the CF gene.

### A Genomics Approach to the CF Gene

If you wanted to home in on the CF gene today, you would not need to perform a chromosomal walk. You could use a process similar to what we will do here. The next part of your *Study Guide* reading is very interactive and will use information from the human genome project. Because these databases are updated frequently, this part of the reading (and questions) can be found on the web (rather than in print here). Go to this web page and follow the instructions there to complete the reading and exercise.

http://www.bio.davidson.edu/people/macampbell/Hope/DQ/DQ9.html

Near the end of this web reading you will need to use the information you've uncovered to draw the structure of the CFTR protein. The membrane diagram below is a good place to start for this drawing:

*extracellular*

*plasma membrane*

*cytoplasm*

**📖 Focused Reading**

- p 359 "16.3 How are new…" to "Vectors carry..."
- p 363 "cDNA libraries…" to "DNA can be…"

We now have the CF gene sequence, but it is so big we'd rather work with just the exons. Furthermore, we'd like to compare the exons for a wild-type person with the exons of a person suffering from CF. How can we isolate just the exons for CFTR from people with and without CF? Well, all cells contain all genes but each cell; type (liver, retina, and muscle) uses only specific genes of the genome. Because we know that CF patients have problems in their lungs, pancreas, and sweat glands, these cells are a good place to find CFTR exons. 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 CFTR, they must contain mRNA for this protein (*i.e.,* these cells must 'use' (or "express") the CFTR gene). After isolating the mRNA from these cells, investigators incubated the mRNA with all four DNA nucleotides (dNTPs) 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 unusual enzyme later when we discuss HIV.) Using CFTR cDNA, investigators were able to compare wild-type and mutant CFTR cDNA sequences.

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 264 of your textbook). Just as you did from the Genome Browser, the original investigators noted that CFTR contains long stretches of hydrophobic amino acids alternated with long stretches of hydrophilic amino acids (we suspect we are looking for a channel protein). 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.,* evolutionarily conserved proteins in different species are called **orthologs**). 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 wild type cDNA as a probe to isolate cDNAs from CF patients and they sequenced this gene from both wild type people and CF patients. After comparing the DNA sequence from the wild type gene to the sequences of the same gene in people having CF the researchers 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 one-letter abbreviation for the amino acid phenylalanine is "F." Thus, they called this mutation ΔF508 -- a deletion (Δ) of phenylalanine (F) at position 508.

So, it appears as though investigators have found the allele that causes CF, at least in 70% of the cases. Unfortunately, the remaining 30% of CF cases are caused by over 900 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. (Note: if you want to learn more about the wide variety of CF mutations (optional) check out www.genet.sickkids.on.ca/cftr/ This web site has a table of all CF mutations.)

Clearly, several mutations in this gene were associated with CF. 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), and/or if this protein is embedded in the membrane due to regions of hydrophobic amino acids. Computer assisted prediction of protein conformation is a rapidly growing field, though predictions for large proteins are still fairly crude.

❖ **Study Questions:**

1. What is cDNA and how is it produced?
2. Why would you want to sequence cDNA instead of a gene?
3. How many alpha helices do they predict are in the cytoplasmic part of CFTR? How many beta strands?
4. How many transmembrane domains are in CFTR?
5. How many ATP binding sites? Phosphorylation sites? Glycosylation sites?
6. What feature of CFTR is closest to the amino acid F508?

The DNA encoding the wild-type allele at the CF locus is 189 kb long (huge!) and contains 27 exons. After processing, the final mRNA is 6129 bases long. (Thus, more than 180,000 bases in the allele are in introns.) This mRNA is translated into a protein that is 1480 amino acids long.

As you discovered using the Genome Browser, the protein 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 to ADP and $P_i$. CFTR's 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 transporter. The regulatory domain can be phosphorylated by a cAMP-dependent protein kinase (PKA - sound familiar?). When a serine or threonine in the regulatory domain 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** ("conductance" being referred to here is chloride ion conductance). CFTR 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.

[Note: Sometimes a protein gets named accurately for a function it performs and then later scientists find out that the protein also serves other functions. For example bone morphogenetic protein (BMP) was first implicated in bone development as the name suggests. Later, scientists found that there are numerous, related BMP genes and these BMPs are involved in the growth of many other types of cells such as neurons, but the BMP name stuck.]

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 106-107 "Osmosis is…" to end of p 107
- p 107 fig. 5.9 (Osmosis Can Modify…)

Unlike sodium or calcium, water is not a leader but a follower -- a lamb in a world of Marys (as in Mary had a little lamb). Think of ions as Mary; 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. Osmosis 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 thick and dry.

✎ 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 cannot move sodium into the cell from airway surfaces. Using what you know about osmosis, why would a defective $Na^+$ channel 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.) [*Molec Med Today* 5:462.]

✎ NEWS ITEM: *Karenia brevis* is a marine single-celled organism underlying red tide events along Florida's coast. *K. brevis* makes a toxin called brevetoxin. It also makes an antidote molecule called brevenal that binds to sodium channels and protects them against the toxin. Brevenal may end up being a treatment for CF; in a sheep CF model, brevenal blocked bronchoconstriction and enhanced the clearance of mucus. [*Science* 316: 1561]

❖ **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. 5.9 page 107).

Now, back to our understanding of CF. Where does the ΔF508 mutation appear in the CFTR? It is near the first ATP-binding site. Aha! Good place for a mutation that seriously impairs protein function. One hypothesis would be that maybe the mutated CFTR gene produces a protein that cannot bind ATP and therefore cannot 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 still 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 move $Cl^-$ successfully.

💻 **Web Reading**

- *In situ* methodology
  www.bio.davidson.edu/courses/genomics/method/insitu.html

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 (PKA) 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 via *in situ* hybridization, 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 to the cell membrane. You could use immunohistochemistry to look for the protein on the cells of CF patients. (Consult this Study Guide's Appendix C for more information on immunostaining techniques.) 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 *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 in the cytoplasm or nucleus become radioactive when the cDNA hybridized to the mRNA. Cells not expressing this mRNA will not become radioactive (because the probe found nothing to bind with and was subsequently rinsed away). These *in situ* hybridization 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 in the cells that are producing clinical symptoms.

Do these *in situ* hybridization results 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" (and why it is incorrect to use the word "proof" when talking about even the strongest scientific evidence).

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 after 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 wild type 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. This type of experiment is often called a "rescue" experiment.

---

📖 **Focused Reading**

• p 391-392 "Gene therapy…" to end of p 392

---

There are several ways to transfect cells with DNA. 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. Fig. 16.16 (page 367) contains a diagram of an **expression vector**—a plasmid that allows you to express a foreign gene.

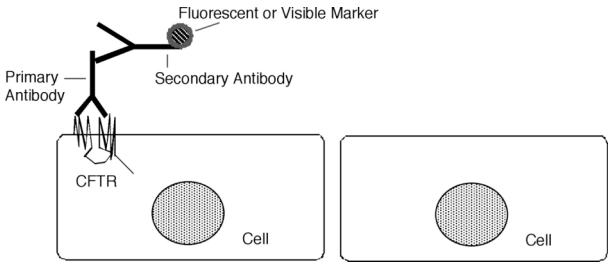
The CFTR cDNA with its artificial promoter is combined with the expression vector to make an **expression construct** (the cDNA plus promoter is the **insert**). This construct is incubated with CF respiratory cells in tissue culture. Under certain conditions, the cells will take up DNA and begin to express this foreign gene as if it were their own. The process by which transgenes are put into eukaryotic cells is called **transfection**. Some types of expression constructs remain in plasmid form, leading a **transient transfection**. The plasmid is not propagated if the cell divides. On the other hand, some types of expression constructs trigger the incorporation of the insert into an existing chromosome. This change is permanent for the cell, and the transferred gene is called a **transgene**. The cell containing the transgene is called a **transgenic cell.**

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). The results of these transfection experiments provide pretty strong evidence that the CFTR 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.

**Web Reading**
- Immunofluorescence Methodology
  www.bio.davidson.edu/old_site/student/IMF.html

If our initial hypothesis is not supported, we have to modify it. Another possibility is that the ΔF508 mutation influences 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). For immunocytochemistry, you need a way to 'see' or detect a protein. Before you can see proteins in cells or tissues, 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?). Antibody-antigen 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** .




Now you have tagged the CFTR with the primary antibody but you need a way to 'see' your antibody tag. So, you then apply a **secondary antibody**—one that 1) has been produced to recognize the primary antibody; and 2) has been covalently bound 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.

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 immunocytochemistry, you can determine the density of the protein in the membrane, and/or the protein's precise subcellular localization.

When investigators used **immuno-histochemistry** 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, missing a single amino acid causes the CFTR to be inappropriately sorted - it never reaches the plasma membrane and this mutation is the cause of 70% of all CF cases.

✎ NEWS ITEM: Turcumin, a component of the spice turmeric, affects CFTR protein localization. When cells are treated with curcumin, more ΔF508-CFTR protein makes it to the plasma membrane where it can do its job. However, this spice is not a 'cure,' since the bioavailability of curcumin is low— it may be there but the cells cannot use it efficiently. Labs working in this area are focusing on ways to increase the curcumin's bioavailability to hopefully develop this treatment. [*Nature Reviews Molecular Cell Biology* 7:426]

❖ **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 DNAse 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?). Because $Cl^-$ 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 is a very useful protein isolated from jellyfish; it is useful because it glows. Anywhere this CTFR-GFP is located, researchers could see it glowing under the microscope. The conclusion: cAMP acted like a switch to open the channel which was already located in the plasma membrane. [*JBC* 273:21759-68.]

🕮 **Focused Reading**

- p 353-354 "Restriction…" to end of p 354
- p 359-361 "Vectors can…" to "Reporter genes…"

We now know that the binding of ATP at site #1 converts the channel from a locked mode to an unlocked mode, but this interaction does not open the channel. ATP binding at site #2 opens the channel, but only if there is ATP already bound to site #1, and the cytoplasmic domain is phosphorylated. The process seems complicated, but even this explanation is a simplified version of a process we do not fully understand.

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 viruses and liposomes. 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 wt version of 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 hope sounds really straightforward, but a long journey from an idea to the finished product still lies ahead. We do not 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 for CF is available (*Science* 256: 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 overexpression of CFTR toxic? One problem with transgenes is that they do not 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 destroys the respiratory cells. This process is called **autoimmune disease**.

In addition, the immune system can potentially react to the viral vector itself. In 1998, Jesse Gelsinger, an 18 year old participant in a gene therapy clinical trial for a different disease, died when his immune system mounted an overwhelming attack on the adenovirus vector, related to a common cold virus. Nowadays, adenovirus is not used as a vector for potential gene therapies.

6. Can safety be insured? Safety is always a question with bioengineered organisms such as the viral vectors in this approach. People unfamiliar with the technique might fear that the viral vector woud "get loose" in the population. However, the engineered virus is not able to duplicate itself. The
I

---

❖ **Study Questions:**

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

2. What are some technical barriers that must be overcome 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. These results 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 export of $Cl^-$ ions, which would draw water into the mucus, dilute the salt concentration and allow the lung's naturally produced bactericide to function. An "alternative" $Cl^-$ ion channel has been identified. It is a calcium-activated chloride-ion channel that can be stimulated to open when ATP or UTP is administered to the outside 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. (Welsh - *Cell*)

✎ NEWS ITEM: Gerald Pier and 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. Instead bacteria live and reproduce in the lungs. Therefore, CF patients are hypersusceptible to infection by *P. aeruginosa.* [*Science* 271: 64-67.]

## Huntington's – A Dominant Neurodegenerative Disease

**Web Reading**

As the last part of this genetics unit, we will look briefly at the quest for the gene that causes Huntington's disease (HD). HD is a neurodegenerative disease, which means that neurons in particular brain regions die. Consequently, normal mental function is compromised as neurons are lost. Unlike CF or SC that can be observed in babies, the first signs of HD do not appear until a person is in her/his 30s, 40s, or 50s. The early signs of HD can be subtle grimaces, absentmindedness, and involuntary gestures. As the disease progresses, these involuntary movements become more pronounced and people with HD often move in a fashion that is easily confused with drunkenness. These involuntary movements become so severe that HD patients cannot dress or feed themselves and must often be restrained in bed to protect them from the falls and head injuries that often claim their lives. Mentally, as HD progresses, patients lose their ability to complete simple tasks and make plans. Eventually dementia and psychosis result. The progress of HD is slow but very steady – patients often spend 15-20 years experiencing this degeneration. Families and health care professionals can provide no cure and very little treatment for HD. Because HD is a dominant disease, all children of HD patients painfully watch their mothers or fathers deteriorate knowing all the time that they each have a 50:50 chance of the same fate.

While the fine points of identifying the gene responsible for HD 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, how the mutated protein identified in HD patients actually causes HD is practically a mystery even though scientists have been working to understand it for about ten years now. The HD gene was identified in 1993. Because HD is less common than CF in the human population, the odds of finding a person who is homozygous for HD are very low.

A real breakthrough in HD research came when a team of scientists including Dr. Nancy Wexler, a clinical psychologist whose mother was dying of

HD, found an area in Venezuela where the incidence of HD was unusually high (maintained through marriage within the isolated town). Wexler recommended that geneticists study the inheritance pattern of the disease in this group. A team of scientists 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 research produced the richest source of familial genetic information for HD that has ever been assembled. In a total population of 12,000 people, 258 had the disease and all were direct descendents of a woman who lived in the 1800s. It is believed that this woman had the misfortune of having a spontaneous ('new') mutation that was not present in her parents that caused HD in her and her descendents.

A large consortium of scientists called the Huntington's Disease Collaborative Research Group was begun, headed by Dr. James Gusella. After collecting genetics samples from families in this Venezuelan town and identifying individuals that were very likely to be homozygous for HD (the child of two parents with HD), the consortium investigators returned to their labs and began looking for RFLPs that were linked to the HD gene. Usually this process takes years, but the team 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 HD became available. The availability of this test forces children with an HD parent to make an agonizing decision. Should the child have the test or not? If one parent had HD, each child stands a 50% chance of having HD themselves. Most of these children have watched the chronic deterioration of body and mind caused by this disease as their parent dies and they know there is no effective treatment for HD. These children must now make a choice about whether or not they want to know if they will also deteriorate from HD even though there is absolutely nothing they can do to prevent HD or slow its progress. This situation brings into sharp focus the impact of biotechnology on our lives. Because the genetic test now exists, children of Huntington's patients 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 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, *etc.* Having to make difficult decisions about revealing our genetic fates are certainly significant by-products of the biotechnology revolution that makes genetic tests possible.

But back to the quest for HD. As it turned out, the euphoria about how quickly the HD gene would be discovered evaporated as it became apparent that the search would be long and arduous because 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 region, 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 mRNAs expressed by respiratory or sweat gland cells because they knew those cell types malfunctioned in CF. In contrast, HD investigators had little idea which neurons might be making the normal or abnormal versions of the HD protein.

HD 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 research narrowed the DNA segment of the search to a 500 kb region (three 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 4p mapped segment with cDNA. So now what? The HD investigators began by creating a set of overlapping fragments (called contiguous fragments of DNA) that were mapped, as was done for CF.

The investigators who eventually cloned the HD gene in 1993, used a novel approach that led them to the correct gene. However, because the human genome sequence is freely available online, we will utilize this database to look more closely at the HD gene. Rather than using the Genome Browser as we did for CFTR, we will use a sequence search engine called BLAST. There are two types of BLAST searches, one for protein sequences (BLASTp) and one for nucleotide sequences (BLASTn). We will use both to find the HD gene.

The next part of your *Study Guide* reading is very interactive. Because the on-line databases are updated frequently, this part of the reading (and questions) can be found on the web (rather than in print here). Go to this web page and follow the instructions there to complete the reading and exercise.

www.bio.davidson.edu/people/macampbell/Hope/DQ/DQ10.html

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 repeated between 11 and 34 times in this region. In themselves, repetitive codons are not so unusual. Many functional genes contain trinucleotide repeats. However, the HD gene from afflicted individuals contains from 38 - >100 copies of "CAG" in this region. This increase in the number of codons is a type 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. (You saw something like this during the VNTR lab measuring the number of 16 base pair repeats at your D1S80 locus).

While identifying the trinucleotide repeat doesn't give us much help in understanding the protein defect in the HD gene, this type of mutational change is also found in several other genetic diseases: myotonic dystrophy, fragile X syndrome (a form of mental retardation), spinal bulbar muscular atrophy (see page 383 and fig. 17.10), and more. Therefore, this trinucleotide repeat expansion, a form of insertion mutation, is a type that has been shown to produce at least twelve different genetic diseases. This correlation greatly strengthens the evidence that investigators actually found the HD gene.

Researchers have found the HD protein huntingtin in mice, fruit flies, and yeast, giving them the important ability to perform experiments with huntingtin gene and protein. In 1995, researchers made a startling discovery. When the protein encoded by the HD gene (later named **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 binds 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 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 immunocytochemistry 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. Recently, scientists have learned that the huntingtin protein enhances a neuron's ability to transport growth factor peptides and uses HAP1 to help move these growth factor peptides within neurons. It is thought that when neurons cannot move growth factor molecules properly out to the ends of their axons and/or dendrites that the neurons might die prematurely.

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 a treatment or cure for the biochemical defect that causes Huntington's Disease.

Note: If you want to learn more about Huntington's Disease you might consider reading *Mapping Fate* by Alice Wexler (Nancy Wexler's sister). This book describes the quest to identify the genetic basis of Huntington's and how the Wexlers cared for their mother dying of HD while they were influential in organizing the scientists that identified huntingtin. The December 2002 issue of *Scientific American* also contains an easy-to-read article summarizing the difficult quest to understand how the mutant huntingtin protein causes this disease. [Cattaneo *et al.* (2002) *Scientific American* 93-97.] If you want to hear Dr. Wexler speak about her family's decision not to be tested for HD check out the NPR story in the web listening below.

💻 Optional Web Listening (~10 min)
- Reading Genes for Disease, Part 3: Huntington's
  www.npr.org/templates/story/story.php?storyId=1897199

---

✎ NEWS ITEM: A research team has developed and tested suppressor polypeptides in *Drosophila*. The suppressor polypeptides bind to the mutant huntingtin protein (with all the repeats) and reduce huntingtin protein aggregation, neuron degeneration, and death in fruit flies. By interfering with the protein interactions that cause aggregations in neurons, researchers may be able to design therapies to treat HD. [*Nature Genetics* 30: 367-76.]

---

❖ **Study Questions:**

1. 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?

2. 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?

3. 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?

---

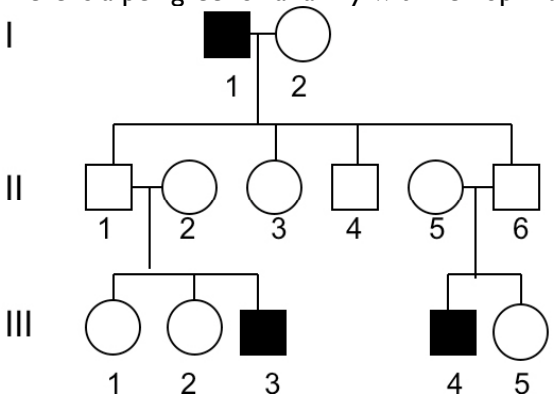
## Sex-linked genetic disorders

And finally, a note about sex-linked genetic disorders such as color blindness, hemophilia A, and Duchenne's muscular dystrophy…

**📖 Focused Reading**

- p 226-228 "Genes on sex…" to "10.4 Recap"
- p 227 fig. 10.23 (Eye Color is a Sex-Linked Trait)
- p 228 fig. 10.24 (Red-Green Color Blindness…)

Here is a pedigree for a family with hemophilia:




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 Xhomologue, 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 I.1, a male, has the disease. The disease is carried on his X chromosome. Therefore, he <u>cannot</u> pass the disease on to his sons because they must receive his Y chromosome in order to become male. However, <u>all</u> of his daughters will inherit his X chromosome (that is what makes them girls, they must inherit an X from both parents). Individual III.4 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 because 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 I.2 is homozygous normal? Assuming individual I.2 is a hemophilia carrier?
2. How did individual III.3 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 <u>thoroughly</u> explaining the newspaper article below 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*:

**Most common form of Alzheimer's linked to cholesterol-processing gene**

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 four 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 copies."

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 cholesterol through the bloodstream.

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 effect.

---

# *Unit III: Bioenergetics*

**📖 Overview Reading**

- Chapter 2 • The Chemistry of Life
- Chapter 3 • Macromolecules…
- Chapter 4 • Cells: The Working Units of Life
- Chapter 5 • The Dynamic Cell Membrane
- Chapter 6 • Energy, Enzymes, & Metabolism

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, O.K., dead creatures do not move, or vocalize, or breathe, or eat -- *i.e.,* they cannot 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 do not sense anything. They do not see, hear, feel, or taste. That is, they have no sensory functions. And, again, as far as we living types know, dead creatures do not think about anything or have any emotions -- they do not remember, plan, enjoy, problem-solve, love, hate, or do homework.

If you look for the common denominator in all these activities that make an organism alive -- moving, sensing, thinking, and feeling -- you find that such activities all require ENERGY. Only living creatures can use energy to accomplish these activities, these characteristics of life. A biological creature can have a unique structure, but without energy, it cannot be alive.

Energy is an unusual and sometimes challenging concept to study. Energy is 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 vibrations of air at certain frequencies (500-16000 Hz: sound waves). However, we have no sense organs for many forms of energy such as radio waves (your radio receiver can detect these waves, but you cannot), radioactivity (a Geiger counter can detect these, but you cannot), or neutrinos (they are passing through you right now, but you cannot sense them).

What is energy anyway?

**📖 Focused Reading**

- p 30-31 "How do atoms…" to "2.3 Recap"
- p 119-126 "6.1 What…" to "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 is 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, the definition of energy 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 kayak? 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 energy in your own terms.

2. How is energy measured? In what units? 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 bioenergetics 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 might kill you.

## QUESTION #1: The US Government vs. 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 is 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.*)

---

📖 **Overview Reading**

- Chapter 7 • Pathways that Harvest…
- Chapter 8 • Photosynthesis: Energy from the…

---

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. Paraquat kills almost all plants (except a few plants that are resistant to it). How does paraquat kill plants? Why might paraquat 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.

Plants are unsung heroes. They 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 plants far 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.

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 cycle.** Both processes occur in the leaves of plants. A typical plant leaf is illustrated in fig. 8.17 (page 173) 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 cannot 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 the reactions that keep them alive.

---

🕮 **Focused reading**

- p 163-165 "8.2 How does…" to "Light Abs…"

---

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 a wavelength of 600 nm will give us a yellow sensation. The relationship between colors and wavelengths is illustrated on page 164, fig. 8.5.

Visible light has some of the properties of waves and therefore can be described by 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 wavelength = 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. ]

Green plants harvest the energy of photons. 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 **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 photons 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 165-166 "Light absorption …" to "Reduction…"

**Chlorophyll a** and **b** are green pigments and **carotenoids** are shades of yellow and orange (as in carrots and fall leaves). Because 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 color variation 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. Therefore, it reflects green light. So, chlorophyll a does not absorb green photons. Because colors are so complex, however, it is really hard to say what colors are absorbed. Color absorption has to be measured using a **spectrophotometer**. The **absorption spectrum** for chlorophyll a and b are illustrated in fig. 8.6 on page 165. 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 the spectrum 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. Chlorophyll a in green plants harvests the energy of the blue/violet and orange/red photons. In addition to chlorophyll a, most plants have accessory or secondary pigments (*e.g.,* chlorophyll b and the carotenoids) that absorb photons at other wavelengths. Therefore, plants can frequently harvest photons across nearly all of the 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 precise order. The process of photosynthesis is something like assembling an automobile; you cannot put in the stereo before you have assembled the dashboard; you have to do things in order. Instead of having the molecules involved in photosynthesis floating around haphazardly in the cytoplasm, important photosynthetic molecules 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 83 "Plastids…" to "Several other…"

Look at the picture of chlorophyll a on page 165 (fig. 8.7). 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 ion 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 (rather than have them float about the cytoplasm).

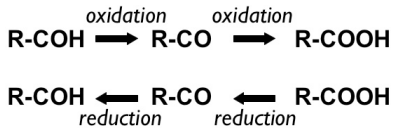
3. Describe the structure of the chloroplast including the structure and location of thylakoids, grana, and stroma. Describe the location of chlorophyll a in the chloroplast, and explain how the molecule is anchored into the membrane.

### 🕮 Focused Reading

- p 164 fig. 8.4 (Exciting a Molecule)

Each pigment has a particular photon energy level that "fits" it perfectly such that photons of that energy level can be absorbed very effectively, while photons of other "misfitted" energy levels cannot be absorbed. When a pigment molecule absorbs light, photon energy is transferred 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 photon 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 by a photon, this primary electron acceptor takes excited electrons away from the excited chlorophyll before they have a chance to fall back to ground state.

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 140-141 "Redox reactions..." to "The coenzyme…"
- p 141 fig. 7.3 (Oxidation and Reduction are…)

### ❖ 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.

Chlorophyll harvests light energy by passing along excited electrons before they have a chance to fall back down to their ground state. Understanding how chlorphyll harvests light energy is not enough information to understand how paraquat kills plants and how it endangers humans. In order to understand paraquat's herbicidal effects, 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.

Molecules contain potential chemical energy. We say that the potential chemical energy is "in the molecule's bonds," but this statement is misleading in many ways. 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 molecule moves toward a stable configuration occurs is the **heat of the reaction** (**$\Delta$H, H** is referred to in your text 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, because 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 $\Delta$H of the reaction is designated as negative. Thus, a reaction that **gives off energy** (*e.g.,* burning fuel) has a **-$\Delta$H** 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 $\Delta$H of the reaction is positive. Thus, a reaction that **requires the input of energy** (draws energy from the environment) has a **+$\Delta$H** and is said to be **endothermic.**

### 🕮 Focused Reading

- p 122-123 "Chemical reactions:…" to "6.1 Recap"

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 ($\Delta$**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** $\Delta$**S**, while reactions (*e.g.,* refining gasoline) in which entropy is decreased have a **negative** $\Delta$**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 ($\Delta$H) and changes in entropy ($\Delta$S). Usually, reactions that give off energy to do work (exothermic or -$\Delta$H reactions) also involve an increase in entropy (have a +$\Delta$S). 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 (-$\Delta$H), and the entropy of the molecules is dramatically increased (+$\Delta$S).

Conversely, reactions that absorb energy (endothermic or +$\Delta$H) usually involve a decrease in entropy (have a -$\Delta$S); 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 (+$\Delta$H) and the entropy of the molecules is dramatically decreased (-$\Delta$S). 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 $\Delta$H of reactions because $\Delta$**H** 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, the **free energy of the reaction (**$\Delta$**G)** is also very important to biologists. $\Delta$**G** 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 **-**$\Delta$**G** while non-spontaneous reactions are said to be **endergonic** and have a **+**$\Delta$**G**. Usually, exothermic reactions are exergonic and endothermic reactions are endergonic, but not always. If an endothermic reaction (takes heat from the environment, +$\Delta$H) involves a large increase in entropy (+$\Delta$S), then it may be spontaneous (have a -$\Delta$G) 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 +$\Delta$H), but results in a dramatic increase in entropy (+$\Delta$S) as ice goes from an organized crystal to a disordered liquid form.

**Quick reference chart**

-$\Delta$S = product has less randomness
+$\Delta$S = product has more randomness
-$\Delta$H = rxn gives off energy (exothermic)
+$\Delta$H = rxn takes in energy (endothermic)
-$\Delta$G = rxn is spontaneous (exergonic)
+$\Delta$G = rxn non-spontaneous (endergonic)

The friendly relationship between $\Delta$H and $\Delta$S is:

$$\Delta G = \Delta H - T\Delta S$$

*(where T is the temperature in degrees Kelvin)*

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 energy terminology to confuse you, although it may feel confusing. Rather, we face a dilemma in teaching you. In chemistry, you are learning (or will learn) about $\Delta$H, 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 power biological creatures). Thus, we have to introduce $\Delta G$, and we have to tell you how it relates to $\Delta H$ 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, $-\Delta H$, $+\Delta H$, endergonic, exergonic, $+\Delta G$, $-\Delta G$, spontaneous, non-spontaneous, entropy, $+\Delta S$, $-\Delta S$.

3. What determines whether or not a reaction will proceed without an input of energy from the cell? Give 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.

✎ NEWS ITEM: Think about how fossil fuels (natural gas, petroleum, and coal) are created. These energy sources are fossil fuels because they are made from decayed plants and animals (full of proteins, lipids, etc.). Millions of years of heat, pressure, and bacterial processing is thought to change the complex organic (carbon-containing) molecules such as proteins into simpler carbon skeletons of these fossil fuels. Functional groups and double bonds are lost through chemical reactions. The result is saturated hydrocarbon skeletons, which are the raw materials of oil. It has been long assumed that bacteria carry out the reduction reactions that make these hydrocarbons. In the summer of 2006, a group of scientists examined lake sediment to demonstrate that inorganic, abiotic reactions (not bacterial processing) occur in the first stages of sedimentation. They showed that sulfides can reduce double bonds in hydrocoarbons under geochemical conditions in sediment. Conventional wisdom describing the pathway of carbon processing in nature may now need to be revised to include abiotic steps independent of bacteria. [*Science* 312: 1627]

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. Said another way, living is an energy-absorbing activity. Living is endergonic and endothermic. So, you have a problem because endergonic reactions do not 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), ATP is the direct source of energy for most endergonic reactions in living things.

📖 **Focused Reading**

• p 123-125 "6.2 What is the role…" to "6.2 Recap"

❖ **Study Questions:**

1. Describe and draw the reactions converting ATP to ADP, and *vice versa*. What is the $\Delta G$ of each reaction?

2. Be able to describe the process of energy coupling by phosphate transfer outlined in fig. 6.7 on page 125.

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 using ATP's energy. Consequently, cells must continuously replace this lost ATP. Marijuana plants can restock this ATP supply directly by using solar energy (photosynthesis) or indirectly by burning fuel molecules (cellular respiration). Animals are not photosynthetic, they can create ATP 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 & Generating $O_2$: The Light Reactions

**📖 Focused Reading**
- p 161-163 "8.1 What is…" to "8.2 How does…"
- p 166-169 "Excited chlorophyll…" to "8.3 How is…"

**💻 thelifewire.com Reading**
- Tutorial 8.2 • Photophosphorylation

**💻 Web Reading**
- Animation of Photosynthesis
  www.bio.davidson.edu/courses/Bio111/Photosynth/PS.html
- Diagram of $NADP^+$ conversion to NADPH
  www.bio.davidson.edu/courses/Bio111/NADPH.html

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). Lipid, amino acid, and nucleotide synthesis processes are highly endergonic and the energy to power these non-spontaneous processes 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:

$$6\ CO_2 + 12\ H_2O + \text{light energy} \rightarrow C_6H_{12}O_6 + 6O_2 + 6\ H_2O$$

This overall reaction is actually a **redox reaction**. The light reaction component is as follows:

$$12\ H_2O + \text{light energy} \rightarrow 6\ O_2 + 24\ e^- + 24\ H^+$$

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) on the electron in orbit around it. Thus, **electronegative** molecules (that have high affinity for electrons and attract them very strongly) can readily take hydrogen's electron away from the hydrogen nucleus, rather than sharing the hydrogen 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 their hydrogen nucleus – hydrogen electrons 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 six molecules of $O_2$, which the plant releases into 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 (two electrons per $NADP^+ \rightarrow NADP^-$) and 12 protons (one proton each, $NADP^- \rightarrow$ NADPH). [To form reduced NADPH, $NADP^+$ (the oxidized form) picks up two 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** (chlorophyll 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 that enter the chlorophyll molecule, and protons that 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 chlorophyll 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 do 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.

## Synthesizing Food: The Calvin Cycle

📖 **Focused Reading**

- p 162-163 "Photosynthesis…" to "8.2 How does…"
- p 169-172 "8.3 How is…" to "8.4 How do…"
- p 171 fig. 8.13 (The Calvin Cycle)

💻 **thelifewire.com Reading**

- Tutorial 8.3 • Tracing the pathway of $CO_2$

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:

$$24e^- + 24H^+ + 6CO_2 + \text{energy (ATP)} \rightarrow C_6H_{12}O_6 + 6H_2O$$

In this reaction, $CO_2$ is reduced to glucose; the carbon atoms have 12 hydrogens added and six 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 G3P is shown on p 147 of your text. G3P is a three-carbon sugar and it is the starting molecule (**precursor**) for the synthesis of lipids and several sugars (including glucose). Amino acids used for protein synthesis can also be made using G3P as a precursor, in addition to a source of reduced nitrogen.

The overall reaction of the Calvin cycle is:

**$3CO_2$** (from the air or water)
**9 ATP** (from the light reactions)
**6 NADPH** (from the light reactions)

↓

**1 G3P**
**9 ADP + 8 $P_i$**
**6 $NADP^+$**

The complete cycle is outlined in fig. 8.13 on page 171.

---

❖ **Study Questions:**

1. What role does "reducing power" play in photosynthesis? What molecules provide reducing power directly to the Calvin 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 cycle? Why is this molecule of pivotal importance in the life of the plant?

4. **Rubisco** is one of the most important and abundant enzymes in the entire biological world. What does rubisco do that is so impressive? What does rubisco's full name tell you about its functions?

5. Explain the Calvin 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 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 direct conversion would save six ATP per cycle and would be of great adaptive advantage to the plant. Use an energy diagram to explain your hypothesis.

---

✎ NEWS ITEM: In 2006, three investigators from Australian National University in Canberra determined that Rubisco can bind to $CO_2$ instead of $O_2$. Our atmosphere currently has 25 times more $O_2$ than $CO_2$, but Rubisco binds to $CO_2$ 100 better than it binds to $O_2$. Every second, each Rubisco molecule can fix about five $CO_2$ molecules into sugars. It turns out the active binding site appears to be largely determined by an intermediate metabolite that preferentially binds $CO_2$ over $O_2$. Based on their proposed mechanism, the investigators hypothesize that Rubisco will become less efficient at binding $CO_2$ as global temperatures increase, unless evolution selects for mutations that are more efficient at elevated temperatures. Unfortunately, humans have little control over the outcome of plant evolution, but our survival will be dependant upon favorable results. [*Proc. Natl Acad. Sci. USA* 103: 7246]

✎ 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. [*Naturforsch.* 45c:633-637.]

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 do not 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 any process that involves the transfer of energy, 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. Energy transfers 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.

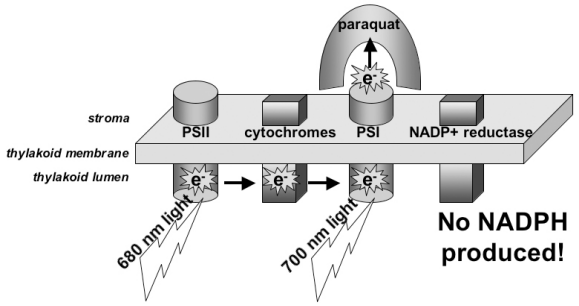
   D. As a precursor for amino acids and nucleic acids made in the cells of the root.

See fig. 8.19 (page 176) for a diagram of how the Calvin cycle fits in plant metabolism.

✎ NEWS ITEM: Petroleum (oil) is a key ingredient in many products such as plastics. As you know, petroleum supplies are expensive and in short supply. As a consequence, chemical companies are investigating using renewable resources such as plant-based materials as oil alternatives. In the summer of 2006 a research team described a new process that can convert fructose (a sugar in fruit) into a key plastic precursor called 5-hydroxymethyl furfural (HMF). There have been previous

methods for turning sugar into HMF, but this new process may be a more efficient, method to turn plants into plastics. [*Science* 312:1861]

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, electrons are "stolen" by paraquat. As a result, no NADPH is produced even in the presence of sunlight. Obviously NADPH is essential for the plant. Thus a shortage of NADPH means the Calvin cycle cannot create glucose to store energy and the plant will eventually die if it cannot store energy. Further, paraquat also damages plants by producing free radicals. Paraquat is very electronegative, but not more than $O_2$, so paraquat transfers electrons to $O_2$, producing free radicals (superoxide ($O_2^-$) and hydroxyl ion ($OH^+$)). As long as the plant absorbs light, the paraquat will continue to transfer electrons from photosystem I to $O_2$ and produce more destructive free radicals. Free radicals are highly reactive and particularly destructive to membranes such as the thylakoid membrane. Without an intact thylakoid membrane, the plant cannot sustain photosynthesis.




But human cells do not have chloroplasts, so why should you worry if you are exposed to paraquat? Later in this unit, we will see why paraquat might be harmful to humans.

✎ NEWS ITEM: The marine cyanobacteria (phytoplankton that used to be called blue-gree algae but are really prokaryotes) *Prochlorococcus* and *Synechococcus* are the smallest and most numerous photosynthetic cells in the oceans, and probably fix more $CO_2$ than all the land plants combined. In 2006, Sallie Chisholm at MIT and her collaborators at Dalhousie University in Canada have sequenced the genomes of many cyanobacteria and the viruses that infect them. They have discovered that these viruses carry a variety of genes that provide central components of photosystem II. By comparing DNA sequences, Chisholm *et al.* have discovered the viruses can shuttle genes between each other and their hosts. This result reveals that photosynthesis can evolve quickly through horizontal gene transfer (non-sexual transmission of genes across species). [PLoS Biology 4:]

## ❖ 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 targeted?

✎ NEWS ITEM: Why aren't plants black, absorbing all (or most) wavelengths of light? Absorbing more light is not necessarily better because light in large quantities generates reactive chemical species, such as superoxides and radicals. These reactive chemical species are associated with the electron-transport system in the chloroplast and can damage the plant, causing photoinhibition. Black plants might be destroyed by relatively low levels of light. [*New Scientist* 139:47.]

an excerpt from: *The New York Times Magazine* (11/19/1978)

### Poisonous Fallout From The War On Marijuana
by Jesse Kornbluth

Mexican marijuana growers had learned that paraquat-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 unforeseen 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 widespread confusion. Secretary 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 than 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 two percent of this laboratory's "contaminated" samples were paraquat-positive, PharmChem reexamined 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 potentially dangerous as Secretary Califano originally indicated, this news went almost unreported.

## Question #2: Why is Cyanide Poisonous?

### 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 that 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 require oxygen. While all nutrient molecules can be burned to obtain energy, **glucose** is by far the molecule is most frequently used for this purpose. Glucose is the predominant sugar in human blood, maintained at about 80 mg per 100 ml of blood.

---

### ❖ Study Question:

1. If a normal blood glucose level is 80 mg/100 ml, what is this concentration in % (w/v)?

---

### 🕮 Focused Reading

- p 131 "6.5 How are…" to "Enzymes can…"
- p 140 fig. 7.1 *(note relationship between auto- & heterotrophs)*
- p 49-53 "3.3 What are the…" to "Chemically…"

---

The process of photosynthesis harvests light energy and stores this energy as glucose molecules. Animals 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 – as you know from unit I). 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 cellulose contains **beta-glycosidic** linkages, we cannot break down cellulose to glucose. Humans 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 cannot get at these calories because you cannot break down the primary bulk of the vegetables -- cellulose. So cellulose simply passes through your body as "roughage."

The overall equation for cellular respiration is:

$$C_6H_{12}O_6 + 6O_2 + 6H_2O \rightarrow 6CO_2 + 12H_2O + \text{energy (ATP + Heat)}$$

You will immediately recognize this equation as the reverse of photosynthesis. Photosynthesis is an endergonic reaction with a $\Delta G$ of +686 kcal/mole. Conversely, cellular respiration is an exergonic reaction with a $\Delta G$ 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, cellular respiration is an extraordinarily efficient process.

In addition to being highly exergonic, cellular respiration, like photosynthesis, is a **redox reaction**.

📖 **Focused Reading**

- p 139-140 "7.1 How does…" to "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:

$$C_6H_{12}O_6 + 6H_2O \rightarrow 24e^- + 24H^+ + 6CO_2$$

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 chemical equation is as follows:

$$24e^- + 24H^+ + 6O_2 \rightarrow 12H_2O$$

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 ($H^+$) from one molecule to another using a dinucleotide (*e.g.,* NAD) as an intermediate. Nature often uses particularly good ideas over and over again in slightly different ways.

Electrons are not all equal in energy level. When an electron shares a covalent bond between hydrogen and carbon (as in glucose, amino acids, lipids, *etc.*), it has a relatively high energy level. When electrons share a covalent bond with oxygen, as in water, they have relatively low energy levels. 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 used to build ATP and also releases some heat 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 cellular functions. 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?

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 105-106 "Diffusion is…" to "Osmosis is …"
- p 108-110 "Diffusion may..." to end of p 110

The model for the glucose transport protein is very much like the ion transport proteins you studied in Unit 1. 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. Because 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 because the 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. Once the glucose is inside the cytoplasm, it can be oxidized for energy.

### ❖ 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 is a different glucose transporter, a symporter, that uses the $Na^+$ gradient to power glucose uptake. (See fig. 5.15 page 113) Why is this symporter necessary?

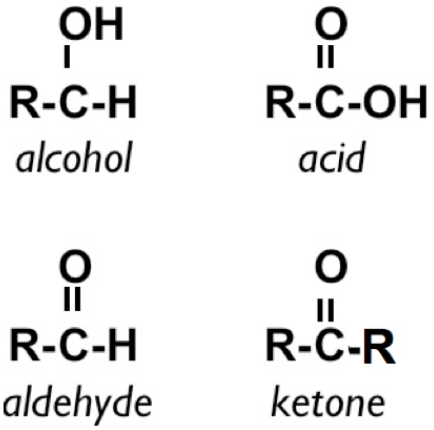
### 🕮 Focused Reading

- p 82 "Mitochondria" to "Plastids"
- p 39-40 "Functional groups…" to "Isomers…"
- p 40 fig 3.1 (Some functional groups…)
- p 142-147 "7.2 What are…" to "7.3 How is…"
  *(Find the step catalyzed by mitochondrial IDH)*

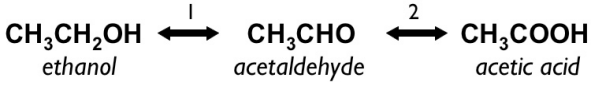
Look over all the steps in glycolysis and the Krebs cycle and try to understand each one. While you do not have to memorize every step, you will understand the overall concepts a lot better if you have some understanding of the individual steps in the process.

General chemical rules that will help you understand glycolysis and the Krebs cycle.

- 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.)




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:



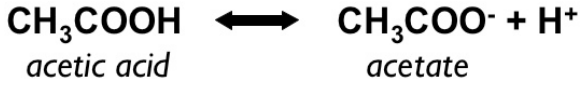

The forward reaction is oxidation while the reverse is reduction. In reaction 1, the oxygen in the **hydroxyl group** breaks its bond with hydrogen and the carbon breaks its bond with one

of its hydrogens, and carbon and oxygen form a double bond. This carbon/oxygen double bond (C=O) is a **carbonyl group**. (If this bond 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 loss of hydrogen is an example of oxidation.

In reaction two, 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:




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 do not 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 spending ATP?

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. [*Science* 271:1233-1234.]

✎ NEWS ITEM: 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. [*Molec Biol Cell* 8:1665.]

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 as well.

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.

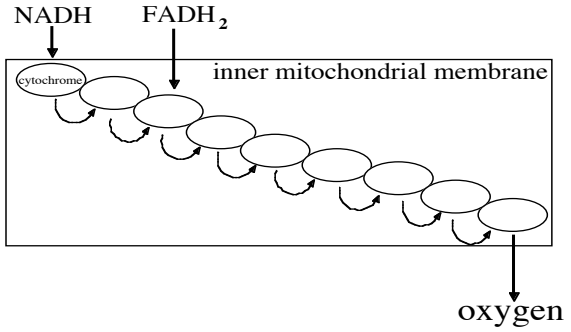
The majority of ATP is synthesized by the cell from the energy stored in the high-energy electrons found in NADH and $FADH_2$. The process of oxidative phosphorylation harvests this energy.

### 🕮 Focused Reading

- p 148 "7.4 How does…" to "The electron…"
- p 150-151 "Proton diffusion…" to "Experiments…"

- Mitochondria section of CancerQuest
  http://www.cancerquest.org/index.cfm?page=44

ATP synthesis in photosynthesis and oxidative phosphorylation are very similar processes. Thus this method of generating ATP must be very ancient, having evolved before plants and animals separated during evolution. In fact, because bacteria also synthesize ATP this way, 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 disease tuberculosis, caused by the bacterium *Mycobacterium tuberculosis*, is reemerging as a global problem because of multi-drug resistance. New antibiotics must be developed to address this issue. An anti-tuberculosis drug currently under development binds and inhibits the *M. tuberculosis* version of ATP synthase, which differs enough from human ATP synthase that normal cellular respiration in the patient would be unaffected. This new drug is effective against both the drug-resistant and drug-sensitive forms of tuberculosis. [*Science* 307: 214]

### ❖ 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 names 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 bodies 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 can get energy from several molecules, neurons 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 reduced 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 conversion 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 two 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. Burning amino acids 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 154-155 "7.6 How are…" to "Metabolic…"
- p 54-55 "3.4 What are..." to "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?

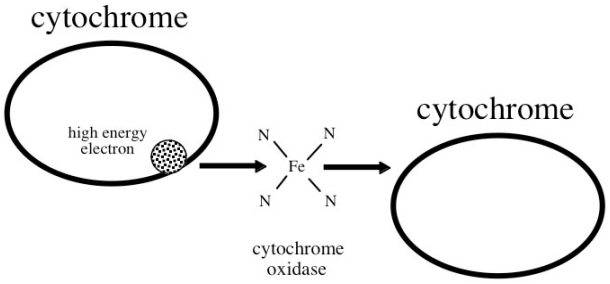
2. How is glycerol burned for fuel? Where does it enter the cellular respiration pathway?

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 do not 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 155-157 "Metabolic…" to "Chapter Summary"
- p 156 fig. 7.19 (Allosteric Regulation of…)

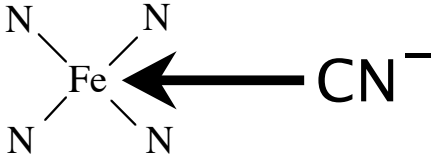
❖ **Study Questions**:
1. Explain in chemical terms how the rate of glucose oxidation is controlled by environmental conditions. Why is this evolutionarily adaptive?

2. Describe the structure of phosphofructokinase. How many binding sites 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?



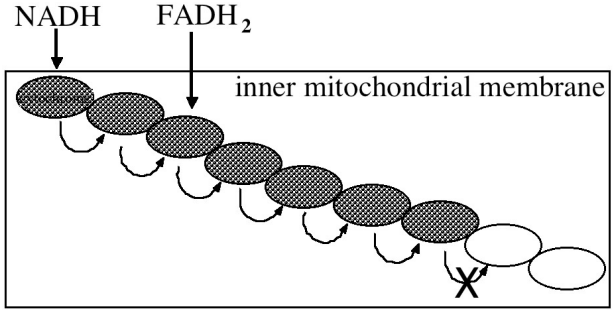

Now that we understand how cells get energy from sugar, we can understand why cyanide is so lethal, and so popular with extortionists. Cyanide (its chemical formula is $CN^-$) has a negative charge, as the name indicates because it ends with the suffix "-ide". As you know from your understanding of basic chemistry, negative ions (anions) 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, cytochromeoxidases 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 the diagram above.

As a result, our very lives depend upon cytochrome oxidases being able to carry high-energy electrons temporarily. Cyanide has the unfortunate ability 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 protons 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? Because 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. In 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 rapidly clogs up the entire electron transport pathway.




Therefore, no protons are transported into the intermembrane space. Without a $H^+$ gradient, there can be no 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 and paraquat action in plants, you should be able to come up with a molecular explanation for how paraquat could also harm human cells. Why do you think that lung cells are most sensitive to paraquat? Why would damage to lung cells be particularly dangerous?

2. How does cyanide kill?

3. **What do photosynthesis and cellular respiration have in common? How do they differ?**

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.**

---

✎ NEWS ITEM: 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 that encodes one subunit of the mitochondrial ATP-synthase. The mothers of each patient contained mixtures of wild type and mutant mitochondrial DNA. [*Biochem Biophys Acta*. 1271:349-357.]

---

### 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.

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.

## Question #3: Why Update Your 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, a brief introduction to the metabolic diversity of bacteria will help broaden your understanding of the variety of ways organisms can acquire energy.**

✎ NEWS ITEM: **Microbes provide us with additional calories.** We know that we can convert the energy present in the food we eat into ATP via the citric acid cycle and cellular respiration. But our cells do not produce enzymes capable of digesting everything that we ingest. What happens to these items? Researchers at Washington University School of Medicine showed in 2006 that bacteria present in our gut act together to provide us with additional calories. By examining previously germ-free mice, researchers showed that mice colonized with two prokaryotes (*Bacteroides thetaiotaomicron* and

*Methanobrevibacter smithii*) gained more weight than mice colonized with either species alone. Apparently, the two species collaborate to increase the calories harvested from certain carbohydrates. [*PNAS* 103: 10011]

**📖 Focused Reading**

- p 565-569 " Prokaryotes…" to "Nitrogen and…"
- p 579 "A small…" to end of p 579

We will focus on one bacterium, *Clostridium tetanii*, the organism that causes tetanus. In the fall of 2005 Hurricane Katrina overwhelmed the levees in New Orleans and the Mississippi River flooded the city, causing billions of dollars in property damage and catastrophic losses for thousands of people. The Red Cross and many other agencies responded to this natural disaster with water, food, clothing, and tetanus vaccines. Why, in the midst of chaos and misery, did the Red Cross spend time and money delivering this vaccine (and what does immunization 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 the diaphragm 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**). LPS serves as an alarm for the immune system, but 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 *Clostridium tetanii* are growing somewhere in the body and is releasing the tetanus toxin, the toxin is carried throughout the body by the blood. When the toxin reaches the nervous system, it binds to and inactivates components within the membranes of motor 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 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 require oxygen (that's why they are called 'cellular respiration'). So how do these anaerobic 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 (two ATP per glucose) and NAD is regenerated so it doesn't build up in the cell.

---

📖 **Focused Reading**

- p 147 "7.3 How is Energy…" to "7.3 Recap"

---

❖ **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 <u>obligate anaerobes</u>, 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 dominant 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 that can survive with or without oxygen. 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 that has had its blood supply interrupted. Because *Clostridium tetanii* lives in soil and the intestinal tract of animals, open wounds that come in contact with dirt or animal feces are particularly susceptible to the development of tetanus. This danger explains why people who work with animals professionally or as a hobby must keep their tetanus vaccinations up to date. The common practice of bleeding a wound, especially a deep puncture wound, is a good one because 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 floodwaters can stir up the tetanus bacteria normally inaccessible in 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 and travels by retrograde axonal transport (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 in synaptic vesicles that facilitates neurotransmitter release. (Remember way back when we talked about neurotransmitter release? It might be helpful to go back and review the process of synaptic vesicle exocytosis now.) By blocking the release of inhibitory neurotransmitters, no muscles get the message to relax and, consequently, 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 non-toxic, denatured 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 that neutralizes the real toxin, should it ever be encountered, before it reaches the nervous system.

✎ NEWS ITEM: Botulinum toxins A and B bind to proteins at your axon terminals. The receptor for toxin A was discovered in 2006. Specifically, toxin B binds to synaptotagmin and gets internalized along with toxin A which is bound to another protein in the synaptic vesicle called SV2. After the vesicle has released its contents, it reforms a vesicle and internalizes anything bound to extracellular surface of the membrane proteins of the vesicle. Once inside the cell, toxin B forms a channel that allows toxin A to enter the nerve's cytoplasm. Toxin A binds to and cleaves SNAP-25 and toxin B cleaves synaptobrevin. It might be possible to create a drug that can block the binding of toxins A and B and thus prevent the potentially lethal effects of food poisoning. [*Science* 312: 592]

❖ **Study Questions:**

1. What kind of paralysis is caused by tetanus toxin? How does tetanus toxin cause paralysis?

2. Are humans capable of anaerobic metabolism? If so, when and where? How?

3. What kind of toxin would cause "limberneck" and why would it be fatal? (see news item below)

---

✎ NEWS ITEM: During the summer of 1997, millions of water birds 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* 278: 1019.]

As you know, we are capable of burning glucose anaerobically for short periods of time. It has been determined that our muscles contain about five millimoles of ATP per kg. 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 causes your muscles burn with extensive exercise. Eventually, this oxygen debt must be repaid so your muscles can return to aerobic metabolism, which is why we are obligate aerobes.

## Question #4: Why Is Tofu A Good Source of Protein?

Tofu is made from soybeans and soybeans are excellent sources of protein. Soybeans harvest the sun's energy, give off oxygen that we gratefully consume, AND store nutrients in its seeds that contain an unusually high concentration of protein. As you know, proteins give us our structure, which allows us to function. Without proteins, we cannot produce any chemical reactions, pump any ions, phosphorylate any substrates, or send any electrical impulses (just 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, meat is an inefficient source of protein. It takes at least ten 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 a notable exception.

**📖 Focused Reading**

- p 788-790 "36.4 How does…" to "Plants and…"
- p 1072 "Food provides…" to end of p 1072
- p 1072 fig. 50.5 (A Strategy for Vegetarians)

A note here about "complete" and "incomplete" proteins: All twenty 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 fig. 50.5 on page 1072). From these eight, you can biosynthesize the other twelve--the **non-essential amino acids**--thus giving you all twenty. [FYI, the essential amino acids are isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, and valine.]

Because most animals (especially vertebrates) are composed of much the same proteins you are, if you eat animal muscles (or milk or eggs), you will automatically take in the right balance of amino acids for your dietary needs. [We'll ignore fats and vitamins for now, but there are essential fats and other molecules that must also be included in our diets.] However, plants are quite different from animals in terms of the relative proportions of the twenty amino acids contained in their fruits, seeds, leaves, and stems. Thus, plants high in protein can have too much of some particular amino acids and too little of others. Eating only wheat, corn, or rice will cause your diet to have "incomplete protein." Therefore, if you are vegetarian, it is good to eat combinations of plants such as legumes (peanuts, garbanzo beans, navy beans, kidney beans, pinto beans, *etc.*) AND grains (wheat, rice, oats, corn, *etc.*) that together provide a "complete" protein mixture. Very few plants produce all of the essential amino acids needed for a human diet. Soybeans and the grain quinoa are two examples.

**❖ 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?

Sugars and lipids contain only carbon, hydrogen, and oxygen. 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: synthetic reactions are endergonic and require an energy source in the form of ATP.) However, proteins (composed of amino acids) contain carbon, hydrogen, oxygen, and nitrogen. Therefore, in order to synthesize amino acids the plant must have a source of nitrogen.

**📖 Focused Reading**

- p 155 fig. 7.17 (Coupling Metabolic Pathways)
- p 176 fig. 8.19 (Metabolic Interactions in a…)
- p 569 "Nitrogen and…" to "26.4 How can…"

The equation for nitrogen fixation catalyzed by nitrogenase is:

$$N_2 + 8H^+ + 8e^- + 16ATP \rightarrow 2NH_3 + H_2 + 16ADP + 16P_i$$

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 a costly endeavor for biological creatures. Nitrogen fixing bacteria contribute about $2 \times 10^8$ 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* microorganism residing in its root nodules. Sucrose is transported to the plant roots where it is converted to $\alpha$-ketoglutarate (you should recall alpha-ketoglutarate

from the IDH labs). Root cells can then synthesize the amino acid glutamate by combining α-ketoglutarate and ammonia.

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 the garden burger served at the Union.

✎ NEWS ITEM: The symbiotic nature of legumes and nitrogen-fixing bacteria allow soybeans to be grown with relatively little fertilizer. Crops such as corn and wheat, however, require farmers to apply fertilizer to their fields regularly. Fertilizer production is an expensive and ecologically unfriendly process. Two international teams independently identified the same gene expressed by legumes that encourages nitrogen-fixing bacteria. One team named the gene *SYMRK* (for symbiosis receptor-like kinase) and the other named it NORK (for nodulation receptor kinase). (Note: it is not unusual for two separate teams of researchers to identify the same gene; eventually one name will be commonly adopted). Both groups show that this receptor kinase initiates an intracellular signaling cascade that leads to nodulation. While this research is an important step in identifying the genes necessary to encourage symbiosis in non-legume crops, NORK/SYMRK is not sufficient to give a non-legume the ability to host nitrogen-fixing bacteria. More molecules in the nodulation signaling pathway remain to be identified. If scientists can genetically modify non-legume crops to develop nodules for hosting nitrogen-fixing bacteria, the need for fertilizer could be significantly reduced. Minimizing fertilizer dependence would not only be an important economic advantage for poor farmers, but would also reduce the ecological impact of fertilizer production and use on our fragile planet. [*Nature* 417: 910 – 11.]

✎ NEWS ITEM: Nitrogen runoff from agricultural fertilizer threatens many ecosystems. The excess nitrogen allows excessive growth of selective plant species, choking out others. For example, in the Gulf of Mexico, algal growth is suffocating other marine life. An international program, the Global Nitrogen Enrichment group, is performing studies and making recommendations on this issue in an attempt to help preserve biodiversity hotspots. [*Nature* 433: 791]

---

❖ **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?

---

# *Unit IV*
# *Current Topics: Cancer, HIV/AIDS, & Genetic Engineering*

## Cancer

In the United States one person in four 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. However, we are still a long way from winning "the war on cancer" that President Nixon declared in the 1970s.

---

📖 **Focused Reading**

- p 386-387 "17.4 What…" to "Some Cancers…"

💻 **Web Reading**

- Cancer Biology Video: Normal Control of Cell…
  http://www.cancerquest.org/videoseries1_english.html
- Cancer Biology Video: Tumor Cell Division
  http://www.cancerquest.org/videoseries1_english.html
- Hallmarks of Cancer: Growing Uncontrollably
  http://insidecancer.org/

---

Cancer is defined as the presence of a **malignant tumor** in the body. Cancer arises from a single cell that is growing out of control. A **neoplasm** (new growth) or **tumor** is a relentlessly growing mass of abnormal cells that are dividing in defiance of normal restraints on growth. Cancers are caused by cells that divide too frequently. 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 and spinal cord -- 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 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**. This spreading process is called **metastasis**. Metastasis is the hallmark of cancerous tumors. While benign and pre-metastatic tumors are relatively easy to cure by surgery and/or localized radiation, metastasized tumors are very difficult to treat because they have spread far and wide throughout the body and require chemotherapy. In order to cure metastasized cancer, every single cancerous cell in the body must be destroyed. Destroying every cancerous cell is virtually impossible in widely disseminated cancers.

The type of cell that becomes cancerous determines the name of the cancer. Names can be a bit tricky because tumors are often classified according to the original site of the cancer (even if it has spread). For example Tour de France champion Lance Armstrong had testicular cancer that reached his brain. Animal tissues come in four varieties: 1) Epithelia lines the inside and outside surfaces of the body (*e.g.,* skin, lungs, blood vessels, stomach, intestine lining, *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, that 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 (glial cells).

**📖 Focused Reading**

- p 857-858 "Physiological…" to "Connective…"
- p 860 fig. 40.7 (Tissues from Organs)

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: basal cell carcinomas being very easily treated and are rarely fatal while melanomas are 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.

At right you see 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 lung cancer. 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, lung cancer is by far the most preventable. The vast majority of lung cancer patients smoked cigarettes and/or lived with a heavy smoker. Like AIDS, lung cancer is a preventable disease that continues to claim lives needlessly. Every year, lung cancer kills three times as many people as died in the Vietnam War. When you compare the incidence of lung cancer worldwide you will notice that the highest rates occur in developed countries where people can afford the luxury of cigarettes.

| Site of Cancer | New Cases in 2006 (both sexes) | Deaths in 2006 (both sexes) | Male Deaths in 2006 | Female Deaths in 2006 |
|---|---|---|---|---|
| *All sites* | *1,399,790* | *564,830* | *291,270* | *273,560* |
| Oral Cavity | 30,990 | 7,430 | 5,050 | 2,380 |
| Digestive System (incl. colon) | 263,060 | 136,180 | 75,210 | 60,970 |
| Respiratory System (incl. lung) | 186,370 | 167,050 | 98,820 | 73,230 |
| Bones & Joints | 2,760 | 1,260 | 730 | 530 |
| Soft Tissue | 9,530 | 3,500 | 1,830 | 1,670 |
| Skin | 68,780 | 10,710 | 6,990 | 3,720 |
| Breast | 214,640 | 41,430 | 460 | 4,0970 |
| Genital System (incl. prostate) | 321,490 | 56,060 | 28,000 | 28,060 |
| Urinary System | 102,740 | 26,670 | 17,530 | 9,140 |
| Eye & Orbit | 2,360 | 230 | 110 | 120 |
| Brain & Nerv. Syst. | 18,820 | 12,820 | 7,260 | 5,560 |
| Endocrine System | 32,260 | 2,290 | 1,020 | 1,270 |
| Lymphoma | 6,6670 | 20,330 | 10,770 | 9,560 |
| Myeloma | 16,570 | 11,310 | 5,680 | 5,630 |
| Leukemia | 35,070 | 22,280 | 12,470 | 9,810 |
| Other | 27,680 | 45,280 | 24,340 | 20,940 |

[Adapted from American Cancer Soc. Cancer Facts & Figures 2006]

You probably know that sunlight exposure causes most skin cancers and that fair-skinned people are more susceptible. There are other environmental factors that contribute to different cancer rates in different countries. Liver cancer is correlated with Hepatitus B viral infections and aflatoxin, a mutagen released by mold on peanuts and other foods that have been stored improperly. Liver cancer is very rare in the US, but more frequent in Mongolia, China, Africa, South America, the former Soviet Union, and some Western European countries. Stomach cancer is strongly influenced by diet and most prevalent in Asia. It is thought that the pickled and smoked foods common in Asian diets contribute to stomach cancer. Breast cancer is most common in the US and Canada and other affluent areas. A woman's lifetime exposure to estrogen is related to her breast cancer risk. Girls reach puberty earlier if they have access to good nutrition. In addition, a high fat diet may also contribute to breast cancer. Similarly, colon and rectal cancers are associated with high fat, low fiber diets in affluent nations. As

colonoscopies become more routine, deaths from colon cancer are declining. Cervical cancer is associated with sexually transmitted human papillomavirus (HPV) infections and poor health care for women. Routine pap smears can detect precancerous cells, making it relatively easy to detect in early stages. Finally, most cancers are more frequent in men than in women. (A map of various cancer rates around the world is at insidecancer.org--see the overview under causes and prevention.)

✎ NEWS ITEM: A New Cancer Vaccine. In 2006, the FDA approved the first vaccine for the prevention of a specific type of cancer. In a study of 21,000 women, the vaccine, being marketed by the drug company Merck under the name Gardasil, was nearly 100% effective at preventing infection by human papillomavirus (HPV) 16 and 18. HPV is the most common sexually transmitted disease in the US, with an estimated 6.2 million Americans becoming infected each year. HPV strains 16 and 18 are responsible for approximately 70% of all cervical cancer cases. According to Susan Crosby, president of Women in Government, "The FDA's decision marks an historic milestone in the fight against cervical cancer and should be celebrated by women and health advocates around the world."

To begin looking at the causes of cancer, here are a few things we know from observations:

1. Cancers tend to run in families. Very few cancers, however, demonstrate Mendelian inheritance ratios indicating that most are heavily influenced by non-genetic factors. Thus, cancer is not usually inherited in an obvious dominant or recessive fashion (with a few, rare exceptions such as retinoblastoma).

2. Exposure to certain environmental agents (chemicals, irradiation, *etc.*) is associated with the development of cancer. Any agents demonstrated to cause cancer are termed carcinogens.

3. If we perform the Ames test for mutagenicity, we find that most 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:
   i) They divide too frequently
   ii) They leave their normal tissues and take up residence in areas of the body that are completely foreign to them.

At a minimum, based on this information, we should be able to hypothesize that:

A) Cancer is caused, or enhanced, by changes in the DNA that may be
   a. inherited mutations (because predisposition for cancer runs in families); and/or
   b. new mutations (because carcinogens cause mutations in the DNA)

B) 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

Obviously cells need to divide – and at certain times and places cells need to divide rapidly. Cells in a recently fertilized egg can go through mitosis as fast as every 15 minutes. All cells (except red blood cells) retain the genetic instructions for making the proteins necessary for cell division. The expression of proteins that control cell division is strictly controlled by the cell's age and environment. The decision for a cell to divide is not made lightly – many molecules regulate this process.

The genetic changes could occur in the molecules that control communication concerning where and when to divide. The mutations could occur 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 signals that control cell division in normal cells.

---

🕮 **Focused Reading**
- p 184-187 "9.2 How is…" to "9.3 What happens…"
- p 184 fig. 9.3 (The Eukaryotic Cell Cycle)
- p 186 fig. 9.6 (Cyclin-Dependent Kinases)

---

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. There are lots of picky details in setting up the system but

tissue culture is a great way to understand how a cell really works without dealing with an entire pesky organism! 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 progresses 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 that arrest cell division at a given stage. As each cell enters this stage, it gets stuck there. Because 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 bother 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 the 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 **$G_o$**--get it a cell must pass 'go.') To test the hypothesis, scientists 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." "Double cells" have 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 a soluble factor in the cytosol for the rest of the cycle or if the factor was destroyed after S. So they fused $G_2$ cells with $G_1$ cells. This fusion did <u>not</u> 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 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 that 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 cells 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.

Many yeast cell cycle genes are very similar to vertebrate cell cycle genes, suggesting that cell division is a very old idea that evolution has preserved. The investigation of the cell cycle in wild type and mutant yeast is a powerful tool to investigate the cell cycle. The 2001 Nobel Prize in Medicine was awarded to Lee Hartwell, Paul Nurse, and Tim Hunt for their work using yeast to determine key regulators of the cell cycle. Following in the footsteps of researchers like these many cancer researchers study yeast because many different mutant yeast strains are deficient in

different proteins required at different stages of the cell cycle. These yeast strains are called **cell-division cycle (cdc) mutants**. Therefore, by determining which protein a given yeast strain is missing, and correlating the protein with the stage of cell division that is eliminated or dysfunctional in that strain of yeast, the role of various proteins in the process of cell division can be determined. Thus, understanding single celled organisms has provided very important information even for complicated diseases in multicellular organisms.

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**, small proteins that act as extracellular ligands to stimulate cell division. The following is a list of some of the major growth factors and the types of cells that responds to each.

| Growth 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 (IGF)I and II | Collaborates with PDGF & EGF; simulates connective tissue cell division |
| Transforming growth factor ß (TGF-ß) | Increases cell sensitivity to other growth factors; controls differentiation |
| Fibroblast growth factor (FGF) (>20 types) | Stimulates cell division in many cell types |
| Interleukin-2 (IL2) | Stimulates cell division in T lymphocytes |
| Nerve growth factor (NGF) | Allows neurons to survive and differentiate |
| 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 cells use to produce two genetically identical daughter cells during cell division. While two daughter 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 the 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.
   - Cell synchronization in culture
   - Cell fusion
   - 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?

What do we know now about cell division that will help us figure out what causes cancer?

A. Cell division is carefully synchronized and controlled by many proteins (that obviously are encoded by genes).

B. 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 signal transduction system mediated by proteins that are ultimately controlled by genes.

External signals include:

a. the presence of adequate nutrients
b. the presence of specific growth factors (in some cases)
c. the degree of contact with neighboring cells (how "crowded" the cells "feel")
d. 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 place)

A defect in any of these processes that control when and where cells divide may cause a cancerous transformation. And because cancers arise in various tissues or organs have very distinct characteristics, cancers generally have very unique causes.

What experimental 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 identified, the **Rous sarcoma virus**, (discovered by Dr. Rous) causes connective tissue tumors in chickens. Several other tumor viruses have also been identified and characterized.

Note that the only way you can be sure that a virus causes cancer is to inject the virus into an organism and then evaluate the tumors that grow in these experimental animals (including proper control injections of non-cancerous viruses of course). We have no way of definitively determining if the human viruses listed above actually cause or contribute to the development of cancer because scientists obviously cannot ethically inject suspected cancer viruses into human subjects. Further, viruses are species specific, that is most viruses only infect one type of animal. However, we do know that people who are infected with the viruses listed in the table below have an increased likelihood of developing certain types of tumors. On the other hand, infection with these viruses does not guarantee the development of a tumor -- infection only increases the likelihood.

| Virus | Species | Tumor |
|---|---|---|
| Rous Sarcoma | Chicken | Connective Tissue |
| FBJ osteosarcoma | Mouse | Bone |
| Simian sarcoma | Monkey | Connective Tissue |
| Abelson murine leukemia | Mouse | Leukemia |
| Avian erythroblastosis | Chicken | Bone Marrow |
| Harvey murine sarcoma | Mouse | Connective Tissue |
| MC29 myelocytomatosis | Chicken | Bone Marrow |

| Human Virus | Tumor |
|---|---|
| Papillomavirus (HPV) | 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 | Kaposi's Sarcoma (AIDS-related opportunistic infection) |

Human tumor viruses contribute only minimally to the overall incidence of cancer in humans. However, tumor viruses have been exceptionally important to 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 cells have been transformed into tumor cells.) 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 normal cells in the process of transformation that makes them divide inappropriately? Well, the answer 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 weak spots in a garden hose.)
   C. Increased mobility of the plasma membrane proteins
II. Adherence abnormalities
   A. Diminished adhesion to substrates and other cells
   B. Disorganization of the cytoskeleton
   C. High production of proteases 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 ability to divide when unattached 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 plastic petri 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 culture media (with sugars, amino acids, *etc.*), they will continue to divide indefinitely. We would have absolutely no trouble filling Dana Science Building with the offspring of just one, well-fed tumor cell in a surprisingly short period of time.

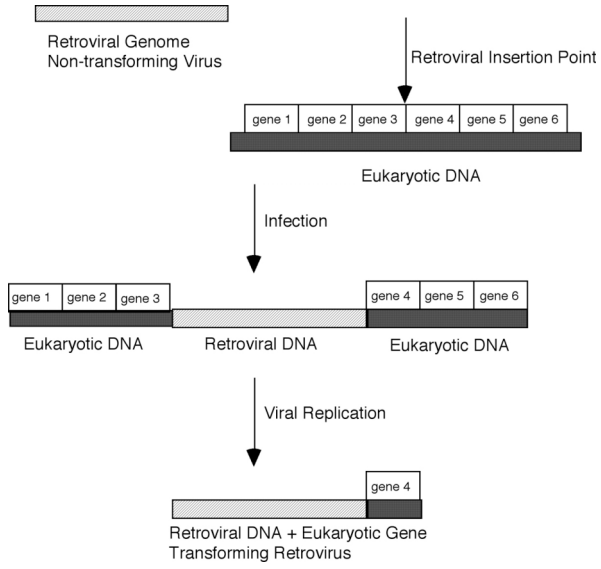
A puzzle for fun:
How long would it take to fill Watson Science Building starting with a single cell? Assume the following parameters:
Average cell volume = 125 pL;
1 picoliter = 0.001 nanoliter, and 1 nL = 0.001 μL
Watson Science Building volume =~50,000 $m^3$
Average cell cycle = 1 division every 12 hours

**Focused Reading**
- p 283-284 "13.1 How do…" to "Bacteriophages..."
- p 288 fig. 13.6 (The Reproductive Cycle of HIV) *as example of a retrovirus*
- p 387 "Some cancers…" to "Most cancers…"
- p 387 Table 17.4 (Human Cancers Known…)

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** because retroviruses actually insert viral genes into the cell's genome that are then passed to the next generation of cell. Thus viral 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 RSV strain that seemed like a perfectly competent virus (it was able to infect cells, insert its DNA, and induce production of new viruses), but did not transform the cells. When the investigators looked for the difference between this non-transforming RSV and the transforming version of the virus they found the non transforming RSV 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 question, investigators created a radioactive DNA probe that was complementary to the *src* gene and probed the DNA of normal chicken 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 similarity. 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 Nobel laureates) first characterized 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" for "viral") and the eukaryotic cellular version of the gene ***c-src*** ("c" 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 100 oncogenes have been identified overall. It is worth mentioning that the src protein is a kinase that often phosphorylates growth factor receptors. The viral kinase is about 20 times more active than the proto-oncogene cellular kinase, which helps explain why some viruses can contribute to cancers.




As an aside, you might be wondering why a virus would contain a gene that causes cancer. These viral oncogenes do not appear to confer any survival value whatsoever to the virus. In the case of a retrovirus, the virus' direct ancestor probably picked up the src gene from an animal 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 adjacent human DNA, the viral genome will contain a copy of the host's gene. It is thought 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 valuable approaches 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, tumor viruses are responsible for only in a few cases of animal cancer. However, these viral oncogenes have led us to their normal gene counterparts, the cellular proto-oncogenes. It is thought that most cancers are caused when these normal proto-oncogenes become mutated in a manner that increases their ability to stimulate cell division, thus becoming oncogenes.

Investigators think 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 S phase

While most cancers are caused by mutations, not all mutations cause cancer. Most mutations probably do not cause cancer. Some mutations for instance, cause cystic fibrosis, or color-blindness, or a predisposition to heart disease. Remember that some mutations have no influence on a protein's ability to do its job if the mutation is silent, a conservative amino acid substitution, occurs in a non-coding or non-regulatory region of

the genome, or involves an amino acid substitution in a non-critical region of the protein. It is assumed that what distinguishes cancer-causing mutations from other mutations is that cancer-causing mutations occur in proto-oncogenes. A proto-oncogene is a normal gene that directly or indirectly plays a role in regulating cell division. Thus, investigators have focused intensively on understanding proto-oncogenes.

❖ **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 demonstrate that viruses cause cancer in humans.
2. What is a tumor virus? What is cellular transformation? Tumor viruses do not cause most human tumors. Explain why tumor viruses have been the focus of such intensive research efforts even though viruses are not the predominant cause of most cancers. 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 cultured cells, 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 non-carcinogenic mutations. What are these agents? How do carcinogenic mutations differ from non-carcinogenic 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 a gene name is preceded by a "v?" By a "c"?

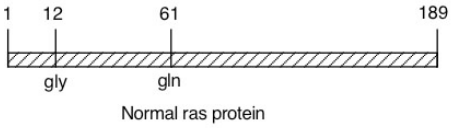
In order to illustrate some of the normal functions of proto-oncogenes, let's look at some specific examples of proto-oncogenes that have been fairly well characterized. First, let's look at the *ras* proto-oncogene (first identified in a rat sarcoma). The proto-oncogene *ras* encodes a G-protein. (Remember G-proteins?) G-proteins transduce signals from transmembrane receptors (for a hormone or, in this case, probably a growth factor) and adenylate cyclase (AC) or phospholipase C (PLC).

📖 **Focused Review Reading**

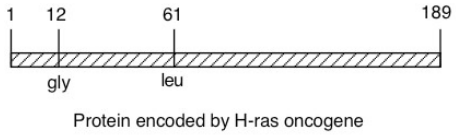
- p 336 "How do signal…" to "Receptors can…"
- p 339 "Protein kinase…" to "Second messengers…"
- p 341 fig. 15.10 (A protein kinase cascade)

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.

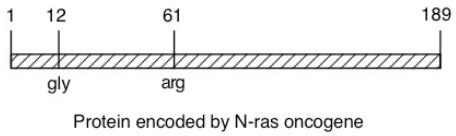
Turn to page 264 in your textbook and compare the amino acids in the different versions of the ras protein. H-ras has a leucine instead of a glutamine at position 61. Leucine is non-polar while glutamine is polar. This mutation could change the folding pattern of the molecule significantly. The other two mutations (glutamine to arginine in N-ras: and glycine to arginine in K-ras) also change the characteristics of the amino acid significantly.



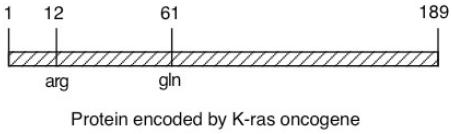
Normal ras protein



Protein encoded by H-ras oncogene



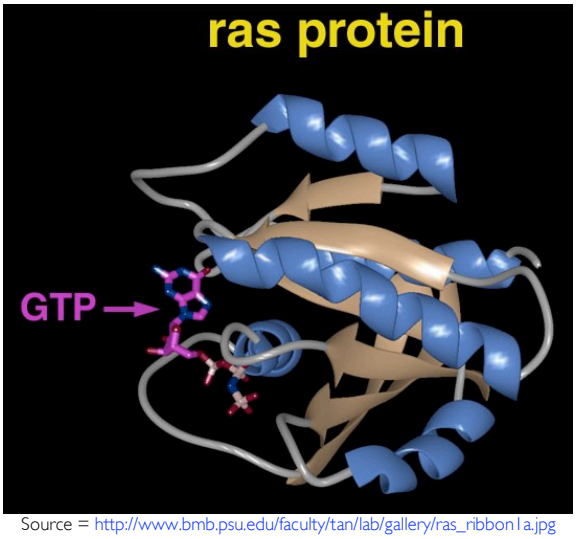
Protein encoded by N-ras oncogene



Protein encoded by K-ras oncogene

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 mutant 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 division.

The drawing below is an illustration of a computer-generated structure for the normal *ras* G-protein. The molecule changes shape when GTP is hydrolyzed into GDP. 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. A signal transduction molecule that is always on is called **constitutively active**. Again we see that protein structure is very important in determining protein function.



Source = http://www.bmb.psu.edu/faculty/tan/lab/gallery/ras_ribbon1a.jpg

Further evidence of the linkage between the ras proteins, growth factors, and control of cell division comes from intracellular antibody binding studies. Normal cells will divide when growth factor signals are received by their receptors. When anti-ras antibodies (against the normal version of the protein) are injected into the cytoplasm of normal cells bathed in growth factors, the antibodies will bind to the ras proteins and prevent ras from transducing the growth factor signal. Consequently, these antibody-treated cells are unable to divide in response to growth factors. Thus, this experiment demonstrates that the ras protein forms an important 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 cause cancer?

3. Describe the intracellular antibody binding studies that link the ras protein to the response of the cell to growth factors.

| Proto-oncogene | Type of Protein Product | Specific Protein Product |
|---|---|---|
| *sis* | Growth Factor (GF) | Platelet-Derived Growth Factor (PDGF) |
| *fms* | GF Receptor | Colony-Stimulating Factor-1 Receptor |
| *erbB* | GF Receptor | Epidermal Growth Factor Receptor (EGFR) |
| *neu* | GF Receptor | Protein with similar structure to EGFR |
| *erbA* | GF Receptor | Thyroid Hormone Receptor |
| *src* | Signal Transducer | Tyrosine kinase, required for entry into G2 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) |

To illustrate the types of growth-related proteins that can be altered in cancerous changes, above 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.

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 times, and is repeated over and over again in tandem sequences. Each gene is active, and therefore, the protein product of such gene amplification is overexpressed and therefore overstimulates 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 because 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 four types of protein products that proto-oncogenes are known to encode. Give an example of each.

✎ NEWS ITEM: Small transcripts called micro RNAs (miRNAs) are synthesized from many regions of the genome but do not encode proteins. Instead, they regulate other genes by helping to trigger breakdown of mRNAs that contain similar sequences. This represents a huge emerging body of knowledge about gene regulation. Mutations in some miRNAs have been shown to be associated with cancer. For example, the ras gene is regulated by the miRNA let-7; in some tumors, let-7 expression is too low, and so this miRNA can't turn down the expression of Ras; extra Ras protein enhances the cancerous phenotype. [*Cell* 120: 635.]

### 🕮 Focused Reading

- p 387-390 "Most cancers…" to end of p 390

### 💻 Web Reading

- Cancer Biology Videos: Genes II: Oncogenes
  http://www.cancerquest.org/videoseries1_english.html
- Cancer Biology Video: Genes III: Tumor Supressor…
  http://www.cancerquest.org/videoseries1_english.html

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 fig. 17.19 on page 390 for the steps required for colorectal cancer, the cancer 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.) Frequently 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?

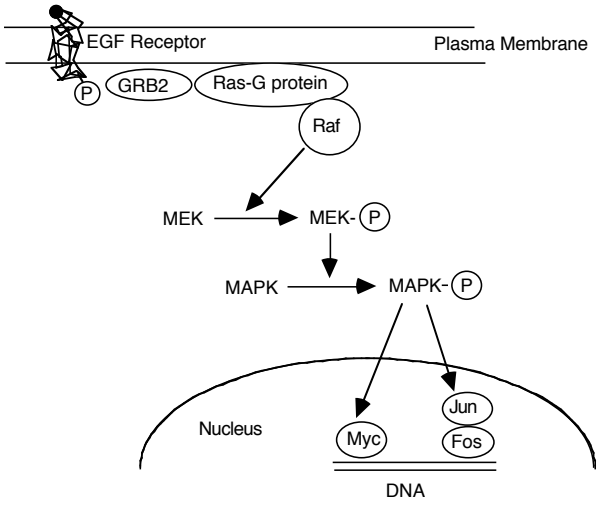
2. What is wrong with the comment, "Some day we may find *the* cure for cancer"?

✎ NEWS ITEM: A group at the Scripps Research Institute in La Jolla, CA have synthesized a bacterial compound called Epothilone A that 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. [*Science* 274: 2009.]

### Web Reading

- MAPK Signal Transduction
  www.bio.davidson.edu/courses/Immunology/Flash/MAPK.html

In a flurry of scientific papers recently, investigators have outlined the relationship between the *ras* protein and cell division. By examining this pathway, you should be able to get a clearer idea of the link between the development of cancer and changes in G-proteins, tyro sine 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 (EGFRs.) These receptors are membrane-bound tyrosine kinases that work in pairs (two EGFRs). When EGF binds to its receptor, one EGFR **phosphorylates**, that is, it adds a phosphate group to the partner's tyrosine residues. This phosphorylation 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, 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.

Thus proteins in the signaling pathway initiated by EGF are all products of proto-oncogenes. Changing any one of these proteins in a way that hyperactivates the protein could cause abnormally high cell division signals, thus producing a cancer cell dividing out of control.

✎ NEWS ITEM: In the summer of 2006 researchers at UC Berkeley and Johns Hopkins learned important information about how EGFRs are activated. It had been known for a long time that when EGF (ligand) binds EGFRs the recptors dimerize (pair up) and that this dimerization activates the, kinase portion of the EGFRs, but it wasn't well understood how this happens. Researches now know that the kinase region of the EGFR is normally in the off state (inactive). Mutations that active EGFRs in cancer patients have kinase domains that are 20X more active. Using x-ray crystallography to examine EGFR structures, the researchers now propose that an activated EGFR in close proximity to an inactive EGFR can switch on the inactive receptor. This activation method was more complex than what the scientists expected. They propose that such specialized and complicated signal transduction mechanisms allows evolutionary fine-turning so that different combinations of EGFRs can transmit a wide variety of signals. This idea may also explain why EGFRs play powerful roles in a diverse array of important cellular processes such as cell division, maturation, and movement. [*Cell* 125:1137]

### ❖ 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) altered the EGF-receptor to become constitutively active
B) altered the MAPK constitutively active
C) altered the transcription factors constitutively active
D) all of the above
E) none of the above

Be able to explain your answer.

---

✎ NEWS ITEM: A new finding by Pascal Goldschmidt-Claremont from Ohio State Univ. suggests that *ras* also uses superoxides (an oxygen molecule with an extra electron – free radical) 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. [*Science* 275: 1567-68.]

✎ NEWS ITEM: The National Cancer Institute has a 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. http://www.ncbi.nlm.nih.gov/ncicgap/

✎ NEWS ITEM: Many people think that the hottest area in cancer research focuses on an enzyme called telomerase. Telomerase is the DNA polymerase that replicates the telomeres (tips) 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.

✎ NEWS ITEM: A new type of cancer-causing mutation was found. 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. [*Science* 277: 1201.]

In focusing on mutations in the genes that control cell division, we shouldn't forget about the second criterion for malignancy, the ability to metastasize. In order to spread, cancer cells must be able to break free from the tissue 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 adult body that can normally move all about the body. (During embryonic development many cells make long and complicated movements – a time when many cells are also dividing rapidly. Thus, understanding the basic mechanisms of how cells move and divide is important to both cancer biologists and developmental biologists.) Most adult cells do not move; they just sit in their place doing their job. Metastasis requires changes in cellular motility based on changes 1) in the cytoskeleton and 2) in the secretory products of the cell because cells have to digest their way across barriers. All these changes are caused by mutations (inherited or new) in genes controlling the cytoskeleton, enzyme secretion, cell adhesion, and receptors that allow information to be exchanged between cells. Thus, even if an individual cell acquires an oncogene mutation and loses function of some tumor suppressor genes, these combined mutations may not be enough to form a dangerous cancer. Cancers are most life threatening when cells that have lost control of the cell cycle also gain the ability to leave the original tumor and invade new tissues, forming more tumors. Such cancer cells must acquire further mutations that allow them to metastasize, which makes cancer a very difficult disease to understand and treat.

✎ NEWS ITEM: 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 to which cells often attach. The 'laminase' may be a target to block metastasis. [*Science* 277: 225.]

✎ NEWS ITEM: We know that tumors become much more difficult, if not impossible, to treat after 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 to contain cancer.

---

### 💻 Web Reading

- Cancer Biology Videos: Angiogenesis
  http://www.cancerquest.org/videoseries1_english.html
- Hallmarks of Cancer: Processing Nutrients
  http://insidecancer.org/

Some tumors grow to large sizes because they are very good at recruiting new blood vessel formation. Blood vessels bring important nutrients to tumors and clear out wastes. Cancer cells that do not get enough $O_2$ delivered or $CO_2$ picked up may not be able to produce enough energy to survive. The formation of blood vessels is called **angiogenesis**. Many research labs are trying to understand how blood vessels develop and find drugs to block angiogenesis.

**Web Reading**

- Cancer Biology Videos: Cell Death Via Apoptosis
  http://www.cancerquest.org/videoseries1_english.html
- Hallmarks of Cancer: Evading Death
  http://insidecancer.org/

Healthy cells are usually quite good at detecting when things are going wrong due to abnormalities such as mutations. When a cell detects such an abnormality it undergoes a process called programmed cell death or **apoptosis**, essentially a form of cell suicide. Many cancer cells have deficiencies in apoptosis such that mutated cells are not removed by apoptosis and then may begin to divide and propagate the mutation, forming a tumor. As you might expect, there are many labs investigating the molecular and cellular mechanisms of apoptosis that may lead to treatments for cancer and other diseases.

**❖ Study Question:**

Test your understanding of experimental design, oncogenes, and tumor suppressors via case studies at the Cancer Cell Biology web site (www.ibiblio.org/pmabs/cancer.html).

(Note: you do not need to submit verification to your professor, but you will have to provide your last name so the computer can address you properly.)

✎ NEWS ITEM: The human *IDH1* gene (encoding a version of isocitrate dehydrogenase, which we studied in lab) appears to be a tumor suppressor! In gliomas (cancers of the supporting cells of the brain), mutations in *IDH1* cause a redution in the amount of $\alpha$-ketoglutarate (why?). Without enough $\alpha$-ketoglutarate, the transcription factor HIF1-$\alpha$ is more stable. This transcription factor then turns up the expression levels of genes that contribute to tumor formation. [Zhao et al, 2009, *Science* 324:261]

## HIV & AIDS

**Overview Reading**

- Chapter 12 • From DNA to Protein
- Chapter 13 • The Genetics of Viruses and…
- Chapter 16 • Recombinant DNA & Biotech…
- Chapter 17 • Genome Sequencing, Molecular…
- Chapter 18 • Immunology: Gene Expression…

**Focused Reading**

- p 261 "RNA viruses…" to "12.3 How is the…"

**Some Definitions**

Since its identification in 1981, Acquired Immune Deficiency Syndrome (AIDS), a preventable but uncurable sexually transmitted disease (STD), has claimed the lives of over 28 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 asymptomatic period is called the **clinical latency period** and it is one reason that HIV is so dangerous. People can carry, and spread, the virus for many years without having any symptoms of the disease to inform 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 2001, an estimated 900,000 people (0.6% of the population) are living with HIV/AIDS. Roughly 10,000 of these people are children under the age of 15 and 180,000 are women. Since 1981, approximately 500,000 people in the US have died of AIDS; 15,000 AIDS deaths occurred in 2001. Recent statistics show that women account for an increasing proportion of people with HIV/AIDS in the US. Also, various racial/ethnic groups are disproportionately affected by HIV/AIDS. According to a recent report by UNAIDS, 62% of men recently diagnosed with HIV/AIDS are non-Hispanic black or Hispanic and 81% of women recently diagnosed are non-Hispanic black or Hispanic. This report also estimates that 41% of recent infections in the US resulted from male-to-male sexual contact, 22% resulted from heterosexual contact, and 30% resulted from injection drug use.

Worldwide, over 42 million people are living with HIV/AIDS and 25 million people have died of this disease since 1981. Half of the people living with HIV/AIDS are women and three million are children under the age of 15. 50% of new infections are in people between the ages of 15 and 24. In 2001, there were five million new infections (14,000 per day) and three million deaths. An estimated 14 million children are classified as AIDS orphans, indicating that their parents have died of AIDS. The region of the world hardest hit by the HIV/AIDS pandemic is sub-Saharan Africa. Roughly 30 million people with HIV/AIDS are living in this region. In several countries, over 30% of the population is HIV+ (Botswana: 38.8%; Lesotho: 31%; Swaziland: 33.4%; Zimbabwe: 33.7%). Another 5.6 million people with HIV/AIDS live in South and Southeast Asia. Currently, Eastern Europe is experiencing the fastest growing epidemic in the world.

HIV is spread when bodily fluids containing the virus contact the blood of an uninfected individual. The bodily 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 (*e.g.,* mouth, rectum, and vagina). The breach can be microscopic -- well below the size one would detect normally. High-risk behaviors include sharing needles during injection 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. Similarly, vaginal intercourse also presents a high risk 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. 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. Unprotected oral sex represents a potential risk of infection, but is safer than unprotected anal or vaginal sex. Receiving oral sex is considered to be very low risk. Giving oral sex (to a man or woman) represents a greater risk, but still represents an inefficient means of transmission. The virus also is present in fairly high levels in breast milk, resulting in frequent transmission of the virus from mother to child during breast-feeding. 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. Of course, relatively simple steps can be taken to reduce the chance of infection. The use of clean needles, latex gloves, condoms, and dental dams dramatically reduce transmission of HIV.

In $HIV^+$ individuals, the virus can be detected in saliva, tears, and urine. Because of the low levels of virus in these fluids, and other inactivating agents present in these fluids, transmission via these fluids is extremely unlikely. Thus, touching, hugging, kissing, and sharing utensils are not risky activities. Also, HIV cannot be transmitted by insects, according to epidemiological data. HIV, unlike West Nile virus, does not replicate in insects. It should be noted that AIDS is <u>not</u> a "gay disease." That common misconception arose because in the United States the disease initially spread most rapidly in the gay male population.

Today, however, all sectors of society are affected. HIV makes absolutely no distinctions based on gender or sexual orientation. Heterosexual women make up one of the fastest growing HIV+ subgroups in the US. Worldwide, 50% of people living with HIV/AIDS are women. Any type of risky behavior, by anyone, increases your risk of contracting HIV/AIDS.

While HIV/AIDS is a major health threat, it is a preventable disease. We do not yet have a vaccine to prevent HIV infection, but 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 tissue abrasions too small to detect, it is simply not enough to make sure that the blood or semen of another does not come in contact with an open wound. Rather, only complete protection from contact with the blood or semen of another person will guarantee safety. As mentioned earlier, the proper use of latex gloves, condoms, and dental dams offer excellent protection. It should be noted, though, that oil-based lubricants can decrease dramatically the strength of latex condoms. Also, recent studies indicate that spermicides, such as Nonoxyl-9, may increase the risk of transmission. Numerous studies have shown that needle exchange programs can reduce the rate of transmission of HIV among injection drug users (IDUs). Despite the proven effectiveness of condoms and needle exchange programs in reducing the rate of transmission of HIV, many people (including government officials!) in the US oppose providing condoms to teenagers or clean needles to IDUs.

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.

**United States HIV statistics**

| | 2002 | 2003 | 2004 | 2005 | 2006 |
|---|---|---|---|---|---|
| AIDS diagnoses | 38,132 | 38,538 | 37,726 | 36,552 | 36,828 |
| Deaths of persons with AIDS | 16,948 | 16,690 | 16,395 | 16,268 | 14,016 |
| Persons living with AIDS | 350,419 | 372,267 | 393,598 | 413,882 | 436,693 |

(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 two to three Davidson students (student population ≈1,700) to be infected with the virus. Why are there relatively high levels of HIV infections among college students? In a recent survey of 8500 college students, only 43% of the students reported that they always use condoms during sex. Despite knowing how to protect themselves, many students, it appears, do not protect themselves.

There is an especially disturbing trend in the rise of HIV/AIDS among young people in the United

States. In 2006, about 15% of newly diagnosed HIV patients were 24 years old or younger. In addition, 25% of new diagnoses in 2006 were among people in the 25-34 age group, many of whom were initially infected when they were in high school or college. The Centers for Disease Control reports that the rise in HIV infection in young people is due to "lack of recognition of their partners' risk factors" and lack of concern about being infected, leading to insufficient use of condom A recent newscast reported that new infections in teenage girls are 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 is rising.

## 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 all remember the news about the first cases and the drama that surrounded identifying HIV.

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 usually are 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 determine what all these men had in common. Four characteristics emerged, which were called the "Four Hs" -- being a male Homosexual, Haitian, IV drug user (Heroin), 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 and/or semen 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 later was 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 and of itself, is not a risk factor.

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 accusations of foul play. Now, Luc Montagnier, from the Pasteur Institute in Paris, and Robert Gallo of the National Institutes of Health (USA) officially share credit for the discovery of HIV as the causative agent for AIDS. A legal battle over patent rights has been settled only recently.

Note: For a very good history of the HIV/AIDS epidemic, you may be interested in reading *And The Band Played On*, by Randy Shilts (or watching the film adaptation). Also, *Science Fictions: A Scientific Mystery, a Massive Cover-up, and the Dark Legacy of Robert Gallo*, by John Crewson and *Virus Hunting: AIDS, Cancer, and the Human Retrovirus: A Story of Scientific Discovery*, by Robert Gallo provide very different accounts of the discovery of HIV.

### ❖ 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 clinical 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 was the disease originally so prevalent among gay men in the United States?

5. How can the spread of AIDS be prevented?

---

**Focused Reading**

- p 283-289 "13.1 How do…" to "Many plant…"
- p 288 fig. 13.6 (The Reproductive Cycle of…)
- p 421-423 "AIDS is an…" to end of chapter

**Web Reading**

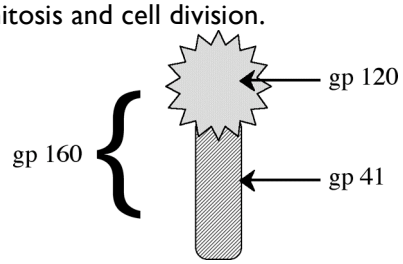
- Life Cycle of HIV Attachment
  www.bio.davidson.edu/courses/HIVcellsalive/hiv0.htm

---

**Structure of HIV**

From your focused reading, you can see that 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). Fig. 13.6 (page 288) 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 in the diagram below) and **gp41** (the stalk in the diagram below). gp120 is the protein that specifically binds CD4 on human cells, allowing attachment and infection (more about CD4 in a few paragraphs). The genome consists of two identical strands of ssRNA (single stranded RNA), which contain nine genes. Each ssRNA strand is bound to a molecule of **reverse transcriptase (RT)**, a viral enzyme with three separate functions required to convert the viral ssRNA into dsDNA. Because human cells never execute this type of conversion, they do not contain RT; therefore every virus particle must contain RT. The two identical copies of viral ssRNA are **reverse transcribed** (the first function of RT) into an RNA:DNA heteroduplex. The RNA portion of this heteroduplex then is hydrolyzed by RT (the second function of RT) and the resulting ssDNA is converted into dsDNA (the third function of RT). The viral dsDNA then is transported into the nucleus and inserted into the host genome by another viral enzyme, integrase. Once the viral DNA has become integrated, it is a permanent part of the host cell's genome. Thus, the viral DNA is transmitted to progeny cells during mitosis and cell division.




Following integration, the viral DNA can be transcribed by cellular RNA polymerases. These new strands of viral RNA can serve one of two functions. Some strands bind to viral proteins and become the genomic ssRNA of new viral particles. Other strands function as messenger RNAs and are translated (again, by the host cell machinery) to form viral proteins. The fate of a given molecule of viral RNA (genomic RNA vs. mRNA) is determined by a series of complex processing events that occur within the cell. The resulting viral genomic RNA molecules and viral proteins assemble into new viral particles that bud from the cell membrane.

One more important feature of HIV biology is that when the nine genes of the HIV genome are transcribed and translated, all of the encoded proteins are not made individually; a few of the genes are translated as a single polypeptide. In order for the individual protein components to be functional, they must be cut free from each other. One of HIV's genes encodes for a **protease** that acts like a molecular scissors to cut the multi-protein structures into their proper and functional subsections. This proteolytic cleavage occurs after the new viral particles are formed and exit the host cell by budding off of the cellular membrane. Once the polypeptide is cleaved, viral maturation is complete and the resulting viral particle can infect another cell.

Based on this short description of HIV replication, it should be clear why the viral reverse transcriptase, protease, and integrase have been the subject of a lot of research.

❖ **Study Questions**:

1. Describe the general structure of a virus.

2. In general, how do viruses replicate? 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.

📖 **Overview Reading**

• Chapter 18 • Immunology: Gene Expression…

📖 **Focused Reading**

• p 407-413 "How Does Specific..." to "Monoclonal..."
• p 412 fig. 18.9 (The Structure…)
• p 414-417 "18.5 What is..." to "MHC proteins..."
• p 416 fig. 18.15 (Phases of the Humoral...)

💻 **thelifewire.com Reading**

• Tutorial 18.3 • Humoral Immune Response
• Tutorial 18.4 • Cellular Immune Response

💻 **Web Reading**

• Cytotoxic T-cell Killing Its Target
www.bio.davidson.edu/misc/movies/CTL.mov

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 yet some microbes still get in. These resourceful microbes that make it through all of your body's 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 very quickly, 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.

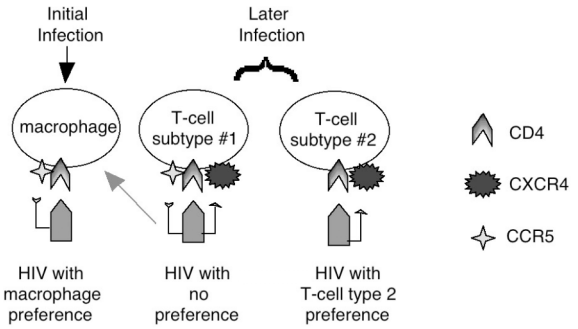
Lymphocytes come in two varieties -- **T cells** (mature in the thymus) and **B cells** (mature in the bone marrow.) B cells make **antibodies**, the same specific proteins you have encountered in looking at the method of immunocytochemistry or immunohistochemistry (recall that you studied this technique at the end of Unit I and in the localization of the CFTR to the ER). 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** (**$T_c$**). $T_c$ cells 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 punch holes in the membranes of other cells. These "holes" make it impossible for the host cell that contains the pathogen to maintain any ion gradients across its plasma membrane and consequently the infected cell dies. $T_c$s kill virally infected cells, some 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. $T_h$s also help $T_c$s become capable of killing. $T_h$s perform both helping functions by secreting various **cytokines** 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 $T_c$s require their help, the $T_h$s play a pivotal role in all immune responses. Unfortunately, **$T_H$s are primarily targeted by HIV**. Thus, by interfering with the function of $T_h$, HIV cripples the entire immune capacity of the individual.

✎ NEWS ITEM: In the disease lupus, B cells inappropriately make antibodies that recognize the patient's own proteins. Symptoms of this autoimmune disease include rashes, hives, ulcers, itching, easy bruising, hair loss, muscle aches, joint pain, and fatigue. In the summer of 2006, scientists identified two genes that may influence B cell responses and susceptibility to lupus. They used an unusual strain of mice in which males have much more severe forms of the lupus because a particular segment of X-chromosomal DNA (called the Y-linked autoimmune accelerator (*Yaa*)) was duplicated and transposed to the Y chromosome. One of the genes duplicated in Yaa is the innate immune receptor TLR7, which normally discriminates between microbe (bacteria) and self. However, with increased TLR7 expression, B cells are more sensitive. Another research team found that a particular variant of a signaling protein causes B cells to multiply inappropriately and to make self-reactive antibodies. These research results highlight that the immune system treads a very fine line between providing us with critical antimicrobial immunity yet preventing harmful autoimmune reactions against ourselves. [*Science* 312:1606]

Viruses target certain cells based on specific binding between proteins in the virus' envelope and proteins in the cell's membrane. For example, the influenza virus binds specifically to proteins on the surface of the respiratory tract and the various hepatitis viruses bind to proteins on the surface of hepatic (liver) cells. These virus-cell interactions are specific, similar to the specificity 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 structures.




*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.*

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. $T_h$s, macrophages, and some supporting cells in the brain express CD4. The story is more complicated, however. CD4 is necessary for HIV binding, but not sufficient for infection. 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 made significant progress in understanding HIV infection (*Science* 272: 809; *Science* 272: 1740; *Science* 274: 502). There are at least two types of molecules (coreceptors) that also are required for HIV infection: **CXCR4** and **CCR5**. As shown in the figure above, 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 variants of HIV that infect different types of $CD4^+$ cells at different times during a person's HIV infection. One variant infects macrophages during the first phase of infection, and another variant 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 are affected first. A plausible explanation is that the strain of HIV that is responsible for the initial infection requires CCR5 as a coreceptor, but 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.

✎ 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 deficiency 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. [*Science* 280:825.]

As is always the case, these membrane proteins that bind viruses are <u>not</u> in the membrane for that purpose (this function 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. 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 its job?
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 two major 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 HIV's reproductive cycle in detail. Understand what happens in each of the steps shown in the web 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 reproductive cycle? How is it involved in this stage?

---

✎ NEWS ITEM: For many years it has been known that some people are exposed to HIV but exhibit increased resistance to developing AIDS. This observation led some researchers to hypothesize that HIV is not the cause of AIDS. New data have shed light on why a person can be HIV+ and show increased resistance to developing 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 32 base pair deletion (note that it is not multiple of three) that caused a nonsense mutation. The encoded truncated protein misfolds and 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 of northern European descent than in people of other ethnic groups (approximately 1% of people of northern European descent are homozygous for this mutation and 17% are heterozygous). Some researchers have speculated that one reason for this higher allele frequency among certain populations may indicate that the altered form of CCR5 provided protection to some pathogen that affected these populations many (hundreds) of years ago. This selective advantage (survival of an epidemic) would explain why Caucasians have a higher frequency of the resistance allele. [*Science* 273: 1797-98; *Nature*. 382: 722; *Cell* 86: 367.]

✎ NEWS ITEM: A French group has discovered that another molecule (US28) can act as a coreceptor for HIV. Surprisingly, 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. 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. For example, CMV might be able to infect cells that lack CCR5 or CXCR4 and thus provide a new host cell for HIV. [*Science* 276: 1794.] But, more recent research indicates that this story is much more complex. Various groups have shown that the coreceptor capabilities of US28 differ greatly in different cell types. In some cells, many types of cells, US28 does not function as a coreceptor for HIV [*AIDS Res Hum Retroviruses* 16:27.] Furthermore, a recent report indicates that CMV infection of various cell types leads to a decrease in CCR5 expression in these cells, thereby making these cells less susceptible to HIV infection. [*Microbes Infect.* 4: 1401.]

### Treatments for HIV and AIDS

So, how can HIV's replication 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 damaging 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** that function by inhibiting the action of proteins or enzymes that are <u>peculiar to bacteria</u> 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 an agent that can selectively destroy HIV, 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 HIV knocks out our defenses, leaving us unable to fend off the virus or, as the disease progresses, any other microbe. Defenseless against microbial attack, AIDS victims are ultimately killed by microorganisms, such as *Pneumocystis carinii*, growing out of control in the body.

---

❖ **Study Questions:**

1. In general, what are antibiotics and how do they work? Why don't antibiotics work against viral infections?

2. 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.

<u>Example #1</u> Interfering with viral binding to T cells: $T_c$ 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 is 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-a**, and **MIP1-b**. (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. (*Science*

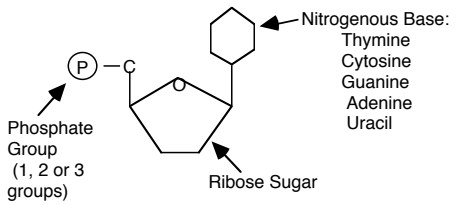
275: 1261-64.) Currently, though, no such treatments are FDA-approved.

**Web Reading**

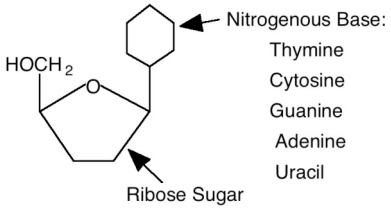
- Life Cycle of HIV - Reverse Transcriptase
  www.bio.davidson.edu/courses/HIVcellsalive/hiv1.htm

Example #2: Interfering with reverse transcription of viral RNA: A second therapeutic approach interferes with the viral replication cycle, specifically the action of the viral enzyme reverse transcriptase. As of 2009, the FDA (Food and Drug Administration) has approved 17 of these drugs for treating HIV infection. The first drug treatment for AIDS was **AZT**. This drug's chemical name is 3'-azido-2', 3'-deoxythymidine. AZT and twelve of the other approved drugs are **nucleoside analogs** and are referred to **nucleoside reverse transcriptase inhibitors (NRTIs)**. 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). Compare the structures of nucleotides and nucleosides.

**Nucleotide**




**Nucleoside**

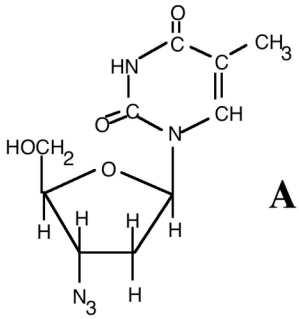



Nucleotides have three components: nitrogenous base, a ribose sugar and one, two or three phosphate groups. ATP, GTP, CTP, TTP, ADP, and AMP are all nucleotides. A nucleoside is simply a nucleotide without any phosphate groups.

In making the nucleotides that will ultimately be part of DNA and RNA, 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.)

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 and this deoxyribose is the $N_3$ group (the azido group is $N_3$, similar to the sodium azide ($NaN_3$) we used in the Ames test) on carbon **3'** in AZT. If you look at the chemical name of the compound, it is actually 3'-azido-2',3'-deoxythymidine. The name 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 more fully).




**AZT**

Because the thymine part of the molecule is identical in thymidine and AZT, DNA polymerase mistake AZT for thymidine. Thus AZT functions as

a **thymidine analog** in the cell. While you certainly could 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 (RNA uses uracil instead) and therefore the normal processes of transcription 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 (ddNTPs) prevented strand elongation.

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 replication cycle. If viral reverse transcription is inhibited, viral replication will be blocked and the virus will not be able to replicate. In light of this description of the action of AZT, one might ask why the drug is not toxic to our cells? In other words, why doesn't AZT inhibit DNA replication in our cells. There are two general answers to this question. First, AZT is toxic (see below). Second, the DNA polymerase in our cells is "smarter" than the viral reverse transcriptase. The DNA polymerase in eukaryotic cells is better able to differentiate nucleotides from nucleosides. Thus it is less likely than RT to incorporate a nucleoside analog into a growing DNA strand. Additionally, eukaryotic DNA polymerases are better able to correct mistakes than RT. As a result, a nucleoside analog that is incorporated into a growing DNA strand may be removed and replaced by a true nucleoside.

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 four hours; 3) its ability to extend the life of the person with HIV/AIDS 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.

Another 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 a selective pressure on the microbial population (evolutionary selection at a microscopic level). If a mutation occurs that allows the microbe to remain infectious in the presence of the drug, then the widespread use of the drug will give this mutant a competitive advantage over non-mutated microbes that were inactivated by the drug. Thus, the widespread use of AZT is undoubtedly favoring the development of AZT-resistant variants of HIV. Such variants certainly exist and may be responsible for some of the cases in which AZT has lost its effectiveness in certain individuals.

✎ NEWS ITEM: Sometimes "failed" inventions can have unexpected second lives. In 1964 Chemist Jerome Horwitz synthesized AZT, hoping that it would halt tumors. He was disappointed that neither AZT nor 50 other compounds that he and his colleges synthesized in 14 years of work showed any promise in treating cancer. Thinking of AZT as a "failure," he refocused his research and didn't think about AZT. Then, in 1985 a colleague showed him a paper that demonstrated that AZT could slow the replication of HIV. AZT was eventually approved by the FDA in 1997 as an AIDS treatment. Dr. Horwitz's lab continued to study cancer until he retired in 2005. [*Chron Higher Ed* Aug 12, 2005:A25]

The three remaining drugs that block reverse transcription are **nonnucleoside reverse transcriptase inhibitors (NNRTIs).** Unlike NRTIs, these drugs are not nucleoside analogs. Rather, these drugs bind to reverse transcriptase and alter its shape (does this sound familiar?). The altered conformation of RT makes it inactive. As a result, RT is unable to convert the viral RNA into DNA and the virus does not replicate.

---

### 💻 Web Reading

- Life Cycle of HIV - Viral Protease
  www.bio.davidson.edu/courses/HIVcellsalive/hiv4.htm

Example #3: Interfering with viral protease activity. As mentioned in the discussion of HIV replication, several of the viral proteins initially are translated as a single, inactive polypeptide. A viral protease must cleave this inactive polypeptide into functional proteins. The newly formed viral particles are not infectious until these cleavage events occur. Thus, researchers quickly hypothesized that inhibition of the viral protease could be an effective means of stopping viral replication and delaying the onset of disease. The first such drug was approved for use in 1995. Currently, seven **protease inhibitors (PIs)** have been approved by the FDA. Like the NNRTIs, these drugs work by altering the shape of a viral enzyme. PIs have an affinity for the viral protease. By binding to the protease, the drugs alter its conformation, making it inactive.

The positive effects of these drugs are dramatic. When PIs first were approved for use in the mid-1990s, many people with HIV/AIDS experienced remarkable recoveries. There are numerous reports of people with AIDS who were extremely sick before beginning protease inhibitor treatment, and then experienced remarkable recoveries. Of course, like the other drugs discussed, PIs have limitations: 1) they are not a cure; 2) they are expensive (most PIs cost approximately $600-700 per month); they have a number of side effects; and 3) drug resistant HIV mutants have been identified.

Example #4: Combination drug therapy. Until 1995, physicians prescribe one anti-HIV drug at a time to people with HIV/AIDS. This **monotherapy** contributed to the development of drug-resistant mutants of HIV. Beginning in 1995, though, David Ho (an HIV/AIDS researcher) and others began recommending that people with HIV/AIDS take two or three different anti-HIV drugs simultaneously. Such combination drug therapy has two main advantages. First, it decreases the amount of virus present within a person (**viral load**) more dramatically than monotherapy. Second, drug resistant mutants are less likely to develop with combination drug therapy. The development of this **highly active anti-retroviral therapy (HAART)** is what made David Ho *Time* magazine's Man of the Year for 1996 and *Science*'s Breakthrough of the year 1996. Currently, it is recommended that people with AIDS take a **triple drug cocktail** of two reverse transcriptase inhibitors and a protease inhibitor.

Despite the great benefits of HAART, downsides do exist. First, drug resistant mutants still can develop. If such mutants develop within a person, then the effectiveness of HAART for that person will decline. As a result, physicians must constantly monitor the viral loads in people with HIV/AIDS and change the drug cocktail if the current treatment becomes ineffective. Second, the required drug regimen is difficult to follow. Multiple pills must be taken every day, and at various times throughout the day. Pharmaceutical companies are working on combination pills that can be taken only once or twice a day. Third, unusual side effects have been observed in people receiving HAART. Most noticeably, people on HAART experience a fat redistribution, resulting in increased waist size and development of a fat deposit between their shoulders. The exact cause and long-term effects of this fat redistribution still is under investigation.

It should be noted that all of these advances in anti-HIV drug development are beneficial to people with HIV/AIDS only if the drugs are available to them. As mentioned previously, nearly three quarters of the people with HIV/AIDS worldwide live in sub-Saharan Africa. Most of these people do not have access to the drugs.

✎ 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 life-saving drug, it is difficult to watch the control group get worse, knowing that you might be able to prolong their lives if they had access to the new medication. [*Science* 276: 520-523.]

✎ NEWS ITEM: Researchers at U Mass have made double-stranded RNA that can degrade specific RNA sequences. By designing small interfering RNAs (siRNAs) that target HIV's genomic RNA they were able to reduce HIV replication in cultured human cells 30-50 fold in the first 24 hours of infection.

Their work suggests that RNA interference may provide a new way to prevent or reduce viral replication. [*Nature* 418: 435-8.]

✎ NEWS ITEM: Another type of antiretroviral drug may be developed as a result of the identification of an HIV entry inhibitor. The 20 amino acid peptide called VIRIP, or virus inhibitory peptide, binds to gp41 and keeps HIV from entering cells. [*Cell* 129: 263-275]

### 💻 Web Reading

• Treatments for HIV
http://highered.mcgraw-hill.com/olc/dl/120088/treatmentHIV.swf

• Optional supplementary article: "Can HIV Be Cured?" by Mario Stevenson, 2008
Download from the Scientific American Archive Online (see http://www.bio.davidson.edu/courses/ bio111/topics.html for the link to the Archive, or find it through the library web site.)

### ❖ 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 HIV/AIDS?

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, former 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 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 shown to be 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." [*Science*. 274: 2005.]

In order to understand how vaccines are developed, we need to return to the immune system and see how vaccinations protect against disease.

### 🕮 Focused Review Reading

• p 409-410 "Immunity &…." To "Animals…"

Before vaccines were developed, the only way for a person to become immunized to an infectious disease was to get exposed to the pathogen and survive its effects. Given the nastiness of some infectious diseases, this scenario was not ideal, and many individuals died in their youth of an infectious diseases. If an individual contracts 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 <u>specific</u> 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 are 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 concept is the very first vaccine ever developed - the vaccine against small pox. Smallpox 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 rarely contracted smallpox, 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 smallpox (most notably open skin lesions). The cow version of the disease was called "cowpox." Suspecting that milkmaids were in 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 smallpox (from an open sore of a small pox victim). The boy did not become sick from smallpox. (Biomedical ethics committees would have locked Jenner up for doing such a thing today.) Thus Jenner discovered a way to protect against smallpox. He called this potion a vaccine (after "vacca", Latin for "cow"). (This story 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 studying animal diseases!) It should be noted that, while Western culture credits Jenner with the development of the smallpox vaccine, there is evidence that a similar approach was utilized by the Chinese as early as the 10th century!

✎ NEWS ITEM: Smallpox has been completely eradicated from the human population. The smallpox virus is thought to be present in only two known places on earth -- in a vial at the Center for Disease Control in Atlanta, and in a vial in a comparable institution in Moscow. But many people believe that other sources may exist - and that terrorist organizations may have access to these sources. Since the 1970s, people have not been routinely immunized against smallpox because the disease was thought to have been eradicated. If smallpox is released by a bioterrorist, most of the people in the world would be susceptible to the virus. Many officials believe that we may need to begin vaccinating people against smallpox again. In 2002 the US government instituted smallpox vaccination for US military personnel and heath care workers (first responders) so they would be immune in the event of terrorist release of the virus.

We have come a long way since Jenner scraped cells from 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 HIV vaccines have been engineered and tested in humans. Here are some examples:

1. Attenuated virus. These vaccines are infectious viruses that have been altered in some way to make them non-pathogenic even though they remain capable of replicating (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 replication). 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. In the case of HIV, however, 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 attenuated viruses. However, so far, alternative methods have failed to produce an effective vaccine, and, in December 1992, a group of investigators reported that they could prevent infection by Simian immunodeficiency virus (SIV) using an attenuated SIV with one gene removed. These results are intriguing and may cause the HIV/AIDS research community to rethink their resistance to the use of 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. Some researchers feel that such a *nef*- virus may be useful as a vaccine.

✎ NEWS ITEM: Dr. Baltimore's group has 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 event explains why the *nef*- strains of HIV are not as potent as their wild-type relatives. [*Science* 276: 1196-97.]

2. Cloned Envelope Glycoproteins (also called subunit vaccines because they contain only a subunit of the virus, not the entire virus) These vaccines are the safest vaccines because there is no virus present to cause an infection. By applying genetic engineering techniques (many 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 then should be able to "see" the natural gp160 or gp120 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 gp120) 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. It is this type of immunity that is primarily responsible for eliminating viral infections. For that reason, research is underway to attempt to bind gp160 and gp120 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 1998, VaxGen, a biotechnology company in San Francisco, received FDA approval to begin a large-scale human trial of a subunit vaccine. This vaccine, which consisted of two forms of gp120, was administered to approximately 5000 uninfected homosexual men in the US and 2500 uninfected IDUs in Thailand. Their HIV status was checked periodically over the next three years. In February of 2003, the results of this trial were announced. VaxGen officials claimed that the vaccine showed promise in certain racial/ethnic groups. Many scientists, however, disagreed with this optomistic conclusion. [*Science* 299: 1290-91]

3. Viable vector vaccines. Many researchers now are investigating the effectiveness of viable vector vaccines. In this approach, HIV gene(s) coding for major structural proteins (*env* gene that encodes the surface protein and *gag* protein that encodes the core protein, for instance) are cloned into a non-pathogenic microbe (currently, canarypox virus is being used most extensively as the 'carrier' microbe). This engineered microbe then can be administered to people. The HIV proteins will be produced, and the immune system will mount a response against them, but HIV replication will not occur. Many researchers believe

that such a vaccine may be as effective as a typical attentuated vaccine.

Currently, over 20 potential HIV vaccines are in human trials. 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.

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" because 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 strain 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 the previous 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.]

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.

The lack of a suitable animal model for the disease is also a significant problem. 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 such changes 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 approach is as 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 and it is 82% identical to the human protein and yet HIV cannot bind to the mouse CCR5. [*Science* 274: 1924-26.]

It's been nearly three decades since AIDS became known to the medical community, but we still do not have a good vaccine. Part of the problem may have been recently elucidated. In 2005 scientists coaxed cultured immune cells to make neutralizing antibodies against HIV, but these antibodies are also reactive against cardiolipin, an important phospholipid in the body. This result suggests that a broadly HIV-neutralizing immune response in the human might be suppressed by the mechanisms limiting self-reactivity, or autoimmunity. See *Science* 308: 1878. Meanwhile, research continues.

**Web Reading**

• Optional supplementary article: "The Vaccine Search Goes On" by David I. Watkins, 2008
Download from the Scientific American Archive Online (see http://www.bio.davidson.edu/courses/bio111/topics.html for link to the Archive)

❖ **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.

## Diagnosis of HIV+ Individuals

**Web Reading**

• ELISA for HIV
www.biology.arizona.edu/immunology/activities/elisa/elisa_intro.html

A blood test for HIV infection has been available for since 1985. 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 do not engage in any high risk behaviors and 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 after infection 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 bottoms 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 into 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 (the enzme was isolated from the horseradish plant). 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.

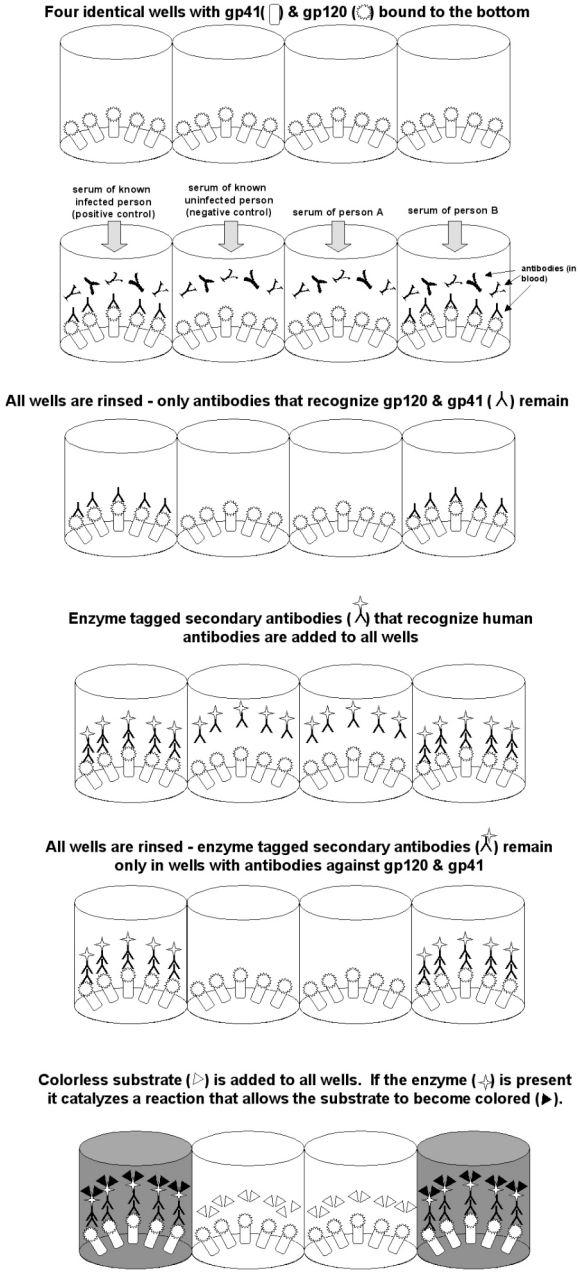
The change in color is measured by a **plate reader** (just like the spectrophotometer you used

for the IDH labs) 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, although it is considerably more expensive due to the time involved.

You have already encountered the **Southern blot** in Unit II. In Southern blots, restriction fragments of DNA are electrophoresed and then transferred to a piece of nitrocellulose paper where the DNA is hybridized with a probe. Two other types of blots are based on similar ideas. 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 separated into its individual protein molecules. These molecules are electrophoresed and separated by molecular weights and blotted onto nitrocellulose paper. 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 also will 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 (ELISA) and confirmatory (Western blot) tests examine only the presence of antibody to the virus. There is also a PCR test available that detects the presence of the virus inside T helper cells. You should recall that polymerase chain reaction (PCR - the same method you used in lab) amplifies specific sequences of DNA. The PCR test is used in situations where it is absolutely essential to know whether or not someone is $HIV^+$. The PCR test is not used to screen the general public because of its expense.

When PCR is used to amplify 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 HIV, 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) that are complementary to a base sequence present only in the HIV 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 <u>one</u> 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.

An entirely different HIV testing approach called **transcription-mediated amplification (TMA)** is used to screen donated blood. This technique was developed by Davidson alumnus Larry Mimms '75 and is used worldwide. See http://www.bio.davidson.edu/Courses/Bio111/TMA/TMA_Method.html for a step-by-step explanation of this technique.

---

❖ **Study Questions**:

1. What is seroconversion? Why is it called seroconversion? 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?
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. Compare how PCR and TMA are used to detect the presence of HIV. Why are these tests far more accurate than the Western blot?

---

Note: Another detection method that 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. At thelifewire.com Tutorial 18.2 goes over how pregnancy tests work and how they have been designed to include the all-important 'control.'

**Future Directions**

Despite all the research on HIV, it is not clear 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 people with AIDS have very low $T_h$ cell counts and acquire infections that are caused by the lack 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 ($T_c$s?) to kill $T_h$ cells
3. HIV causes $T_h$ cells to commit suicide

For years, theory #1 was assumed to be true, but several years ago scientists 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 observation caused several investigators to wonder how HIV could be directly responsible for T-cell death. 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 that 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 classic 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 $T_H$ cell a target for destruction by $T_c$s or some other immune system cell. Thus, according to this theory, the virus simply marks the $T_H$ for destruction, but does not destroy the cell itself.

In support of the third hypothesis, $T_H$ 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 $T_H$ cells, but rather it programs them in some way to kill itself at a later time. Of course, these three theories are not mutually exclusive, and all three processes may be acting to destroy $T_H$ 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 a 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 kills neurons and leads to the development of neurological symptoms that affect up to one third of all AIDS patients. 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 scientists think that infected macrophages may induce apoptosis in astrocytes and maybe $T_H$ cells. [*Science* 274:1464-65.]

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, 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 experimentation at the laboratory bench.

Given our understanding of how the virus infects cells and new treatments, is the AIDS epidemic slowing down? At the 14th International AIDS conference in Barcelona in 2002, a United Nations (UN) report predicts that the AIDS epidemic has not yet peaked. Only 1.75% of HIV+ people receive anti-HIV treatments, and 70% of these treated patients live in high-income countries. Obviously, AIDS goes untreated for the vast majority of patients in underdeveloped countries. If more countries do not significantly expand their prevention programs, the UN forecasts that AIDS will claim an additional 65 million lives by 2020 (more than three times the number of patients who died in the first 20 years of the epidemic). Obviously we have a long way to go.

**Major Stories to Follow:**

As mentioned previously, a number of anti-retroviral drugs have been developed, but most of the people with HIV/AIDS worldwide can not afford these drugs. Many countries are considering making cheaper, generic versions of these FDA-approved drugs. Opponents of this plan claim that such a practice could expose people to incorrect formulations and/or dosages, thereby increasing the development of drug resistant viral variants.

Recently, President George W. Bush proposed providing $15 billion to combat HIV/AIDS in Africa and the Caribbean. This initiative represents the largest allocation of funds by the US to the worldwide HIV/AIDS pandemic. Many people have applauded this announcement; others have criticized it, claiming it was too little money. Others are worried that the money will not be used for programs known to work - sex education, condom distributions, and free needles.

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* 276: 1197]

Antisense therapies are making a comeback in many areas, and HIV/AIDS is no exception. Antisense technology is fairly simple in theory, but has many practical obstacles. Because 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 might be able to prevent HIV replication. 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. 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.

## 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 429-432 "Is Cell..." to "Pluripotent stem…"

**Web Reading**

- How Cloning Works
  science.howstuffworks.com/cloning1.htm

**Cloning Organisms**

Plants are very easy to clone; in fact, many plants clone themselves naturally. When a plant sends out a runner and establishes a new individual without reproducing sexually, it is cloning itself. Cloning is the creation 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 some commercially important plants by starting with single cells and growing them in

tissue culture. The big news in 1997 was that a cloned mammal had been born.

A simpler and less controversial form of mammalian cloning, embryo splitting, had previously been performed. In this technique, a single embryo made up of a few cells is split into two smaller clumps of cells, and the clump then mature into genetically identical individuals, which develop simultaneously, as with human identical twins. In contrast, cloning from a donor already at an advanced stage of development has been difficult. Amphibians were cloned in the 1970's but only from early embryonic donor cells. It was not until Dolly stunned the world in 1997 that a clone was generated from an adult mammal donor. Dolly was the product of combining a mature nucleus (from an adult sheep's mammary gland) with an undeveloped oocyte (egg cell) cytoplasm. This procedure is called somatic cell nuclear transfer (SCNT).

A potential problem with SCNT stems from the fact that your chromosomes, like batteries, are designed to keep going for a set length of time but eventually will expire. (Even the Energizer Bunny will die at some point.) The telomeres of chromosomes are the limiting factor. Every time your chromosomes replicate, a short stretch of your telomeres is lost. Eventually the telomeres are gone, and important genes start getting lost, so cells die. Many people think this process contributes to aging. An individual's telomere "batteries" are recharged during the normal processes of sperm and egg development and subsequent fertilization, but the danger is that a cloned individual may be born with already partially depleted telomere "batteries." When Dolly's telomeres were examined, they were indeed shorter than those of normal sheep her age. Dolly developed arthritis at a young age, though it is unknown whether that condition stemmed from shorter telomeres. Dolly was euthanized at age six (young for a sheep) in February, 2003, because she had a severe lung infection. She is now stuffed and on display in the Royal Museum in Edinburgh, Scotland.

Would a human clone have unusually short telomeres and perhaps age faster than normal? Maybe, maybe not. When mice and cows were cloned, scientists found that somehow the telomeres in those clones (unlike in Dolly) had been regenerated. So, it is not clear what would happen with a cloned human.

To date, cloning is primarily used for plants and certain commercially important animals, though several pet cloning ventures have started up and folded in recent years.

✎ NEWS ITEM: Some people wonder if cloning technology could help save endangered species from extinction. This proposition leads to a hot debate on where conservation money should be spent, but cloners do have a unique argument. In animals such as cheetahs where the gene pool is too small for long-term survival of the species, 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* 276: 1329.] For example, two different kinds of endangered Asian cattle (the gaur and the Javan banteng) and an endangered sheep (European mouflon) have been cloned. Scientists in several countries are trying to clone many other species.

## Transgenic Organisms

### 🕮 Focused Reading

- p 369-372 "DNA manipulation…" to "16.6 Recap"

### 💻 Web Reading

- Methodology for Making Transgenic Mice
  www.bio.davidson.edu/courses/Bio111/topics.html
- A Portable Gene Gun
  www.bio.davidson.edu/courses/Bio111/genegun.html
- Fixing Food
  www.sciam.com/article.cfm?articleID=000143E5-AD76-1E0C-8B3B809EC588EEDF

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 a tube

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 an 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 **germline** 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 that cause her to "superovulate" -- that is, release many, many eggs. You then harvest the eggs just before they burst from the surface of the ovary and place them into a dish with sperm collected from the male of the species. The egg and sperm join and form a zygote. (This procedure is called ***in vitro* fertilization**; in humans, at least thus far, the embryos that are created are implanted without genetic modification in a woman who has sought fertility treatment.)

In experimental animals, at this stage the transgenes--which have been constructed with a promoter that turns on the genes at the appropriate time 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 (artificially treated previously with hormones at levels associated with pregnancy). When the offspring come to term, they are tested by either a Southern blot or by PCR to see which of them carry the transgene. Given all the steps in this procedure in which something could go wrong, the chances of producing a transgenic offspring are about one in ten births, and much worse odds if you count every implanted embryo.

The primary animal that has been used for transgenics has been the mouse. For example, one mouse strain has a human immune system, enabling experimental exploration of human immune function in the context of various diseases and conditions. Another modified mouse has twice the normal amount of skeletal muscle. This mouse could be used to understand and perhaps treat muscle diseases such as 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 common experimental goal is the **knockout mouse**, in which both alleles of a particular gene have been disabled. The 2007 Nobel Prize in Physiology and Medicine was awarded to Mario Capecchi, Oliver Smithies, and Martin Evans for advances that led to the development of this technique. The phenotype of a knockout mouse gives us important information on the function of the protein whose gene has been disrupted. Making a knockout mouse is much more complicated than the simple "gene addition" approach described in the previous paragraph. See page 364-365 and fig. 16.13 for an overview of the knockout technique. Pharmaceutical companies and academic laboratories are making many varieties of knockout mice that serve as models for human genetic diseases; scientists can test new therapies on these mouse models. Cystic fibrosis, Alzheimer's Disease, muscular dystrophy and sickle cell anemia are some of the many human diseases for which mouse models have been generated.

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 plant 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 given a transgene in one of two ways: 1) infection with *Agrobacterium*, a bacterium which has the ability to introduce plasmid DNA into plant cells. The naturally occurring *Agrobacterium* plasmids cause tumors in the plant, but scientists can engineer the plasmid so that it carries a gene of interest instead of the tumor-causing genes. 2) delivery by gene gun, which shoots tiny DNA-coated gold particles into cells.

Most transgenes tested so far confer resistance to viruses, insects and herbicides. Herbicide-resistant soybeans, as well as pest-resistant cotton and corn, have been approved for cultivation in the United States. In 2001 about two thirds of the total US soybean and cotton crop (and about a quarter of the corn crop) consisted of genetically modified varieties. Scientists are working to develop plants with a wide range of genetic modifications—examples include vitamin-enriched rice, naturally caffeine-free coffee, allergen-free peanuts, and turf grass that needs little fertilizer.

In North Carolina the tobacco industry is under fire, and farmers 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 the plants are given the correct DNA. Likewise, they can produce other **ph**armaceutical products like growth hormone, blood clotting factors, and insulin (review page 369-371 in your text for more information). Instead of a few dollars per bushel, these plants may well be worth their weight in gold, if not more!

The US Department of Agriculture regulates the field trials of transgenic crops and livestock. The following table lists many of the crops that scientists have modified genetically.

| Plant | Trait conferred by transgene(s) |
|---|---|
| Alfalfa | Herbicide tolerance, virus resistance |
| Apple | Insect resistance |
| Oilseed rape | Herbicide tolerance, insect resistance, modification of seed oils |
| Cantaloupe | Virus resistance |
| Coffee | Decreased caffeine production |
| Corn | Herbicide tolerance, insect and virus resistance, wheat germ agglutinin |
| Cotton | Herbicide tolerance, insect resistance |
| Cucumber | Virus resistance |
| Melon | Virus resistance |
| Papaya | Virus resistance |
| Peanut | Reduced allergenicity |
| Potato | Herbicide tolerance, virus & 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, beta carotene production |
| Soybean | Herbicide resistance, modified seed protein storage, reduced allergenicity |
| Squash | Virus resistance |
| Strawberry | Insect resistance |
| Sunflower | Modified seed protein storage |
| Tobacco | Herbicide tolerance, insect resistance, virus resistance |
| Tomato | Virus resistance, herbicide tolerance, insect resistance, modified ripening, frost resistance, saline resistance |
| Turfgrass | Drought resistance, need for less fertilizer |

partially adapted from Kareiva (1993) *Nature* 363:580

✎ NEWS ITEM: Scientists have developed transgenic tomatoes that are resistant to high soil salinity. The plants sequester salt in the leaves, leaving the tomatoes themselves unaffected. Salt-resistant plants would enable farmers to cultivate crops on land that would otherwise be unusable, minimizing the need for clearing new farmland. The challenge these farmers face is lack of public acceptance of genetically modified crops; currently there is little market for these tomatoes. (*Current Opin. Biotech.* 13: 146-50. )

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 show secondary phenotypic effects that create imbalances in ecosystems, threatening other species. 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 with AZT treatment for HIV-infected people, giving one variety of an organism (such as AZT-resistant viruses) an artificial advantage may change the selection pressures on the population, altering 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 issue raises all kinds of problems. For instance, what if I buy a transgenic mouse (or hog or goat) from someone that holds the patent and 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, because it is a good bet that this controversy will be raging for years to come.

✎ NEWS ITEM: In 2002 the Canadian Supreme Court decided that transgenic mice could NOT be patented. This decision puts Canadian policy at odds with US statutes. The particular mouse at the center of the proceedings was the Harvard "oncomouse," which is modified to be predisposed to cancer, and which has been protected by a patent in the United States for many years. The Canadian justices decided that "higher life forms" could not be covered under the Federal Patent act of 1869, although genetically modified plants and single celled organisms are still protected. See the following news story from the Canadian Broadcasting Company's web site: http://www.cbc.ca/stories/2002/12/05/scc_mouse021205

---

❖ **Study Questions:**

1. Define a transgenic organism. Compare the creation of a transgenic organism with the application of gene therapy for a disease like 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. Be able to cite examples of transgenic organisms and the product they are designed to produce.

6. How might gene therapy be used to generate a T cell-mediated vaccine for AIDS?

7. 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?

8. Describe how the famous sheep Dolly was created.

---

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? (Of course, it is a fallacy to think there's "a gene" for something like IQ--all of the qualities mentioned above result from the cumulative action of many genes—but the possibility remains that genetic modification could alter these traits in humans.)

Anyway, what do we mean by "normal"? What pressures will parents be under to ensure that their offspring are genetically "normal?" If you do not 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 40 years ago. If this area is interesting to you, you should take Genetics as well as 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 do not 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.