Pantothenic acid is necessary in the biosynthesis of coenzyme A. Coenzyme A is an exceedingly important substance in many biological processes because this coenzyme transfers acyl groups.

The wonders of vitamin B_{12}

Vitamin B_{12} is the only known natural organometallic compound. It does not occur in higher plants, and apparently only bacteria are capable of synthesizing it — bacteria who live in their hosts in a symbiotic relationship. Unfortunately, higher animals including human beings do not have these types of bacteria. Thus, it is necessary to obtain vitamin B_{12} from food. The name *cyanocobalamine* refers to the presence of cyanide. The cyanide is an artifact of the isolation of the compound and is not naturally present. Vitamin B_{12} is necessary to the formation of two coenzymes: methylcobalamin and 5'-deoxyadenosylcobalamin. The structure of methylcobalamin is shown in Figure 10-8.

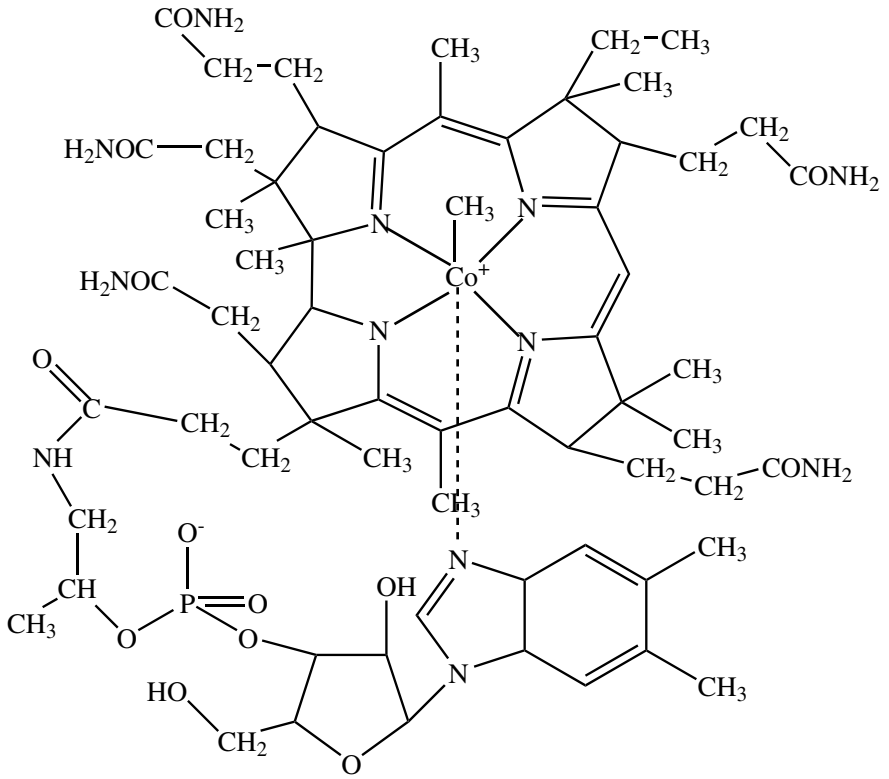


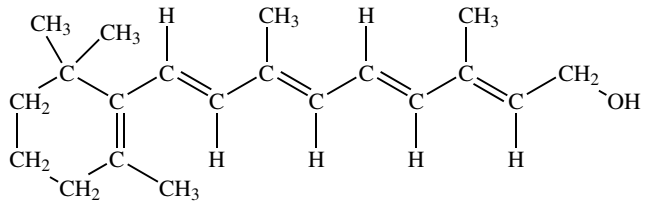
Figure 10-8: Structure of methylcobalamin.

Both coenzymes assist in reactions involving rearrangements. Methylcobalamin is useful in methyl transfer reactions. The coenzyme 5'-deoxyadenosylcobalamin works in some rearrangement reactions where a hydrogen atom and a group attached to an adjacent carbon exchange positions.

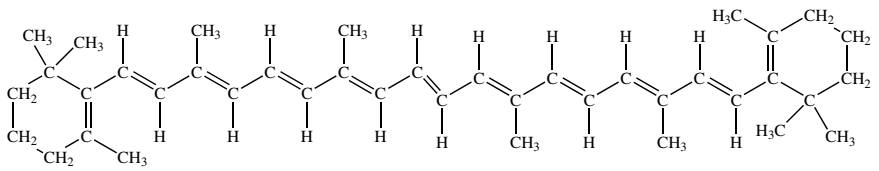
Pernicious anemia usually results from poor absorption of vitamin B_{12}. Normal stomach cells produce a glycoprotein that aids in the absorption of the vitamin in the intestine. It is the lack of this intrinsic factor that leads to the vitamin deficiency and not the lack of the vitamin in the diet. Elderly people may have difficulty in generating sufficient quantities of the intrinsic factor, and strict vegetarians also may develop symptoms. The symptoms of pernicious anemia include lesions on the spinal cord leading to a loss of muscular coordination and gastrointestinal problems. The blood contains large, fragile, and immature red blood cells. Dietary sources include meat, eggs, milk and cereals. This vitamin is stable to cooking.

Vitamin A

Vitamin A is not a single compound — a number of compounds are biologically active, that is they undergo biological reactions within the organism. The parent compound is 11-trans-retinol, found in milk and eggs. Vitamin A is exclusive to animals, and the plant pigment β-carotene can serve as a precursor (Figure 10-9). As a precursor, it is a *provitamin.* Cleavage of β-carotene yields two vitamin A active species. Any β-carotene that doesn't become vitamin A is used as an antioxidant.



11-trans-retinol



β-carotene

Figure 10-9: Structures of 11-trans-retinol and β-carotene. Note that carbon 11 is the fifth from the right in the main chain.

Vitamin A is especially important to vision. Part of the vision process involves the absorption of light. This absorption causes the geometry on the double bond between carbon atoms 11 and 12 to change from cis to trans. The isomerization triggers a series of events, giving rise to a nerve impulse. An enzyme reverses the isomerization so the molecule may be reused. In addition to being directly involved in vision, vitamin A also promotes the development of the epithelial cells producing the mucous membranes, which protect the eyes and many other organs from infections and irritants. Vitamin A also helps in the changes in the bone structures that occur as an infant matures.

A deficiency in vitamin A begins with night blindness, followed by other eye problems, which could lead to blindness. An extreme deficiency may lead to *xerophthalmia,* inflammation of the eyelids and eyes, which can cause infections and blindness. Young animals require vitamin A for growth, and adults are capable of storing several months' supply of it, primarily in the liver. The livers of some animals, such as polar bears and seals, may have such a high vitamin A concentration that they are toxic to humans. Excessive dosages of vitamin A may lead to acute toxicity, and as a fat-soluble vitamin, it is not easily eliminated. Symptoms include nausea, vomiting, blurred vision, and headaches. Large doses have been linked to birth defects and spontaneous abortions. The provitamin, β-carotene, is not toxic.

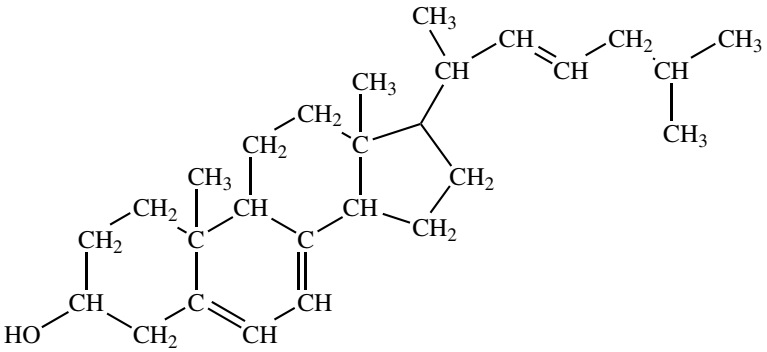
Vitamin D

Vitamin D is sometimes called the *sunshine vitamin*. It can be produced in the body through the action of sunlight, which is ultraviolet radiation. Individuals walking around outside nude or semi-nude normally have very little trouble with vitamin D deficiency. The rest of us, however, depend on vitamin D-fortified foods, especially milk.

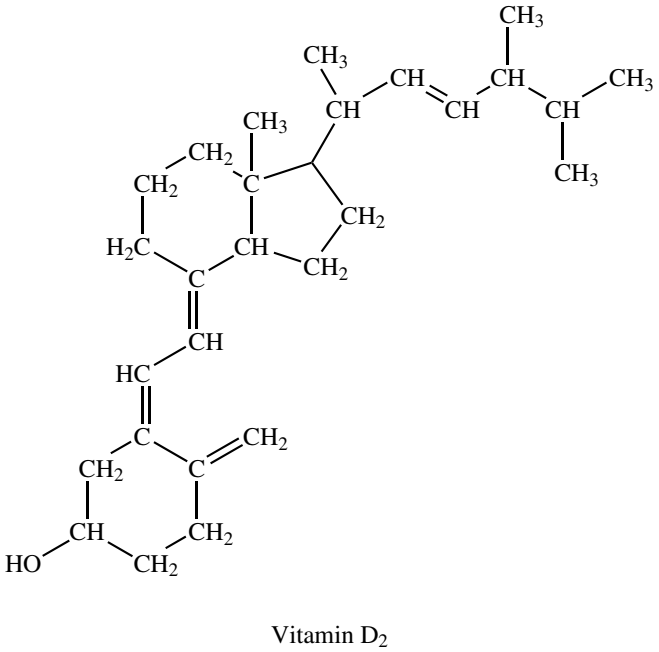
Several compounds exhibit vitamin D activity. Only two of them — actually provitamins — occur commonly in food: ergosterol and 7-dehydrocholesterol. Irradiation with ultraviolet light converts ergosterol into vitamin D_2, *ergocalciferol.* Ultraviolet irradiation, particularly in the skin of animals, converts 7-dehydrocholesterol into vitamin D_3, *cholecalciferol.* (A little confusingly, vitamin D_1 is a mixture of vitamin D_2 and vitamin D_3.) The structures of ergosterol, vitamin D_2, 7-dehydrocholesterol, and vitamin D_3 appear in Figure 10-10.

The body's ability to absorb calcium and phosphorus is tied to vitamin D. Teeth and bone have large amounts of these two elements and are the first parts of the body affected by a vitamin D deficiency. Osteomalacia, a condition in which a softening of the bones may lead to deformities, may also result. (In infants and children, osteomalacia is called rickets.) A vitamin D

deficiency is more serious in children than in adults because growth requires larger quantities of calcium and phosphorus. Persons with some portion of their skin routinely exposed to sunlight seldom develop a deficiency.

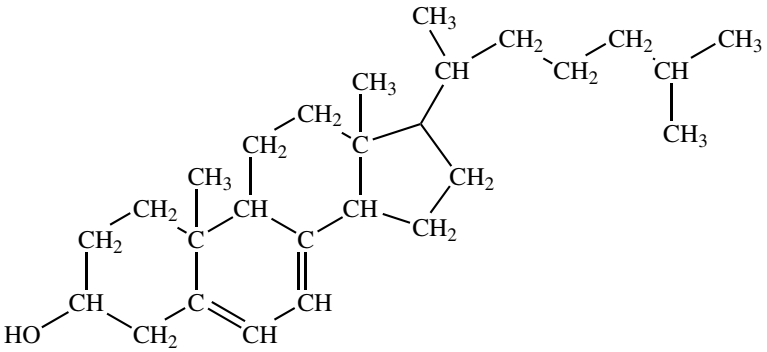


Ergosterol

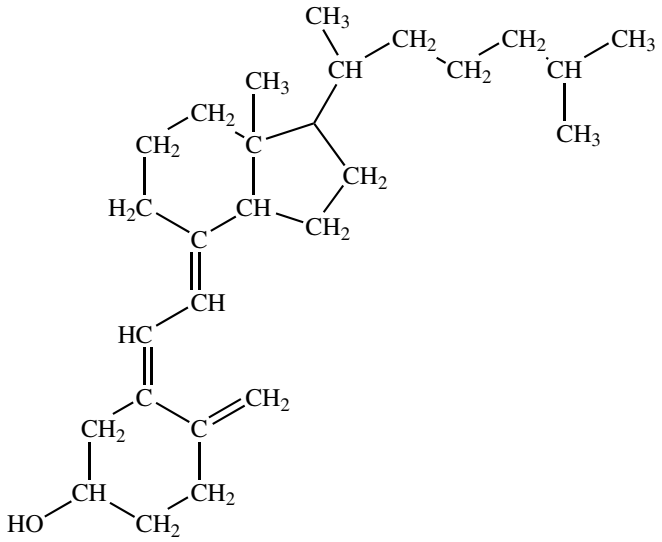


Vitamin D_2

Figure 10-10: Structures of ergosterol, vitamin D_2, 7-dehydrocholesterol, and vitamin D_3.



7-Dehydrocholesterol



Vitamin D_3

Figure 10-10: (continued)

Excess vitamin D is toxic. It is not easy to eliminate this fat-soluble vitamin. Symptoms of excessive amounts of vitamin D include nausea, diarrhea, kidney stones and other deposits, and sometimes even death.

Vitamin E

The *tocopherols* are a group of compounds that exhibit vitamin E activity. The most effective is α-tocopherol (see Figure 10-11). Vitamin E comes from a number of sources, vegetable oils, nuts, whole grains, leafy vegetables, to name a few. Deficiencies are rare except in individuals on a no-fat diet or who, for medical reasons, cannot efficiency absorb fat. Cystic fibrosis may interfere with fat absorption.

Vitamin E serves as an effective anti-oxidant. *Anti-oxidants* are necessary to minimize the damage caused by oxidants present in the body — many problems associated with aging are apparently due to oxidants. Vitamin E may also help prevent cholesterol deposits in the arteries. There are no well-documented problems with the use of large doses; however, some recent studies, although disputed, warn against taking mega doses of vitamin E.

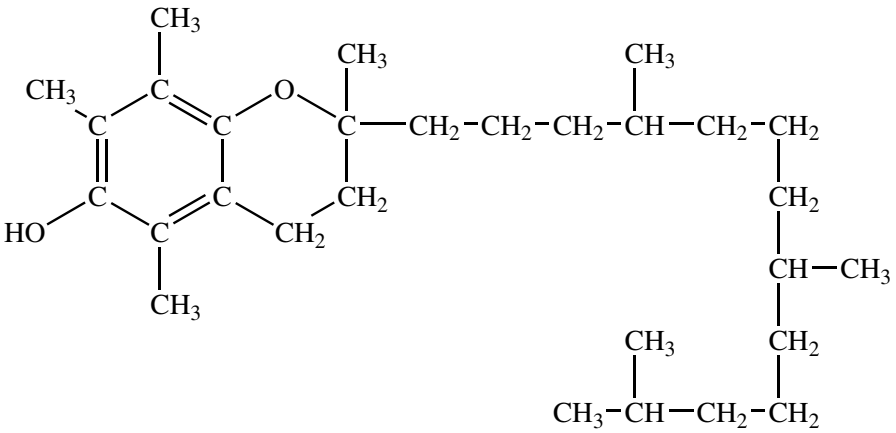


Figure 10-11: Structure of α-tocopherol (vitamin E).

Vitamin K

Vitamin K_1 (Figure 10-12) is one of many compounds that exhibit vitamin K activity and differ in the side-chains attached to the ring system. One chain is usually a methyl; the other typically has at least 20 carbon atoms.

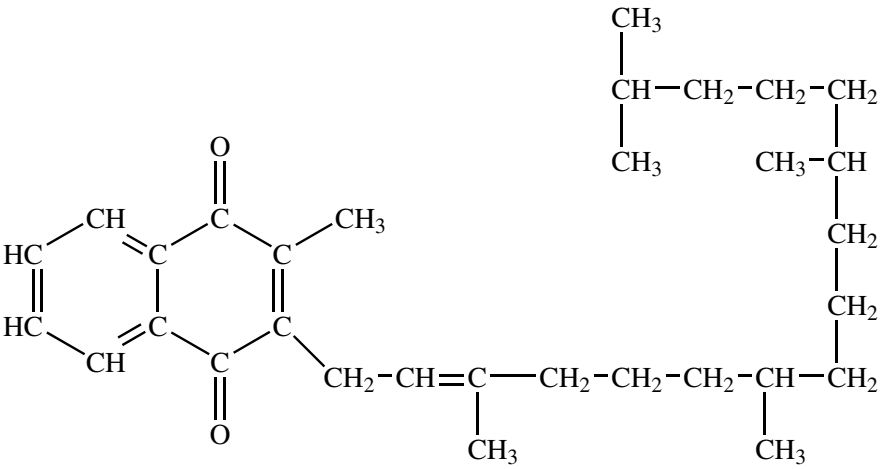


Figure 10-12: Structure of vitamin K_1.

Vitamin K is necessary to produce the proenzyme *prothrombin,* which helps blood clot. A vitamin K deficiency is uncommon because intestinal bacteria normally produce sufficient quantities, although several foods are also good sources, including green leafy vegetables, cauliflower, broccoli, organ meats (love that liver!), milk, soybeans, avocados, and bananas. Two tablespoons of parsley contains almost twice your recommended daily amount of vitamin K. Prolonged use of antibiotics can decrease the number of these vitamin K-producing bacteria and lead to a reduction in vitamin K in the body. One symptom of a deficiency is an increase in the time necessary to form a blood clot, and such individuals are prone to develop serious bruises from even minor injuries. Infants with a deficiency have been known to die from brain hemorrhage. Increasing the vitamin K intake of the mothers decreases the likelihood of this occurrence.

Vitamin C

Vitamin C is another name for *ascorbic acid* (Figure 10-13). Dehydroascorbic acid also has vitamin C activity. Vitamin C is water-soluble — thus the body can readily eliminate excess, and large doses are not toxic.

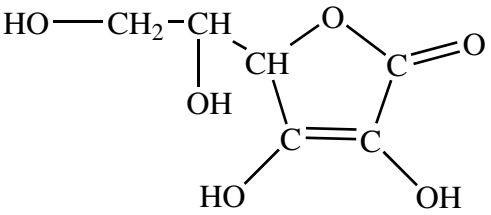


Figure 10-13: Structure of vitamin C.

A deficiency in vitamin C leads to the disease *scurvy,* symptoms of which include a weakening of the collagen — an important protein in connective tissues such as ligaments and tendons. Many foods contain vitamin C, especially plants and citrus fruits, so it is easy to prevent scurvy. For years, British ships carried limes as a source of vitamin C (leading, incidentally, to the slang term *limey* to refer to a British sailor). Many mammals (other than humans) synthesize vitamin C from glucose. Cooking, especially prolonged cooking, destroys vitamin C. Vitamin C is an antioxidant. Like vitamin E, it helps prevent damage produced by oxidants. It also helps in the absorption of iron, and keeps the iron in the +2 state. Vitamin C helps convert some of the proline in collagen C to hydroxyproline, which stabilizes the collagen.

Chapter 11

Be Quiet: Hormones

In This Chapter

- Examining the structures of hormones
- Finding out about some important hormones
- Discovering how hormones function

We *know* with this chapter title you are expecting several jokes, but we'll try to restrain ourselves and play it straight.

Hormones are materials produced in one area of the body and used in a different area. They are molecular messengers that are created in certain glands in the body and then travel through the bloodstream to the target organ. Other substances, called *paracrine factors* or *growth and differentiation factors* (GDFs), also convey biochemical information within a particular organ (a bit like passing a note in class). This conveyance is accomplished by simple diffusion over a small distance. Some biochemical substances may be both a hormone and a paracrine factor.

The endocrine glands produce most — but not all — hormones. Endocrine glands include the hypothalamus, pituitary, pancreas, adrenal, liver, testes, and ovaries. Now surely that got your attention! Some glands produce a single hormone, whereas others produce more than one. The simplified viewpoint (and we are all about keeping it simple) is that the pituitary gland acts as the central control for the endocrine system. Hormones from the pituitary gland do cause other glands to produce hormones that affect other systems — however, there are glands that have the same effect on the pituitary gland.

Structures of Some Key Hormones

There are three groups of hormones:

- Proteins, such as insulin
- Steroids, materials derived from cholesterol
- Amines, such as epinephrine

These materials allow one part of the body to influence what occurs elsewhere. These molecules are so efficient that only very low concentrations, typically 10^{-7} to 10^{-10} M, are necessary. That's a really small amount! The low concentrations make identification and isolation of these substances difficult.

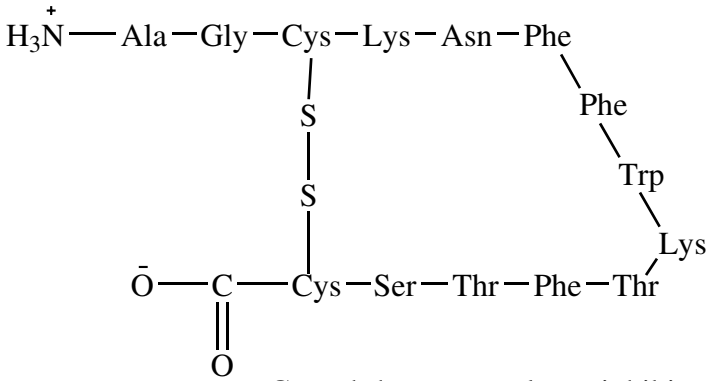
Proteins

The *protein,* or *polypeptide,* hormones, typically produced by the pituitary and hypothalamus glands, vary greatly in size — from simple tripeptides to larger molecules with more than 200 amino acid residues. Protein hormones are a diverse collection of molecules, including insulin (the structure of which you can see in Chapter 5).

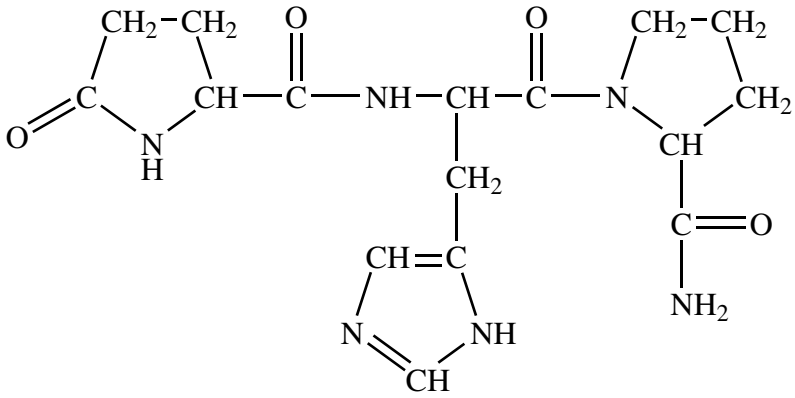
Others include the *thyrotropin-releasing factor,* which induces the release or production of a biochemical (thyrotropin, in this case). The thyrotropin-releasing factor hormone is a tripeptide containing glutamine (modified), histidine, and proline (modified). Another one is the *growth-hormone-release-inhibitory factor,* which inhibits the release or production of a chemical species. Together (Figure 11-1), these types of hormones provide a mechanism to start and stop an action. The idea is to maintain a tight biochemical control of biochemical processes, such as growth.

Steroids

You have no doubt read about steroid use among athletes, where it is used to increase muscle mass — to "pump-up," in other words. *Steroid* hormones, produced by the body's ovaries, testes, and adrenal glands, are cholesterol derivatives of about the same size as the parent molecule. They include the *estrogens* (female sex hormones), the *androgens* (male sex hormones), and the *adrenal cortical* hormones, such as aldosterone and cortisol. The estrogens and androgens are responsible for the development of the secondary sex characteristics of both females and males, respectively. These characteristics include enlargement of the breasts of females and development of facial hair in males.



Growth-hormone-release-inhibitory factor



Thyrotropin-releasing factor

Figure 11-1: Structures of the growth-hormone-release-inhibitory factor and thyrotropin-releasing factor.

The adrenal cortical hormones (Figure 11-2), which include the glucocorticoids and the mineralocorticoids, have a variety of functions. The glucocorticoids, such as cortisol, are important to several metabolic pathways. The mineralocorticoids, such as aldosterone, are important to the transport of inorganic species, such as sodium or potassium ions.

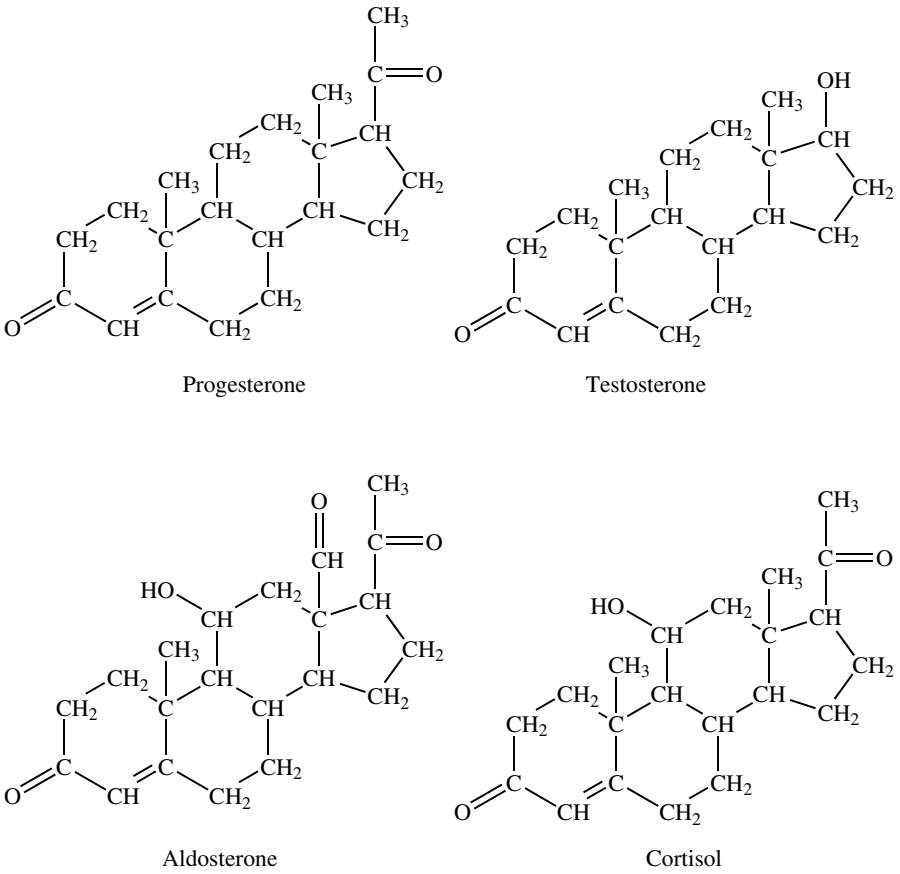
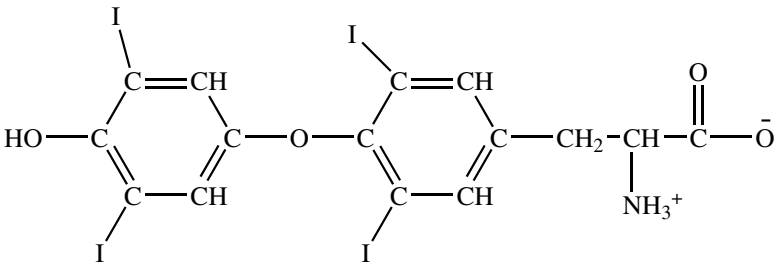


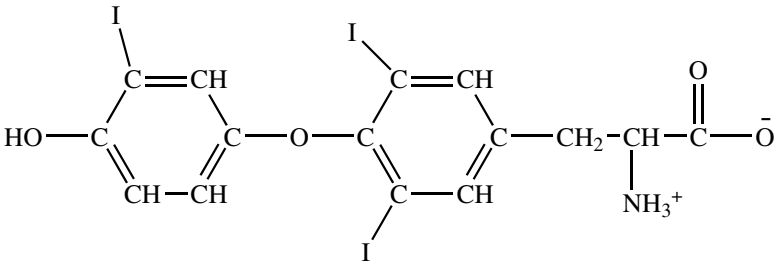
Figure 11-2: Structures of progesterone (an estrogen), testosterone (an androgen), and aldosterone and cortisol, adrenal cortical hormones.

Amines

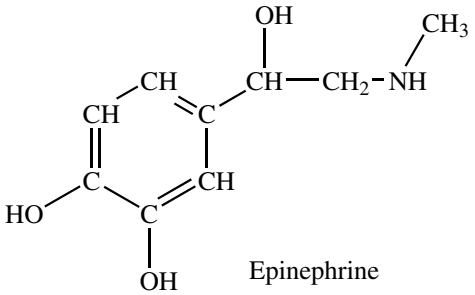
The *amine* hormones, typically produced by the thyroid and adrenal glands, are small molecules, many of which are derivatives of tyrosine. These hormones include *thyroxine* and *triiodthyronine,* produced by the thyroid gland, and *epinephrine* and *norepinephrine* produced by the adrenal gland. Figure 11-3 illustrates the structures of these hormones. Thyroxine and triiodothyronine are important metabolic-rate regulators. In fact, thyroxine is one of the most important substances in the body. It influences carbohydrate metabolism as well as protein synthesis and is involved in cardiovascular, brain, and renal function. Epinephrine and norepinephrine are important to control heart rate, blood flow, and metabolic rate.



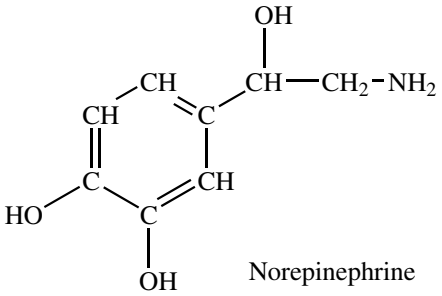
Thyroxine



Triiodothyroxine



Epinephrine



Norepinephrine

Figure 11-3: Structures of thyroxine, triiodothyronine, epinephrine, and norepinephrine.

Before and After: Prohormones

The synthesis of some hormones, like some enzymes, does not begin by producing the molecule in its active form. Instead, a *prohormone* forms, which remains unreactive and dormant until activated — sort of like us in the morning until we get our first cups of coffee. This process allows the body to build a store of a hormone for quick activation. Activating the prohormone requires less time than would the total synthesis of the molecule.

Proinsulin

Proinsulin is an example of a prohormone. *Insulin* is the hormone responsible for controlling blood sugar levels. Too much insulin results in a low blood sugar level (hypoglycemia), whereas too little insulin leads to elevated blood sugar levels (hyperglycemia). Your body needs to have a supply of insulin readily available for when you eat a piece of candy, such as a large chocolate-hazelnut truffle. If all of this insulin were already in your bloodstream, upon eating the candy you would become hypoglycemic. If the insulin were not present at all, you might become hyperglycemic until your body was able to synthesize sufficient insulin from the individual amino acids. Both hypoglycemia and hyperglycemia can lead to serious medical problems. The presence of a quantity of inactive insulin, ready to jump into action at a moment's notice, is the solution.

Bovine insulin (insulin produced from cows) contains two polypeptide chains, A and B, linked by disulfide linkages, with a total of 51 amino acid residues. Bovine proinsulin has 30 more amino acid residues than insulin does. Proinsulin is a single polypeptide chain with the insulin disulfide linkages. By removing a polypeptide sequence from the central region of this chain (residues 31–60), insulin forms. The excised portion originally connected one end of the A chain of insulin to the end of the B chain. The conversion of proinsulin to insulin requires the cleavage of two peptide bonds.

Angiotensinogen

Angiotensinogen is the prohormone of angiotensin II, a hormone that signals the adrenal cortex to release aldosterone. (In addition, angiotensin II is the most potent known vasoconstrictor.) The conversion of the prohormone to the hormone requires two steps. The first step uses the enzyme rennin. This peptidase, produced in the kidney, specifically cleaves a peptide bond between two leucine residues, the result of which is the decapeptide angiotensin I.

The second step utilizes the peptidase known as the angiotensin-converting enzyme. This enzyme, which occurs primarily in the lungs, cleaves the C-terminal dipeptide from angiotensin I to yield angiotensin II. These biochemical reactions can occur very rapidly, ensuring that the hormone can be quickly activated when needed by the body.

Fight or Flight: Hormone Function

The endocrine system, which generates the hormones, consists of a number of apparently unrelated organs: the liver, the ovaries or testes, the thyroid, the pancreas, and a number of other glands — components that are part of a complex, integrated network. A malfunction of one affects others.

Opening the letter: Hormonal action

Several mechanisms lead to the regulation of hormones. A *control loop* is the simplest. In many cases, one hormone stimulates the production of others so that many actions may occur before some type of control occurs.

Simple control loops

We are all familiar with *control loops.* You study for a test, but get a so-so grade. So you study harder for the next exam. Your grade provides *feedback,* causing your study habits to (hopefully) change. In the body, a control loop process begins with an external stimulus signaling a gland to generate a hormone. This hormone then influences its target site. Action by the target leads to a change, which signals the gland to stop. The action of the hormone causing the stop signal provides a negative feedback. An example of this type of loop is the production of insulin by the pancreas. The presence of high glucose levels in the bloodstream signals the pancreas to release insulin. The released insulin lowers the glucose level in the bloodstream. The reduced glucose level signals the pancreas to stop releasing insulin. The low glucose level is the negative feedback. This is a simplification; other factors may trigger the release of insulin. In addition, high glucose levels can trigger other biochemical functions, such as the synthesis of glycogen in the liver.

Hypothalamus-pituitary control

The hypothalamus-pituitary system is a very complex example of the other extreme of hormone control. The hypothalamus and the pituitary glands are in such close proximity that they behave almost as a single unit.

Initially, the central nervous system signals the hypothalamus to release a hormone called a *hormone-releasing factor,* which signals the pituitary. The pituitary, thus signaled, releases another hormone into the bloodstream. This hormone may target a specific organ or signal another part of the endocrine system to secrete yet another hormone. The presence of this final hormone serves as a negative feedback signal to the hypothalamus to stop secreting the hormone-releasing factor to the pituitary. Again, this is a simplistic view of a complicated system. An analogy might be your parents seeing your so-so exam grade. They freak out and force you to study harder. You are being influenced by an external force, in this case, your parents.

Figure 11-4 gives a more detailed representation of this system. Although the pituitary gland is known as the "master gland," this figure indicates that it is, in fact, the hypothalamus that deserves this honor.

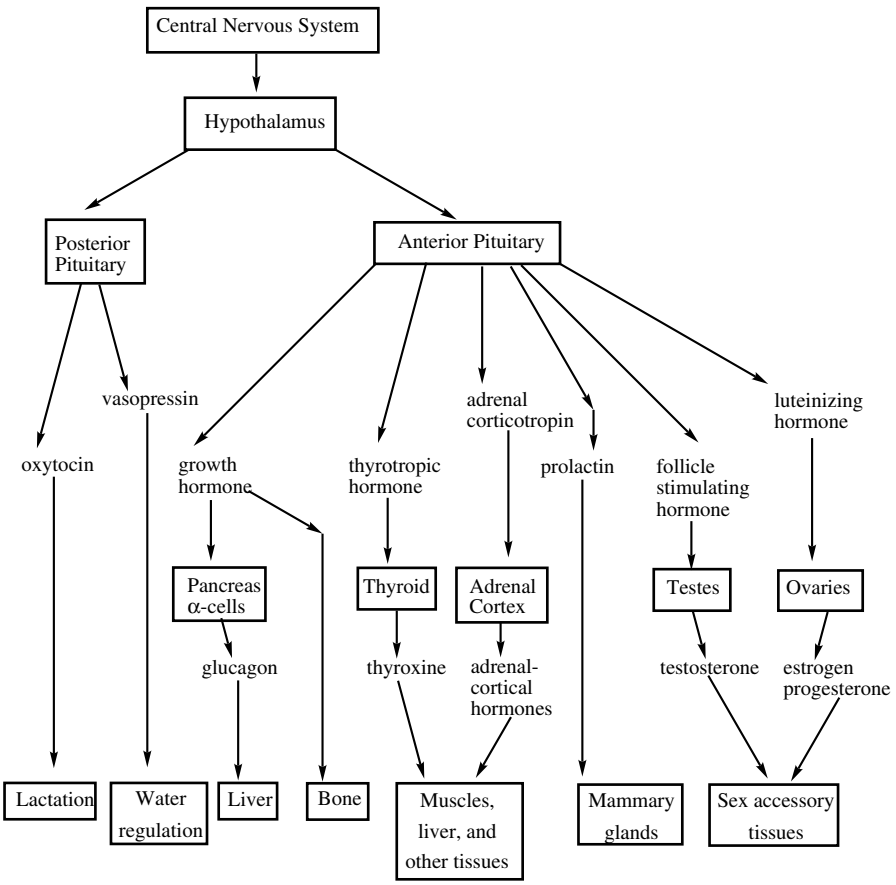


Figure 11-4: Schematic of hormone control in the body.

Models of hormonal action

Two models have been proposed to account for the molecular action of hormones. The first is the *two-messenger hypothesis,* which applies primarily to polypeptide and amine hormones. The other, *steroid hormonal action,* applies primarily to steroids. We use a simplistic approach (the KISS rule: Keep It Simple, Silly) for each model to emphasize their basic concepts. The actual processes involve many more changes.

The two-messenger model: Like the mail

Studies into the hormonal action of epinephrine (adrenaline) led to the development of this model. Later work indicated that the model applies to other hormonal systems as well. In the two-messenger model, a hormone binds to a receptor site on the exterior of a cell. This binding induces the release of another agent within the cell. The hormone is the first messenger, and the other agent in the second messenger.

For example, the adrenal medulla releases epinephrine, the "fight or flight" hormone, in vertebrate animals. This release initiates a number of responses, including glycogenolysis, the breakdown of glycogen. Glycogenolysis releases glucose for use in rapid energy production. As with other hormones, the concentration of hormone required is very low. For epinephrine, it is about 10^{-9} M. The released epinephrine acts as the first messenger (the extracellular one). Molecules of epinephrine bind to specific receptor sites on the surface of the target cells — primarily the skeletal muscles and, to a lesser extent, liver. The binding of epinephrine to the outside of liver cells induces the enzyme adenylate cyclase, bound to the interior of the cell membrane, to synthesize cyclic AMP (see Figure 11-5). Cyclic AMP, or cAMP, is the second messenger (the intracellular one). The second messenger initiates a series of events terminating in the release of glucose (glycogenolysis).

Initially, the cAMP binds to the regulatory subunit of protein kinase, and this activates the membrane-bound enzyme. The released protein kinase then activates phosphorylase kinase. This process requires calcium ion and ATP. (Muscular action releases calcium ion, which aids the process.) Phosphorylase kinase, with aid of ATP and magnesium ion, converts inactive phosphorylase b to active phosphorylase a. The increased presence of this enzyme accelerates the breakdown of glycogen with the release of D-glucose-1-phosphate. Phosphoglucomutase then converts D-glucose-1-phosphate to D-glucose-6-phosphate. Finally, D-glucose-6-phosphatase catalyzes the loss of the phosphate to release glucose, which may be used in the cell or, more importantly, may enter the bloodstream. Whew!

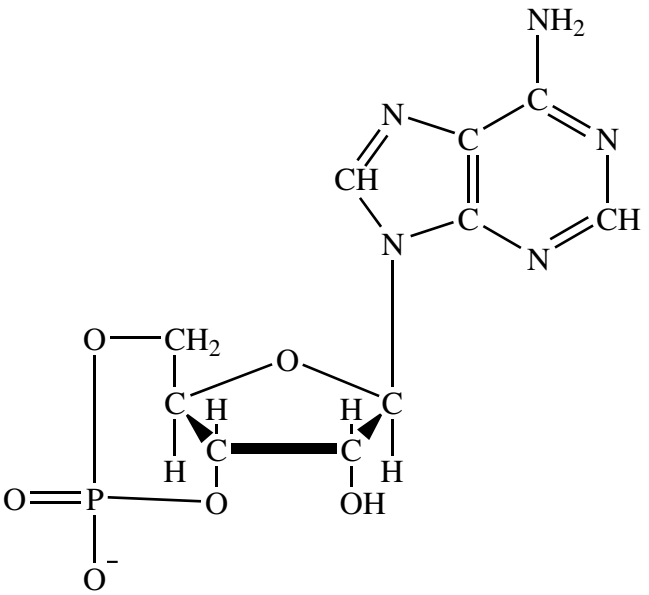


Figure 11-5: Structure of cyclic AMP.

The enzyme protein kinase also catalyzes the conversion of glycogen synthase (active) to phosphor-glycogen synthase (inactive). Thus, while the level of protein kinase is high, the production of new glycogen ceases. The inhibition of glycogen synthesis also means that more glucose will be available for rapid actions, such as running away from an angry pit bull.

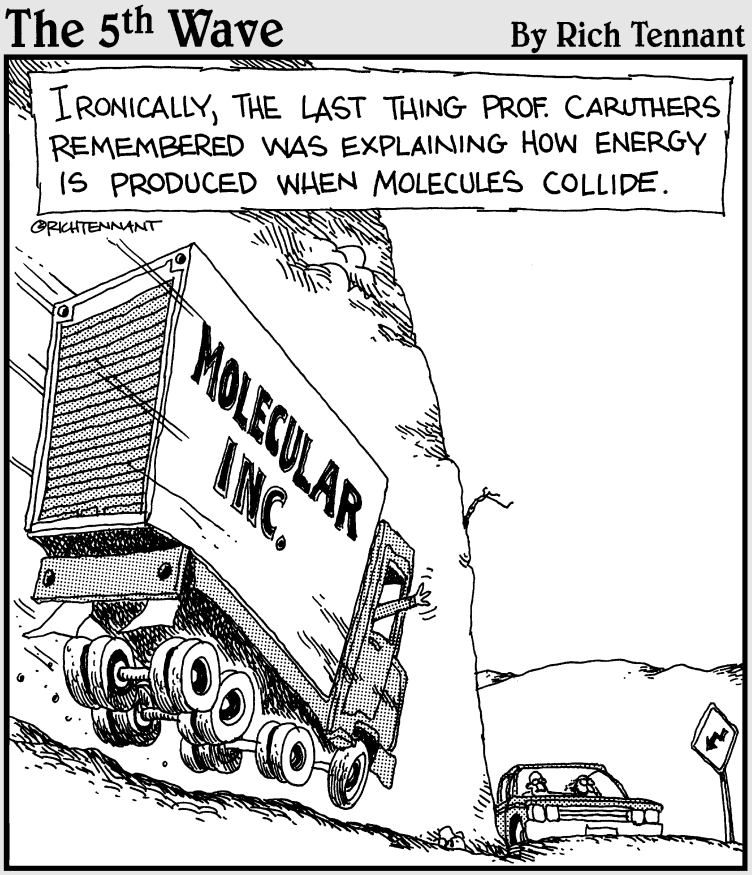
Steroid hormonal action

Unlike hormones in the two-messenger system, steroid hormones cross the membrane and enter the cell. This mechanism applies to other hormones as well, such as thyroid hormones, in addition to the steroid hormones.

The first system described by this model was the action of estradiol on uterine tissue in mammals. The estradiol, an estrogen, enters the cell where it binds to an estrogen-receptor protein. The binding does not involve covalent bonding, but induces instead a conformational change in the protein. The change in the shape of the protein allows it to pass through the "door" into the cell nucleus. The hormone-protein complex then enters the cell nucleus where it binds to a specific site on a chromosome. This binding to the chromosome stimulates transcription to produce mRNA, which, in turn, exits the nucleus and synthesizes protein molecules through translation.

Three basic factors differentiate the steroid system from the two-messenger system. First, in the steroid system the hormone enters the cell. Second, there is a specific receptor molecule within the cytosol, the fluid inside the cell. Finally, the hormone action is at the chromosome level.

Part IV
Bioenergetics and Pathways



In this part . . .

For anyone to do anything requires energy, and this is where we focus on the way life obtains and uses it. Here we take a gander at energy needs and follow the trail of where that energy goes and why. The main character in this part is your good buddy ATP, and running through this episode is where you'll find the citric acid cycle. Finally, we tackle nitrogen chemistry.

Chapter 12

Life and Energy

In This Chapter

- Learning about ATP and energy
- Visiting the nucleoside triphosphate family
- Considering AMP, ADP, and ATP
- Going without food

The chapters in this part examine metabolism — all the processes involved in maintaining a cell. Metabolism has two components: catabolism and anabolism. *Catabolism* deals with the breaking down of molecules, whereas *anabolism* deals with the building up of cells. Both processes take place in the mitochondria. All metabolic processes involve energy: They either absorb energy (*endergonic*) or produce it (*exergonic*).

The key energy molecule is *adenosine triphosphate,* abbreviated *ATP,* which forms as a product of the common catabolic pathway.

ATP: The Energy Pony Express

Determining the basic reaction processes involved in the production and use of energy is called *bioenergetics*. This study has developed bioenergetic principles that allow us to examine energy at the microscopic level.

Fortunately, ATP is recycled within the body. The typical daily requirement for an adult is over 140 pounds of ATP per day. However, the amount of ATP present in your body at any one time is only about one-tenth of a pound. That means each ATP molecule in your body is recycled about 1,400 times each day. Now that is effective recycling — and you don't even have to put anything into a blue container.

ATP and free energy

The *free energy content* (G) is the intrinsic energy present in a molecule. In a reaction, the change in this energy is written as ΔG. The change in energy is equal to the energy of the products minus the energy of the reactants. The value of ΔG is the key: If a reaction produces energy, ΔG represents the maximum possible amount of energy that the reaction may produce. If a reaction requires energy, ΔG represents the minimum possible amount of energy that a reaction will require. Reactions producing energy have a negative value of ΔG and are *spontaneous.* Reactions requiring energy have a positive value of ΔG and are *nonspontaneous.*

Spontaneity bears no relation to speed. Spontaneous reactions may be very rapid or very slow.

The conditions under which a reaction occurs may alter the value of ΔG. (The "ideal" or standard value of ΔG is ΔG°.) The formula for modifying the free energy for the equilibrium reaction A ⇌ B is:

$$\Delta G = \Delta G^\circ - RT \ln [B] / [A] = \Delta G^\circ - RT \ln K$$

According to this relationship, the free energy change, ΔG, comes from a modification of the standard free energy value. R is the universal gas constant ($8.314\ J \times mol^{-1}K^{-1}$ or $1.987\ cal \times mol^{-1}K^{-1}$). T is the absolute temperature. K is the equilibrium constant found by dividing the concentration of the product, [B], by the concentration of the reactant, [A].

In many bioenergetic studies, *calories* are the unit instead of joules (J). The relationship is 1 calorie = 4.184 J (exactly) or 1 kilocalorie = 4.184 kJ.

In research, it is often better to use ΔG°'. This modification of ΔG stems from the use of the biologically more realistic value of pH = 7 ($[H^+] = 10^{-7}$ M) instead of the standard pH = 0 ($[H^+] = 1$ M). Some relationships between K and ΔG°' are shown in Table 12-1.

Table 12-1 Relationships Between ΔG°' and K

$\Delta G^{\circ\prime}\ kJ \times mol^{-1}$	K
–17.1	1,000
–11.4	100
–5.7	10
0	1

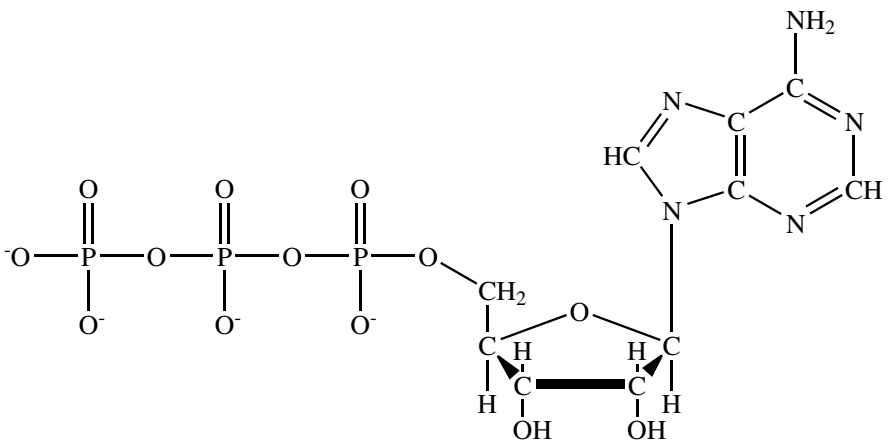
$\Delta G^{\circ\prime}$ kJ × mol^{-1}	K
5.7	0.1
11.4	0.01
17.1	0.001

Table 12-1 shows that the larger K is, the more *exergonic* (spontaneous) the reaction. For example, if K = 1000, the concentration of the product, [B], is 1,000 times that of the reactant, [A], and 17 kJ per mole will be released. It is important to remember that, in biological systems, variations in [A] and [B] must be taken into account in addition to $\Delta G^{\circ\prime}$. For example, increasing the reactant concentration promotes the reaction, whereas increasing the product concentration inhibits the reaction.

ATP as an energy transporter

Cells utilize exergonic processes to provide the energy necessary for life processes, and the key supplier of this energy is ATP (Figure 12-1). ATP supplies the energy required to force endergonic reactions to take place, to provide mechanical energy (muscle movement), light energy (in fireflies), and heat energy (to maintain body temperature).

Hydrolysis of the terminal phosphate of ATP yields ADP and inorganic phosphate, indicated as P_i. The structure of ADP is shown in Figure 12-2. This hydrolysis releases 30.5 kJ × mol^{-1}.



Adenosine triphosphate (ATP)

Figure 12-1: Structure of ATP.

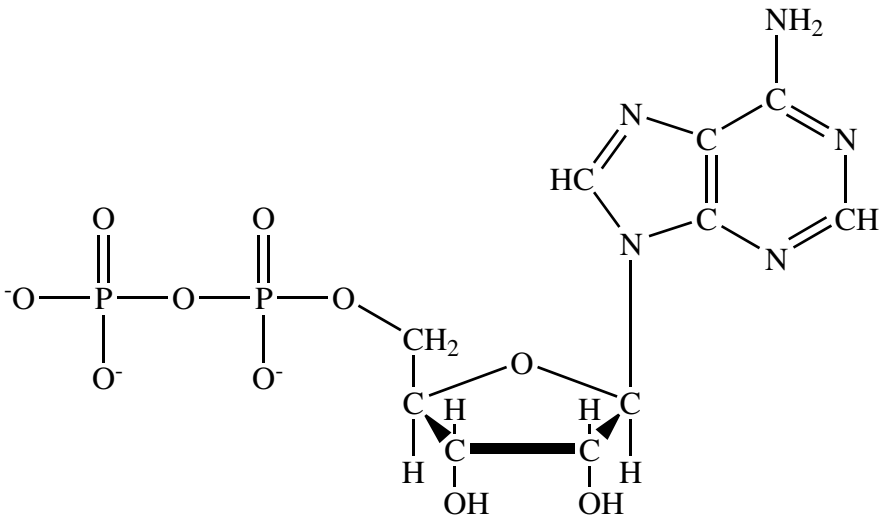


Figure 12-2: Structure of ADP.

Adenosine diphosphate (ADP)

Concentration variations lead to changes, usually minor, in energy.

Hydrolysis of the terminal phosphate of ADP yields AMP and inorganic phosphate, indicated as P_i. The structure of AMP is in Figure 12-3. This hydrolysis also releases 30.5 kJ × mol^{-1}. (This reaction is of less biological importance than the ATP to ADP hydrolysis.)

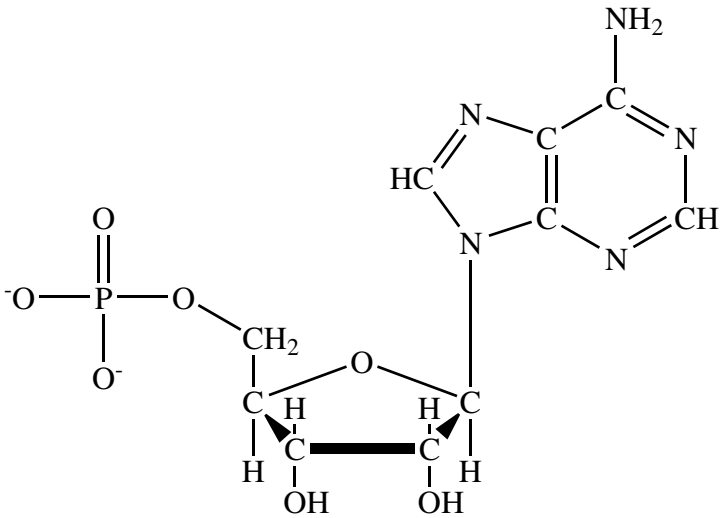
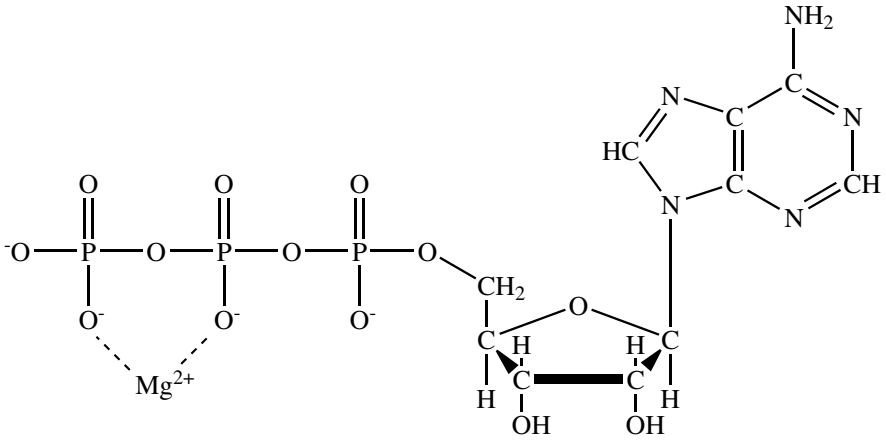


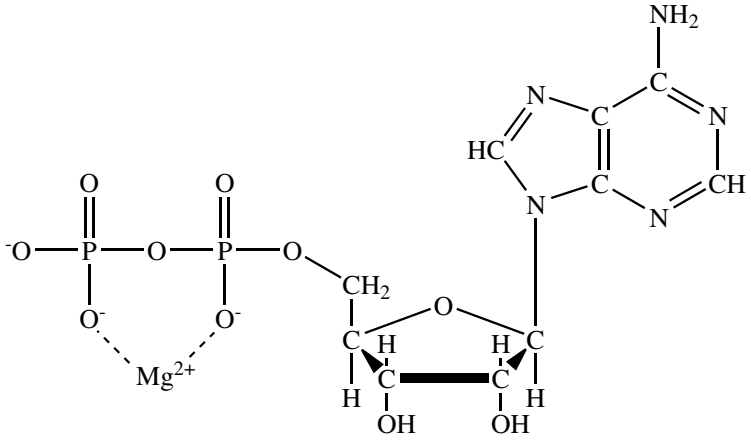
Figure 12-3: Structure of AMP.

Adenosine monophosphate (AMP)

It is also possible to go directly from ATP to AMP, cleaving a pyrophosphate, $P_2O_7^{4-}$, from the phosphate chain. Biochemists use PP_i to indicate pyrophosphate. This furnishes slightly more energy then a simple hydrolysis to release P_i (about 33.5 kJ × mol^{-1}). Under physiological conditions, the phosphate portions of ATP and ADP form a complex with magnesium ions. In certain circumstances, manganese (II) ions, Mn^{2+}, may take the place of Mg^{2+} ions. Figure 12-4 depicts the magnesium complexes with ATP and ADP.



Adenosine triphosphate (ATP)-Mg^{2+}



Adenosine diphosphate (ADP)-Mg^{2+}

Figure 12-4: Magnesium complexes with ATP and ADP.

The removal of the last phosphate involves the loss of the least amount of energy (14.2 kJ•mol^{-1}). This hydrolysis involves the cleavage of an ester bond instead of an anhydride bond. In general, the hydrolysis of an ester bond involves less than half the energy of the hydrolysis of an anhydride bond.

It's Relative: Molecules Related to ATP

A few other biomolecules can provide energy equivalent to that which comes from the hydrolysis of ATP. GTP is an example of such a molecule. There are also a few molecules that supply *more* energy. Table 12-2 compares some of the high-energy molecules to ATP, and Figure 12-5 shows their structures.

Table 12-2 Energy Released ($\Delta G°'$) by Some High-Energy Molecules Related to ATP

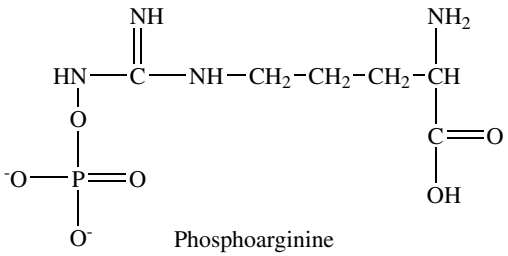
Biomolecule	*Energy released (kJ × mol^{-1})*
ATP	30.5
Phosphoarginine	32.2
Acetyl phosphate	43.3
Phosphocreatine	43.3
1,3-Bisphosphoglycerate	49.6
Phosphoenolpyruvate	62.2

Phosphopyruvate, 1,3-bisphosphogylcerate, and acetyl phosphate are important for the transfer and conservation of chemical energy. Phosphoarginine and phosphocreatine are important molecules for storing metabolic energy. Phosphocreatine is stored in muscles and can be quickly converted to ATP to give energy for muscle contraction. Production of phosphocreatine occurs when ATP concentration is high — high ATP concentration is needed to overcome the energy deficit of 12.8 kJ × mol^{-1}. The reverse, phosphate transfer to form ATP from ADP, occurs at low ATP concentrations. Phosphoarginine behaves similarly in certain invertebrates.

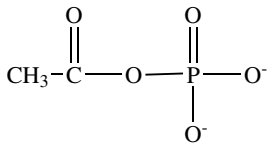
The nucleoside triphosphate family

The predominant energy transfer molecule, as we have been saying, is ATP. But other nucleoside triphosphates (such as CTP, GTP, TTP, and UTP) may also serve this energy transfer function. These five molecules also supply part of the energy necessary for DNA and RNA synthesis. All the nucleoside

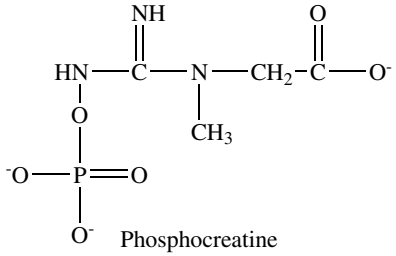
triphosphates have about the same energy yield. (Note that ATP is necessary for the synthesis of the remaining nucleoside triphosphates.)



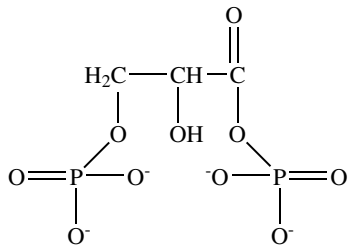
Phosphoarginine



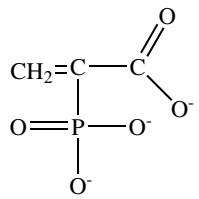
Acetyl phosphate



Phosphocreatine



1,3-Bisphosphoglycerate



Phosphoenolpyruvate

Figure 12-5: Structures of some high-energy molecules.

The biosynthesis of the ribonucleoside triphosphates, in general NTP, begins with the production of the appropriate monophosphate, NMP. The stepwise addition of the next two phosphate groups requires two enzymes of low specificity. These enzymes are nucleoside monophosphate kinase and nucleoside diphosphate kinase. (The term *kinase* refers to a transferase enzyme that transfers a phosphate group of a nucleoside triphosphate.) The general reactions are shown in Figure 12-6.

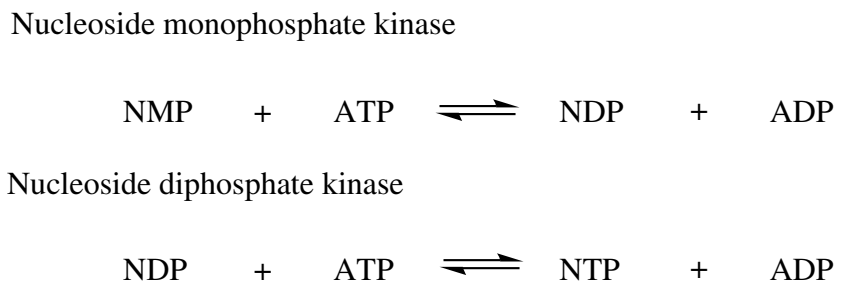


Figure 12-6: Two of the reactions catalyzed by the kinase enzymes.

The formation of the deoxyribonucleoside triphosphates, dNTP, follows two different paths. In one path, a multienzyme system converts the appropriate nucleoside diphosphate to the corresponding deoxyribonucleoside diphosphate. Then nucleoside diphosphate kinase catalyzes the formation of the deoxyribonucleoside triphosphate. The other path occurs in certain microorganisms where there is a direct conversion of NTP to dNTP.

As easy as 1, 2, 3: AMP, ADP, and ATP

It is possible to hydrolyze ATP either to ADP plus phosphate, P_i, or to AMP plus pyrophosphate, PP_i. (The pyrophosphate will undergo further hydrolysis to two phosphates, 2 P_i.) ADP and P_i are the immediate precursors for the reformation of ATP. To produce ATP starting with AMP utilizes the enzyme adenylate kinase. This enzyme catalyzes the transfer of a phosphate group from an ATP to an ADP. This reaction results in the formation of two ADP molecules. (Adenylate kinase also catalyzes the reverse reaction.)

The easy transfer of phosphate groups between nucleotides creates a metabolic network for the transfer of energy. The key to this network is the intercellular production of ATP.

Where It All Comes From

One of the purposes of the food we eat, of course, is to supply energy, with carbohydrates and fats being the major sources of energy. Digestion breaks polysaccharides into glucose and other monosaccharides, whereas fats are broken into glycerol and fatty acids. Catabolism converts these energy sources primarily to ATP. Proteins are broken into amino acids, which usually do not serve as energy sources. (We explain the details of these reactions later in this book.) Glucose produces 36 ATP molecules. This is an average of 6 ATPs per carbon. The step-by-step energy change for glucose is in Table 12-3. Other carbohydrates give a similar yield.

Table 12-3 ATP Yield for Each Step in the Metabolism of Glucose

Chemical Steps	*Number of ATP Molecules Produced*
Activation (conversion of glucose to 1,6-fructose diphosphate)	-2
Oxidative phosphorylation 2(glyceraldehyde 3-phosphate $\rightarrow$ 1,3-diphosphoglycerate), producing 2 NADH + H^+ in cytosol	4
Dephosphorylation 2(1,3-diphosphoglycerate $\rightarrow$ pyruvate)	4
Oxidatie decarboxylation 2(pyruvate $\rightarrow$ acetyl CoA), producing 2 NADH + H^+ in mitochondrian	6
Oxidation of two C_2 fragments in citric acid and oxidative phosphorylation common pathway, producing 12 ATP for each C_2 fragment	24
Total	36

Each fat molecule hydrolyzes to a glycerol and three fatty acid molecules. Glycerol produces 20 ATPs per molecule. The energy production from a fatty acid will vary with the identity of the particular acid. Stearic acid, $C_{18}H_{36}O_2$, produces a total of 146 ATPs per molecule. This amounts to an average of 8.1 ATPs per carbon. The step-by-step energy change for stearic acid is shown in Table 12-4. Other fatty acids give a similar yield.

Table 12-4 ATP Yield for Each Step in the Metabolism of Stearic Acid

Chemical Steps	*Happens*	*ATP Molecules Produced*
Activation (stearic acid → stearyl CoA)	Once	-2
Dehydrogenation (acetyl CoA → transenoyl CoA), producing $FADH_2$	8 times	16
Dehydrogenation (hydroxyacyl CoA → keto acyl CoA), producing $NADH + H^+$	8 times	24
C_2 fragment (acetyl CoA → common catabolic pathway), producing 12 ATP per C_2 fragment	9 times	108
Total	146	

What happens if you stop eating?

Starvation is the total deprivation of food. Here is what happens during starvation: Initially, the body utilizes its glycogen reserves. Then it moves on to its fat reserves — the first ones are those around the heart and kidneys. Finally, the body relies on the reserves found in the bone marrow. Early in a total fast, the body metabolizes protein at a rapid rate. The amino acids are converted to glucose, because the brain prefers glucose. These proteins come from the skeletal muscles, blood plasma, and other sources in a process that produces a quantity of nitrogen-containing products, which need to be excreted. Excretion requires large quantities of water, and the resulting loss of water may lead to death by dehydration. If the starvation continues, the brain chemistry adjusts to accept fatty acid metabolites, which uses the last of the fat reserves. Finally, the body resorts to structural proteins, systems begin to fail rapidly, and death follows quickly.

Chapter 13

ATP: The Body's Monetary System

In This Chapter

- Checking out carbohydrate metabolism and examining the citric acid cycle
- Finding out about electron transport and oxidative phosphorylation
- Seeing how biosynthesis takes place

Here we examine a number of general processes that either produce or consume energy. Breaking down molecules often produces energy. The breakdown of one molecule is often coupled with the synthesis of another, and this other synthesized molecule is often adenosine triphosphate, or ATP. *Catabolism* is the breaking down of molecules to provide energy. *Anabolism* is the building of molecules. These two processes combine to give metabolism. *Metabolism* comprises all reactions in biological systems.

As you can see in Chapter 12, the "currency" in biological systems is ATP. There are other energy-containing molecules, but the rate of exchange to ATP is the reference. The breakdown of certain molecules produces the currency of ATP, and there is a cost involved in the synthesis of other molecules. Polysaccharides and fats are like "banks" that store energy for later use.

Metabolism 1: Glycolysis

The *Embden-Meyerhof pathway,* or *glycolysis,* is a primitive means of extracting energy from organic molecules. The process converts glucose to two lactic acid molecules in an anaerobic (without oxygen) process. Nearly all forms of life, whether a person or a jellyfish, utilize glycolysis. All carbohydrates follow this pathway. Aerobic (utilizing oxygen) processing of carbohydrates uses pyruvate derived from glycolysis. (Alcoholic fermentation also produces pyruvate from glucose. The glucose is converted to two ethanol molecules and two CO_2 molecules.) Glycolysis is a two-part process, which we label Phase I and Phase II. Figures 13-1 and 13-2 help illustrate the upcoming, ahem, rather *involved* discussion. You may want to refer back to these figures as you read.

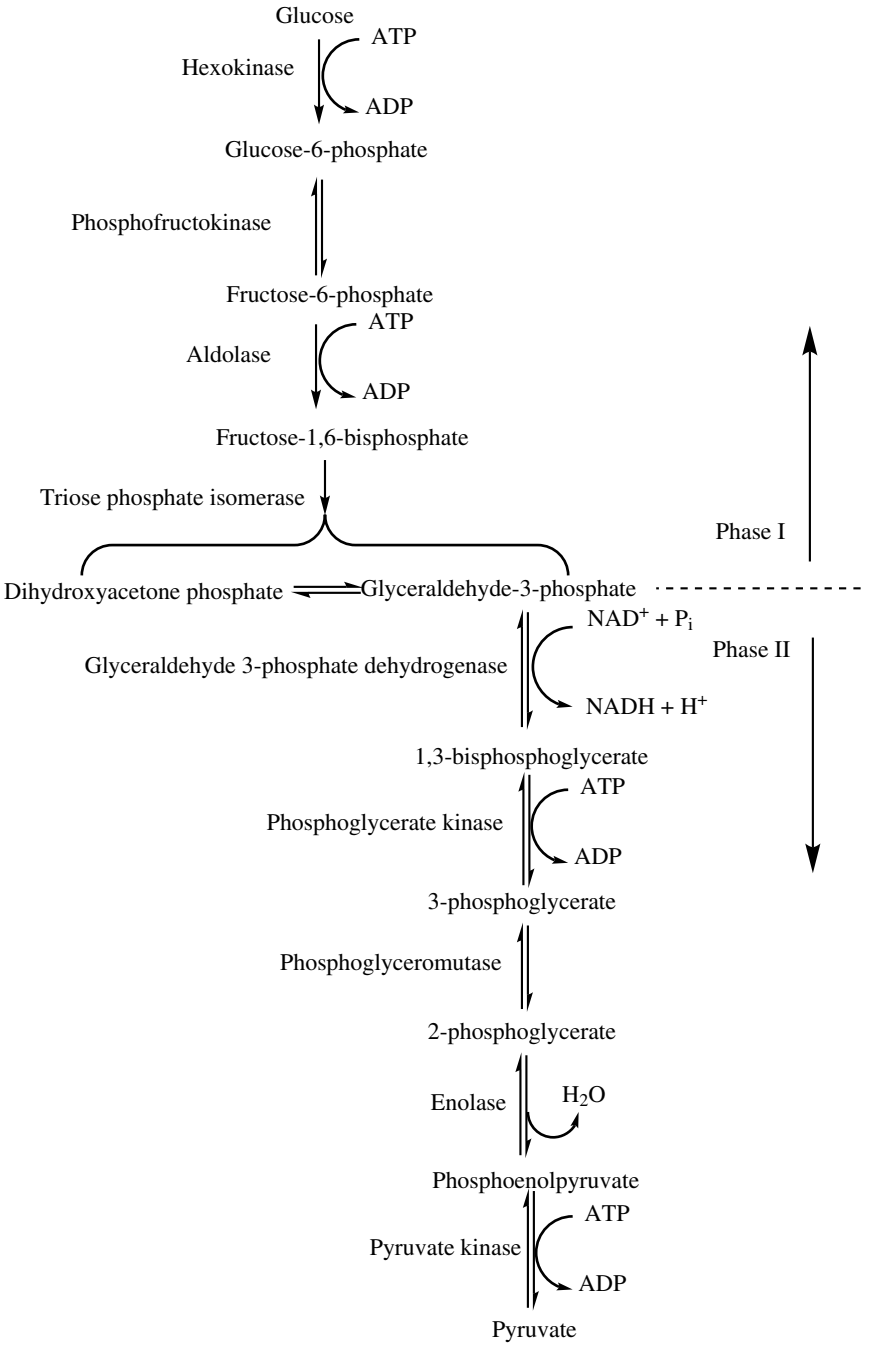


Figure 13-1: Steps in glycolysis.

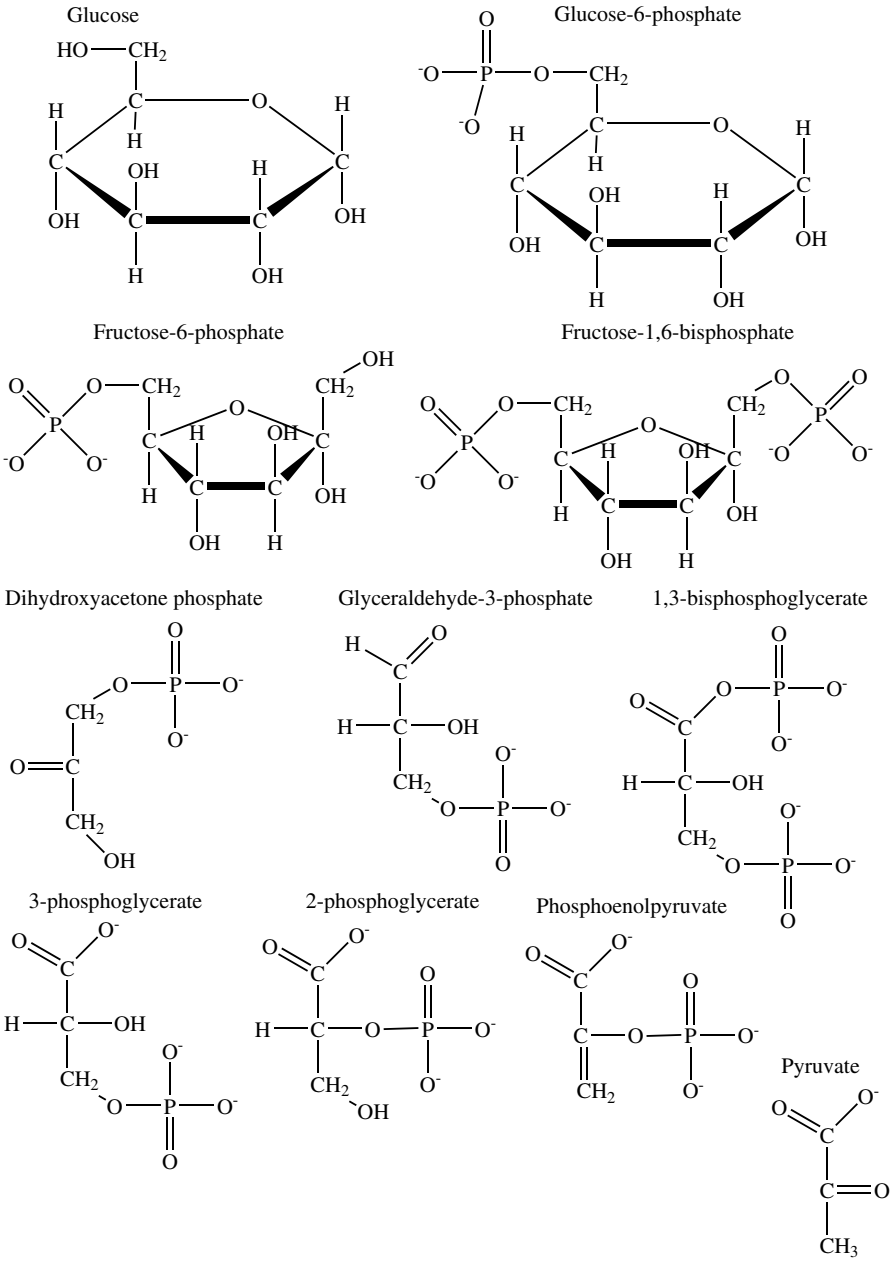


Figure 13-2: Molecules involved in glycolysis.

Glucose: Where it all starts

As we mentioned, glycolysis occurs in two phases: Phase I and Phase II.

Phase 1

As glucose enters the cell, it undergoes immediate phosphorylation to glucose-6-phosphate — the first step in Phase I. The phosphate comes from ATP, and the enzyme hexokinase, with the aid of Mg^{2+}, catalyzes the transfer. Thus, the first step in the production of energy requires an investment of energy, which is necessary to activate the glucose in a reaction that is not easy to reverse. In addition, the presence of the charged phosphate group makes it difficult for this and other intermediates to diffuse out of the cell.

The enzyme phosphoglucose isomerase then catalyzes the isomerization of glucose-6-phosphate to fructose-6-phosphate. This results in a compound with a primary alcohol group, which is easier to phosphorylate than the hemiacetal originally present. Fructose-6-phosphate then reacts with another molecule of ATP to form fructose-1,6-bisphosphate. The enzyme for this step is phosphofructokinase — (try saying that ten times fast!) — and this enzyme requires Mg^{2+} to be active. This is the major regulatory step in glycolysis. ATP inhibits this enzyme, whereas AMP activates it.

Aldolase enzymatically cleaves the fructose-1,6-bisphosphate into two triose phosphates. These triose phosphates are dihydroxyacetone phosphate and glyceraldehyde-3-phosphate. The dihydroxyacetone phosphate isomerizes to glyceraldehyde-3-phosphate to complete Phase I. Triose phosphate isomerase catalyzes this isomerization. (You see why we suggested following along with Figures 13-1 and 13-2?)

The net result of Phase I is the formation of two molecules of glyceraldehyde-3-phosphate, which costs two ATP molecules, and produces no energy.

Phase 11

Phase II begins with the simultaneous phosphorylation and oxidation of glyceraldehyde-3-phosphate to form 1,3-bisphosphoglycerate. Glyceraldehyde-3-phosphate dehydrogenase catalyzes this conversion. Inorganic phosphate is the source of the phosphate. NAD^+ is the coenzyme and oxidizing agent. NAD^+ reduces to NADH.

There is a high-energy acyl phosphate bond present in 1,3-bisphosphoglycerate. Phosphoglycerate kinase, in the presence of Mg^{2+}, catalyzes the direct transfer of phosphate from 1,3-bisphosphoglycerate to ADP. This results in the formation of ATP and 3-phosphoglycerate. Because the formation of ATP involves direct phosphate transfer, this process is called *substrate-level phosphorylation* to avoid confusion with *oxidative phosphorylation* (discussed later). Phosphoglyceromutase then catalyzes the transfer of a phosphate

group from C-2 to C-3, thus converting 3-phosphoglycerate to 2-phosphoglycerate. After that, dehydration occurs to form phosphoenolpyruvate (PEP), which contains a nigh-energy phosphate bond. The enzyme catalyzing the reaction is enolase.

The final, irreversible step is a second substrate-level phosphorylation. Here, an ADP molecule receives a phosphate group from the PEP. The enzyme pyruvate kinase is necessary for this step. This enzyme requires not Mg^{2+}, but also K^{+}. Pyruvate is the other product. Whew!

During Phase II, two molecules of glyceraldehyde-3-phosphate (from Phase I) form two molecules of pyruvate with the formation of four molecules of ATP and two molecules of NADH.

The pyruvate produced by glycolysis has several fates. When there is plenty of oxygen, the pyruvate enters the Krebs cycle, the electron transport chain, and oxidative phosphorylation pathways as Acetyl-CoA. This results in the production of more ATP and the total conversion to CO_2. If oxygen is lacking, vertebrates (you included) convert pyruvate to a related substance, lactate. Other organisms, such as yeast, convert pyruvate to ethanol and CO_2 — and that is why we have beer. These latter two possible fates yield less energy than the oxygen-rich fate.

Miles per gallon? Energy efficiency

Glycolysis is the initial conversion of carbohydrate to energy. After that there is the production of two ATP molecules, two NADH molecules, and two pyruvate molecules. The energy content of the ATP molecules is only 2 percent of the total energy present in each glucose molecule. This shows the relative inefficiency of anaerobic energy production. Fortunately, the pyruvate molecules will undergo further aerobic oxidation to increase this energy yield. The total energy output of anaerobic and aerobic oxidation of glucose is 30–32 ATP molecules, which accounts for about 30 percent of the total energy present in glucose. Much of the remaining energy is available as heat for warm-blooded animals.

Going in reverse: Gluconeogenesis

Gluconeogenesis is a series of reactions that generate glucose from non-carbohydrate sources. This pathway is necessary when the supply of carbohydrates is inadequate (something that is rare in our lives). The non-carbohydrate sources include lactate, pyruvate, some amino acids, and glycerol. In many ways, gluconeogenesis is the reverse of glycolysis. Figure 13-3 summarizes the steps of gluconeogenesis. (The formation of glucose in plants utilizes the process of photosynthesis.)

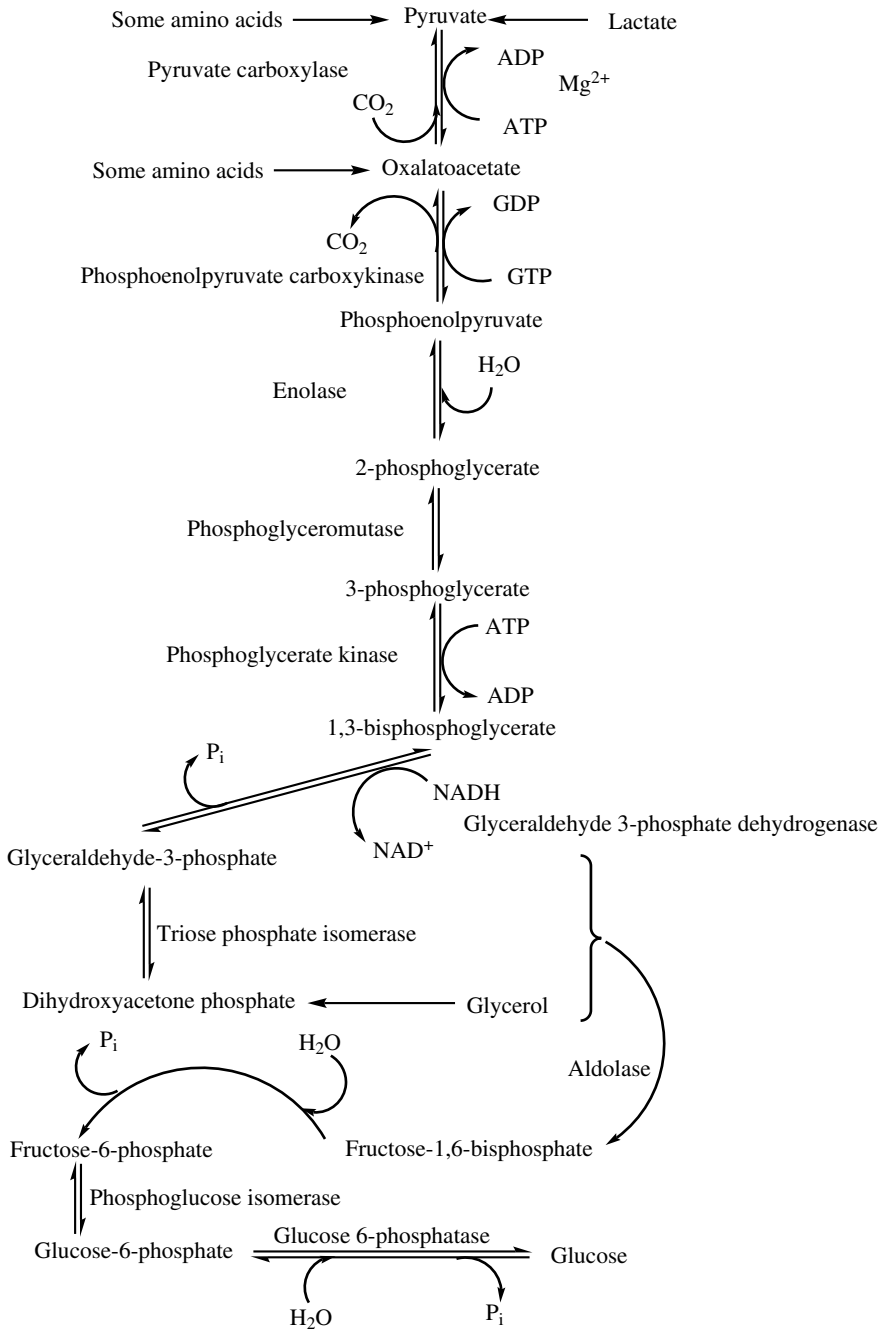


Figure 13-3: Steps in gluconeogenesis.

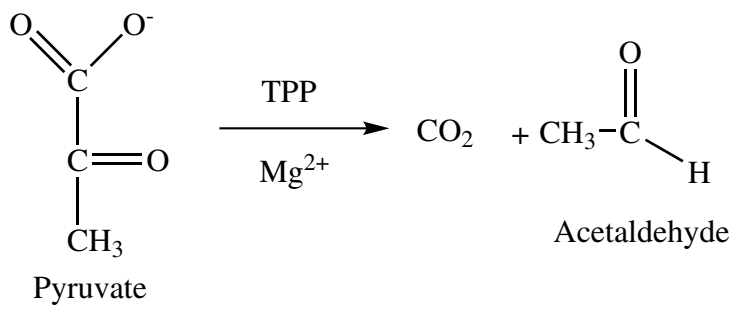
The presence of many of the same intermediates enables the use of many of the same enzymes in both glycolysis and gluconeogenesis. The differences (four enzymes) between the two systems allow regulation, so that the processes don't cancel each other. Regulation is also possible by isolating the two pathways in different organs. Other carbohydrates may also form.

Alcoholic fermentation: We'll drink to that

Under anaerobic conditions, yeast and other organisms convert pyruvate to ethanol and carbon dioxide. This process is accompanied by the oxidation of NADH to NAD^+. The NAD^+ is used in glycolysis. During this process, there is a net generation of two ATP molecules.

The first step in alcoholic fermentation is the decarboxylation of pyruvate to carbon dioxide and acetaldehyde. The enzyme pyruvate decarboxylase, along with the cofactors Mg^{2+} and TPP (thiamin pyrophosphate), catalyze this step. The enzyme alcohol dehydrogenase, along with the coenzyme NADH, catalyzes the conversion of acetaldehyde to ethanol. Makes you really appreciate that shot of tequila, doesn't it? Figure 13-4 summarizes these steps.

1. Pyruvate decarboxylase reaction



2. Alcohol dehydrogenase reaction

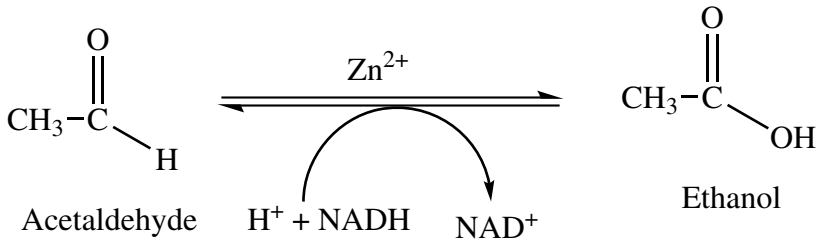


Figure 13-4: Steps in alcoholic fermentation.

Metabolism II: Citric Acid (Krebs) Cycle

The *citric acid cycle* and *oxidative phosphorylation* are the aerobic processes of catabolism that produce energy (ATP). The citric acid cycle is also known as the *Krebs cycle* and also as the *tricarboxylic acid cycle* (TCA). The primary entry molecule for this series of reactions is acetyl-CoA (short for acetyl-coenzyme A). The sources of acetyl-CoA are pyruvate from glycolysis, certain amino acids, or the fatty acids present in fats. The structure of acetyl-CoA is shown in Figure 13-5. Note: these processes take place in the *mitochondria,* the energy factories of the cell.

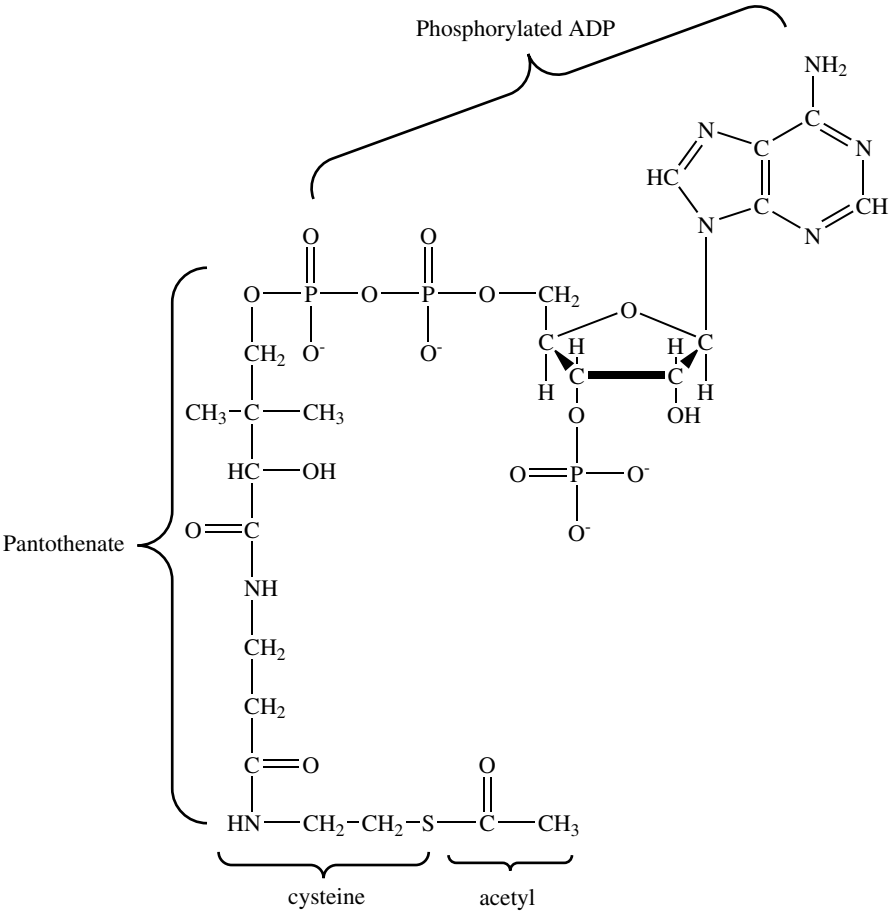


Figure 13-5: Structure of acetyl-CoA.

In addition to being an energy source, acetyl-CoA is the starting material for the synthesis of a number of biomolecules. In the next few sections, we discuss the citric acid cycle. The general cycle is shown in Figure 13-6, and the structures are shown in Figure 13-7.

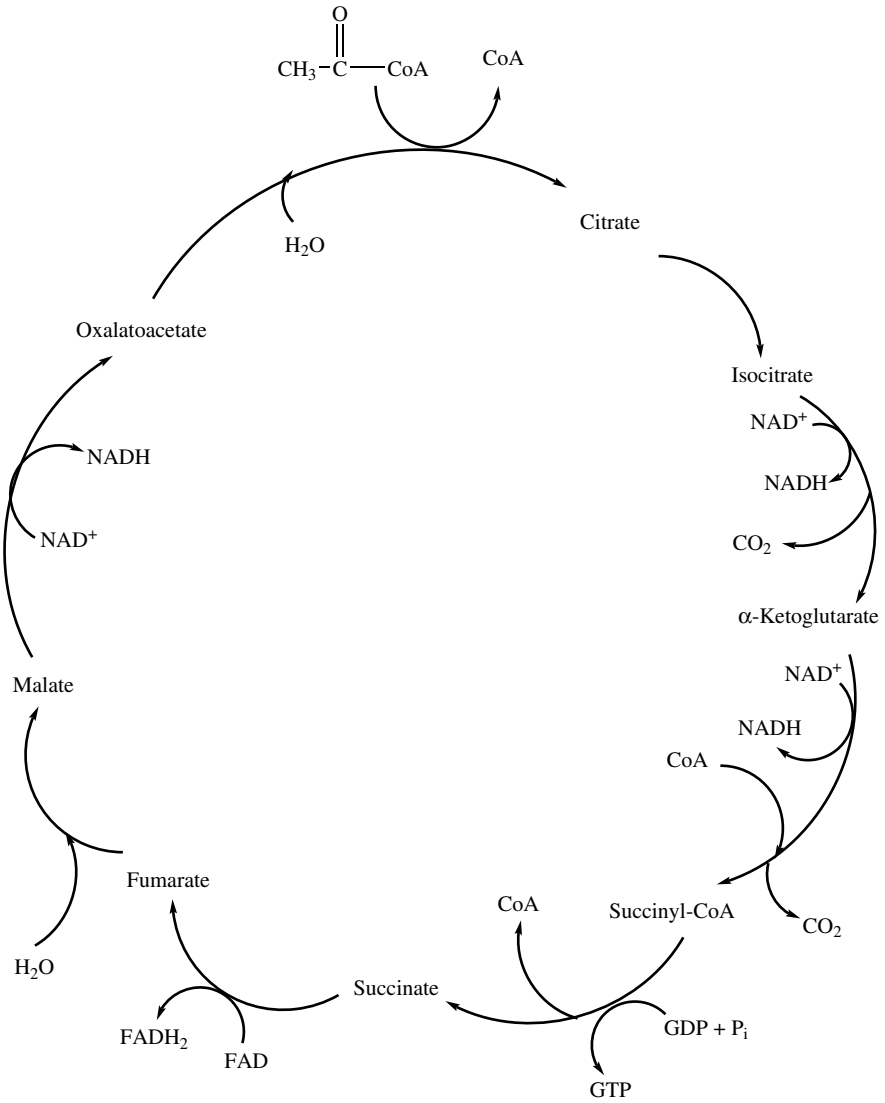
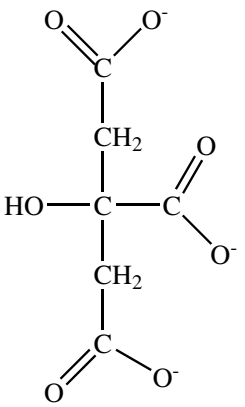
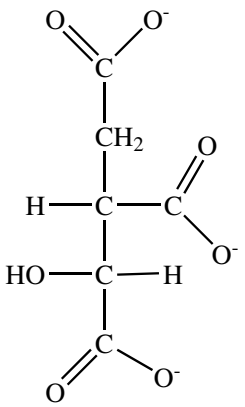


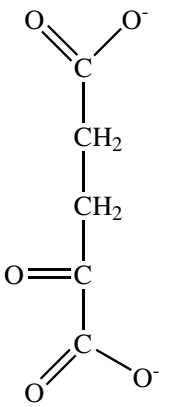
Figure 13-6: Citric acid (Krebs) cycle.



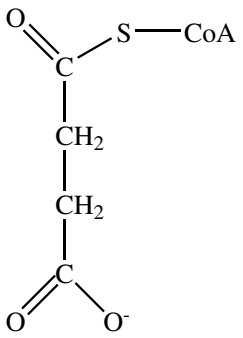
Citrate



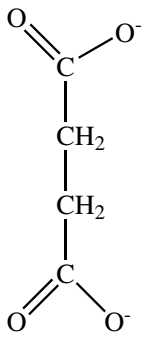
Isocitrate



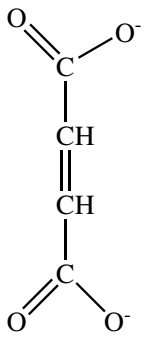
α-Ketoglutarate



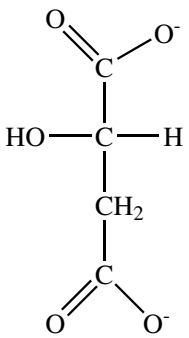
Succinyl-CoA



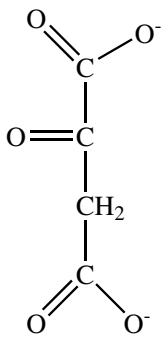
Succinate



Fumarate



Malate



Oxalatoacetate

Figure 13-7: Structures of molecules involved in the citric acid (Krebs) cycle.

Let's get started: Synthesis of acetyl-CoA

The synthesis of acetyl-CoA is a multi-step process. Figure 13-8 shows a simplified version of this process.

These steps are coupled to preserve the free energy produced by the decarboxylation. In the first step, pyruvate combines with TPP (thiamin pyrophosphate) and undergoes decarboxylation. The pyruvate dehydrogenase component of the multi-enzyme complex catalyzes this step. During the second step, the TPP undergoes oxidation, which yields an acetyl group (refer back to Figure 13-8). This acetyl group transfers to lipomide. In this reaction, the oxidant is the disulfide group of lipomide, and acetyllipoamide forms in this step. The pyruvate dehydrogenase component also catalyzes this reaction. In the final step, the acetyl group of acetyllipoamide transfer to CoA to form acetyl CoA. The catalyst for this reaction is dihydrolipoyl transacetylase.

However, the process does not end with the formation of acetyl CoA. It is necessary to regenerate the oxidized form of lopoamide. The enzyme dihydrolipoyl dehydrogenase catalyzes this step. The two electrons from the oxidation transfer to FAD and then to NAD^+. Some of the important intermediates in these steps are shown in Figure 13-9.

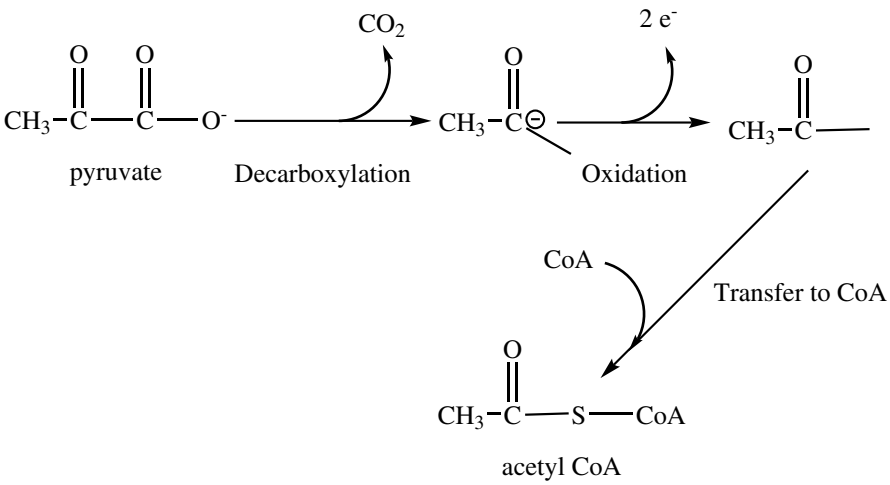
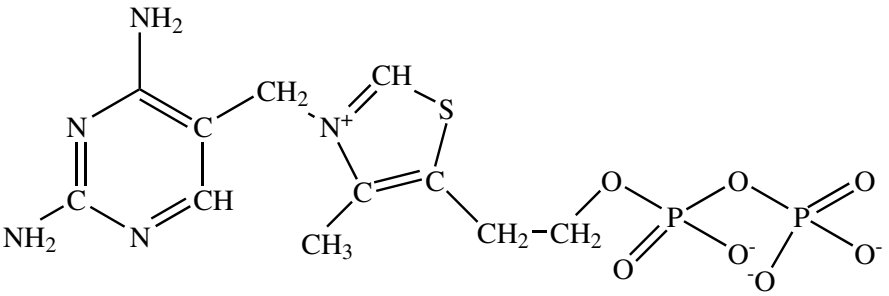
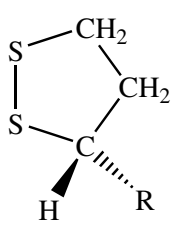


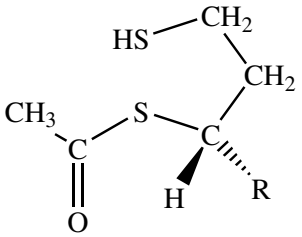
Figure 13-8: Simplified scheme for the formation of acetyl CoA.



Thiamine pyrophosphate (TPP)



Lipoamide



Acetyllipoamide

Figure 13-9: Structures of TPP, lipomide, and acetyllipo-amide.

Three's a crowd: Tricarboxylic acids

When acetyl-CoA enters the citric acid cycle, it interacts, in the presence of citrate synthase, with oxaloacetate. This interaction results in the transfer of the acetyl group to the oxaloacetate to form citrate. The hydrolysis of the thioester linkage of the acetyl-CoA releases a large amount of energy.

The enzyme aconitase, with Fe^{2+} as a cofactor, catalyzes the isomerization of citrate to isocitrate. For a time, cis-aconitate, derived aconitase, was thought to be a part of the citric acid cycle. However, even though the structure of cis-aconitate is related to the other tricarboxylic acids, it is *not* part of the citric acid cycle. The structure of cis-aconitate is in Figure 13-10.

Just a little gas: Oxidative decarboxylation

The next step is the conversion of isocitrate to α-ketoglutarate. The molecule passes through the intermediate oxalosuccinate. The isocitrate binds to the enzyme isocitrate dehydrogenase. During this process, the coenzyme NAD^+ undergoes reduction. Both ATP and NADH are negative factors in the allosteric regulation of isocitrate dehydrogenase, whereas ADP is a positive factor. This is an important mechanism to control the production of ATP.

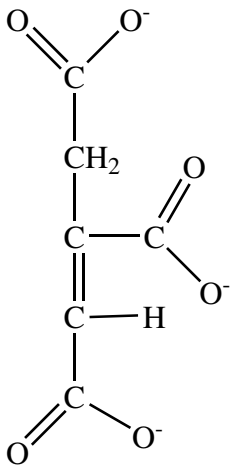


Figure 13-10: Structure of cis-aconitate.

Production of succinate and GTP

The conversion of α-ketoglutarate to succinate requires two steps. The α-ketoglutarate must bind to the enzyme to form an α-ketoglutarate dehydrogenase complex. This reaction requires the same cofactors as needed for the formation of acetyl-CoA. The result of this reaction is the elimination of carbon dioxide and the formation of succinyl-CoA. This process is irreversible under physiological conditions.

In the second step, succinyl-CoA separates to form succinate and release energy, which is harnessed by the conversion of GDP to GTP. This substrate-level phosphorylation is catalyzed by succinyl-CoA synthetase. (GTP contains about the same energy as ATP and can substitute for ATP.)

Oxaloacetate regeneration

The regeneration of oxaloacetate completes the cycle, requiring three reactions which, together, convert a methylene to a carbonyl group. First, a hydrogen atom is removed from each of two adjacent carbon atoms, resulting in the formation of a double bond. Next, a water molecule adds to the double bond. Finally, the removal of two hydrogen atoms yields the appropriate α-keto group. Succinate dehydrogenase catalyzes the first of these reactions. The prosthetic group, FAD, accepts the two hydrogen atoms by covalently binding to the enzyme. Fumarase catalyzes the next step. The final oxidation utilizes the enzyme malate dehydrogenase with the coenzyme NAD^+. The oxaloacetate is now ready to begin the cycle again.

Amino acids as energy sources

Although carbohydrates are the most readily available energy source, there are situations where amino acids can serve as energy sources. This is important for carnivores (like ourselves), who live on a high protein diet. The utilization of amino acids as energy sources is also important during hypoglycemia, fasting, and starvation.

The process begins with the removal of the amino group. This usually occurs through *transamination,* which is the transfer of an amino group from one molecule to another. Any amino acid other than threonine, proline, and lysine will undergo this process. Usually, the amino group transfers to the keto carbon of α-ketoglutarate, oxalatoacetate, or pyruvate to form glutamate, aspartate, or alanine, respectively. Specific transaminases are necessary and the coenzyme pyridoxal phosphate catalyzes this process. A second transamination is involved in the process of transforming aspartate and alanine to glutamate.

Oxidative deamination converts glutamate to α-ketoglutarate. This process, which occurs primarily in the liver, releases an ammonium ion. The reverse reaction, glutamate synthesis, is one of the few reactions that occurs in animals in which inorganic nitrogen is converted into organic nitrogen. The ammonium ion resulting from oxidative deamination may enter one or more biosynthetic pathways or the urea cycle. Most vertebrates convert the ammonium ion to urea, which is excreted in the urine. Most marine organisms, including fish, eliminate ammonia directly, whereas birds, insects, and reptiles convert the ammonium ion to uric acid.

The products of transamination, oxidative deamination, and further modification of the remaining portion of the amino acid produce one of the intermediates in glycolysis or the citric acid cycle. This is the fate of all the amino acids — some of the amino acids go through one intermediate, whereas others require more intermediates. Figure 13-11 shows where each of the amino acids enters glycolysis or the citric acid (Krebs) cycle. Some of the amino acids have more than one entry point.

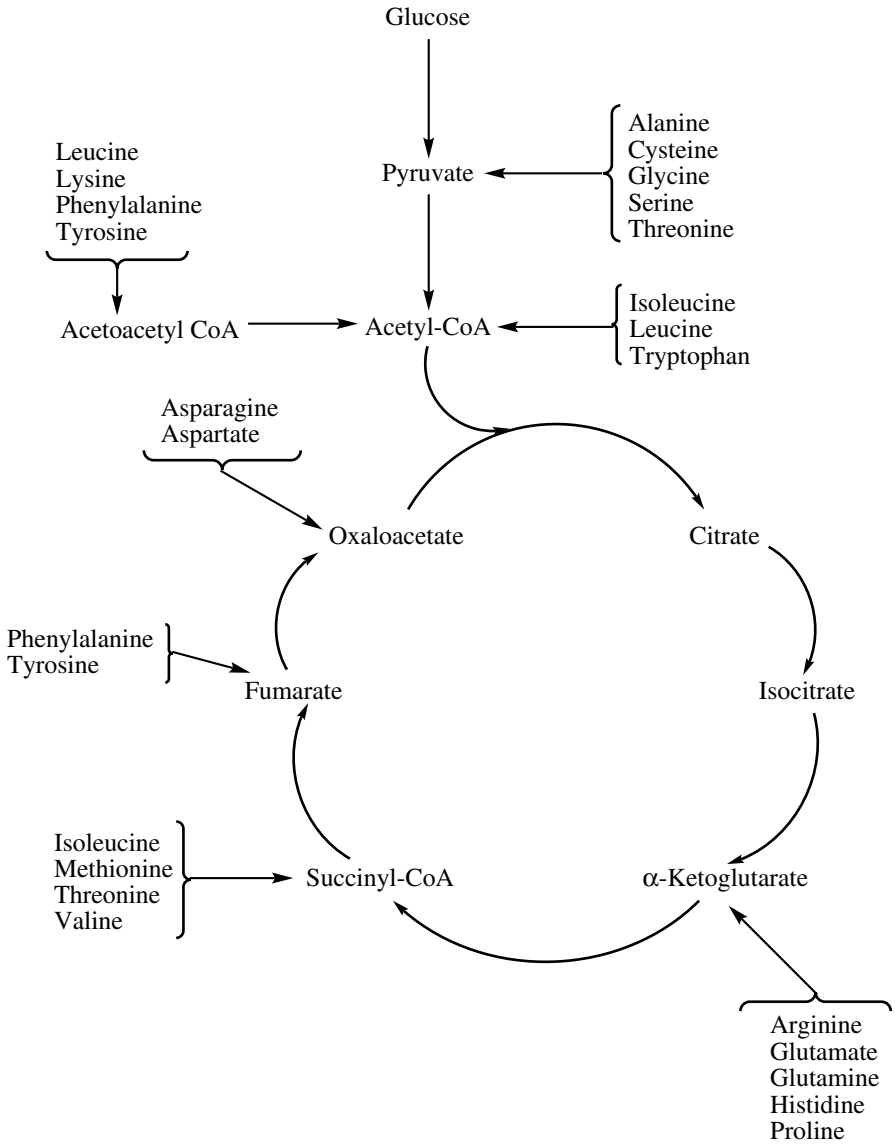


Figure 13-11: Fate of the amino acids.

Electron Transport and Oxidative Phosphorylation

The production of NADH and $FADH_2$ by the citric acid cycle supplies the materials for the next phase: oxidative phosphorylation. These reduced coenzymes transport the electrons derived from the oxidation of pyruvate. The final fate of these electrons is the reduction of oxygen to water.

The details of oxidation phosphorylation are not as easy to study as glycolysis and the citric acid cycle because the processes take place within the mitochondria, where many of the proteins involved are integrated into the walls. In addition, many of the processes are coupled. The separate components of a *coupled* process must not only be in close proximity, but often need to be in a specific arrangement.

The electron transport system

A number of species in the mitochondria must undergo oxidation-reduction reactions. Oxidation involves a loss of electrons, whereas reduction involves a gain of electrons. These processes are coupled in that the electrons lost must equal the electrons gained. The reduction potential indicates how easily a molecule undergoes oxidation or reduction. The molecular players that are important to the electron transport system are the pyridine-linked dehydrogenases, flavin-linked dehydrogenases, iron-sulfur proteins, ubiquinones, and cytochromes.

Off on a tangent: Dealing with reduction potentials

The standard for reduction potentials is the reaction:

$$2\ H^+(aq) + 2\ e^- \quad H_2(g)$$

Under standard conditions (25°C, P_{H2} = 1 atm, and $[H^+]$ = 1.0 M), the standard reduction potential is $E^\circ = 0.00$ V. Under physiological conditions in humans the value is –0.42 V (designated as E'°), because the conditions are not standard.

Table 13-1 lists a number of physiological reduction potentials. We show you how to use these entries later. The values in the table are arranged in order of increasing potential. The higher the value, the better the reaction is at oxidation, and the lower the value, the better the reaction is at reduction.

Table 13-1 Some Physiological Reduction Potentials (E'°)

	E'°(volts)
Ferredoxin-Fe^{3+} + e^- ⇌ Ferredoxin-Fe^{2+}	–0.43
2 H^+(aq) + 2 e^- ⇌ H_2(g)	–0.42
α-Ketoglutarate + CO_2 + 2 H^+ + 2 e^- ⇌ Isocitrate	–0.38
NAD^+ + H^+ + 2 e^- ⇌ NADH	–0.32
FAD + 2 H^+ + 2 e^- ⇌ $FADH_2$	–0.22
Riboflavin + 2 H^+ + 2 e^- ⇌ Riboflavin-H_2	–0.20
Dihydroxyacetone phosphate + 2 H^+ + 2 e^- ⇌ Glycerol 3-phosphate	–0.19
Pyruvate + 2 H^+ + 2 e^- ⇌ Lactate	–0.19
Oxaloacetate + 2 H^+ + 2 e^- ⇌ L-Malate	–0.17
Fumarate + 2 H^+ + 2 e^- ⇌ Succinate	+0.03
Cytochrome b-Fe^{3+} + e^- ⇌ Cytochrome b-Fe^{2+}	+0.08
Cytochrome c-Fe^{3+} + e^- ⇌ Cytochrome c-Fe^{2+}	+0.22
Cytochrome c_1-Fe^{3+} + e^- ⇌ Cytochrome c_1-Fe^{2+}	+0.23
Cytochrome a-Fe^{3+} + e^- ⇌ Cytochrome a-Fe^{2+}	+0.29
Cytochrome a_3-Fe^{3+} + e^- ⇌ Cytochrome a_3-Fe^{2+}	+0.38
1/2 O_2 + 2 H^+ + 2 e^- ⇌ H_2O	+0.82

Each reaction in Table 13-1 is known as a *half-reaction.* It takes two half-reactions — one oxidation and one reduction — to produce a complete (oxidation-reduction) reaction. The electrons lost (oxidation) must equal the electrons gained (reduction). For this reason, electrons only appear in the half-reaction, but never in the overall reaction.

By convention, the reactions in Table 13-1 all appear as *reduction* half-reactions. To convert any of these to an *oxidation* half-reaction, you must do two things. First, reverse the reaction, and then reverse the sign of E'°. In an oxidation-reduction reaction, the overall reaction is created by combining (adding) an oxidation reaction with a reduction reaction. Before adding the two reactions, though, make sure that the electrons in each reaction are equal. This may require multiplying one or both of the reactions by a value to make sure the

electrons are equal. (Multiply the reactions only — do not change the value of E'° [other than a sign change].) For example, look at the following reactions from the table:

$NAD^+ + 2\ e^-$ NADH –0.32

Cytochrome b-Fe^{3+} + e^- Cytochrome b-Fe^{2+} +0.075

Let's now change the first reaction to an oxidation:

NADH $NAD^+ + H^+ + 2\ e^-$ +0.32

If we now want to combine these reactions, we need to multiply the cytochrome reaction by two (so both reactions now involve two electrons):

2 Cytochrome b-Fe^{3+} + 2 e^- 2 Cytochrome b-Fe^{2+} +0.08

The number of electrons lost must equal to electrons gained. Also, notice that only the reaction is doubled, not the voltage. We can now combine these two reactions, canceling the electrons from both sides:

NADH $NAD^+ + H^+ + 2\ e^-$ +0.32 V

2 Cytochrome b-Fe^{3+} + 2 e^- 2 Cytochrome b-Fe^{2+} +0.08 V

NADH + H^+ + 2 Cytochrome b-Fe^{3+} 2Cytochrome b-Fe^{2+} + NAD^+ + 2H^+ +0.40 V

The final reaction will have no electrons. Other species may cancel, if they appear on both sides of the reaction arrow. Any time the sum of the two potentials is positive, the reaction produces energy. Conversely, a negative value means the reaction requires energy. The greater the value of the sum, the greater the amount of energy produced.

Pyridine-linked dehydrogenases

In order for these enzymes to function, the coenzymes NAD^+ or $NADP^+$ are necessary. The coenzymes may be in either the oxidized or the reduced forms. If the general form of the substrate in the reduced form is Z-H_2, and in the oxidized form, it is Z, then the reaction will be:

Z-H_2 + NAD^+ (or $NADP^+$) Z + NADH (or NADPH) + H^+

There are more than 200 pyridine-linked dehydrogenases. The majority of NAD^+-linked dehydrogenases are involved in aerobic respiration. Most of the $NADP^+$-linked dehydrogenases are involved in biosynthesis.

Flavin-linked dehydrogenases

Enzymes (E) of this type require FAD or FMN as tightly bound prosthetic groups or coenzymes. Again, the species may be in either the oxidized or the reduced forms. The general reactions of this type are:

$$Z\text{-}H_2 + E\text{-}FAD \quad Z + E\text{-}FADH_2$$

$$Z\text{-}H_2 + E\text{-}FMN \quad Z + E\text{-}FMNH_2$$

NADH dehydrogenase, which contains the prosthetic group FMN, is the enzyme responsible for transporting electrons from NADH to the next acceptor in the electrons transport chain. There are other flavin-linked dehydrogenases — for example, succinate dehydrogenase.

Iron-sulfur proteins

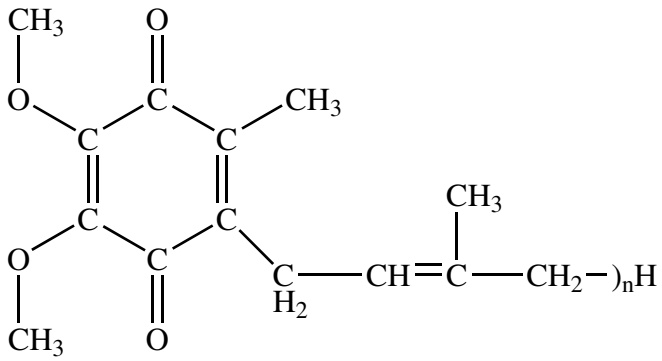
The chief characteristics of iron-sulphur proteins are the presence of iron and sulfur, as S^{2-}. The electron transporting ability of these proteins is the Fe^{2+}/Fe^{3+} couple. Several of these proteins are associated with the electron transport chain, where they are complexed to other respiratory species. Examples include succinate dehydrogenase, with two iron-sulfur centers, and NADH dehydrogenase, with four iron-sulfur centers.

Ubiquinones

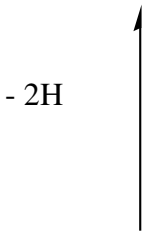
The *ubiquinones* are a group of coenzymes that are fat-soluble. Coenzyme Q (CoQ) is an example of an ubiquinone. The oxidation-reduction center is a derivation of quinine, and the fat-solubility is enhanced by the presence of a long hydrocarbon chain, containing a series of isoprene units. Many of the different ubiquinones differ only in the number of isoprene units present. The oxidized form of coenzyme Q is simply CoQ whereas the reduced form is $CoQH_2$. The general structures of both the oxidized and reduced forms of a ubiquinone appear in Figure 13-12.

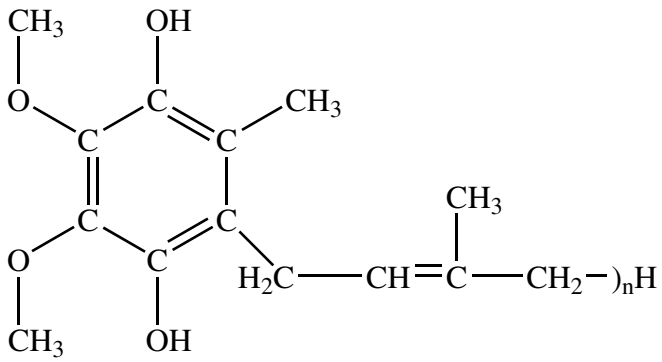
Cytochromes

The *cytochromes* are a group of proteins containing a heme group. Like the iron-sulfur proteins, the oxidation-reduction couple is Fe^{2+}/Fe^{3+}. The three general classes of cytochromes are a, b, and c. The derivation of the class names relates to spectral studies done during the first isolation of these molecules. Cytochromes occur in both the mitochondria and the endoplasmic reticulum. The heme group, present in all cytochromes, is like the heme groups present in myoglobin and hemoglobin. In all cases, the central portion of the group is identical; differences derive from the attachment of side-chains to the heme core. Figure 13-13 shows the heme core and where the side-chains normally attach.



Oxidized ubiquinone





Reduced ubiquinone

Figure 13-12: General structures of the oxidized and reduced forms of a ubiquinone.

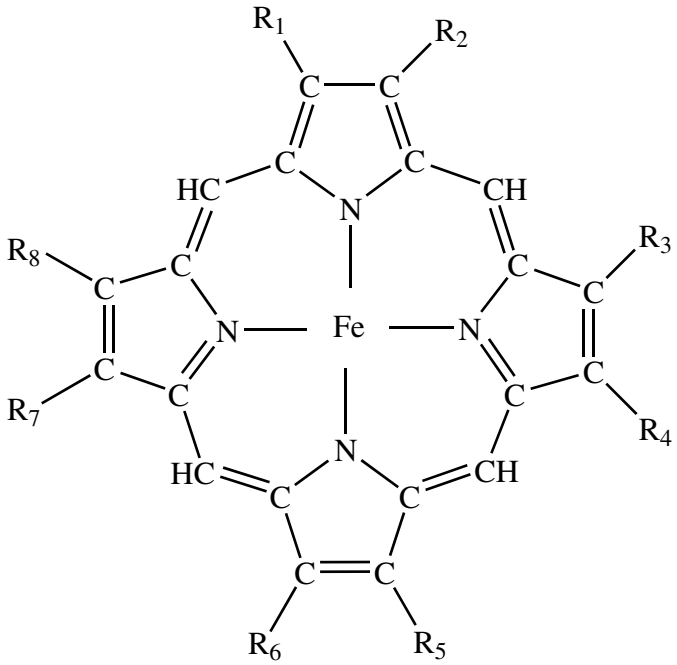


Figure 13-13: The heme core and attachment sites (R).

Five cytochromes (a, a_3, b, c, and c_1) have been identified as part of the electron transport chain of mammals. Cytochrome c, or *cyt c,* is easy to extract from cells, and therefore it is the most studied of the cytochromes. The structure of cytochrome c from different species is important to the study of biochemical evolution. Cytochromes a and a_3, cyt aa_3, occur together as a complex containing not only the expected two heme groups, but also two copper ions. The copper ions are part of another oxidation-reduction couple (Cu^{+}/Cu^{2+}). This complex, known as *cytochrome oxidase,* is the terminal cytochrome, which transfers electrons to O_2.

Interpersonal relationships

The members of the electron transport chain are grouped into four complexes with coenzyme Q (CoQ) and cytochrome c (cyt c) serving as links. One way of indicating the sequence of events in the electron transport chain appears in Figure 13-14. Figure 13-15 illustrates the same sequence emphasizing the cyclic nature of the steps. The processes take place in four complexes with linking CoQ and cytochrome c. These complexes are part of the inner mitochondrial membrane.

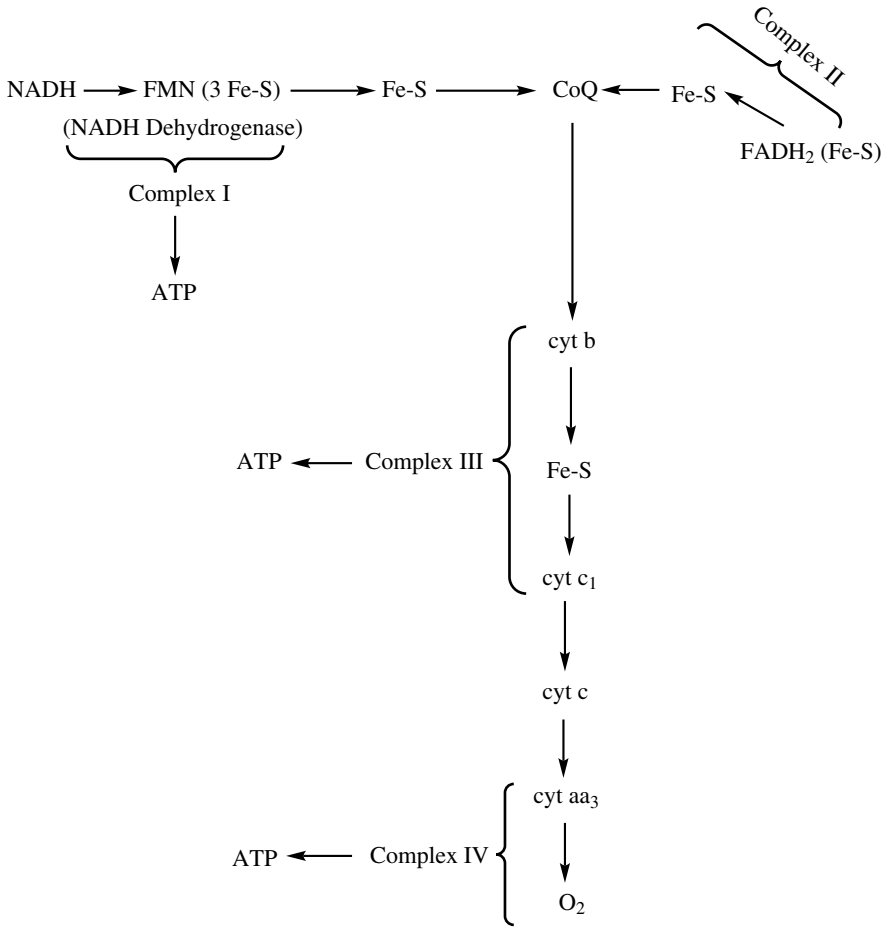


Figure 13-14: Steps in the electron transport chain.

Figure 13-15: Electron transport chain, emphasizing the cyclic nature of each of the processes.

The script: Oxidative phosphorylation

The processes of oxidative phosphorylation and the electron transport chain are closely coupled. It is only possible to oxidize the reduced forms of the coenzymes $FADH_2$ and NADH in the presence of ADP. The oxidations couple with the ADP transforming to ATP (phosphorylation).

If we calculate the oxidation-reduction potentials for NADH and $FADH_2$ reducing oxygen, we find:

$NADH \quad NAD^+ + H^+ + 2\ e^-$	+0.32 V
$\underline{1/2\ O_2 + 2\ H^+ + 2\ e^- \quad H_2O}$	$\underline{+0.82\ V}$
$1/2\ O_2 + H^+ + NADH \quad H_2O + NAD^+$	+1.14 V

And:

$FADH_2 \quad FAD + 2\ H^+ + 2\ e^-$	+0.22 V
$\underline{1/2\ O_2 + 2\ H^+ + 2\ e^- \quad H_2O}$	$\underline{+0.82\ V}$
$1/2\ O_2 + FADH_2 \quad H_2O + FAD$	+1.04 V

In both cases, the combination of the potentials is positive. Positive potentials refer to spontaneous processes, and spontaneous processes produce energy. Each NADH is capable of supplying sufficient energy to produce 2.5 ATP, and each $FADH_2$ can produce 1.5 ATP.

The play: Proposed mechanisms

The current proposed mechanism for oxidative phosphorylation is the *chemiosmotic hypothesis.* This hypothesis assumes that the hydrogen ion gradient is a significant factor promoting the conversion of ADP to ATP. The processes occurring in the four complexes present in the inner mitochondrial membrane result in a net transfer of hydrogen ions across the membrane.

The hydrogen ion transfer results in an increase in the hydrogen ion concentration in the space between the inner and outer mitochondrial membranes. It is necessary to move hydrogen ions back across the membrane. This transfer of hydrogen ions is necessary in the synthesis of ATP.

The box office: ATP production

The reactions from the anaerobic oxidation of glucose (glycolysis) and the aerobic oxidation of glucose result in the production of 32 molecules of ATP from every molecule of glucose. These reactions are:

Anaerobic:

$$\text{Glucose} + 2\ NAD^+ + 2\ ADP + 2\ P_i \rightarrow 2\ \text{Pyruvate} + 2\ NADH + 2\ H^+ + 2\ H_2O + 2\ ATP$$

Aerobic:

$$2\ \text{Pyruvate} + 5\ O_2 + 30\ ADP + 30\ P_i \rightarrow 6\ CO_2 + 34\ H_2O + 30\ ATP$$

Sum:

$$\text{Glucose} + 2\ NAD^+ + 5\ O_2 + 32\ ADP + 32\ P_i \rightarrow 6\ CO_2 + 36\ H_2O + 2\ NADH + 2\ H^+ + 32\ ATP$$

Involving the fats: β-oxidation cycle

Fatty acids may also serve as a source of ATP. Accomplishing this requires a series of reactions, known as β-*oxidation,* or the *fatty acid spiral,* to break down the fatty acid molecule. This series of reactions is a cyclic process. Some of the processes are oxidations, which require the coenzymes NAD^+ and FAD. This process also occurs in the mitochondria. The initiation of fatty acid oxidation requires activation of the relatively unreactive fatty acid molecule. The activated form is analogous to acetyl-CoA. In this case, the coenzyme A binds to the fatty acid to form a fatty acyl-CoA. Activation requires acyl-CoA synthetase and one molecule of ATP. The ATP uses two phosphates and becomes AMP.

At the inner mitochondrial membrane, the fatty acyl-CoA combines with the carrier molecule carnitine. Carnitine acyltransferase catalyzes this transfer. The fatty acyl-carnitine transports into the mitochondrial matrix, where it converts back to fatty acyl-CoA. With the mitochondrial matrix, a series of four reactions constitute the cycle known as β-oxidation. The name of this process refers to the oxidation of the second (β) followed by the loss of two carbons from the carboxyl end of the fatty acyl-CoA. Each trip around the cycle results in the removal of two carbon atoms, and the cycle continues until all the carbon atoms are removed. Figure 13-16 illustrates the general aspects of the cycle.

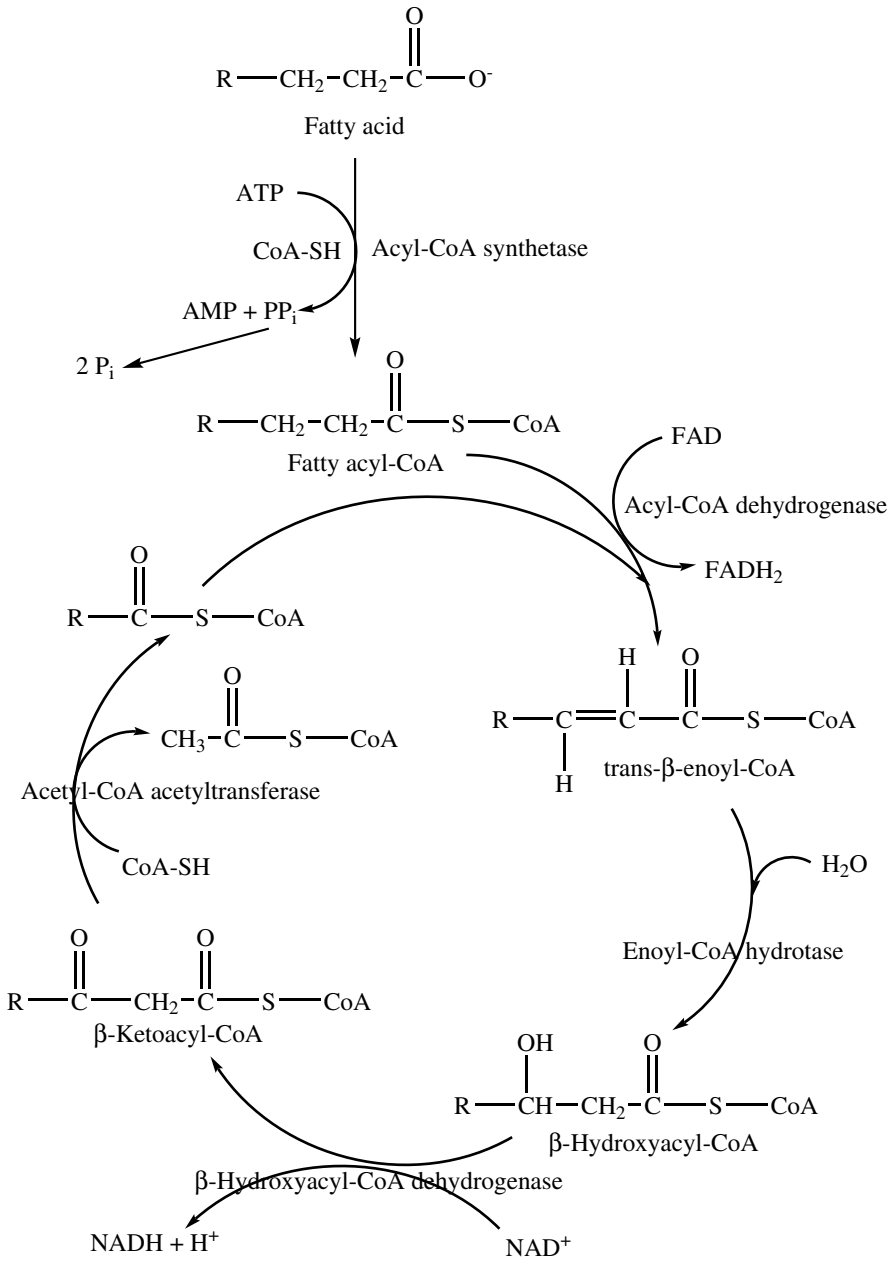


Figure 13-16: General steps in the β-oxidation cycle.

The first step in the cycle is an oxidation, with the catalyst being acyl-CoA dehydrogenase. During this step, coenzyme FAD accepts two hydrogen atoms. One of the hydrogen atoms is from the α carbon, and the other is from the β carbon atom. The process is stereospecific, producing the trans form. Elsewhere, the $FADH_2$ undergoes re-oxidation to FAD with the production of 1.5 molecules of ATP.

The trans-alkene undergoes hydration to form a secondary alcohol in the second step. The catalyst is the enzyme enoyl-CoA hydratase — a stereospecific enzyme yielding only the L isomer. Next, the secondary alcohol undergoes oxidation to form a ketone. The oxidizing agent is NAD^+. The enzyme catalyzing this oxidation is β-hydroxy-acyl-CoA dehydrogenase. The re-oxidation of NADH to NAD^+ via the electron transport chain produces two molecules of ATP.

The final step involves the cleavage of the β-ketoacyl-CoA with a molecule of CoA. This produces acetyl-CoA and a fatty acyl-CoA two carbon atoms shorter than the original. The enzyme from this step is β-ketothiolase (or simply thiolase). The new fatty acyl-CoA goes around the cycle to be shortened by two carbon atoms. An unsaturated fatty acid also goes through similar steps, but needs one or two additional enzymes.

The energy yield from a fatty acid is larger than from glucose. The process begins with the activation of the fatty acid, which costs the equivalent of two ATP molecules. Each trip around the cycle yields ten molecules of ATP, a molecule of $FADH_2$, and a molecule of NADH. The NADH and $FADH_2$ ultimately yield four additional molecules of ATP. Thus, each trip around the cycle produces 14 molecules of ATP. In addition, the final trip around the cycle produces not one but two molecules of acetyl-CoA.

Not so heavenly bodies: Ketone bodies

Some of the excess acetyl-CoA will form a group of relatively small molecules called *ketone bodies.* This is especially important when there is a build up of acetyl-CoA. A build up may occur when the rate of production is too high or if it is not used efficiently. Two acetyl-CoA molecules combine in the reverse of the last step in β-oxidation to produce acetoacetyl-CoA. Acetoacetyl-CoA reacts with water and another acetyl-CoA to form β-hydroxy-β-methylglutaryl-CoA, which in turn cleaves to acetoacetate and acetyl-CoA. Most of the acetoacetate undergoes reduction to β-hydroxybutyrate (a small amount decarboxylates to acetone and carbon dioxide). These steps appear in Figure 13-17.

The other guy

When a fat molecule breaks down, the results are a glycerol and three fatty acid molecules. The fatty acid molecules enter the β-oxidation cycle and produce energy. Catabolism of the glycerol also serves as a source of energy. First, the glycerol is phosphorylated to glycerol 1-phosphate (= glycerol 3-phosphate). This uses one molecule of ATP. Oxidation of glycerol 1-phosphate generates dihydroxyacetone phosphate, which can enter the glycolysis pathway. The net energy production is from 16.5 to 18.5 molecules of ATP.

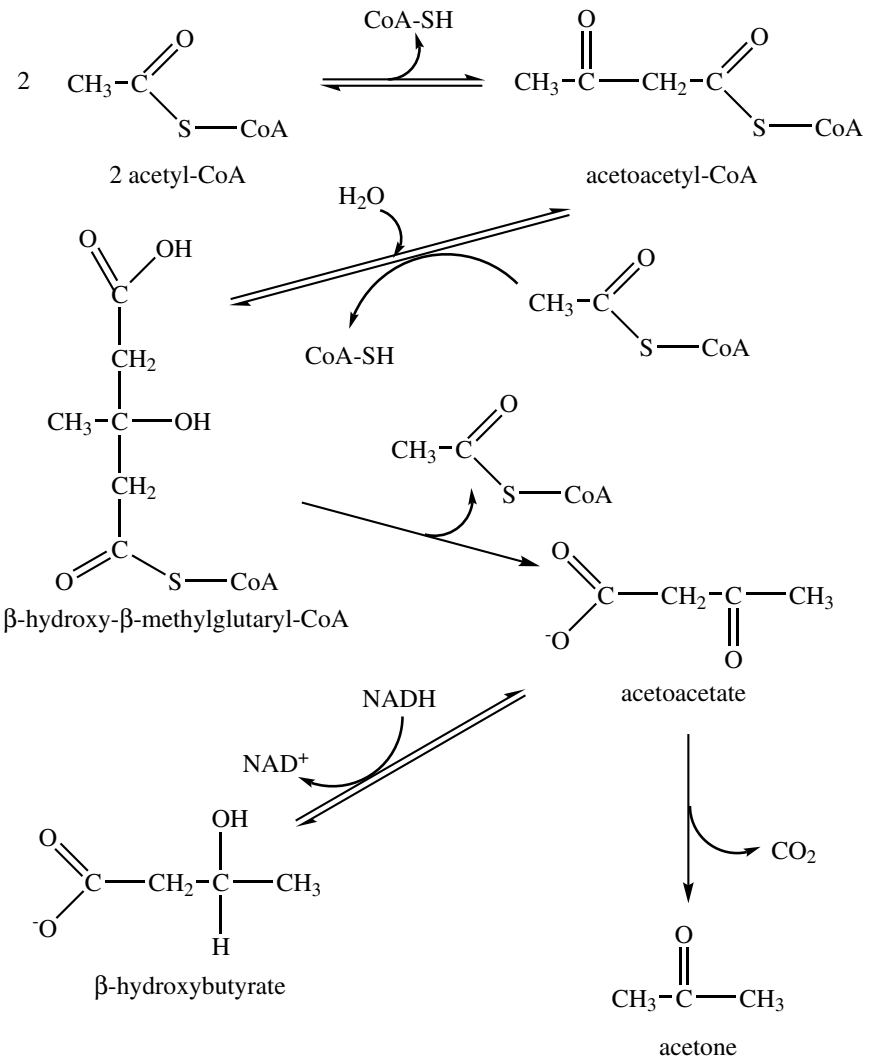


Figure 13-17: Formation of the ketone bodies.

As a group, acetone, β-hydroxybutyrate, and acetoacetate are the ketone bodies.

This process occurs primarily in the liver, and the β-hydroxybutyrate and acetoacetate then enter the bloodstream for use by other tissues. During prolonged starvation, ketone bodies may serve as the major energy source for some tissues. The kidneys excrete excess ketone bodies. Normal blood levels are about 1 mg of ketone bodies per 100 mL of blood.

In starvation or diabetes mellitus, a form of diabetes, cells may not receive sufficient carbohydrate for energy, leading to an increase in the rate of fatty acid oxidation to compensate for the energy deficit. As the amount of acetyl-CoA increases, there is insufficient oxaloacetate in the citric acid cycle available for oxidation of this acetyl-CoA. (The oxaloacetate concentration is lower because of the necessity of using it for glucose synthesis.) This leads to an increase in the production of ketone bodies and an increase of ketone bodies in the bloodstream. At 3 mg of ketone bodies per 100 mL, a condition known as *ketonemia* arises — a high concentration of ketone bodies in the urine. Ketonemia and ketonuria are two aspects of ketosis.

Two of the ketone bodies are in the form of acids. The build up of ketone bodies leads to an overwhelming of the blood buffers. The decrease in blood pH may reach 0.5 units lower than the normal pH (7.4), leading to acidosis, a serious condition, which, among other things, leads to difficulty in oxygen transport by hemoglobin. Dehydration results as the kidneys eliminate large quantities of liquid trying to remove the excess acid. Severe acidosis may result in a coma that may result in death.

Mammals cannot convert acetyl-CoA to carbohydrates. It is possible to convert carbohydrates to fats, but not to do the reverse.

Investing in the Future: Biosynthesis

One aspect of metabolism, catabolism, is to produce the energy required for life. Another aspect, anabolism, is to supply the materials for growth and replacement. Food supplies the raw fuel for metabolism. A number of pathways are available to allow for flexibility. It is necessary to block some pathways to overcome Le Châtelier's Principle, partly because an enzyme will catalyze both the forward and the reverse reaction.

Nearly all intermediates in catabolic processes are also intermediates in anabolic processes. Molecules may easily change from one pathway to another. In general, anabolic processes require the energy produced by catabolic processes. We've already seen one aspect of anabolism — gluconeogenesis. Earlier, we saw how this process, related to glycolysis, could generate glucose and other carbohydrates. We examine other biosynthesis processes in this section.

Fatty acids

Production of the fatty acids is necessary to form the membrane lipids. But the main reason for fatty acid synthesis is to convert excess dietary carbohydrate to fats for storage. The key molecule for this is acetyl-CoA.

The liver is the primary fatty acid synthesis site in humans, and humans can synthesize all the fatty acids but two: linoleic acid and linolenic acid. Linoleic acid and linolenic acid are also essential fatty acids, required components of the diet. Acetyl-CoA from glycolysis or β-oxidation reacts with bicarbonate ion in a reaction (Figure 13-18) powered by ATP and catalyzed by acetyl-CoA carboxylase, forming the three-carbon molecule malonyl-CoA.

The release of insulin triggers a series of steps that result in the activation of acetyl-CoA carboxylase. Release of insulin indicates high food levels. Both glucagon and epinephrine inhibit the enzyme, through a series of steps. In mammals, the enzymes necessary to synthesize palmitic acid from acetyl-CoA and malonyl-CoA are present in a complex known as *fatty acid synthase.* In plants and bacteria, the enzymes are present as separate molecules. Synthesis proceeds two carbon atoms at a time, which is why all the natural fatty acids contain an even number of carbon atoms.

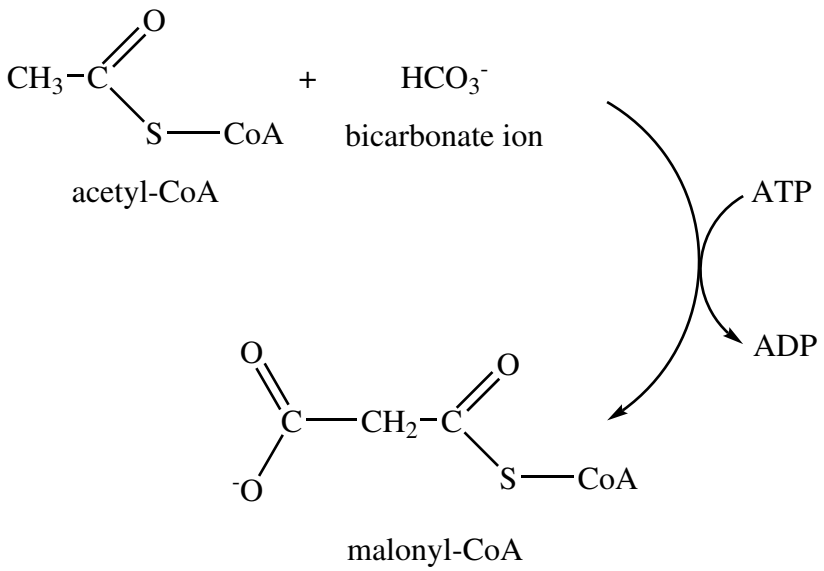


Figure 13-18: Synthesis of malonyl-CoA.

Synthesis begins when a molecule of acetyl-CoA links to an acyl carrier protein, ACP, and a malonyl-CoA does the same with another ACP. The two ACP-linked molecules then join and release a carbon dioxide molecule, an ACP, and an acetoacetyl-ACP. Next are three steps that are the reverse of the first three steps of β-oxidation. First, NADPH reduces the ketone group to an

alcohol. Then dehydration of the alcohol leaves a double bond between the second and the third carbon atoms. The coenzyme NADPH again serves as a reducing agent to produce butyryl-ACP. The sequence repeats with butyryl-ACP replacing the acetyl-ACP. These steps are in Figure 13-19.

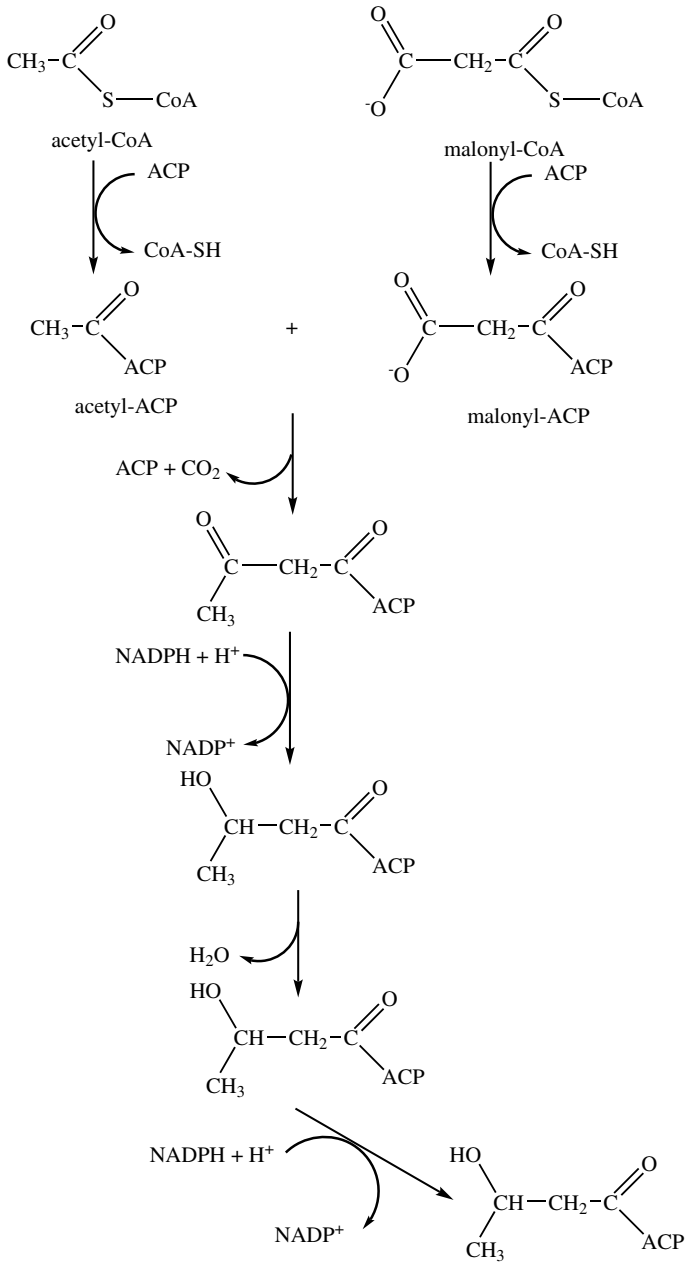


Figure 13-19: Fatty acid synthesis.

The series of synthesis steps continues up to palmitic acid (16 carbon atoms). The overall reaction is:

$$8 \text{ acetyl-CoA} + 7 H^+ + 14 \text{ NADPH} + 7 \text{ ATP} \rightarrow \text{palmitic acid} + 8 \text{ CoA} + 14 \text{ NADP}^+ + 7 \text{ ATP} + 7 P_i$$

Once the palmitic acid forms, additional reactions, where necessary, can lengthen or shorten the chain. These require different enzyme systems. Partial oxidation of a saturated fatty acid yields an unsaturated fatty acid.

Membrane lipids

Like other molecules, it is necessary to synthesize the membrane lipids from their constituents. In the previous section, we explained how to synthesize the fatty acids. These fatty acids need to be activated with acetyl-CoA in order to produce the appropriate acyl-CoA. The reduction of dihydroxyacetone, from glycolysis, yields glycerol 3-phosphate. The glycerol 3-phosphate combines with the appropriate acyl-CoA molecules to yield a phosphatidate (Figure 13-20). The phosphitidate then reacts with an activated serine or an activated choline to form the appropriate phosphoglyceride.

The formation of the spingolipids follows a similar path. In this case, sphingosine replaces glycerol. The synthesis of sphingosine begins with the reaction of palmitoyl-CoA, with serine in the presence of acid. This reaction yields Coenzyme A, carbon dioxide, and the precursor of sphingosine. Oxidation of the precursor yields sphingosine (Figure 13-21).

An acyl-CoA can then add a fatty acid to the amine group to produce N-acylsphingosine (ceramide). The reaction of the alcohol on the third carbon of the ceramide with activated phosphocholine yields sphingomyelin.

The reaction of ceramide with an activated monosaccharide begins the synthesis of the glycolipids. To complete the synthesis, it is necessary to add additional activated monocaccharides (UDP-glucose being one example).

Cholesterol is another membrane lipid. It helps to control the fluidity of cell membranes and is also the precursor of the steroid hormones. The entire synthesis takes place in the liver, where acetyl-CoA molecules are joined. Thus, the cholesterol molecule is built up two carbon atoms at a time.

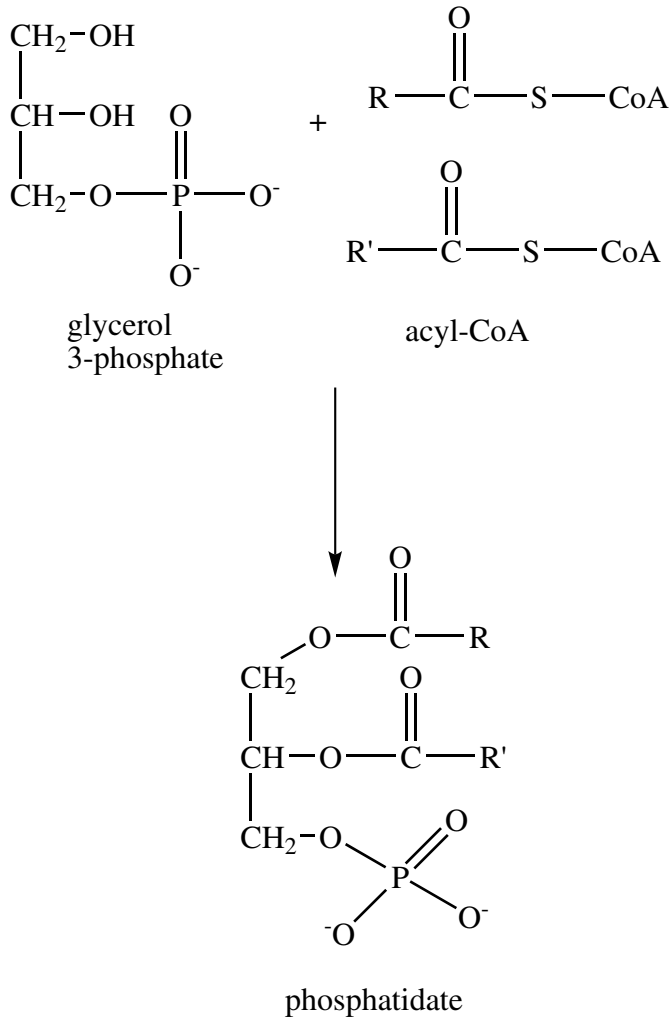


Figure 13-20: Formation of phosphatidate.

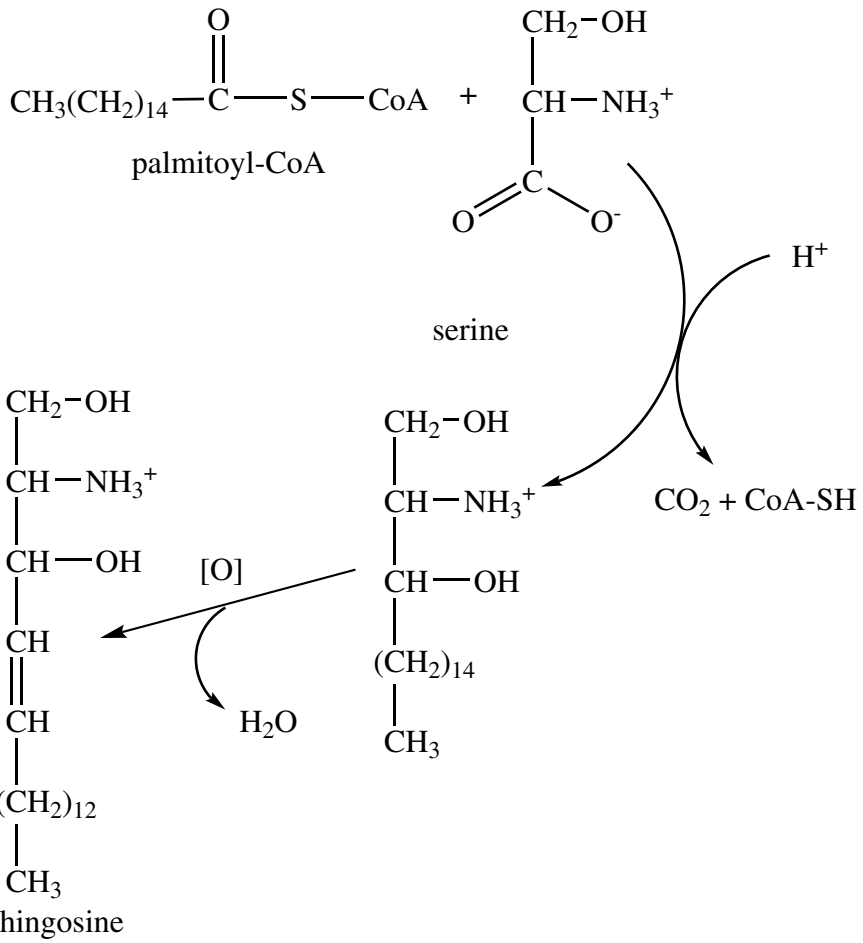


Figure 13-21: Formation of sphingosine.

Amino acids

Synthesis of amino acids becomes necessary when insufficient quantities are present in the diet. Adult humans can only synthesize 11 of the 20 amino acids. The amino acids that humans cannot synthesize are known as the *essential amino acids,* and these are a necessary requirement in the diet. Table 13-2 list the essential and non-essential amino acids.

Table 13-2 Essential and Non-essential Amino Acids

Essential	*Non-essential*
Histidine	Alanine
Isoleucine	Asparagine
Leucine	Aspartate
Lysine	Cysteine
Methionine	Glutamine
Phenylalanine	Glutamate
Threonine	Glycine
Tryptophan	Proline
Valine	Serine

Arginine is essential for children, but not for adults. Tyrosine is non-essential in the presence of adequate quantities of phenylalanine. Glutamate is important to the synthesis of five amino acids. Glutamate may form by the reduction of α-ketoglutaric acid, an intermediate from the Krebs cycle. The process is shown in Figure 13-22.

In the forward direction, this is a synthesis reaction, whereas the reverse reaction is an important oxidative deamination from the catabolism of amino acids. Glutamate, when necessary, serves as an intermediate in the biosynthesis of alanine, aspartate, asparagine, glutamine, proline, and serine. The transamination in Figure 13-23 illustrates the formation of alanine.

Replacing pyruvate in the preceding reaction with oxaloacetate yields aspartate.

It is possible to convert excess phenylalanine to tyrosine by a simple oxidation in the presence of phenylalanine hydroxylase (Figure 13-24).

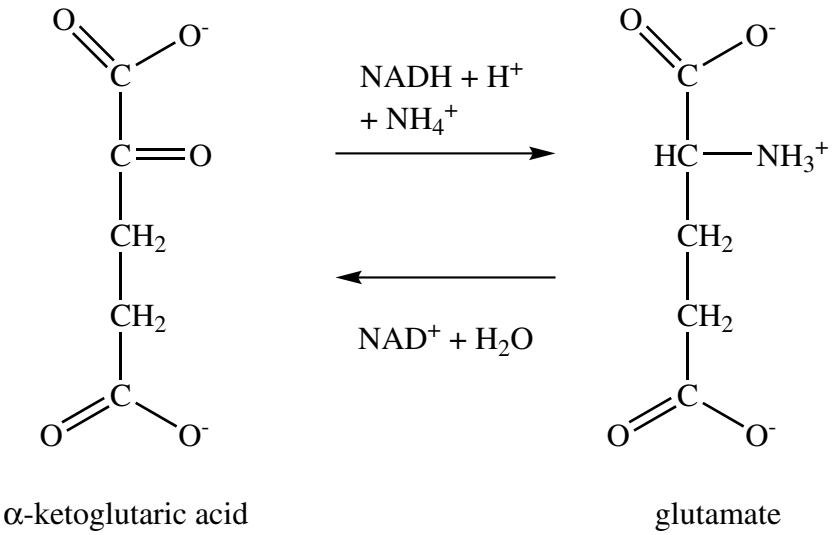


Figure 13-22: Equilibrium between glutamate and α-ketoglutaric acid.

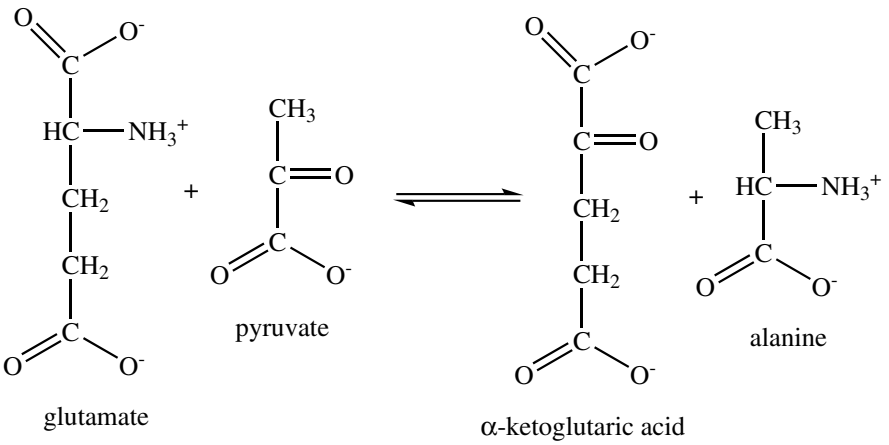


Figure 13-23: Formation of alamine.

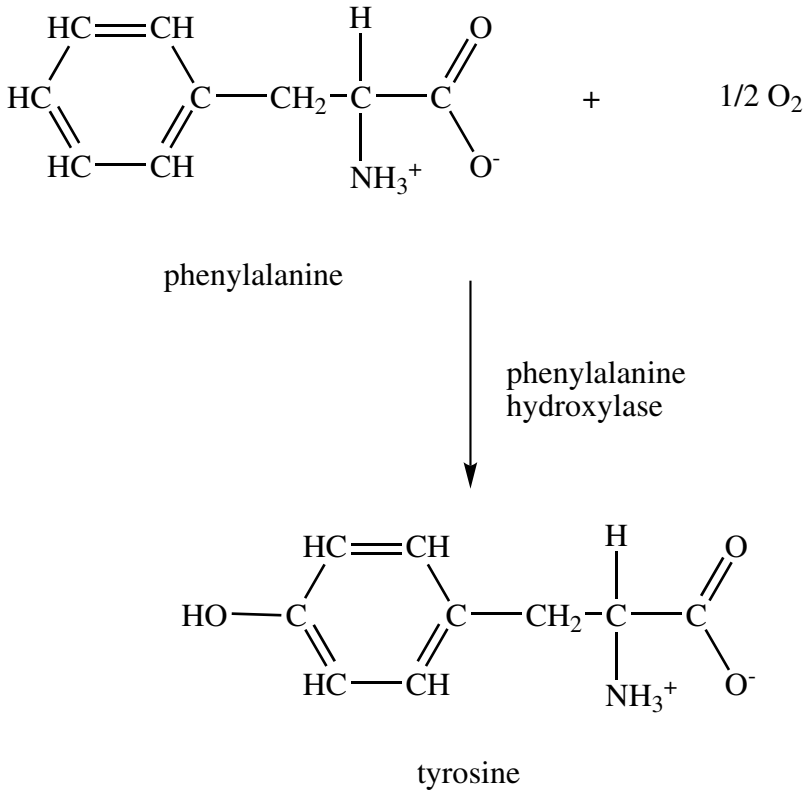


Figure 13-24: Synthesis of tyrosine.

Methionine serves as the source of sulfur for the synthesis of cysteine. Serine serves as the base of the rest of the molecule. *Serine* is the product of a three-step process beginning with 3-phosphoglycerate. The process starts with the oxidation by NAD^+ of the secondary alcohol group. The ketone thus formed undergoes transamination with glutamate to form 3-phosphoserine. Finally, hydrolysis of the phosphate ester yields serine (Figure 13-25).

The formation of proline is a four-step process beginning with glutamate. The process is shown in Figure 13-26.

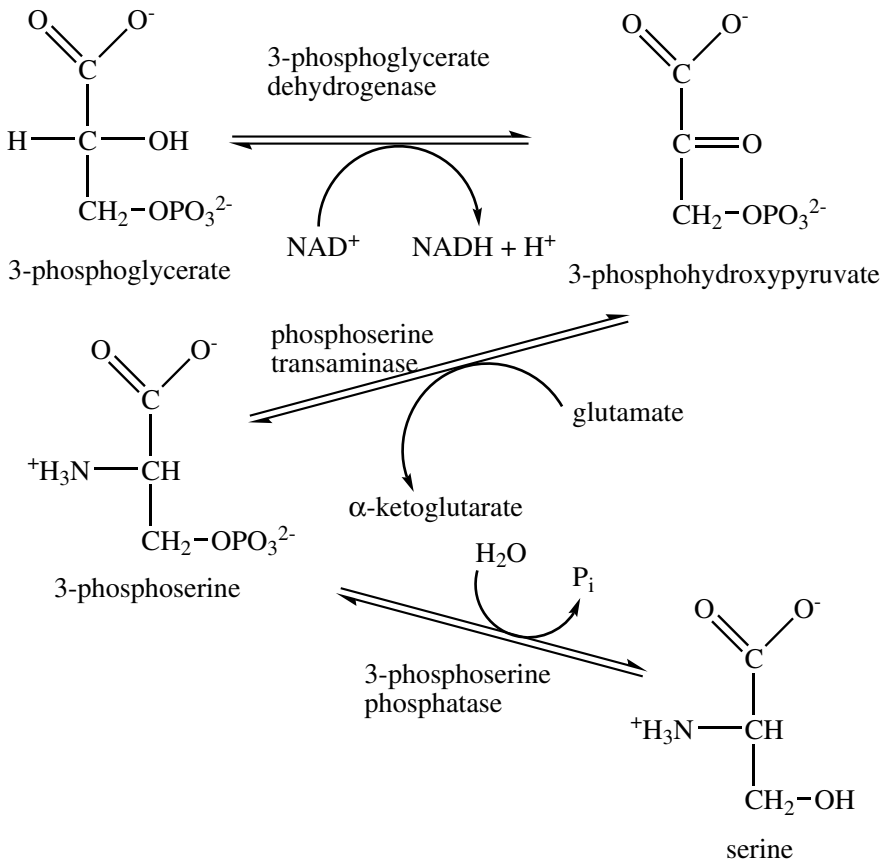


Figure 13-25: Synthesis of serine.

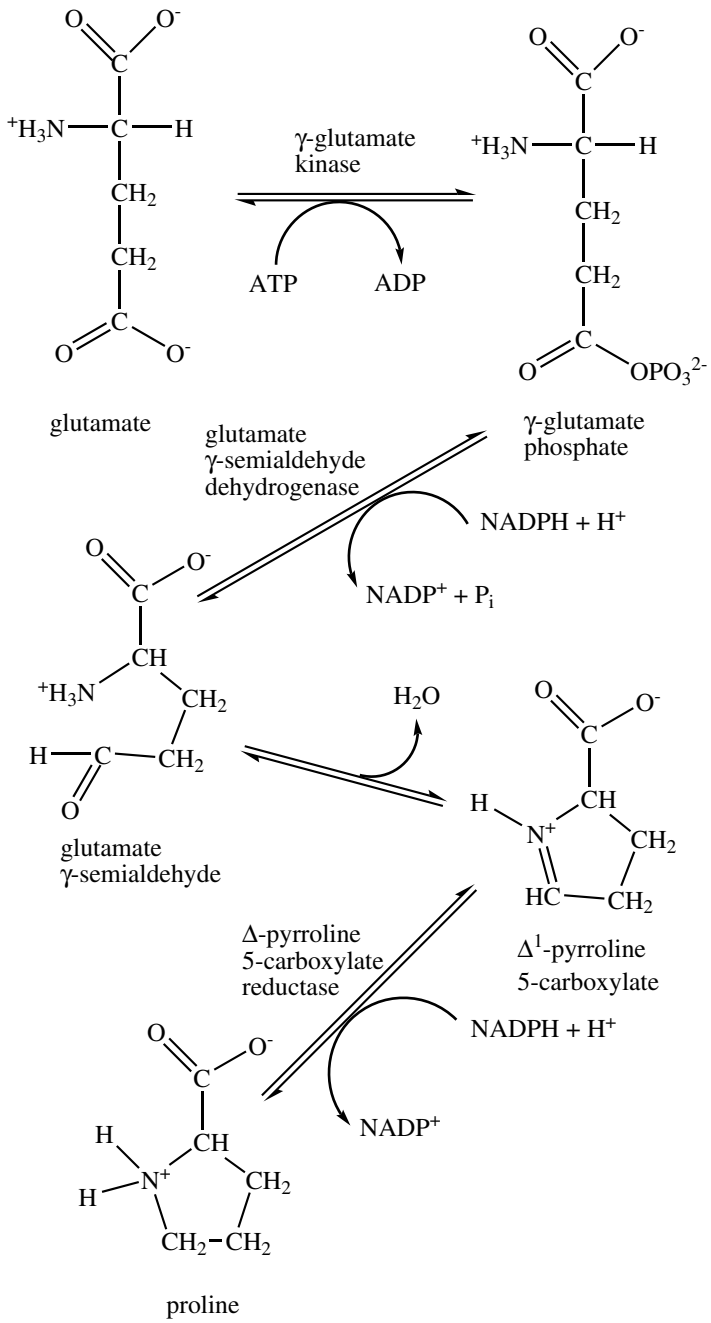


Figure 13-26: Synthesis of proline.

Chapter 14

Smelly Biochemistry: Nitrogen in Biological Systems

In This Chapter

- Talking about purine and pyrimidine
- Examining catabolism and discussing the urea cycle
- Considering amino acids
- Finding out about metabolic disorders

In this chapter, we investigate the role of nitrogen in biomolecules. Nitrogen occurs primarily in the amino acids (proteins) and in nucleic acids (purines and pyrimidines), many of which have a distinctive and generally unpleasant aroma, hence our chapter title. A few other molecules, such as hemoglobin, also contain nitrogen. Humans eliminate nitrogen primarily in the urea.

Ring in the Nitrogen: Purine

Adenine and guanine are nitrogen bases that employ the purine ring system (Figure 14-1). The formation of these molecules is essential to the synthesis of both DNA and RNA. The biosynthesis of the purines generates the molecules in their nucleotide forms instead of the free base form.

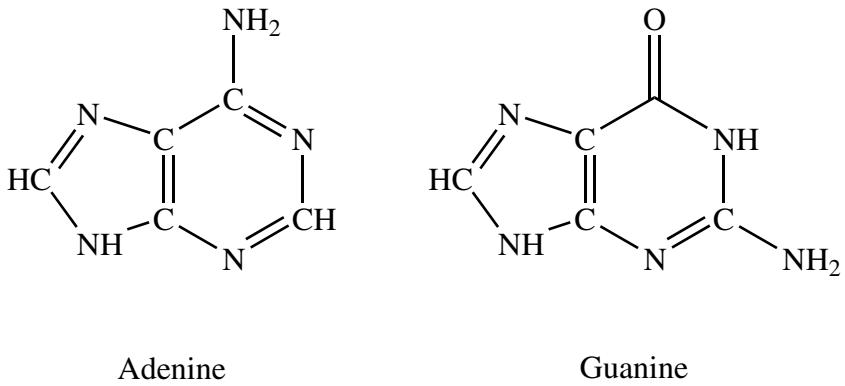


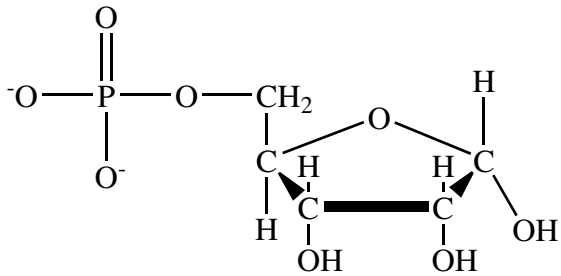
Figure 14-1: Purine nitrogen bases.

Biosynthesis of purine

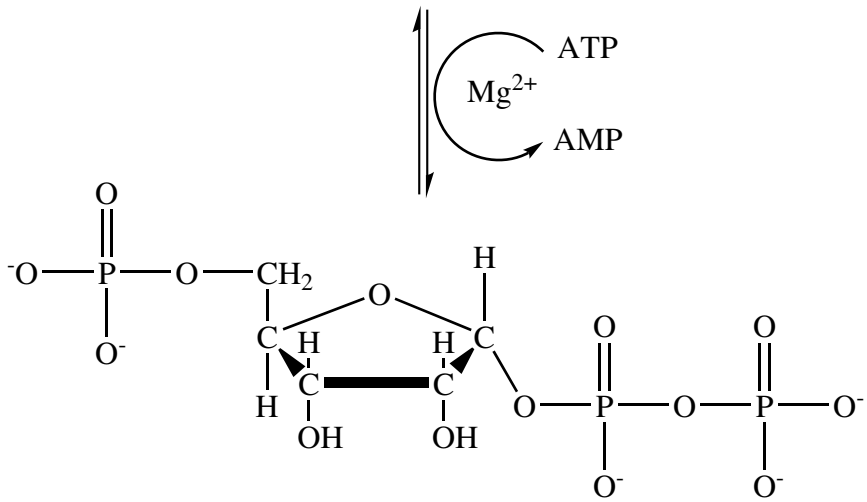
The synthesis of purine begins with the activation of D-ribose-5'- phosphate through pyrophosphorylation. In this reaction a pyrophosphate group from ATP is transferred to C-1 of an α-D-ribose-5'-phosphate. This gives a 5-phospho-α-D-ribose 1-pyrophosphate (PRPP) and AMP. The reaction is unusual because it involves the transfer of an intact pyrophosphate group (Figure 14-2). PRPP is also necessary for the synthesis of pyrimidines.

Inosine synthesis

PRPP goes through a series of ten steps (Figure 14-3) to become inosine 5'-phosphate or inosinic acid (IMP). Notice that throughout these ten steps the D-ribose-5'-phosphate portion of PRPP does not change. The ten enzymes necessary for these steps are in Table 14-1. Two additional, though different, steps are necessary to convert IMP to either AMP or GMP.



α-D-ribose 5-phosphate



5-phospho-α-D-ribose 1-pyrophosphate (PRPP)

Figure 14-2: Activation of D-ribose-5'-phosphate.

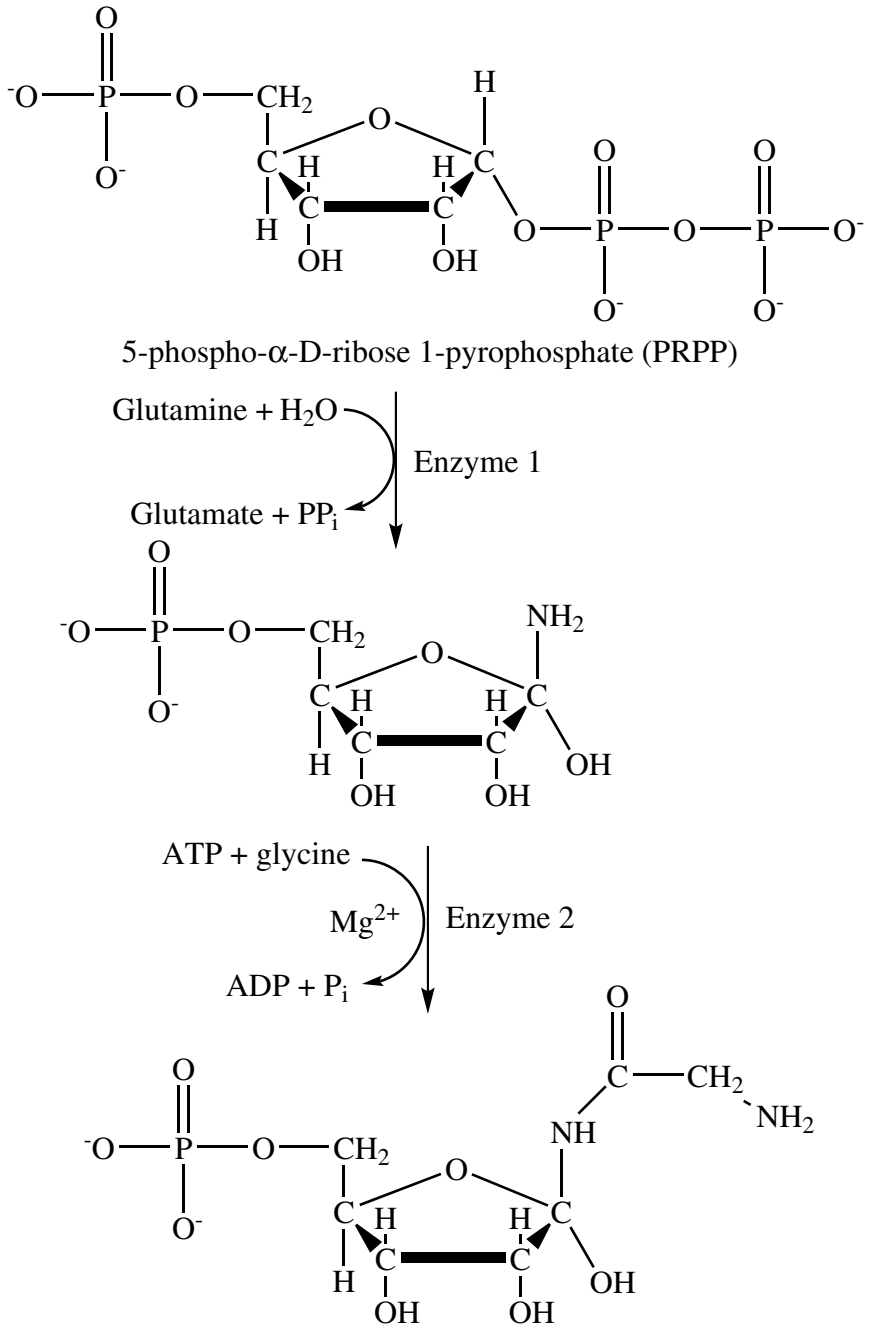
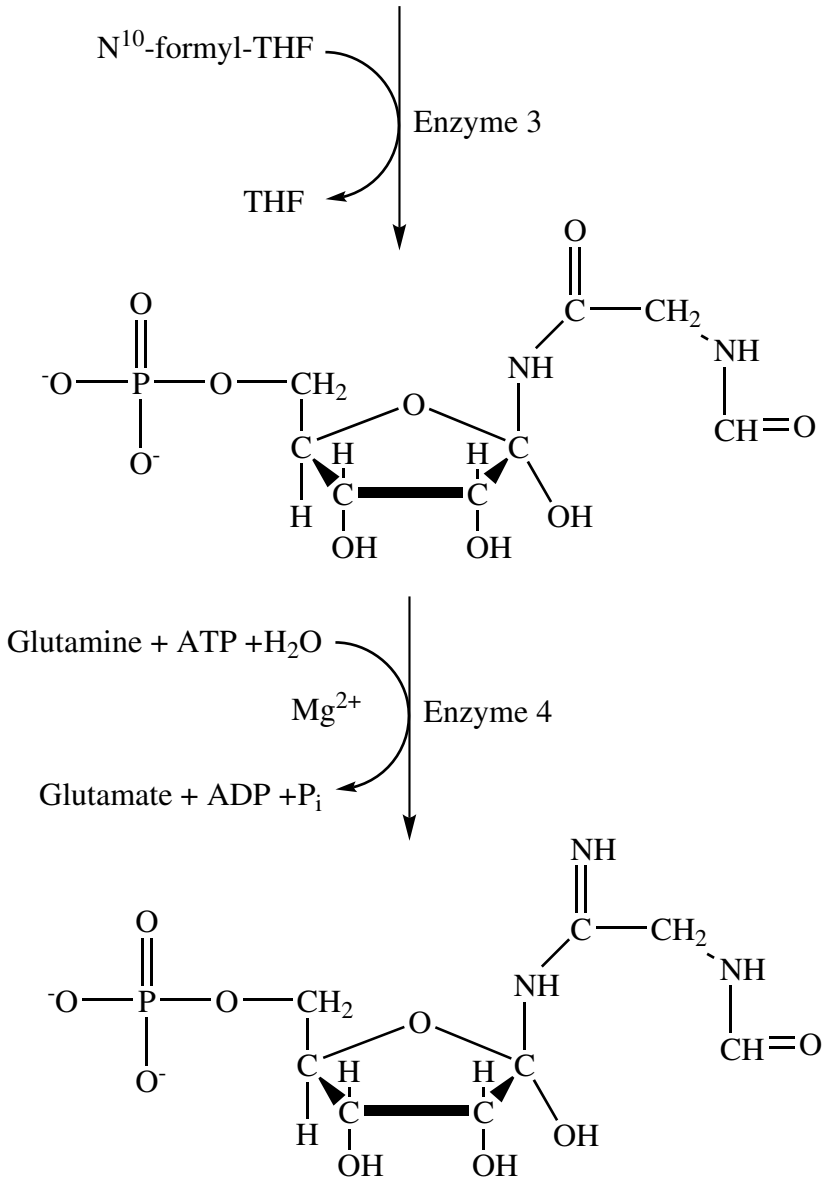
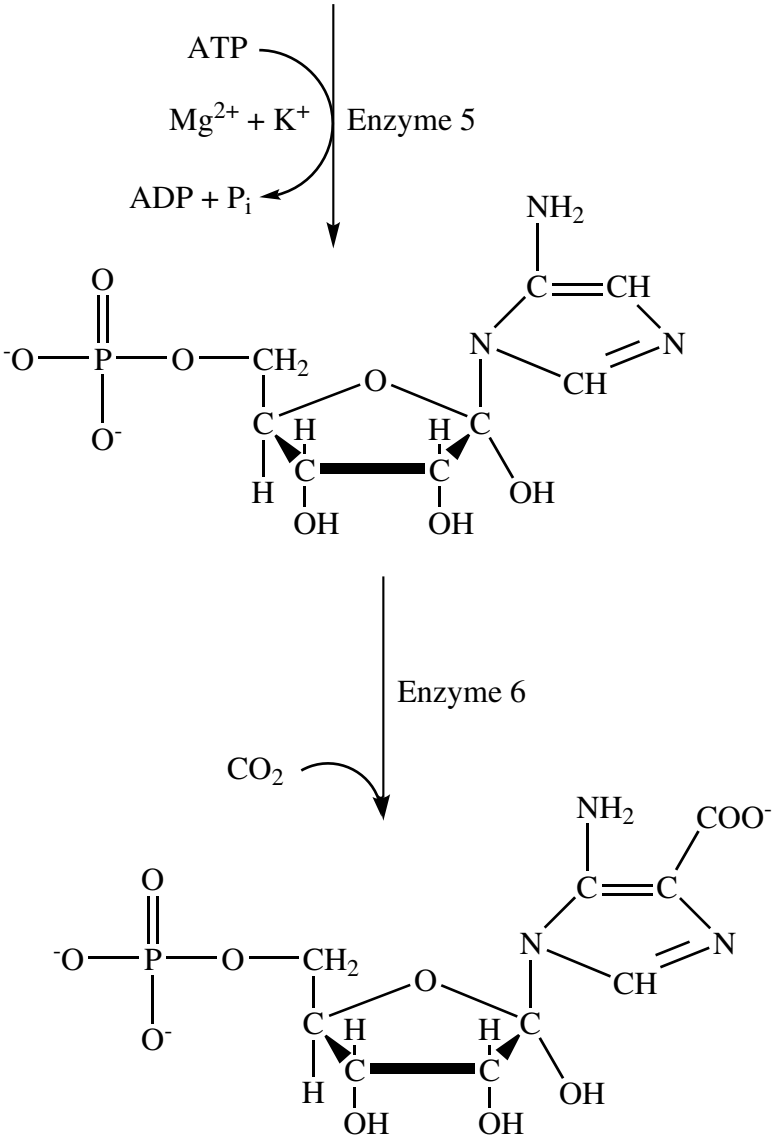


Figure 14-3: The ten steps necessary to convert PRPP (5-phospho-α-D-ribose 1-pyrophosphate) into inosine 5'-phosphate.


N10-formyl-THF
Enzyme 3
THF
Glutamine + ATP +H2O
Mg2+
Enzyme 4
Glutamate + ADP +Pi


ATP
Mg2+ + K+
Enzyme 5
ADP + Pi
NH2
C
CH
N
N
CH
O
-O
P
O
CH2
O
O-
C
H
H
C
H
C
C
H
OH
OH
OH
Enzyme 6
CO2
NH2
COO-
C
C
N
N
CH
O
-O
P
O
CH2
O
O-
C
H
H
C
H
C
C
H
OH
OH
OH

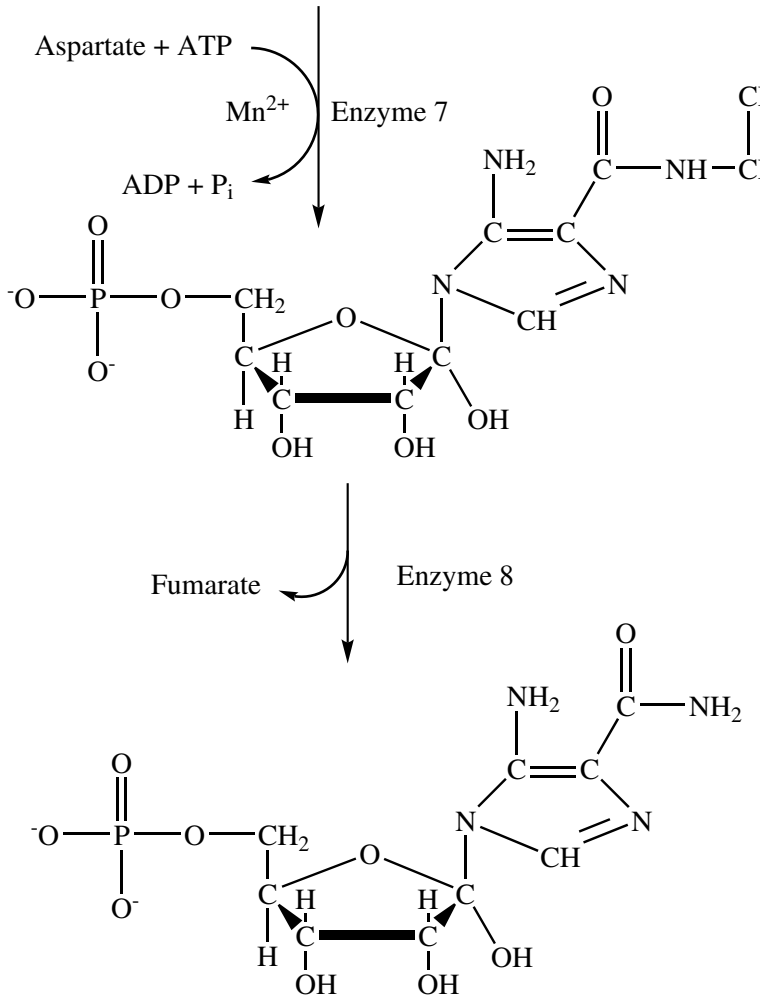

Aspartate + ATP
Mn^{2+}
Enzyme 7
ADP + P_i
Fumarate
Enzyme 8

Table 14-1 Ten Enzymes Necessary for Inosine Synthesis

Enzyme	*Name*
1	Amidophosphoribosyl transferase
2	Phosphoribosylglycinamide synthetase
3	Phosphoribosylglycinamide formyltransferase
4	Phosphoribosylformylglycimamide synthetase
5	Phosphoribosylaminoimidazole synthetase
6	Phosphoribosylaminoimidazole carboxylase
7	Phosphoribosylaminoimidazole-succinocarboxamide synthesase
8	Adenylosuccinate lyase
9	Phosphoribosylaminoimadazolecarboxamide formyltransferase
10	IMP cyclohydrolase

AMP synthesis

To convert IMP into AMP, it is necessary to transfer an amino group from an aspartate. This transfer requires two steps, and the energy to add aspartate to IMP comes from the hydrolysis of a GTP. The process is then completed by the loss of fumarate. The enzyme adenylosuccinate synthetase catalyzes the first step, and the enzyme adenylosuccinate lyase catalyzes the second. Figure 14-4 illustrates the process.

GMP synthesis

The conversion of IMP to GMP begins with the IMP dehydrogenase catalyzed oxidation to xanthosine 5'-phosphate. The coenzyme for this step is NAD^+. GMP synthetase catalyzes the next step — the amine transfer from glutamate. The energy for this step is supplied by the hydrolysis of ATP (Figure 14-5).

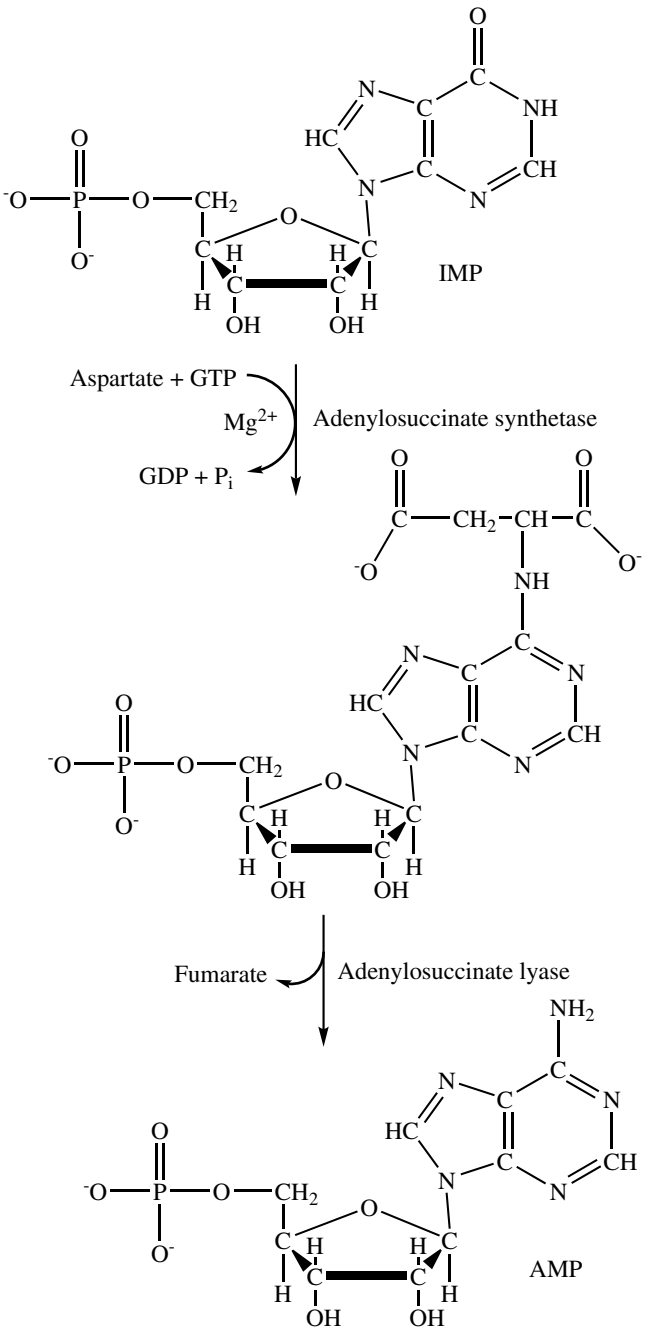


Figure 14-4: Conversion of IMP to AMP.

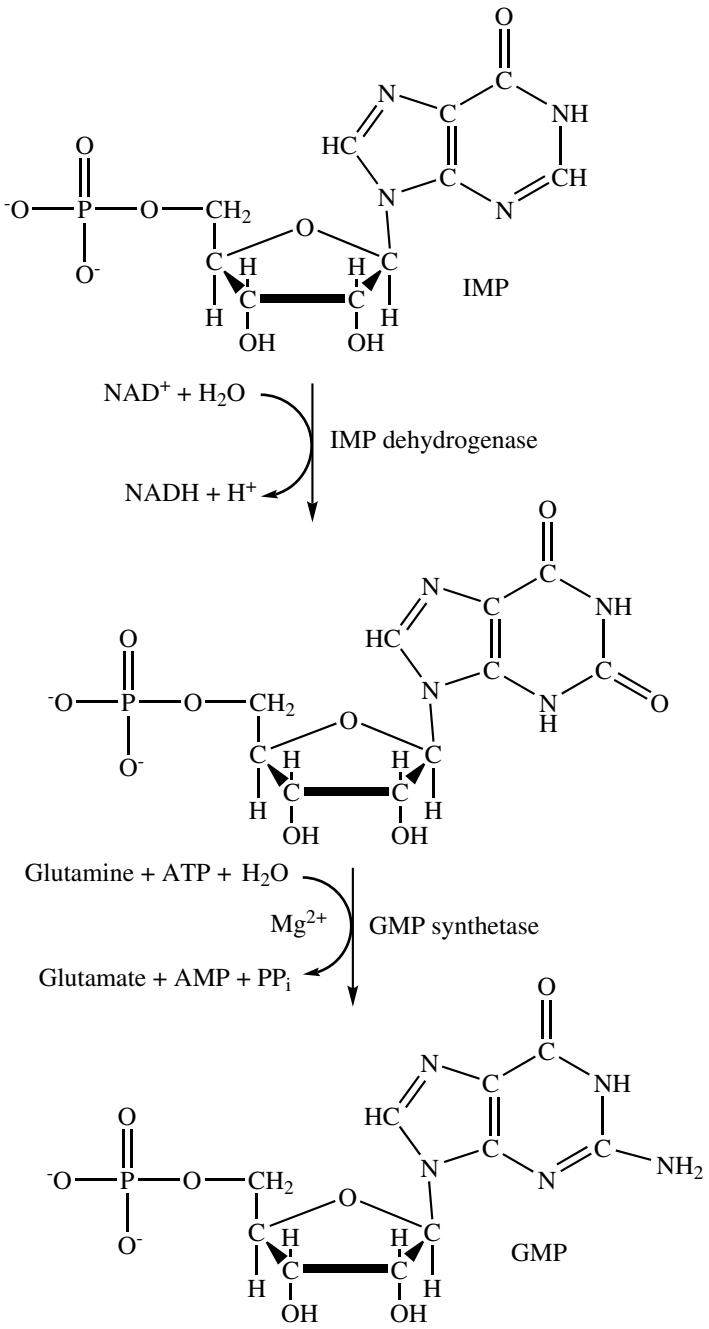


Figure 14-5: Conversion of IMP to GMP.

How much will it cost?

The biosynthesis of both AMP and GMP requires the hydrolysis of several high-energy bonds. To produce IMP from D-ribose 5-phosphate requires the hydrolysis of five high-energy bonds (one PP_i and five ATP). To convert IMP to AMP requires the hydrolysis of one more high-energy bond (from GTP). And to convert IMP to GMP requires the hydrolysis of two high-energy bonds — one ATP and one PP_i.

Anaerobic organisms, such as the bacteria responsible for tetanus or botulism, must oxidize four glucose molecules at two ATP per glucose to meet the energy requirement. An aerobic organism, like you, for example, needs to oxidize only one glucose molecule at 36 or 38 ATP per glucose. The preceding processes require a substantial amount of energy. Sometimes this energy requirement may be lessened by metabolic processes known as the *salvage pathways.* In the salvage pathways, nitrogen bases are recycled instead of synthesized. The nitrogen bases are then converted to nucleotides.

Pyrimidine Synthesis

The biosynthesis of pyrimidines follows a different path from purine synthesis. In this case, synthesis of the base takes place before attachment to the ribose. Ring synthesis requires bicarbonate ion, aspartic acid, and ammonia. Although it is possible to use ammonia directly, it usually comes from the hydrolysis of the side chain of glutamine.

First step: Carbamoyl phosphate

The initial step is to transfer a phosphate from an ATP to a bicarbonate ion to form carboxyphosphate, which in turn undergoes an exchange where ammonia replaces the phosphate to form carbamic acid. Whew! — got that? A second ATP transfers a phosphate to carbamic acid to form carbamoyl phosphate. Figure 14-6 summarizes these steps.

The primary enzyme for the process in Figure 14-6 is carbamoyl synthetase. One region of the enzyme is responsible for the synthesis of carbamic acid, whereas a second region hydrolyzes ammonia from glutamine. A third region completes the process, and a channel connects the three regions.

Next step: Orotate

The next step in pyrimidine synthesis is the formation of orotate, which will be joined to a ribose. It begins with the enzyme aspartate transcarbamoylate,

which joins aspartate to carbamoyl phosphate with the loss of phosphate. This forms carbamoylaspartate. Carbamoylaspartate cyclizes to dihydroorotate, which is oxidized by NAD^+ to orotate (Figure 14-2).

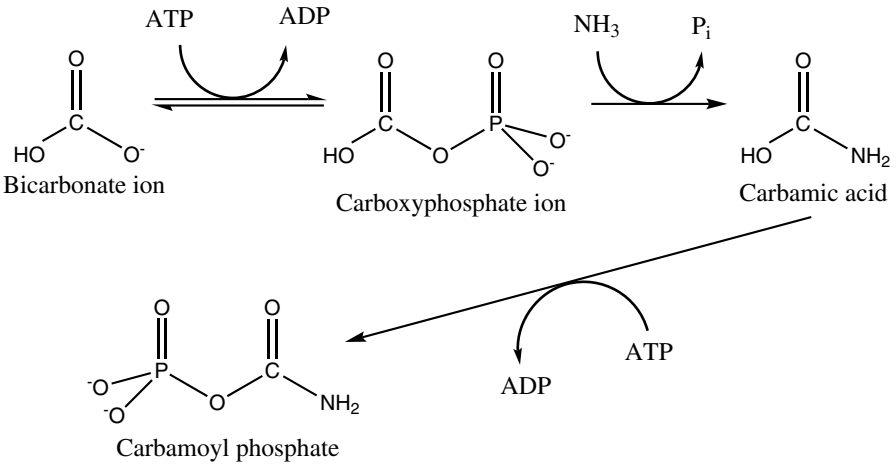


Figure 14-6: Synthesis of carbamoyl phosphate.

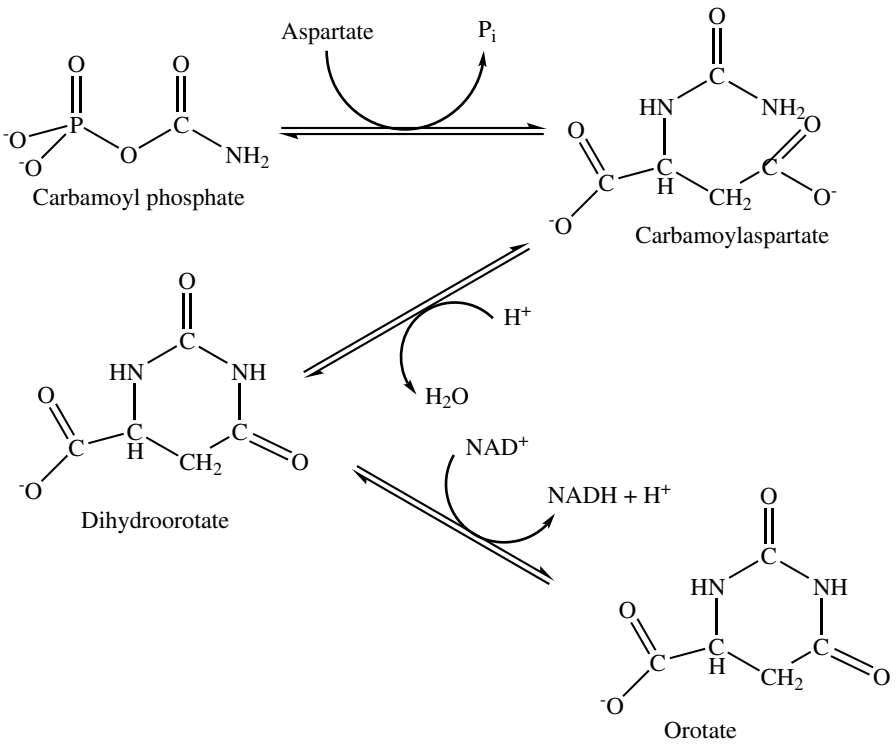


Figure 14-7: Formation of orotate from carbamoyl phosphate.

Orotate joins with 5-phosphoribosyl-1-pyrophosphate (PRPP) to form orotidylate, with pyrophosphate hydrolysis providing the energy necessary. The enzyme pyrimidine phosphoribosyltransferase is responsible for this reaction. The enzyme orotidylate decarboxylase catalyzes the decarboxylation of orotidylate to uridylate (UMP). Figure 14-8 illustrates these steps.

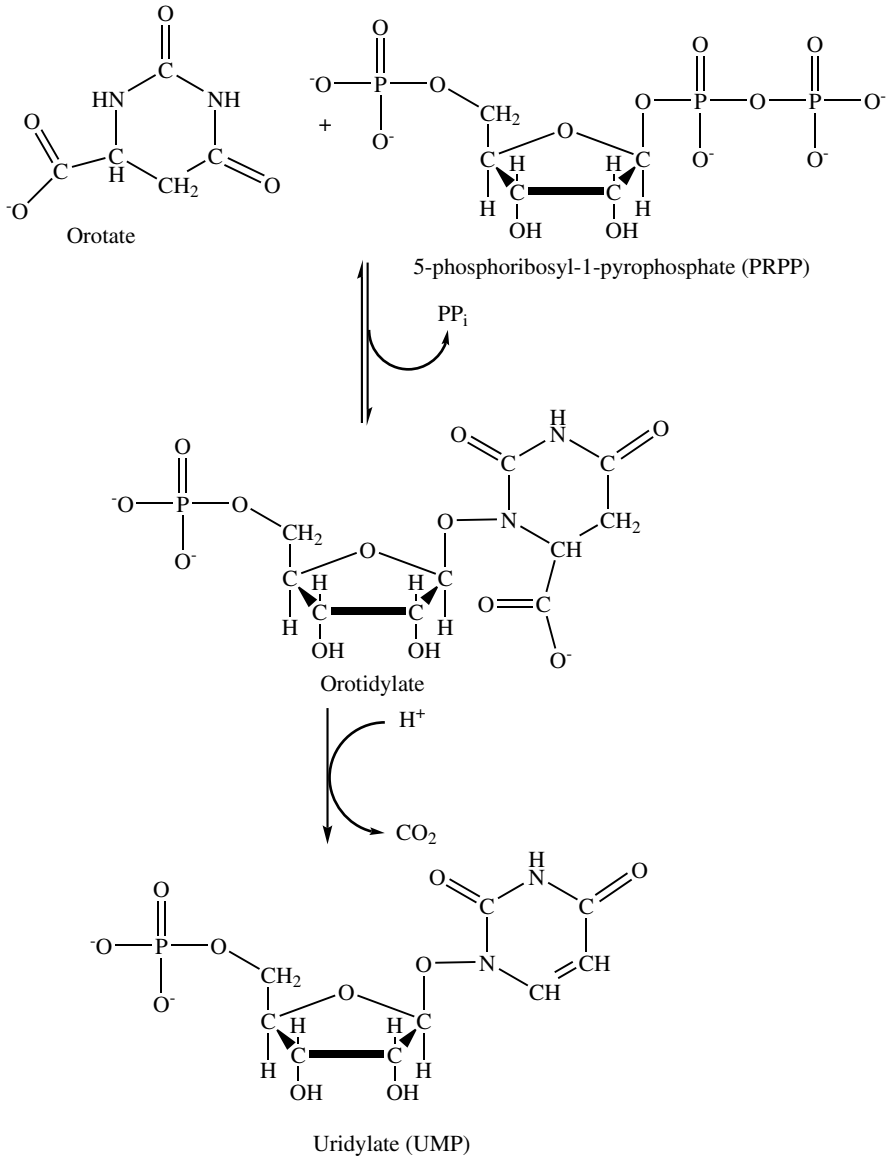


Figure 14-8: Conversion of orotate to uridylate (UMP).

Last step: Cytidine

The final nucleotide, cytidine, forms from uridinemonophosphate (UMP). The first step is to change UMP into UTP. UMP kinase transfers a pyrophosphate from ATP to UMP. Figure 14-9 shows this process.

Back to the Beginning: Catabolism

Catabolism, remember, is the breaking down of molecules to provide energy. In many cases, a complete breakdown is not necessary, because the products from a partial breakdown can be reused when necessary.

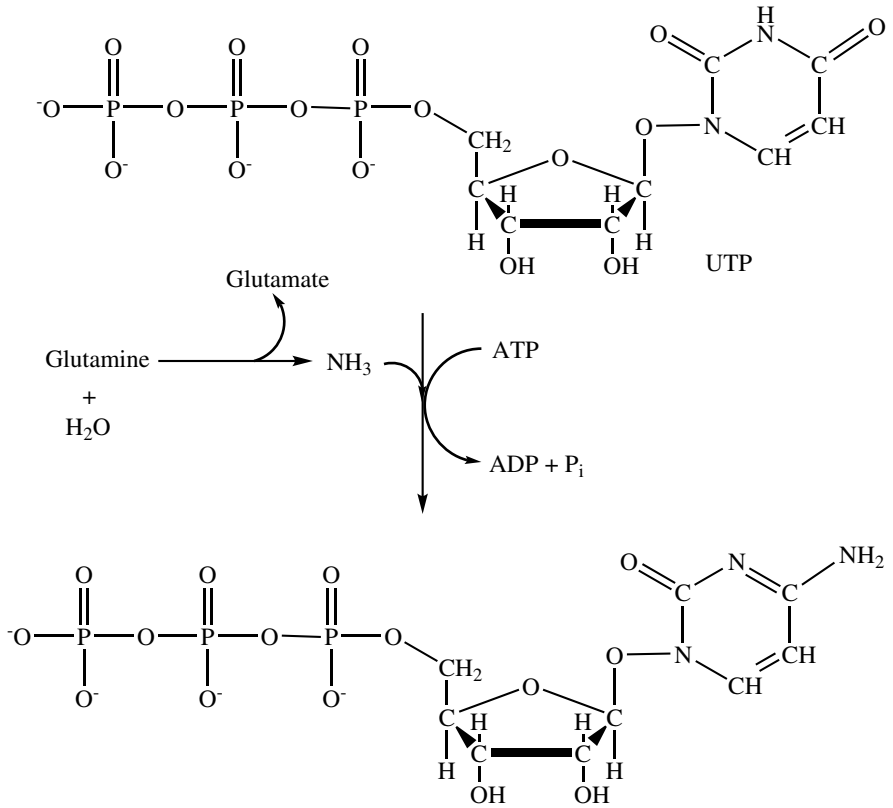


Figure 14-9: Conversion of UTP to CTP.

Nucleotide catabolism

The breakdown of the nucleotides begins with the removal of a phosphate group (from C-5). Next, a phosphate attaches to C-1 to give the sugar-1-phosphate, and the base leaves. In humans and many other species, uric acid (Figure 14-10) is the product of further degradation of purines. Other biochemical species further degrade uric acid into other products.

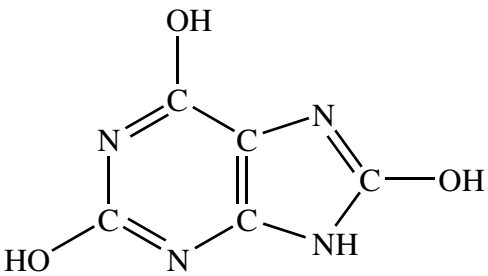


Figure 14-10: Structure of uric acid.

Amino acid catabolism

Hydrolysis of proteins yields the separate amino acids. It is possible to recycle these amino acids, use them in the synthesis of other amino acids, or produce energy from them. Through transamination it is possible to transfer an amino group from any amino acid (other than lysine, proline, or threonine), and an α-keto acid. The general category of enzymes that catalyzes this reaction is a transaminase, and the general reaction is shown in Figure 14-11. Nitrogen destined for elimination transfers to α-ketoglutarate to form glutamate. Transamination is important in the biosynthesis of alanine, aspartate, and glutamate.

Oxidative deamination of glutamate forms α-ketoglutarate (to be recycled), an ammonium ion (to enter the urea cycle) and, indirectly, 3 ATP. Glutamate dehydrogenase and either NAD^+ or $NADP^+$ are necessary for this.

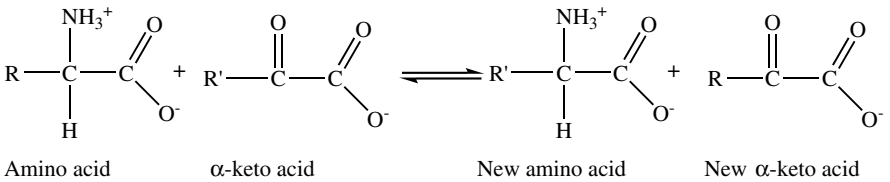


Figure 14-11: General transamination reaction.

The deaminated amino acid (α-keto acid) is further broken down. The α-keto acid may be broken down to pyruvate or some other material the body can

use to form glucose. These acids are called *glucogenic.* The alternative is to break down the α-keto acid to acetyl CoA and acetoacetic acid. These acids are called *ketogenic.* To further confuse you, some amino acids may be both glucogenic and ketogenic (see Table 14-2). These are the two possible fates of the carbon skeleton of the amino acids. The degradation of the amino acid transforms the carbon skeletons into intermediates in the citric acid cycle or into materials convertible to glucose.

Table 14-2 Glucogenic and Ketogenic Amino Acids

Glucogenic: Alanine, arginine, asparagine, aspartate, ccysteine, glutamate, glutamine, glycine, histidine, methionine, proline, serine, threonine, valine
Ketogenic: Leucine
Both: Isoleucine, lysine, phenylalanine, tyrosine, tryptophan

The general process is cyclic, with the various amino acids entering at different points. The basic scheme is shown in Figure 14-12.

Heme catabolism

The other important nitrogen compound in red-blooded organisms is *heme*. This species occurs in both hemoglobin and myoglobin. Hemoglobin is released as aged red blood cells are destroyed. The globin portion hydrolyzes to the appropriate amino acids. The iron separates from the heme and is stored in ferritin. Through a series of steps, bilirubin forms from the heme. The gall bladder temporarily stores bilirubin until the organism eliminates it.

Process of Elimination: The Urea Cycle

The catabolism of nitrogen-containing compounds yields recyclable nitrogen compounds and ammonia. Glutamine serves as temporary storage and transportation of the nitrogen — however, even small amounts of ammonia are toxic to humans. For this reason, ammonia must be converted to a less toxic form for elimination. The first step involves the conversion of ammonia, as the ammonium ion, to carbamoyl phosphate. The enzyme utilized for this conversion is carbamoyl phosphate synthetase. Figure 14-13 illustrates this reaction.

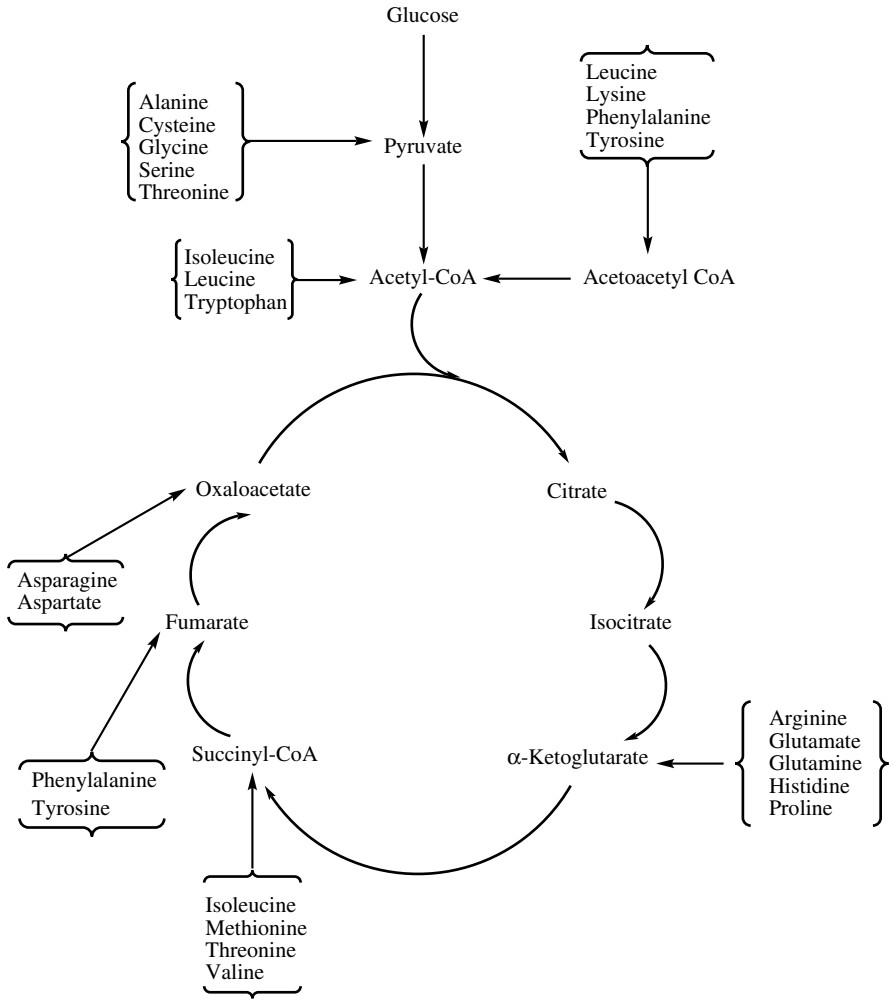


Figure 14-12: Fates of the amino acids.

Carbamoyl phosphate enters the urea cycle by joining to ornithine to produce citrulline, with the enzyme ornithine transcarbamoylase catalyzing this reaction. The enzyme arginosuccinate synthetase, with energy from the hydrolysis of ATP, joins aspartate to citrulline to form arginosuccinate. Arginosuccinase then catalyzes the splitting of arginosuccinate to fumarate

and arginine. The enzyme arginase completes the cycle by cleaving arginine into urea (for elimination) and ornithine (for recycling). The urea cycle and compounds involved in it are shown in Figures 14-14 and 14-15.

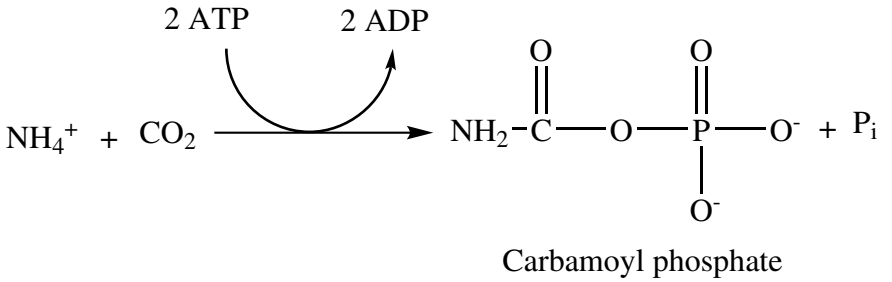


Figure 14-13: Formation of carbamoyl phosphate.

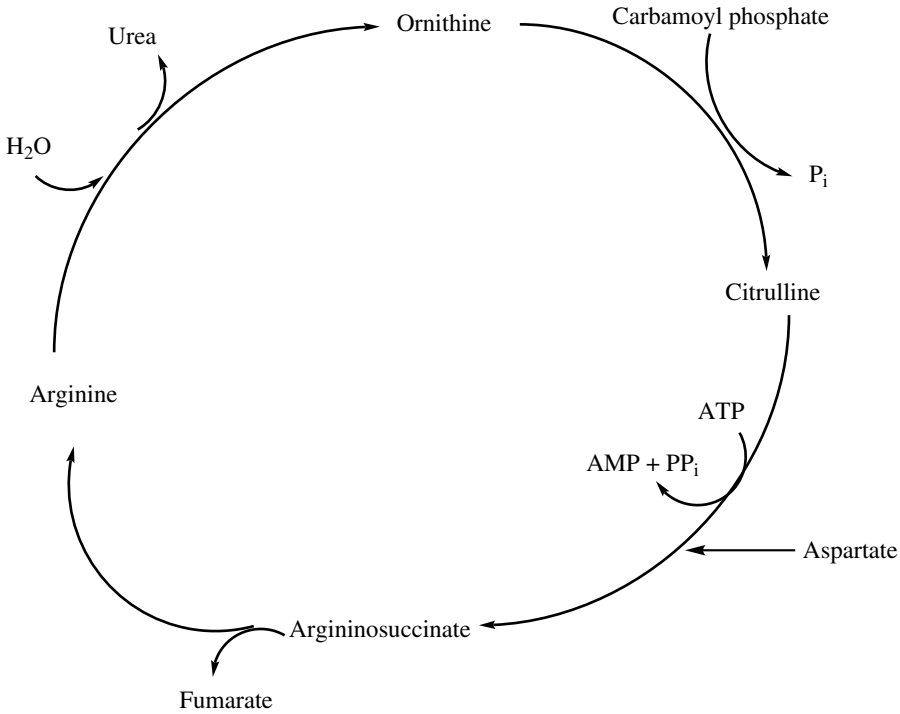


Figure 14-14: Overview of the urea cycle.

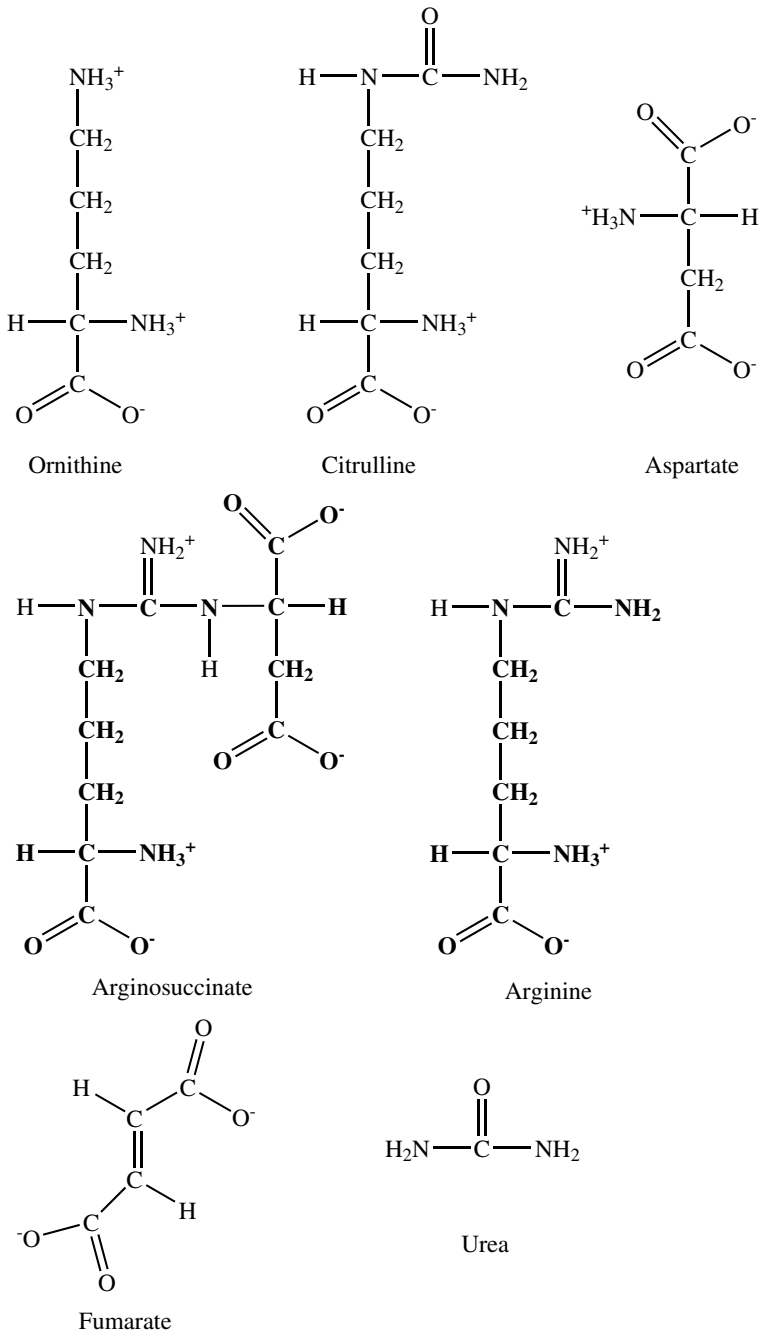


Figure 14-15: Compounds from the urea cycle.

Amino Acids Once Again

The synthesis of proteins requires 20 amino acids. If not readily available, humans can synthesize ten of these amino acids. These are the non-essential amino acids. The remaining ten amino acids, the essential amino acids, must come from the diet. Table 14-3 summarizes these amino acids.

Table 14-3 Essential and Non-Essential Amino Acids

Essential Amino Acids	*Non-Essential Amino Acids*
Arginine*	Alanine
Histidine	Asparagine
Isoleucine	Aspartate
Leucine	Cysteine
Lysine	Glutamate
Methionine	Glutamine
Phenylalanine	Glycine
Threonine	Proline
Tryptophan	Serine
Valine	Tyrosine

** Not essential in adults*

A *complete protein* supplies all essential amino acids. Not all proteins are complete — many are *incomplete proteins*. In order to avoid disorders due to amino acid deficiencies, the human diet must contain complete proteins.

Transamination is important in the biosynthesis of alanine, aspartate, and glutamate. It is easy to convert aspartate to asparagines and glutamate to glutamine. The synthesis of proline requires four steps beginning with glutamate. The synthesis of serine begins with the glycolysis intermediate 3-phosphoglycerate, and after three steps serine forms. It is easy to convert serine to glycine. If sufficient phenylalanine is available, the catalyzed oxidation converts it to tyrosine. If sufficient methionine is available, the body can convert some of the excess to cysteine. Arginine comes from the urea cycle, but infants do not get sufficient quantities from this source.

Metabolic Disorders

When something is out of whack with an organism's metabolism, problems arise that must be treated.

Gout

Gout is the result of overproduction of uric acid, which leads to the precipitation of sodium urate in regions of the body where the temperature is lower than normal (37°C). These low temperature regions are commonly found in the joints of the extremities. Sodium urate may also precipitate as kidney stones. Treatment is partially dietary and partly with drugs. Dietary restrictions include limiting the intake of foods high in nucleic acids (meats) and alcohol, which aggravates the conditions. Doctors often prescribe drugs that inhibit the enzyme that produces uric acid.

Gout may also be the result of faulty carbohydrate metabolism. A deficiency in glucose-6 phosphatase forces phosphorylated carbohydrates to form ribose 5-phosphate instead of glucose. Excess ribose 5-phosphate leads to excess PRPP, which, in turn, stimulates the synthesis of purines. The excess purines cause the production of more uric acid.

Lesch-Nyhan syndrome

Lesch-Nyhan syndrome is another example of defective purine catabolism leading to excess uric acid. Patients with this disorder normally excrete 4–5 times as much uric acid as gout patients do. This is a genetic disease that is a recessive X-linked trait, the trait is carried by the mother and is passed on to her son. There is no treatment for this disease at the present time.

Albinism

Albinism, a recessive trait, is an inborn error of tyrosine metabolism. Tyrosine is the precursor of melanin, the pigment responsible for hair and skin color. In at least one form of albinism, the problem appears to be due a deficiency of the enzyme tyrosinase. A variation of albinism involves a temperature-sensitive form of tyrosinase. The enzyme is only effective at lower than normal temperatures, as found in the extremities. This form of tyrosinase is responsible for the coloration of Siamese cats.

Alkaptonuria

Alkaptonuria is a benign condition that manifests itself as a darkening of the urine. The condition is the result of a problem in the catabolic breakdown of phenylalanine and tyrosine. A defective enzyme leads to an accumulation, and subsequent elimination, of one of the reaction intermediates.

Phenylketonuria

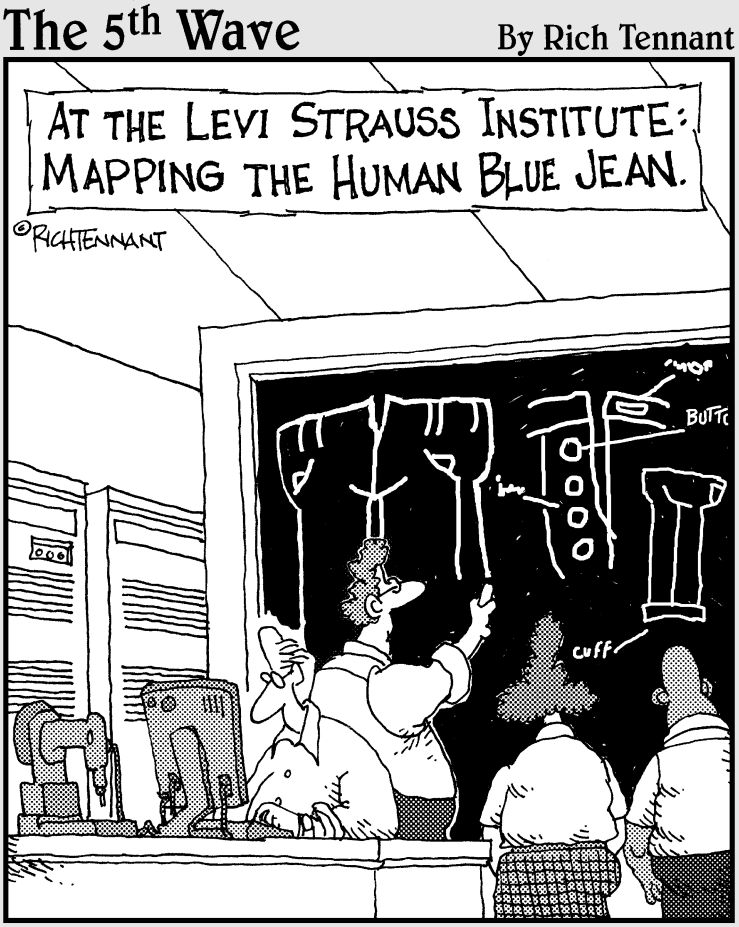
Phenylketonuria, or PKU, is the result of a deficiency in the enzyme phenylalanine 4-monooxygenase, which results in a problem in phenylalanine metabolism. The consequence is an accumulation of phenylalanine in the blood. High levels of phenylalanine enhance transamination to form abnormally high levels of phenylpyruvate. High levels of phenylpyruvate damage the brains of infants with the condition.

The high levels of phenylalanine lead to competitive inhibition of the enzymes responsible for melanin production from tyrosine. Because little tyrosine converts to melanin, afflicted infants have light blonde hair and fair skin (similar to albinism).

Early diagnosis in infants is important to prevent brain damage. One test for PKU is to add $FeCl_3$ to the patient's urine. Phenylpyruvate reacts with iron ions to produce a green color. Another test is to assay for phenylalanine 4-monooxygenase activity. Treatment consists of maintaining a diet low in phenylalanine until at least the age of three.

Part V

Genetics: Why We Are What We Are



In this part . . .

We roll up our sleeves and return to the subject of genes and DNA to look at them much more closely. We cover the way DNA replicates itself and look at a number of applications related to DNA sequencing. Then it's off to RNA transcription and protein synthesis and translation. At the very end we mention some of the goals and questions sought by the Human Genome Project.

Chapter 15

Photocopying DNA

In This Chapter

- Learning about replication
- Checking out recombinant DNA
- Examining DNA ssequencing
- Discussing ethical issues

In 1958, Francis Crick postulated what became the "central dogma of molecular biology." In this postulate, he, and later others, reasoned that DNA was the central source of genetic information and that it passed on some of this information to form RNA, which, in turn, passed this information on to form proteins. This central dogma is an extension of the one-gene one-protein hypothesis. To achieve this, the DNA must be able to pass on its information both to later generations (*replication*), and to RNA (*transcription*). RNA must finish the series by forming the appropriate proteins (*translation*).

Some RNA, especially some viral RNA, can undergo replication and even reverse-transcription — thus, RNA can produce both RNA and DNA. Genetic researchers initially thought this was in conflict with the central dogma; but Crick reasoned that RNA creating DNA was an extension of this postulate.

Many of the viruses capable of reverse-transcription are cancer causing.

The primary structure of DNA consists of two polynucleotide strands held together by hydrogen bonds. Adenine forms hydrogen bonds to thymine, and cytosine forms hydrogen bonds to guanine (Figure 15-1). The sequence of nitrogen bases contains the genetic information. The DNA molecules wrap around a protein called a *histone* — the combination of eight histones with the associated DNA is a *nucleosome.* (We talk more about histones in Chapter 16.)

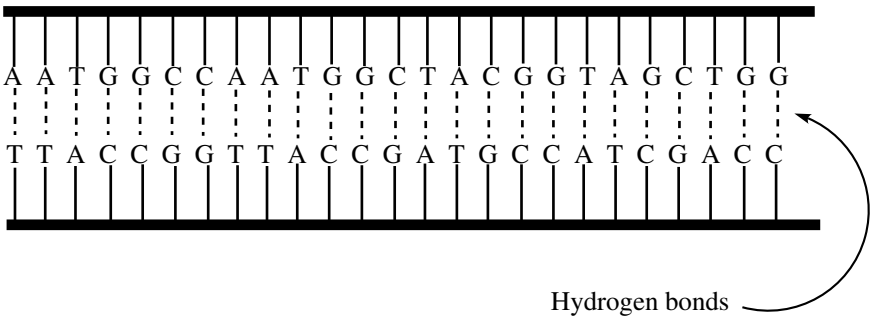


Figure 15-1: A schematic illustration of the base pairs present in a segment of DNA.

A *gene* is a portion of a DNA molecule that carries specific information. The portion of the gene coding for that specific information is called an *exon*. The portion of a gene that does not code for specific information is an *intron*.

Let's Do It Again: Replication

Replication is the process that produces new DNA molecules. One DNA molecule produce two DNA molecules in a process where the DNA must unwind and open — kind of like a zipper. New nucleotides bind to the backbone of each strand of the opened DNA by forming hydrogen bonds to the nucleotides (the zipper's "teeth") that are already present. The process proceeds along the opening DNA strand until each half of the original DNA has a complementary strand hydrogen bonded to it. The result is two DNA double helices each with half old DNA, and half new. It doesn't sound like much fun, but it works for DNA. Replication is illustrated in Figure 15-2.

The specific hydrogen bonding forces the new strands to contain a nucleotide sequence that is complementary to the nucleotide sequence in the old strand. Therefore, it can create an exact duplicate of the original DNA.

This description of replication is a simplification. It barely scratches the surface of this complicated process, but it should give you enough background information in order to understand what comes next.

The first step in understanding replication was the discovery of DNA polymerase from *Escherichia coli*. Subsequent studies showed that this enzyme needed a DNA template and all four deoxyriboside triphosphates (dATP, dCTP, dGTP, and dTTP). In addition, a short section of RNA called a *primer* is also needed. The enzyme prefers a single DNA strand for the template in order to produce a complementary strand.

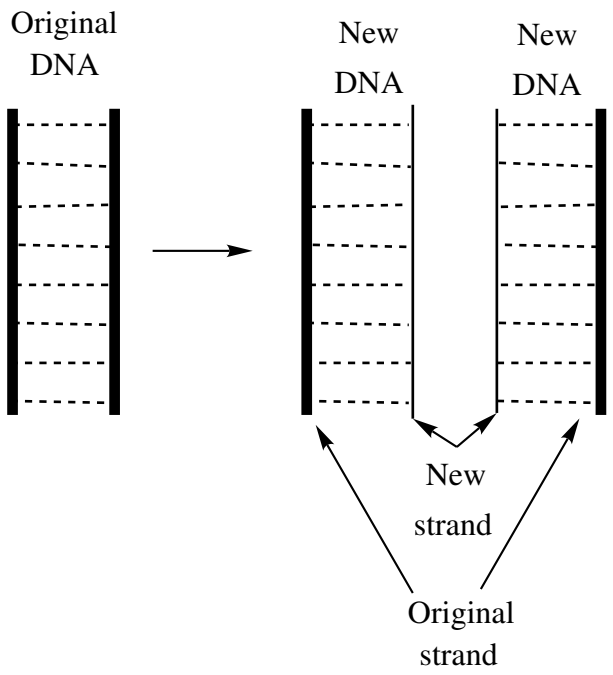


Figure 15-2: A simplified representation of replication.

During replication, simultaneous duplication of the two strands of DNA occurs. Because the two strands of DNA are anti-parallel, the mode of synthesis is different for each strand, but the overall process is the same: moving from one end to the other. For one strand the synthesis is from 5' → 3'. On the other strand it appears to be from 3' → 5', but in actuality it is also 5' → 3'. There is a complication on the 3' → 5' strand (which we discuss later in this section). See Figure 15-3.

The initiation of replication begins at a particular site, and, once initiated, a series of fragments form discontinuously along one strand and continuously along the other strand. These discontinuous fragments, known as *Okazaki fragments,* contain from 1,000 to 2,000 nucleotides. The synthesis of the fragments is always in the 5' → 3' direction. See Figure 15-4. Note that Figures 15-3 and 15-4 appear to be different at first glance. Figure 15-3 represents a simplified view of the overall process, whereas Figure 15-4 illustrates in more detail how this overall process occurs.

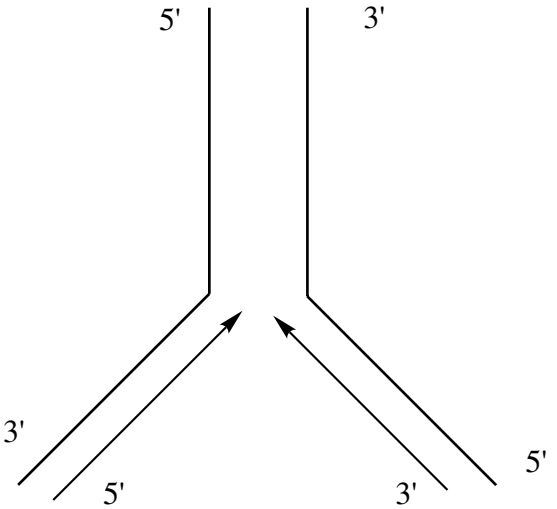


Figure 15-3: A simplified scheme of the replication of DNA.

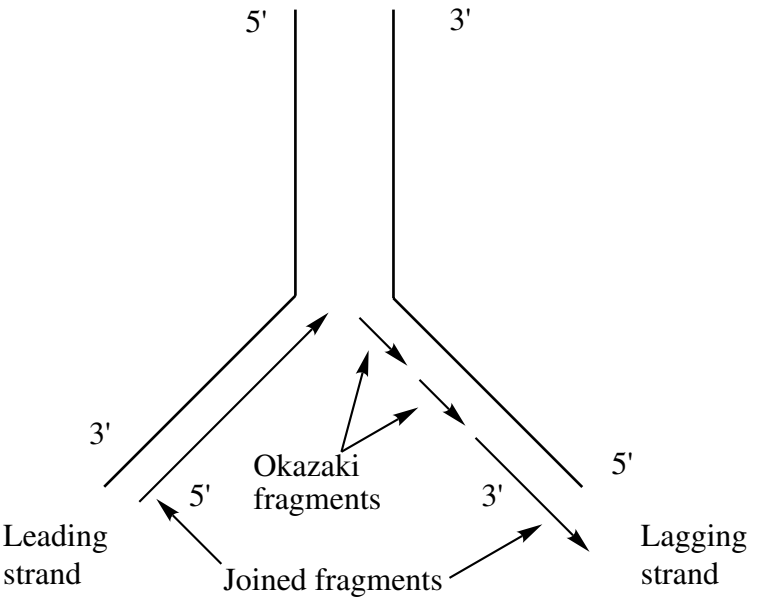


Figure 15-4: Detailed scheme of the replication of DNA.

Researchers unexpected found that RNA synthesis is a prerequisite for the replication of DNA. Initially, an RNA primer, typically 20 to 30 nucleotides in length, forms on a single DNA strand. Once formed, deoxyribonucleotide nucleotides add to the 3' terminus. Later, it is necessary to remove the RNA primer and attach the appropriate DNA fragment to produce the completed DNA.

At least a portion of the double-stranded DNA must be separated before replication can occur, and the separated portions can serve as templates. Enzymes known as *helicases* are responsible for this separation. The energy needed comes from the hydrolysis of ATP. The mechanism of separation is not well understood and is still under investigation. Apparently, the helicase binds more strongly to one strand of the DNA than the other so that the enzyme squeezes in and pushes the other strand away. ATP hydrolysis provides the energy necessary to cause the enzyme to move along the one strand nucleotide by nucleotide. This results in regions of the DNA opening like the afore-mentioned zipper.

DNA polymerases

DNA polymerases are the enzymes responsible for joining the nucleotide triphosphate fragments to produce a strand of DNA, acting as the bricklayers and carpenters in its construction. This process will only occur in the presence of a DNA template (parent DNA). Before the enzyme can connect a nucleotide, the nucleotide must bind to the appropriate site on the template.

There may be more than one DNA polymerase present in a cell. For example, in *E. coli* three different enzymes perform this task. These enzymes may also act as exonucleases. An *exonuclease* has the opposite function as a polymerase; that is, it removes nucleotides from the DNA strand.

The addition of the nucleotides is always to the 3' end of a polynucleotide chain. DNA polymerases cannot start building a nucleotide from scratch — there must be a polynucleotide already present. In contrast, RNA polymerase *can* begin from scratch. RNA polymerase generates the RNA primer, using ribonucleotides, at the beginning of replication. DNA polymerase then takes over the task and adds deoxynucleotides to the RNA primer. The polymerization requires the presence of two metal ions to enable the joining of the nucleotide to the polynucleotide.

Replication of DNA needs to be error free to ensure proper transmission of genetic information, and DNA polymerases are extremely effective in reducing errors. The enzyme binds tightly to the template and to the incoming nucleotide. This nucleotide is initially bound to the template through hydrogen bonding. If the wrong nucleotide is present, the subsequent binding to the polymerase is ineffective, and the nucleotide is "rejected." In addition to this checking, DNA polymerase also proofreads the preceding nucleotide to make sure it is correct. If the wrong nucleotide is present, it does not fit properly, making it necessary to remove the erroneous nucleotide from the polynucleotide so that the correct nucleotide may enter. The exonuclease portion of the polymerases performs this function. The polymerase proofreads the polynucleotide chain as polymerization proceeds. Proofreading is in the reverse direction (3' $\rightarrow$ 5'). There must be a nucleotide already in place before the polymerase can proofread. (We hope that our proofreader is as good as the DNA polymerases.)

The current model of DNA replication

In vitro studies show that in *E. coli,* replication begins when a protein binds a region of the DNA containing four specific binding sites. This is the *origin of replication* site. Once this protein binds, a helicase enzyme attacks the DNA and begins to unwind and separate the two strands. A third protein enters and holds the DNA strands open so that replication can continue. This third protein is the single-strand binding protein. The partially opened DNA and associated proteins are called the *prepriming complex* (Figure 15-5).

It is necessary to expose the DNA templates in this manner. A DNA strand may have more than one origin of replication site — this allows replication to occur in many places at one time. Simultaneous replication allows the cell to replicate the entire strand in less time.

Replication cannot continue until the exposed template is primed. A type of RNA polymerase known as *primase* binds to the prepriming complex in a region known as the *primosome.* Primase synthesizes a short RNA segment of about five nucleotides. Primase is capable of performing this function because its proofreading ability is not as efficient as that of DNA polymerase. For this reason, a nucleotide doesn't need to already be present to be checked. Because the primer consists of ribonucleotides instead of deoxyribonucleotides, it is temporary and will be detected and removed later. Once removed, the appropriate deoxyribonucleotides join to complete the DNA strand (Figure 15-6).

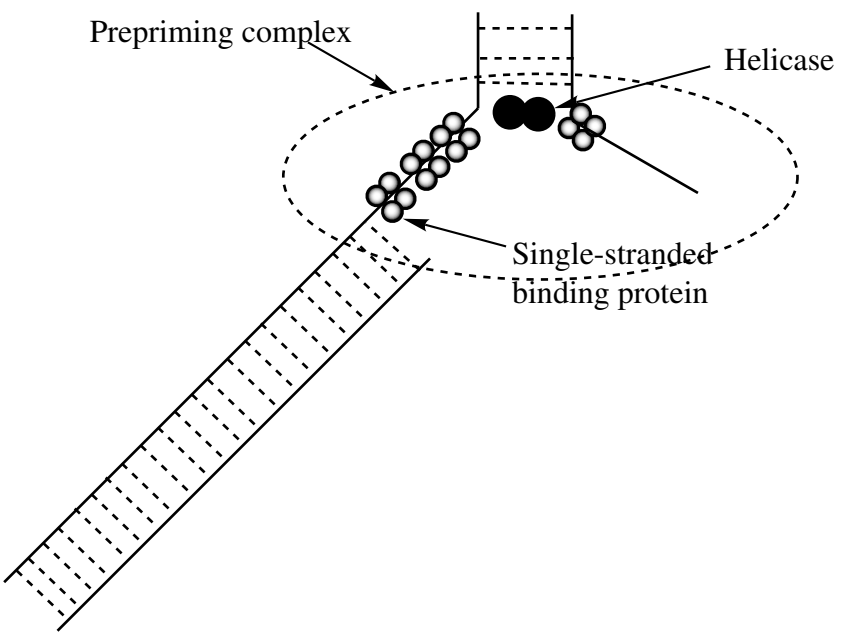


Figure 15-5: A simplified view of the prepriming complex.

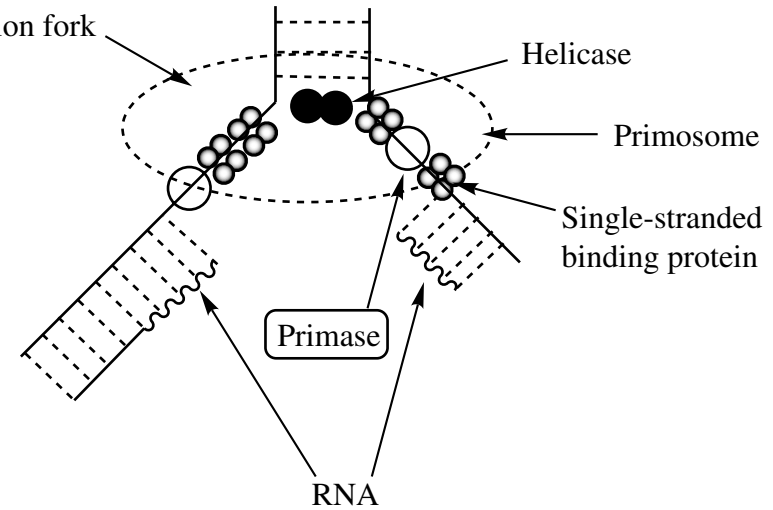


Figure 15-6: Formation of the RNA primer.

Although both strands of DNA serve as templates, the replication process differs on each strand. The point where the strands split and replication occurs is the *replication fork.* Because the two strands are anti-parallel, and DNA polymerase only works in the 5' → 3' direction, direct replication only works on one strand — called the *leading strand.* The other strand is the *lagging strand.*

As the DNA strands separate, eventually there is enough room to begin synthesis in the reverse direction on the lagging strand. (The reverse direction on the anti-parallel lagging strand is still 5' → 3'.) Replication on the lagging strand is discontinuous, and fragments of about 1,000 nucleotides form, called, as we have already noted, Okazaki fragments. DNA ligase then joins the fragments to produce a continuous strand.

DNA polymerase III holoenzyme (complete enzyme) simultaneously produces DNA on both the leading and lagging strands, though the mechanisms on the two strands are different. On the leading strand the process is continuous, whereas on the lagging strand it is discontinuous and more complex. To carry out the polymerization on the lagging strand, this strand loops around so that polymerization in the 5' → 3' direction can take place. After about 1,000 nucleotides — an Okazaki fragment — the polymerase releases the loop and begins a new loop and fragment. Each Okazaki fragment has a RNA primer. DNA polymerase I synthesizes DNA in the gaps between the fragments and removes the primer section. DNA ligase then joins the fragments (Figure 15-7). Wow! John wishes the carpenters who built his new house were that efficient!

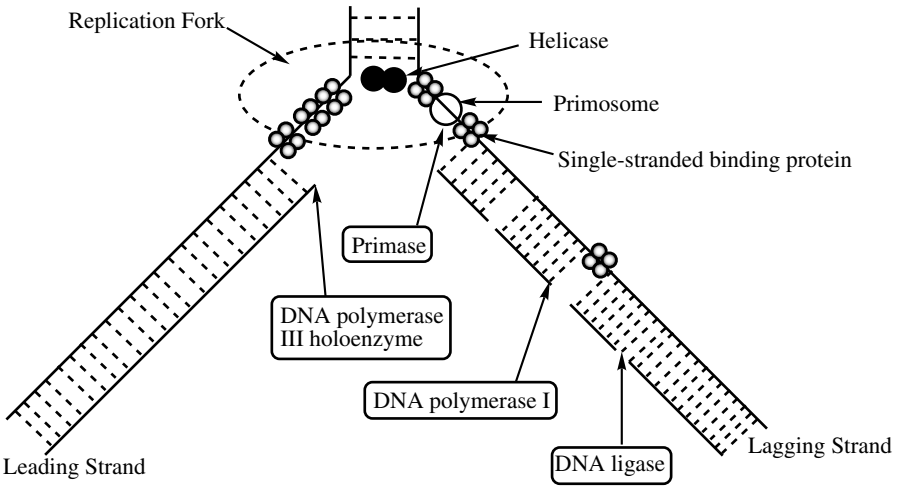


Figure 15-7: An expanded representation of the replication fork.

The ends of the DNA strands require a different procedure than does the majority of the strand, and this procedure is especially important on the lagging strand. If care were not taken, each replication cycle would result in a shorter DNA strand, eventually leading to the loss of important genetic material. To resolve this problem the ends of the DNA strands contain telomeres. *Telomeres* are DNA segments containing hundreds of repeating units. In humans, the repeating units are the hexanucleotide AGGGTT. The enzyme *telomerase,* in humans, detects the primer sequence GGTT and repeatedly attaches the hexanucleotide units, completing the DNA strand.

Mechanisms of DNA repair

All cells have a variety of DNA repair mechanisms, which are necessary to repair defective DNA and ensure retention of genetic information. Damage to DNA may occur during replication or by the action of radiation or chemicals. There is a rare error known as *xeroderma pigmentosum,* which impairs these repair mechanisms. Individuals suffering from it are extremely susceptible to cancers, especially skin cancers. Eventually the skin cancers metastasize, leading to death. The three general types of repair mechanisms are

- Direct repair
- Base-excision repair
- Nucleotide-excision repair

One example of damage needing repair is the formation of a thymine dimer (Figure 15-8) by ultraviolet (UV) light. The *thymine dimer* is an example of a

pyrimidine dimer, and its presence causes distortion of the DNA in the region. Other problems include base mismatches and missing or additional bases.

Direct repair

Here, the correction of the problem occurs in place. The photoreactivating enzyme, DNA photolyase, binds to the cyclobutane ring present in a thymine dimer, using light energy to cleave this dimer into the original bases.

Base-excision repair

In base-excision repair, the correction of the problem involves removal and replacement of the base. This is necessary whenever a modified base is present. There are various causes of modified bases, such as radiation or certain chemicals. The presence of a modified base normally results in a recognizable distortion in the DNA molecule. An enzyme, behaving as a glycosylase, cleaves the glycosidic bond to release the base from the deoxyribose. The result is an AP site (AP meaning *apurinic* or *apyrimidinic*). With apurinic, the purine base is absent, in apyrimidinic the pyrimidine base is absent. An AP endonuclease recognizes this site and cuts the DNA backbone adjacent to the site. Next, a deoxyribose phosphodiesterase completes the removal of the remaining deoxyribose phosphate. DNA polymerase I then inserts a replacement nucleotide to match the nucleotide in the complementary DNA strand. Finally, DNA ligase connects the units to yield the repaired strand — kind of like an electrician cutting out a bad circuit and splicing a good one in its place.

Nucleotide-excision repair

In nucleotide-excision repair, the correction of the problem involves the removal of a segment of DNA around the problem followed by its replacement. When this mechanism occurs, a DNA strand on both sides of the error is cut from the DNA strand. Typically, an exonuclease removes a 12-nucleotide section. DNA polymerase I then synthesizes a replacement segment of the strand. DNA ligase then finishes the repair.

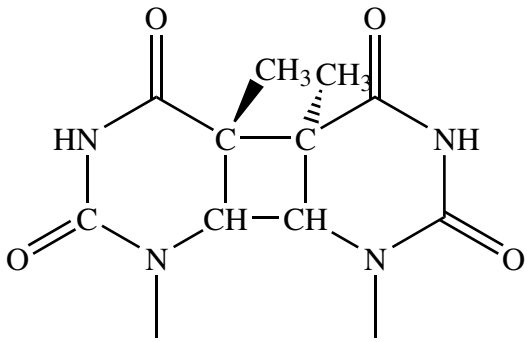


Figure 15-8: Structure of a thymine dimer.

Mutation: The good, the bad, and the ugly

Several types of mutations are known. DNA repair mechanisms try to prevent new mutations — however, such mechanisms are not always effective. Known mutations include the substitution of one base pair for another, the insertion of one or more base pairs, and the deletion of one or more base pairs. Changes, especially subtle ones, may occur during or after replication.

The substitution of one base for another is a common mutation. There are two types of substitutions. The replacement may be of a purine by the other purine (Figure 15-9) or the replacement of a pyrimidine by the other pyrimidine (Figure 15-10). This type of error is a *transition.* The other type of substitution is the replacement of a purine for a pyrimidine or vice versa. This latter type is a *transversion.*

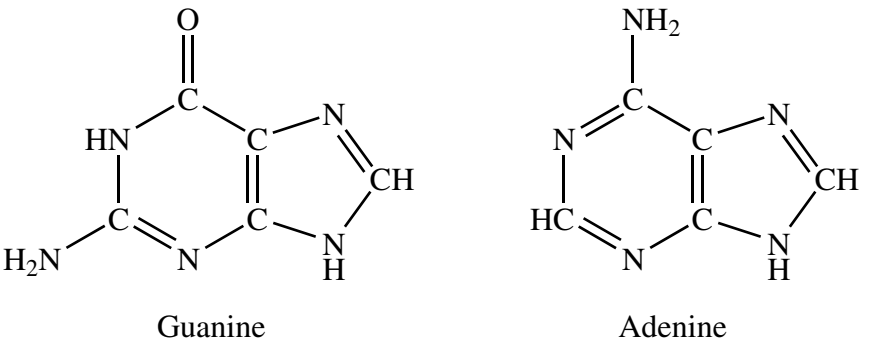


Figure 15-9: The purines.

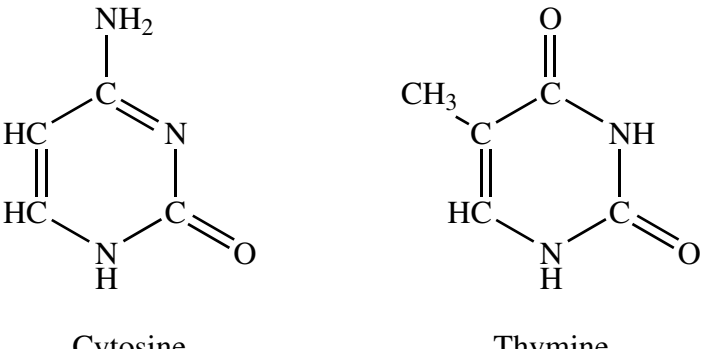


Figure 15-10: The pyrimidines.

Any uncorrected discrepancy in the genetic code will become "normal" in all future generations. The new genetic code is a *mutation.* The change in the base sequence may or may not affect the amino acid for which the codon codes. For example, changing from GTT (coding for leucine) to GTG (also coding for leucine) results in no change. However, if the change results in coding for a different amino acid, the resultant protein will function differently. If the new protein exhibits improved function, the organism benefits from the change. But if the new protein exhibits impaired function — the more likely situation — the organism suffers from the change. Problems from impaired function are genetic diseases. Table 15-1 lists some of these.

Table 15-1 Some Genetic Diseases in Humans

Disease	*Defective Protein*
Acatalasia	Catalase
Albinism	Tyrosinase
Cystic fibrosis	CF transmembrane conductance regulator
Fabray's Disease	α-Galactosidase
Gaucher's Disease	Glucocerebrosidase
Goiter	Iodotyrosine dehalogenase
Hemochromatosis	Hemochromatosis
Hemophilia	Antihemophilic factor (factor VIII)
Hyperammonemia	Ornothine transcarbamylase
McArdle's Syndrome	Muscle phosphorylase
Niemann-Pick Disease	Sphingomyelinase
Phenylketonuria	Phenylalanine hydroxylase
Pulmonary emphysema	α-Globulin of blood
Sickle cell anemia	Hemoglobin
Tay-Sachs Disease	Hexosaminidase A
Wilson's Disease	Ceruloplasmin (blood protein)

Restriction enzymes

Although not directly related to replication, restriction enzymes are important tools in genetic research. *Restriction enzymes,* or *restriction nucleases,* are capable of cutting DNA into fragments. These were first found in prokaryote cells like *E. coli* where these enzymes locate and destroy invading DNA, such as that of a bacteriophage, but leave the cell's own DNA alone. Recent research focuses on the fact that it is possible to manipulate these fragments so that DNA ligases can join the fragments into new DNA. Restriction enzymes are important in vitro biochemical tools that act as very accurate molecular scalpels. Cleavage may leave both DNA strands of equal length or one strand may be longer than the other (a *staggered* cut).

More than 100 restriction enzymes have been identified and are available for research. These enzymes recognize specific regions in the DNA and cleave DNA molecules into specific fragments. Because these fragments are smaller than the parent DNA is, they are easier to manipulate and analyze. Testing a strand of DNA with a series of restriction enzymes can provide a fingerprint of cleaved fragments. In fact, you can map the structure of DNA.

Many times in reading descriptions of genetic determination and modification, you will run across the terms *in vivo* and *in vitro. In vivo* means in the cell, whereas *in vitro* means in a test tube.

Mendel Rolling Over: Recombinant DNA

Recombinant DNA technology allows the synthesis of DNA strands that contain one or more genes not originally present. The addition of new genes enables an organism to produce new biochemicals. For example, *E. coli* has been engineered to produce human insulin. Recombinant DNA technology also allows biochemists to add a gene to compensate for a defective gene.

Restriction enzymes are capable of removing DNA fragments of interest. It is then necessary to join one of these fragments to another DNA strand for replication. The DNA to which the fragment of interest is attached is the *vector.* Common vectors include plasmids. A *plasmid* is a naturally occurring DNA circle. The first step in adding the fragment is to create a staggered cut in the DNA of the vector. The longer end of the staggered cut is a "sticky" or cohesive end. It is possible to attach any DNA fragment to the sticky end if it has the complementary DNA sequence. The complementary sticky end will be present if the same restriction enzyme was used to excise the fragment of interest. DNA ligase completes the joining of the fragment to the vector (Figure 15-11).

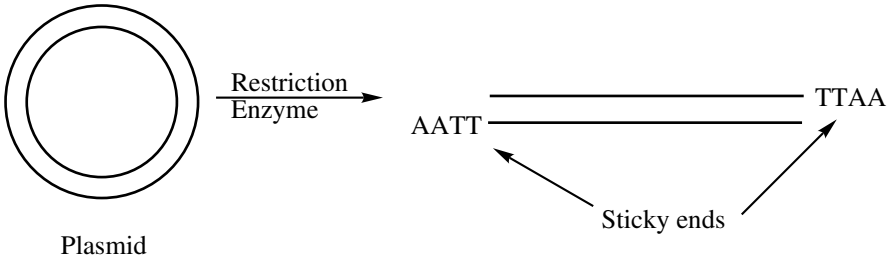


Figure 15-11: Opening of a plasmid by a restriction enzyme such as Eco RI.

It is possible to bond a DNA linker to a DNA molecule to make it susceptible to a particular restriction enzyme. By this method, the cohesive ends characteristic of any restriction enzyme may be added to almost any DNA molecule. The completed DNA can undergo replication.

Plasmids are, to a certain extent, accessory chromosomes. They can replicate independently of the host chromosomes. Thus, there may be multiple copies of a particular plasmid within a cell. This replication, in general, makes plasmids more useful as vectors than host chromosomes. Thus far, these plasmids have only been shown to be relevant in bacterial organisms.

The addition of "new" genes to an organism produces an organism that may be considered a new species. There is somewhat of a risk that these organisms could infect humans and lead to a new disease for which there is no known treatment. To minimize potential risks posed by these organisms, researchers either use enfeebled (weakened) organisms or ones that do not infect humans.

Patterns: Determining DNA Sequences

Restriction enzymes are a major tool in the determination of the base sequence in DNA. The cleaved DNA fragments are significantly smaller than the parent DNA is, making manipulation and analysis significantly easier. To separate the fragments after cleavage, gel electrophoresis is often used.

Gel electrophoresis is a biochemical technique used to separate and purify proteins and nucleic acids that differ in charge, size, or confirmation. The sample is placed into wells within a gel — a polymer that is specifically formulated for the type of analysis. This gel is in the shape of a thin slab. When separating proteins or small nucleic acids (DNA, RNA, and so on) cross-linked polyacrylamide is used. For separating larger nucleic acids, agarose, an extract from seaweed, is used. These gels have the consistency of Jell-O, but probably don't taste nearly as good.

The gel is immersed in a buffer solution, and an electrical current is applied to the ends of the gel. The charged species within the sample migrate toward one or the other of the electrodes. Proteins may have either a positive or a negative charge, but, at the proper pH, nucleic acids have only a negative charge. The positively charged species move toward the negatively charged end of the gel, and the negatively charged species move toward the positively charged end. Normally a buffer adjusts the pH so that all the species of interest have either a positive or negative charge (Figure 15-12).

Different molecules move at different speeds through the gel. When the smaller, faster molecules have about reached the end of it, the process is stopped, and the molecules are stained to make them visible. Sometimes, agents are added to cause the molecules to fluoresce (glow) under UV light. Then a photograph of the gel may be taken as it is exposed to the UV light. When several samples, including a known sample, are run side by side, the molecular weight of a sample component may be determined. This is one step in the identification of unknown components.

The separation of DNA fragments by gel electrophoresis readily distinguishes even minor differences between the fragments. Different gels are useful in separating large fragments than are useful in the separation of small fragments. In some gels, it is possible to distinguish between fragments differing by one base in several hundred. Modification of the electrophoresis method provides further separation. Each type of DNA gives a different pattern, making it possible to distinguish between two different samples. Two samples giving identical patterns must be from the same source or from identical twins.

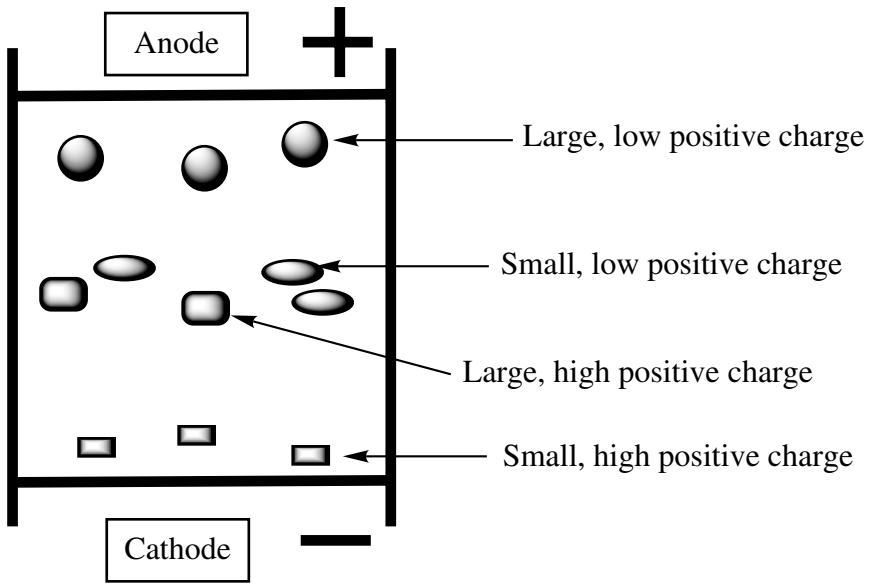


Figure 15-12: Gel electrophoresis.

In the analysis and manipulation of genetic material, it is advantageous to be able to identify whether a certain sequence of nucleotides is present. The general method for finding a particular sequence of nucleotides in DNA was developed by Edwin Southern. It is called *Southern blotting.* This method uses radioactive ^{32}P as a label that is easily detectible. This radioisotope is incorporated into the phosphate in some of the nucleotides. Determination of a particular nucleotide sequence in RNA is achieved through *Northern blotting,* and protein identification through *Western blotting.* (The names Northern and Western do not refer to persons with that name, but are by analogy to Southern. We don't know what happened to Eastern.) Alternatively, Southern, Northern, and Western blotting are DNA, RNA, and protein blots, respectively.

Determining the base sequence

Since the first isolation of DNA, a number of methods have been developed to determine the base sequence. In general, the *Sanger dideoxy method* has replaced all others. It employs the controlled termination of replication with modified nucleotides containing dideoxyribose in place of deoxyribose.

DNA fragments produced by employing restriction enzymes are denatured to give single-stranded DNA. (*Denaturing* typically involves heating a DNA-containing solution to 96°C for a few seconds.) Four samples of this DNA are treated separately to produce double-stranded DNA through replication, with each sample containing a small quantity of a different dideoxy nucleotide. The dideoxy nucleotide contains dideoxyribose (Figure 15-13). The absence of an additional oxygen atom in dideoxyribose means that there is no 3' hydroxyl group available to continue replication. Thus, the incorporation of a dideoxynucleotide terminates the DNA chain.

One of the four samples will contain a small quantity of the dideoxy analog of the nucleotide dGTP. This "defective" unit enters the new DNA strand as the complement to a cytosine base in the original fragment. Separation of the new material from the original strand material gives a set of DNA fragments of varying length. These fragments are then separated by electrophoresis according to length (size). The length of each of these fragments locates the position of each C in the original strand. The other three samples give the positions of all A, T, and G bases in the original strand.

Fluorescence tagging is a useful modification to this method. Each of the dideoxy nucleotides has a different fluorescent tag attached. After attaching the tags, it is possible to conduct all four experiments in one container. Separating the fragments by electrophoresis and examining the tags gives a colored pattern showing all the bases in sequence. This method works for fragments of up to 500 bases.

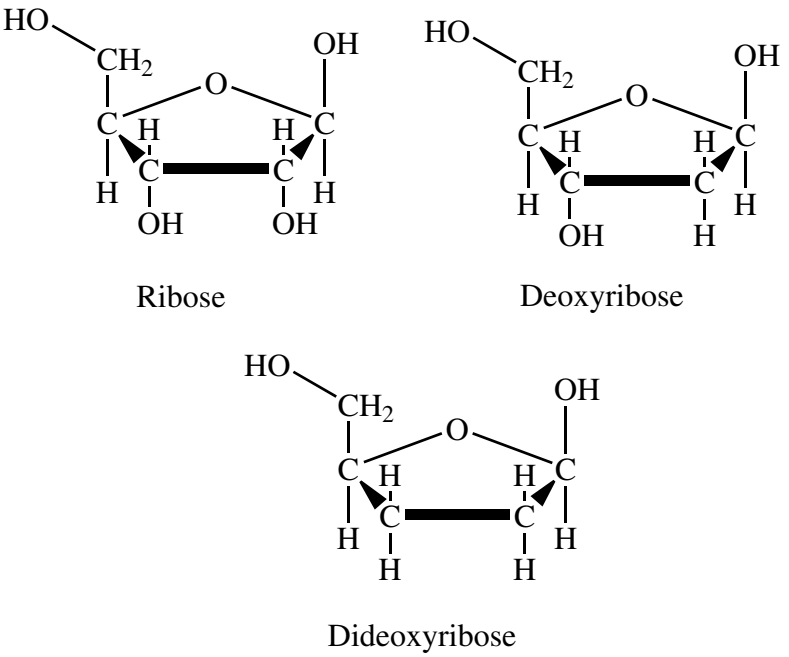


Figure 15-13: Structures of ribose, deoxyribose, and dideoxy-ribose.

To do the studies described, you need a sufficient amount of genetic material. Lack of sufficient quantities of sample has been a problem, especially with forensic evidence. Therefore, ways of quickly duplicating sufficient quantities of identical DNA fragments or producing a number of DNA strands from a very small sample were developed. *Polymerase chain reaction* (PCR), is a useful method to amplify specific DNA sequences. It is an in vitro procedure where it is necessary to know the base sequences, the flanking sequences, adjacent to a particular target sequence. However, it is not necessary to know the base sequence in the target region. Denaturation of a DNA sample provides two separate strands. Two primers are added to the mixture and one primer will attach to the flanking sequence of each strand. DNA polymerase begins replication starting at each of these primers. Repeating these steps quickly generates a large quantity of DNA. After 30 or so cycles, a billion-fold amplification occurs. Thirty cycles take less than one hour.

The butler did it: Forensic applications

Scientists can identify a species by the isolation and examination of the DNA sequences unique to that species. For example, DNA analysis is useful in the identification of organisms, such as bacteria, that may be polluting our water, food, and other samples. It has been used to establish pedigrees for livestock

breeds as well as identify endangered species in the prosecution of poachers. However, the application that has received the most publicity is in the area of *forensics*.

Because an individual's DNA comes from both the mother and father, it is unique to that individual (except in the case of identical twins). Even brothers and sisters, including fraternal twins, with the same parents show some variation in their DNA. This fact makes DNA analysis very valuable in forensics investigations (as anyone who ever watched an episode of *CSI* can attest).

In order to identify an individual, forensic investigators examine 13 regions (markers) of the DNA sample that vary significantly from individual to individual. There is a very small chance that two individuals might have the same DNA pattern at these 13 regions, but it is only about one chance in a billion. The investigation of additional markers can improve the discriminating ability of the procedure. Investigators then combine the results into a DNA profile — also known as a DNA fingerprint — of the individual.

You can isolate DNA samples from blood, hair, bone, fingernails, teeth, and any type of bodily fluid. In a typical crime scene analysis, samples are taken from the evidence and suspects; the DNA is extracted and then analyzed for the specific markers. A match of a single marker does not prove that an individual was at the crime scene, but the matching of four or five markers indicates a very high probability that the individual was present. PCR may be necessary if the sample is very small (see preceding section).

Methods of analysis

Several techniques are used in DNA analysis. The three most common are RFLP, PCR, and STR. In RFLP (*Restriction fragment length polymorphism*), the DNA sample is digested with a specific enzyme, a restriction endonuclease. This enzyme cuts DNA at a specific sequence pattern. The presence or absence of these sites in a DNA sample leads to variable lengths of DNA fragments. Gel electrophoresis then separates these fragments.

RFLP was one of the original forensic DNA analysis techniques. However, it requires relatively large amounts of DNA and samples contaminated with dirt and mold are difficult to analyze with RFLP. It has been somewhat replaced with polymerase chain reaction (PCR) enhancement, followed by STR analysis.

PCR (discussed in the earlier section "Determining the base sequence") is a useful technique that reduces the sample size requirement of RFLP — in essence it is a DNA amplifier. PCR quickly makes millions of exact copies of the DNA sample. Using PCR, DNA analysis can be done on a sample as small as a few cells and on samples that are extensively degraded. After PCR treatment, it is possible to analyze the sample with RFLP or STR.

In STR (*Short tandem repeat*) analysis, the DNA sample is quickly examined for 13 specific regions. The FBI uses this standard STR profile in its CODIS (Combined DNA Index System) program, which links national, state, and local databases of DNA profiles from felons, missing persons, and unsolved crime scenes. CODIS has an index of more than 3 million DNA profiles.

Paternity testing

Along with crime scene analysis, paternity testing is one of the most widely used applications of DNA testing. The procedure begins with the collection of DNA samples from the mother, child, and alleged father(s). The DNA profiles of the child and mother are first determined. The markers not inherited from the mother must have come from the biological father. The alleged father's DNA profile is then compared to the child. If the man's DNA profile contains markers common to the child but not the mother, then the probability that he is the biological father is great. Figure 15-14 indicates that Alleged Father 2 is more likely to be the biological father than Alleged Father 1.

Genetic Diseases and Other DNA Testing Applications

DNA testing always seems to find new ways of being useful. It has been used for a number of years, for example, in determining the gender of athletes. In addition to gender testing, the NFL used a strand of synthetic DNA to mark all the Super Bowl XXXIV footballs as a way to combat fraud associated with sports memorabilia. In a different situation, a section of DNA was added to the ink used to imprint all official goods marketed at the 2000 Summer Olympics Games. This same technology is used to tag original artwork, in addition to sports souvenirs.

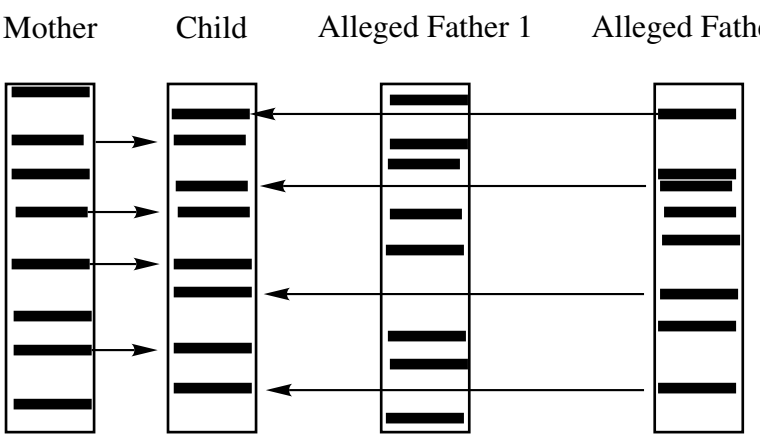


Figure 15-14: Comparison of results for paternity testing.

Genetic diseases are the result of an abnormal pattern in the DNA of an individual. These diseases are inherited, though some individuals are only carriers and not sufferers. Recently quite a bit of research has been done in determining the genetic pattern that is causing the disease, and ways to detect the probability of passing on the disease to offspring. However, methods of treatment for most all of these disease are limited. It is the dream of researchers to find the means of correcting these genetic diseases through genetic modifications. Researchers have investigated several of these genetic diseases in detail. In this section we briefly examine a few of the more well-known genetic diseases.

Sickle cell anemia

Sickle cell anemia is an inherited genetic disease of the blood's *hemoglobin,* a component of red blood cells. Sickle cell anemia is the result of the change of a single amino acid in the protein sequence of hemoglobin. This change involves the substitution of valine (non-polar) for glutamic acid (polar). The condition affects millions of individuals throughout the world, especially those whose ancestors came from Africa, South America, Cuba, Saudi Arabia, and a few other countries. In the United States, it affects about 72,000 people. Sickle cell happens in about 1 in 500 African American births and about 1 in 1,200 Hispanic American births.

Hemoglobin is responsible for carrying oxygen from the lungs to the cells. In an individual with sickle cell anemia, the defective hemoglobin molecules clump together, causing the red blood cells to assume a sickle shape, hence the name. These abnormal cells have trouble squeezing through small blood vessels, causing oxygen depletion in organs and extremities along with episodes of pain. These sickle cells also have a much shorter lifetime in the body, leaving the individual with chronic anemia. Many states now test newborns for sickle cell disease.

Hemochromatosis

Hemochromatosis, one of the most common genetic diseases in the United States, is an inherited disease that causes the body to absorb and store far too much iron. This excess iron is stored in organs, such as the liver, pancreas, and skin (yes, the skin is considered an organ!). It is due to a mutation in the HFE gene, the gene that regulates the absorption of iron from food. If this defective gene is inherited from both parents, then the person will develop hemochromatosis. If the individual inherits the mutated gene from only one parent, the person will be a carrier but will not necessarily develop the disease. About 5 Caucasian people in 1,000 carry both mutated genes, and 1 in 10 is a carrier. Genetic testing can detect it about 90 percent of the time.

Cystic fibrosis

Cystic fibrosis is a chronic and normally fatal genetic disease affecting the body's mucus glands. It targets the digestive and respiratory systems. About 55,000 individuals worldwide have cystic fibrosis. Most of these individuals are Caucasians who have ancestors who came from northern Europe. For the disease to appear it is necessary to inherit the mutated gene responsible for cystic fibrosis from both parents. Estimates are that 1 in 20 Americans carry the abnormal gene. Most of these individuals are not aware that they are carriers. Genetic testing is only about 80 percent accurate.

Hemophilia

Hemophilia is a genetic disorder caused by the lack of the blood-clotting factor stemming from a defective gene on the X chromosome. Females have two X chromosomes, so if there is a defective gene on one, there is little chance that the other one is also defective. However, she will be a carrier. Males, however, only have one X chromosome, so if it is defective, then the individual will develop hemophilia. If a woman is a carrier, she will have a 50 percent chance that her sons will have hemophilia and a 50 percent chance that her daughters will be carriers. Daughters of a hemophilic male will be carriers. Genetic testing can detect the presence of the abnormal gene.

Tay-Sachs

Tay-Sachs is an inherited disease in which a fatty-acid derivative, a lipid called *ganglioside,* accumulates in the brain — the result of a mutation of a specific gene. Although found primarily in the Jewish population, some French Canadians and Louisiana Cajuns also carry the abnormal gene. The symptoms most commonly appear in infants. Death normally occurs before the age of five. Although Tay-Sachs is a very rare disease, it was one of the first genetic diseases for which extensive and inexpensive genetic screening was developed. Screening tests were developed in the 1970s, and Israel offered free genetic screening and counseling. Because of this aggressive testing and counseling, the disease has been almost totally eradicated from Jewish families worldwide.

Ethics of genetic modification and testing

The emerging field of *bioengineering* has raised many ethical questions. One has only to listen to the debates over stem-cell research, gender selection of children, genetic modification to enhance certain traits such as athletic ability, and so on. Public policy decisions related to cost are also being debated, as genetic modification and screening are generally expensive. Is it to be available to only those who can pay, or should there be equal access? Many gray areas concerning genetic modification exist in the field of patent law. There are many questions and concerns but no quick answers.

Although the success in eradicating Tay-Sachs is directly related to genetic testing, such testing is not without its ethical questions. The major concern is one of privacy. DNA samples and profiles can be used to determine parentage and susceptibility to certain genetic diseases. Many people fear that the government, insurance companies, employers, banks, schools, and other organizations could use such information for genetic discrimination. In fact, in the United Kingdom, a man was denied treatment for hemochromatosis because his insurance company claimed it was a preexisting condition. Individuals applying for life insurance have reported other cases of genetic discrimination. Who gets to request the genetic screening and who has access to the results? These are just a few of the questions we will be debating for many years to come.

Chapter 16

Transcribe This! RNA Transcription

In This Chapter

- Finding out what's in your genes
- Breaking the genetic code
- Modeling gene regulation

Cells utilize a number of types of ribonucleic acid — RNA.

Messenger RNA (mRNA), a form that is not very stable, carries information from the cell nucleus (DNA) into the cell and must migrate to the ribosomes. Messenger RNA carries the actual genetic information necessary for the synthesis of a specific protein; however, the other forms of RNA are necessary to complete the process.

Transfer RNA (tRNA) transfers amino acids to the ribosomes for protein synthesis. This is a relatively small form of ribonucleic acid, typically containing from 73 to 93 nucleotides.

The relatively large ribosomal RNA (rRNA) resides in the ribosomes and has a direct influence on the synthesis of proteins. This form of RNA has protein components. There are three types of rRNA (called 23S, 16S, and 5S), and all three must be present in each ribosome.

Finally, small nuclear RNA (snRNA) serves a number of ancillary functions. In this chapter we concentrate upon the synthesis of RNA, which is called transcription.

Protein synthesis begins with *transcription,* the process whereby DNA produces mRNA. First, a portion (a gene) of a DNA double helix opens. Nucleotides can then bind to the exposed DNA nucleotides through a process similar to replication. However, this process differs from replication in that only a portion of the

DNA opens, and the entering nucleotides contain uracil in place of thymine. One gene yields one mRNA, which, in turn, may lead to the synthesis of one or more proteins.

The enzyme RNA polymerase joins the nucleotides to produce RNA in a process that occurs within the cell nucleus. The process begins as an initiation signal toward the 5' end of RNA and goes toward a termination sequence nearer the 3' end.

RNA Polymerase Requirements

Three requirements are needed for RNA polymerase to operate. It requires activated precursors of each of the four ribonucleoside triphosphates (ATP, CTP, GTP, and UTP) from which to produce the new RNA. (See Figure 16-1.) A divalent metal ion, either magnesium or manganese, is necessary. Finally, a template must be present. Single-stranded DNA will work; however, the preferred template is double-stranded DNA. However, the DNA strands must open (separate) in order to allow the RNA polymerase access.

There are many similarities between replication and transcription. In both processes, the direction of synthesis is 5' → 3'. Elongation occurs as the 3'-OH group of the chain attacks the innermost phosphate of the entering nucleoside triphosphate. This is called a *nucleophilic attack.* The hydrolysis of pyrophosphate provides the impetus to drive the process forward. However, there are differences. Unlike its DNA counterpart, RNA polymerase is not capable of "reviewing" its work and then eliminating a mismatched nucleotide. RNA polymerase does not require a primer.

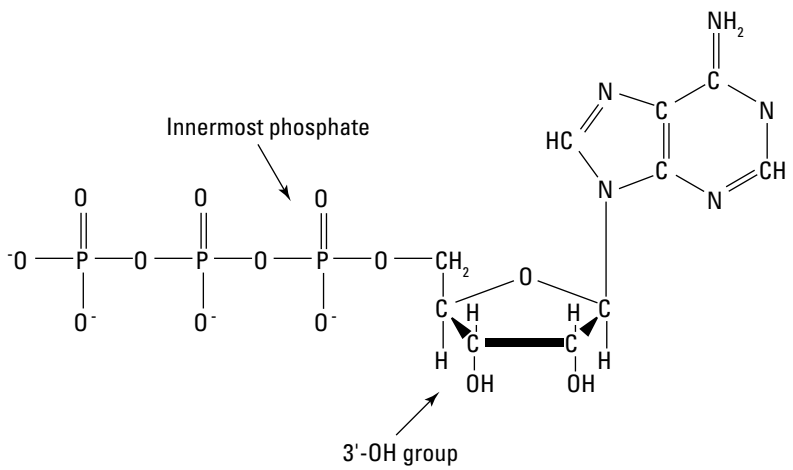


Figure 16-1: Structure of ATP.

In simple organisms, such as *E. coli*, one type of RNA polymerase synthesizes all forms of RNA. In more advanced organisms, like human beings, there are different types of RNA polymerase. Usually at least three different types are present in mammalian cells.

Making RNA: The Basics

The region of a DNA molecule that codes for a protein is a *structural gene.* Other regions are present to regulate the activity of this gene. (We examine these regulatory regions later in this chapter.) To begin transcription, it is necessary for RNA polymerase to detect one particular gene present in a long DNA strand. Detection begins with the enzyme locating a region on the DNA strand known as a *promoter site,* which is "upstream" from the actual gene. (Upstream means on the 5' side.) RNA polymerase tightly binds to the promoter site, and once in place, transcription can begin.

In prokaryotic cells the promoter sites are centered at –10 (the Pribnow box) and in the –35 region. The Pribnow box has the consensus sequence TATAAT centered at –10. The other site has the consensus sequence TTGACA. (Not all organisms have the same consensus sequence.) In eukaryotic cells a promoter is centered at about –25 (the TATA box or Hogness box), and sometimes centered near –75 (the CAAT box). The consensus sequence in the Hogness box is TATAAA. The CAAT box has the sequence GGXCAATCT. In addition, eukaryotic genes may have enhancer sequences up to several thousand bases away from the start site and on either side (Figure 16-2).

The position of sequences along the DNA chain begins at the beginning of a gene. This position is 0. The first nucleotide of the gene is +1. Counting upstream (towards the 5' terminus) is negative. Thus, ten nucleotides before the beginning of the gene would be –10.

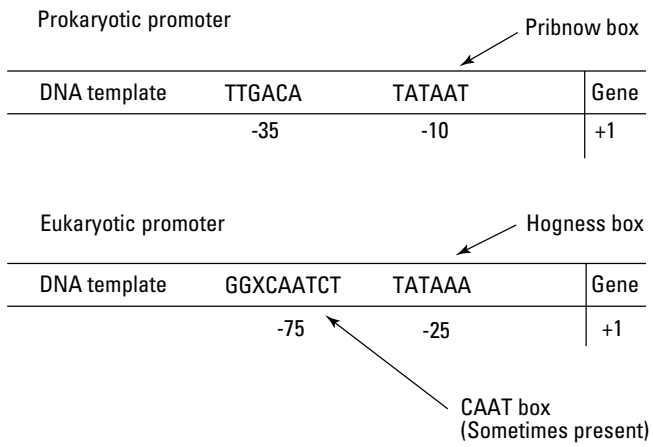


Figure 16-2: Prokaryotic and eukaryotic promoter sites.

Transcription proceeds as an RNA polymerase moves along the DNA strand. Eventually, the enzyme encounters a termination signal. In prokaryotic cells, there are two termination signals. The first is a *base-paired hairpin,* which consists of a self-complementary sequence rich in C and G followed by a sequence of several instances of U. After the sequence forms, the new RNA detaches from the template. The other method uses a *rho protein.*

The termination in eukaryotic cells is not very well understood. In eukaryotic cell, mRNA undergoes further modification after transcription. A "cap" is attached to the 5' end of the RNA, and a poly(A) tail goes onto the other end. These modifications increase the lifetime of mRNA.

The stages in RNA synthesis are *initiation*, *elongation*, and *termination*. To accomplish these tasks, RNA polymerase must perform a series of functions. The enzyme must travel along a DNA strand until it encounters a promoter site. As it "sticks" to the promoter site it unwinds a short segment of the DNA double helix and separates the strands to reach the template. Then the appropriate ribonucleoside triphosphate enters, and hydrolysis of the phosphate occurs in order to supply the needed energy. Each ribonucleoside triphosphate is brought in as the RNA polymerase moves along the DNA strand. (The DNA unwinds as the enzyme passes, and rewinds after the enzymes has passed.) This continues until the RNA polymerase finds a termination signal. The enzyme also must interact with transcription factors or trans-acting factors — proteins that act as activators or repressors — to regulate the rate of transcription initiation.

The best understood operation of RNA polymerase comes from studies of the prokaryotic cells of *E. coli.* Eukaryotic cells behave in a similar, though more complicated, manner. One major difference between the two is that in prokaryotic cells, transcription and translation (protein synthesis) may occur almost simultaneously, whereas in eukaryotic cells there is a gap between the two processes while the mRNA moves from the nucleus to the ribosome. The other major difference is that RNA in eukaryotic cells almost always requires processing after synthesis. Prokaryotic RNA is usually ready immediately after synthesis. Processing includes adding a cap, adding a poly(A) tail, and — in nearly all cases — splicing to remove introns.

Prokaryotic cells

RNA polymerase in *E. coli* contains four subunits that combine to form a holoenzyme designated $\alpha_2\beta\beta'\sigma$. The purpose of the σ subunit is to help find the promoter and to help initiate RNA synthesis. Once synthesis begins, this unit leaves the remainder, the core enzyme. The catalytic site in the core enzyme contains two divalent metal ions, one of which stays with the core and one that enters with the ribonucleoside triphosphate and leaves with

the cleaved pyrophosphate. Three aspartate residues aid in the binding of the metal ions. Although DNA polymerase and RNA polymerase have very different overall structures, their active sites are similar.

In the absence of the σ subunit, RNA polymerase would bind tightly to DNA at any point. When this unit is present, binding at other than a promoter site is significantly lower. Due to its reduced affinity, the holoenzyme can slide along the DNA strand until a σ subunit detects a promoter site. It binds to this site more strongly than to other positions on the DNA strand. The efficiency of this binding is one form of regulation. A number of σ subunits are present, each designed to recognize a different promoter site.

It's almost like tying knots in an anchor rope. A diver could swim upward holding on to the rope, but a knot signals a spot to stop and decompress.

Once the RNA polymerase arrives at a promoter site, it becomes necessary to unwind a 17 base-pair segment of the double helix and to unpair the bases. This unwinding converts a closed promoter complex to an open promoter complex. RNA polymerase is now ready to begin the RNA chain by incorporating the first nucleotide triphosphate. (Unlike DNA replication, no primer is necessary.) This first nucleotide triphosphate is usually a pppG or a pppA, which remains throughout transcription. This tap is at the 5' end of the new RNA molecule, and growth begins when a new nucleotide links to the 3' position (Figure 16-3).

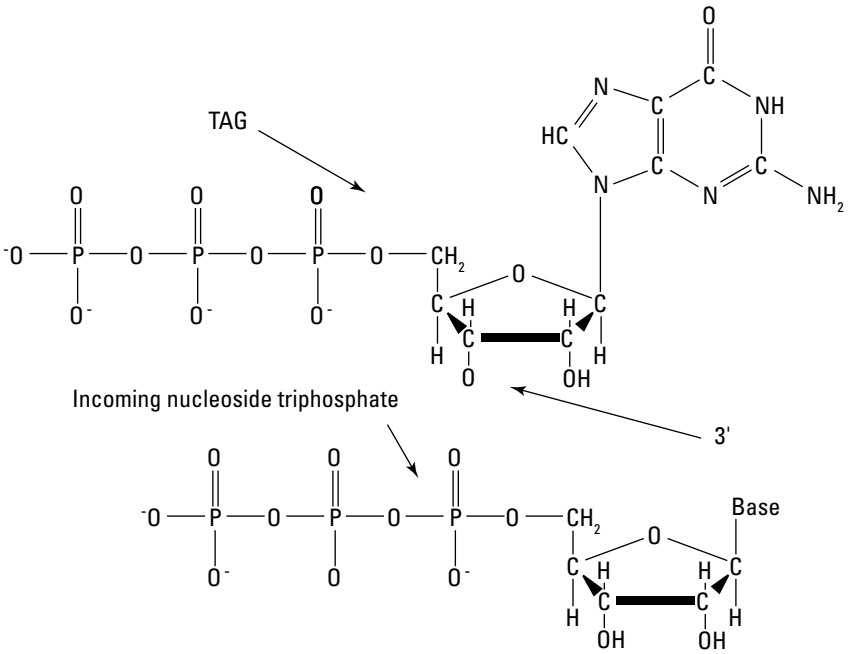


Figure 16-3: Linking of the second nucleotide to the tag, using pppG as an example (top), and linked nucleotides at the beginning of the chain (bottom).

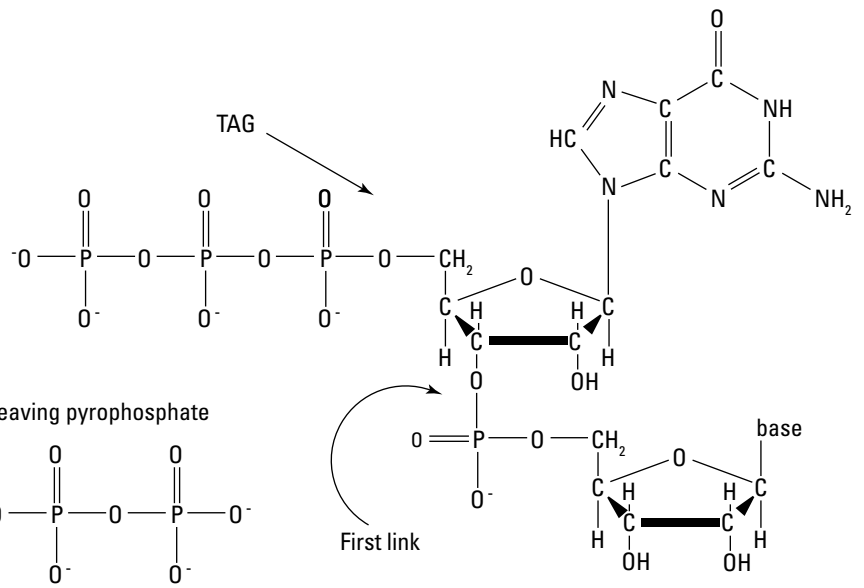


Figure 16-3: continued

Once the first two nucleotides link (through the formation of the linking phosphate diester) the σ subunit leaves. This allows the core enzyme to bind more tightly to the substrate. A transcription bubble now forms that contains the RNA polymerase, the unwound portion of the DNA, and the rapidly forming nascent RNA. Initially a short segment of the new RNA forms a hybrid helix with the DNA. This segment normally consists of about eight base pairs or one turn of the double helix. The growth rate is on the order of 50 nucleotides per second. (Compare this to DNA replication, which proceeds at about 800 nucleotides per second.) It is important to note that RNA polymerase does not "proofread" the new RNA. Thus, errors creep in at a higher rate than in replication. However, because the products do not pass to the next generation, there is no mutation or lasting effect. In any case, the next RNA stand to form lacks this defect and behaves correctly. One bad RNA in several hundred or more copies of the same gene is likely to have a minimal influence on the cell.

Elongation proceeds until the RNA polymerase encounters a termination signal, initiating a series of actions. At this point, formation of new phosphate diesters ceases, the RNA-DNA hybrid separates, the portion of the DNA chain that is still open rewinds, and the RNA polymerase separates from the DNA. There are different termination signals. One simple one is a palindromic (reading the same forward or backwards) GC-rich region followed by an AT rich region. The palindromic region is self-complementry, and these bases hydrogen bond to form a hairpin loop. The AT-rich region results in a number of U_{RNA}-A_{DNA} pairs, which have the weakest hydrogen bond interactions of all types of pairs. The formation of this hairpin and the AT region destabilizes, and the RNA-DNA hybrid and the nascent RNA begin to leave. See Figure 16-4.

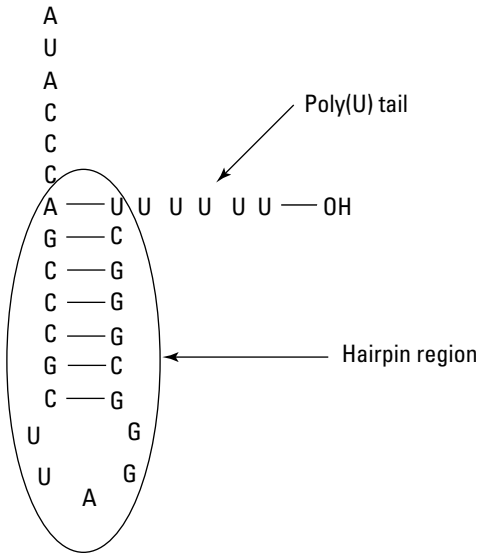


Figure 16-4: The hairpin and subsequent portion of the RNA.

Not all termination signals contain a hairpin and a U-rich segment. In at least some cases, RNA polymerase needs help. Evidence for this came from the observation that in vitro RNA chains were often longer than in vivo chains for the same RNA. Clearly, the in vitro RNA polymerase was unable to terminate elongation. The missing aid was a protein known as the *rho factor* (ρ). This protein wraps about the nascent RNA soon after the RNA exits the transcription bubble. In the presence of RNA, the ρ protein hydrolyzes ATP, which supplies energy. The protein first attaches to an RNA segment that is poor in guanine and rich in cytosine. Rho moves along the nascent RNA until it encounters the transcription bubble. At this point, it breaks the RNA-DNA hybrid and separates the nascent RNA. Other proteins serve a similar function as the rho factor.

In prokaryotic cells, mRNA is either ready or nearly ready to function immediately after release from the transcription (translation may begin before transcription terminates). However, both tRNA and rRNA require cleavage and other modifications of the nascent RNA chain. Various nucleases cleave the RNA in a very precise manner. It is possible to get more than one gene from a long nascent RNA strand. Processing may require the connection of a number of nucleotides — for example, all tRNA molecules need a CCA tail to function correctly. In some cases, there may be modification of the bases or ribose units.

Eukaryotic cells

Unlike prokaryotic cells, transcription and translation occur in different regions of the cell, leading to greater control of gene expression. Another difference is that eukaryotic cells extensively process mRNA in addition to rRNA

and tRNA. After RNA polymerase action, mRNA acquires a cap and a poly(A) tail. Nearly all mRNA molecules are spliced. Splicing involves removal of introns with the remaining exons being connected. Ninety percent of the nascent RNA may be introns.

Eukaryotic cells typically contain three types of RNA polymerase. Type I RNA polymerase (in the nucleolus) produces most forms of rRNA. Type II (in the nucleoplasm) produces mRNA and snRNA. And type III (in the nucleoplasm) produces tRNA and small rRNA molecules. (Actually, these polymerases only produce the pre-RNA forms of these molecules.)

Each of the three polymerases has a distinct type of promoter. These promoters may be in the same upstream sites as in prokaryotic cells, in downstream sites, or within the genes themselves. In addition to promoters, there may be enhancers. Enhancers, though not promoters, increase the effectiveness of a promoter. Enhancers for a single promoter may occur in different positions on the DNA chain and are important for gene regulation. Both promoters and most enhancers are on the same side of the DNA chain as the gene they regulate; for this reason, they are cis-acting elements. The promoters, as discussed earlier, are typically a TATA box (usually between –30 and –100), the CAAT box, and the GC box (both are usually between –40 and –150). Enhancers may appear upstream, downstream, or within the gene about to undergo transcription. Enhancers that are present on the opposite DNA chain are trans-acting factors, known as transcription factors, on the other DNA chain.

The typical series of events is that the transcription factor TFIID binds to the TATA box (TF stands for *transcription factor,* and the II means *RNA polymerase II*). Binding is the result of a small component of TFIID known as TBP (TATA-box-binding protein), which has an extremely high affinity for the TATA-box. When TBP binds to the DNA, substantial changes occur in DNA, including some degree of unwinding.

Other components utilized in transcription later attach to the TBP. These are, in order: TFIIA, TFIIB, TFIIF, RNA polymerase II, and finally TFIIE. This final group is the basal transcription complex. This example illustrates only one of numerous transcription factor initiations.

In eukaryotic cells, nearly all, if not all, products of transcription (precursors) undergo further processing before they reach their final active form. In general, tRNA precursors need to have the 5' leader removed, splicing to remove any and all introns, replacement of the poly(U) tail with a CCA sequence, and possible modification of some of the bases. Each of these processes requires one or more enzymes.

The precursors to the various forms of mRNA normally require the most modification. These precursors need, amongst other things, a 5' cap and a 3' poly(A) tail. The caps are cap 0, cap 1, and cap 2 — the numbers refer to the number of methylated ribose sugars (Figure 16-5). Caps are not present on tRNA, snRNA, or rRNA.

Most mRNA has a poly(A) tail not encoded by DNA. Usually, addition of this tail is preceded by cleavage of an intron portion of the mRNA precursor. The series AAUAAA signals where the cleavage will occur. This series is only part of the signal — the other part is uncertain. After cleavage, a poly(A) polymerase adds about 250 adenylate residues to the 3' end. The exact purpose of the tail is uncertain. It appears to enhance translation and increase the lifetime of the mRNA molecule.

In some cases, it is necessary to edit some mRNA precursors. *Editing* refers to an alteration of the base sequence other than that caused by splicing. An example is to chemically change one base into another. An example of editing occurs in the mRNA that encodes for apolipoprotein B (apo B). The entire protein contains 4,536 residues. However, a related 2,152-residue form is also important. The longer form, synthesized in the liver, is useful in the transport of lipids within the liver. The smaller form, synthesized in the small intestine, interacts with dietary fats. The same mRNA is responsible for both forms of the protein. In the small intestine, a deaminase acts on a specific cytosine and converts it to a uracil, which changes a CAA codon (Gln) to a UAA codon (stop) — which truncates the protein chain to yield the smaller form.

Splicing is a very common form of modification of all forms of RNA. *Splicing* involves the removal of introns and the joining of the exons to yield the final RNA molecule. Splicing must be very precise, as a miss by one base alters the entire sequence of codons present.

A number of different introns need to be removed. In eukaryotic cells, the intron begins with a GU and ends with an AG. Further refinement is present in vertebrates, where GU is the end of the sequence AGGUAAGU. A variety of AG sequences are found in higher eukaryotic cells. In general, one end of the intron loops about and connects to a point (the branch point) on the intron chain. Joining of the exons present then proceeds.

Spliceosomes are important in the splicing of mRNA precursors. These assemblages contain the mRNA precursors, several snRNAs, and proteins known as splicing factors. A group of snRNAs labeled U1, U2, U4, U5, and U6 are important. U1 binds to the 5' end of the splice site and then to the 3' end. U2 binds to the branch point, U4 blocks U6 until the appropriate moment, U5 binds to the 5' splice site, and U6 catalyzes the splicing. There are alternate splicing procedures.

It should be noted that alternate splicing can lead to production of different proteins from the same RNA.

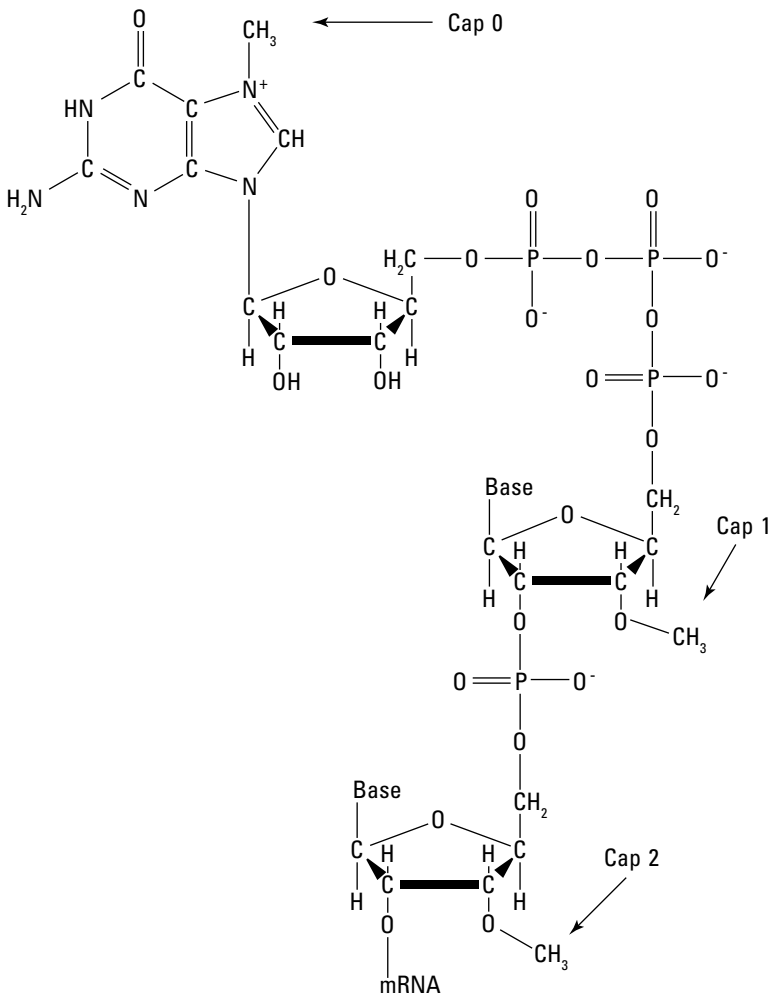


Figure 16-5: The general structure of a mRNA cap.

To Heck with Da Vinci: The Genetic Code

Just as DNA serves as the template for the generation of RNA, mRNA serves as the template for the generation of protein. In order to synthesize the appropriate protein, there needs to be a species that interacts with this template to assure the incorporation of the correct amino acid. The interaction species is tRNA. This relatively small form of RNA has two important regions: a template recognition site and the appropriate amino acid. The template recognition site is an anticodon, which corresponds to a codon on the mRNA. Attachment of the amino acid to the tRNA is by the action of an aminoacyl-tRNA synthetase. Each of the 20 amino acids has at least one

specific synthetase. This enzyme attaches the specific amino acid to the 3' terminal adenosine of the tRNA (Figure 16-6).

Codons

The genetic code contains the information necessary for the synthesis of proteins and consists of a set of three-letter words made from an alphabet containing four letters. Each three-letter word is a *codon.* This vocabulary is universal as it applies to all known living organisms.

The four letters are as follows:

- A, for adenine
- C, for cytosine

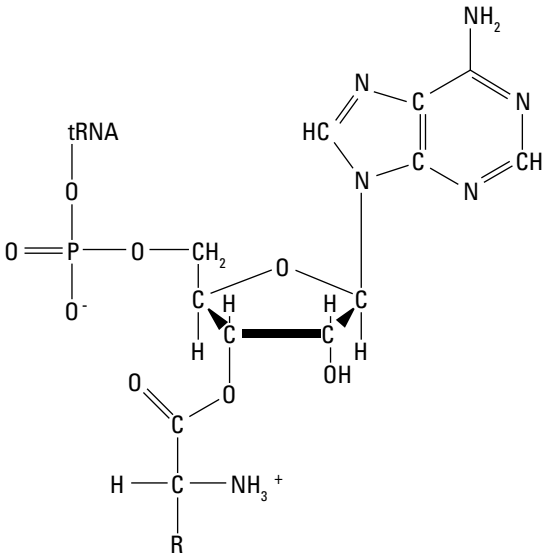


Figure 16-6: The attachment of an amino acid to the terminal adenosine.

- G, for guanine
- U, for uracil

The four letters give a total dictionary containing 64 words. Sixty-one of these words code for specific amino acids, and the remaining three words code for no amino acid. The codons coding for no amino acid are the "stop" signals.

Because there are only 20 amino acids to code for, the presence of 61 codons means that some amino acids can come from more than one codon.

Table 16-1 lists the genetic code.

Table 16-1 The Standard Genetic Code

Codon	*Amino Acid*	*Codon*	*Amino Acid*	*Codon*	*Amino Acid*	*Codon*	*Amino Acid*
AUA	Ile	ACA	Thr	AAA	Lys	AGA	Arg
AUC	Ile	ACC	Thr	AAC	Asn	AGC	Ser
AUG	Met	ACG	Thr	AAG	Lys	AGG	Arg
AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser
CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg
CUC	Leu	CCC	Pro	CAC	His	CGC	Arg
CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg
CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
GUA	Val	GCA	Ala	GAA	Glu	GGA	Gly
GUC	Val	GCC	Ala	GAC	Asp	GGC	Gly
GUG	Val	GCG	Ala	GAG	Glu	GGG	Gly
GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly
UUA	Leu	UCA	Ser	UAA	***Stop***	UGA	***Stop***
UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys
UUG	Leu	UCG	Ser	UAG	***Stop***	UGG	Trp
UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys

Analysis of the genetic code shows that two amino acids — methionine (see Figure 16-7) and tryptophan — only have one codon each. At the other extreme, three amino acids — arginine, leucine, and serine — each have six codons. The remaining 15 amino acids have at least two codons each. Amino acids with more codons are more abundant in proteins. Examining Table 16-1 shows that most *synonyms* (codons coding for the same amino acid) are grouped together and differ by a single base, usually the last base in the codon. Other correlations are present in the table. See what others you can find. The similarity of synonyms limits potential damage due to mutations.

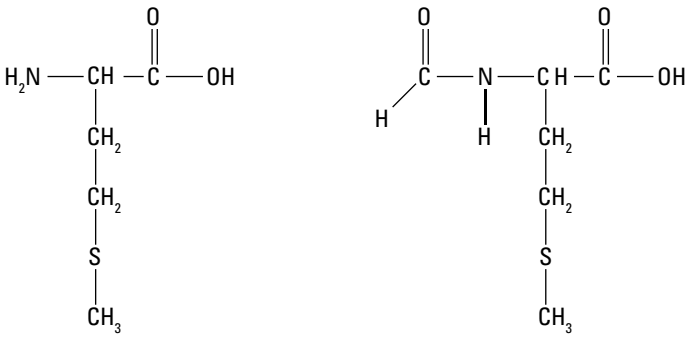


Figure 16-7: Structures of methionine and formylmethionine.

Alpha and omega

Although tRNA does not read the termination sequences, UAA, UAG, and UGA, specific proteins known as *release factors* read them. When a release factor binds to the ribosome, it triggers the release of the new protein, and release of the protein signals new synthesis to begin.

The stop signals are rather obvious on the Table 16-1, but what about the start? What signals the initiation of protein synthesis? The initiation sequence is usually AUG, the codon for methionine (Figure 16-8). In eukaryotic cells, additional factors come into play. In many bacteria, fMet (formylmethionine) is the initial amino acid (refer to Figure 16-7), which AUG usually codes for; however, GUG works sometimes.

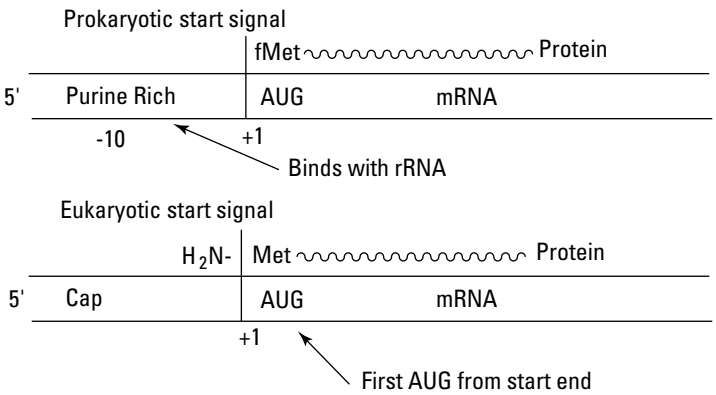


Figure 16-8: The start signals.

The genetic code is nearly universal — the codons correspond to the same amino acid in most cases. A few exceptions are known. For example, the code in mitochondrial DNA has several differences from normal DNA. In mitochondrial DNA, UGA is not a stop signal, but a codon for tryptophan.

In prokaryotic cells, coding for proteins is continuous, but this is not always true in the case of eukaryotic cells. In some mammals and birds, most genes are discontinuous. For example in the gene encoding for β-globin, there are regions that do not encode for a portion of protein. The gene contains about 1,660 base pairs — about 250 pairs on each end, plus an additional 500 pair segment code for the protein. These coding segments are *exons.* Two segments, one of about 120 base pairs and one of about 550 base pairs, do not code for protein. These non-coding regions are *introns.* The entire gene has, in sequence, a 240-pair exon, a 120-pair intron, a 500-pair exon, a 550-pair intron, and a 250-pair exon.

If an mRNA forms from a gene containing introns, it needs to undergo modification before it is of use. It is necessary to cut the intron regions from the mRNA and to splice the exon ends together to form the final mRNA molecule. In most cases, the intron portion begins with a GU and ends with a pyrimidine-rich segment ending with an AG. This combination signals the intron domain.

Models of Gene Regulation

The organism does not need to produce all the different proteins all the time. To control which proteins form at which time requires some form of *gene regulation.* When the organism requires a specific protein, it is necessary to "switch on" a certain gene — and once there is a sufficient quantity of that protein the gene must be "switched off." Control may occur either at the transcription level (gene regulation) or at the translation level.

In this section we examine processes in prokaryotic cells and then move on to the more complicated processes that take place in eukaryotic cells. The examination of the simpler mechanisms in prokaryotic cells gives insight into the processes in eukaryotic cells — the basic processes are similar.

As usual, our prokaryotic example is *E. coli.* Insight on gene regulation came when the diet of the *E. coli* was changed from glucose-rich to lactose-rich. For the cells to utilize this alternate energy source, they must generate the enzyme β-galactosidase. This enzyme is normally available at very low levels — a situation that quickly changes after replacing the glucose with lactose. One clue to the mechanism was that as the levels of β-galactosidase increased, so did the levels of galactoside permease (which transports lactose into the cell) and thiogalactoside transacetylase (which detoxifies other materials transported by galactoside permease). Thus, one change in the environment triggered multiple enzymes. This coordinated triggering of gene expression is

called an *operon*. Francois Jacob and Jacques Monad proposed the operon model to explain gene regulation.

The Jacob-Monod (operon) model

The simultaneous change in the levels of three different enzymes by one change in the environment suggested a link between the control mechanisms, and the operon model was created to account for this link. This model requires a regulator gene that affects a number of structural genes and an operator site. The operator and associated structural genes constitute the operon. The regulator gene is responsible for producing a *repressor protein*. The repressor protein binds to the operator site and prevents expression of the structural genes, as shown in Figure 16-9. The lac operon is one of the better understood operons.

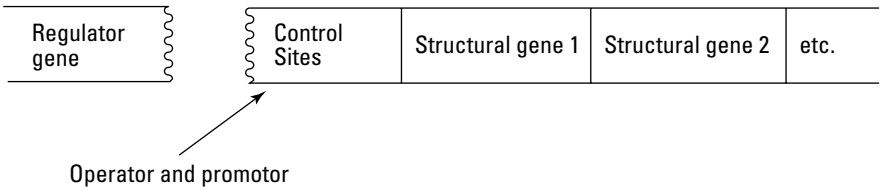


Figure 16-9: Diagram of a generic operon.

The multiple structural genes produce one large mRNA, and this single RNA strand is capable of generating a set of proteins. An mRNA that is capable of encoding for multiple proteins is *polygenic* or *polycistronic*.

The lac operon

The lac operon is the model regulatory system that, since its discovery in 1961, has provided extensive insight into how a cell regulates its genome. Figure 16-10 illustrates the lac operon.

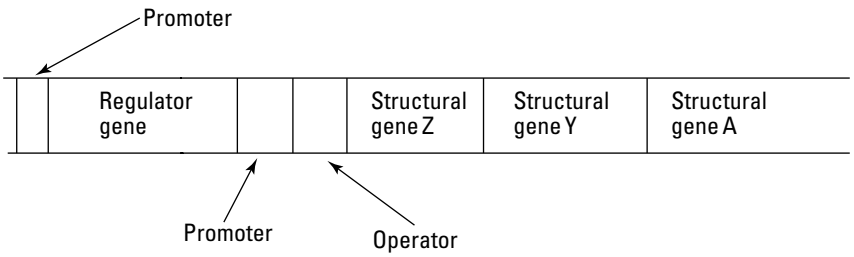


Figure 16-10: Diagram of the lac operon.

The *lac operator* is a palindromic DNA sequence with a twofold symmetry axis. The repeat is not always a perfect palindrome. (Many protein-DNA interactions involve a matching of symmetry.) The lac operator is as follows, with the center axis in bold:

TGTGTGGAATTGTGAGC**G**GATAACAATTTCACACA

ACACACCTTAACACTCG**C**CTAATGTTAAAGTGTGT

The *lac repressor* is a dimeric protein that can join to form a tetramer. In the absence of lactose, the repressor tightly binds to the operator. The presence of the repressor prevents RNA polymerase from unwinding the DNA strand to initiate transcription.

The presence of lactose is not the direct trigger of the lac operon; the trigger is allolactose. Both lactose and allolactose are disaccharides composed of galactose and glucose (Figure 16-11). In lactose there is an α-1,4 linkage, whereas in allolactose the linkage is an α-1,6. Allolactose results when the few molecules of β-galactosidase that are normally present in the cell first encounter lactose. This disaccharide along with a few similar molecules is an inducer of the lac operon. The inducer binds to the repressor and reduces the affinity of the latter to operator on the DNA. With its affinity reduced, the repressor detaches from the operator, and the DNA segment is now open for business.

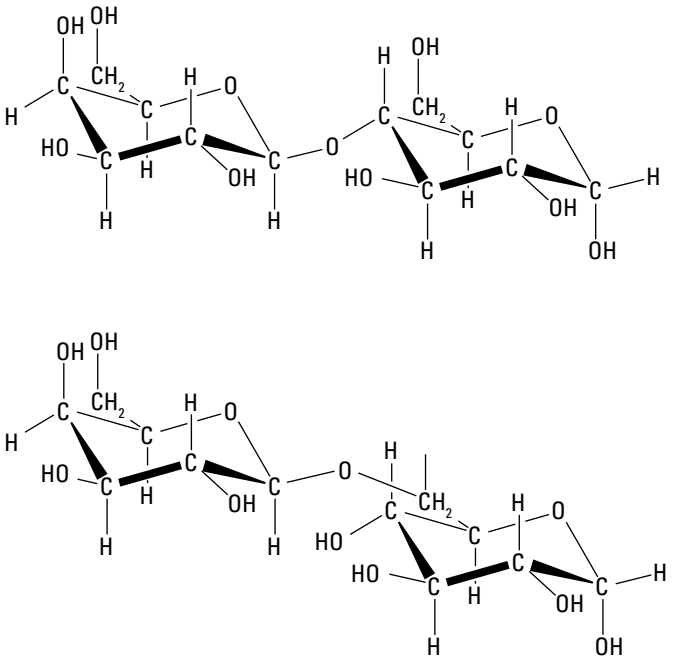


Figure 16-11: Structures of lactose and allolactose.

When transcription begins, all three structural genes become active, and the cell begins producing β-galactosidase, galactoside permease, and thiogalactoside transacetylase. This continues until the lactose and hence the allolactose concentration falls so that the repressor proteins are available to reattach to the DNA.

Other prokaryotic regulators

The *pur repressor* affects the genes responsible for the biosynthesis of purines and, to a lesser extent, pyrimidines. This protein is similar in structure to the lac repressor; however, the pur repressor only binds to the operator after another molecule binds to the repressor. Therefore, while the binding of another molecule releases the lac repressor, the binding of another molecule causes the pur repressor to bind. The other molecule has an opposite affect. In the case of the pur repressor, the other molecule is a *corepressor.*

There are also regulators that stimulate transcription instead of repressing it. The catabolite activator protein (CAP) is one example. This protein interacts with the promoter and, along with two cAMP molecules, interacts with RNA polymerase. This interaction leads to stimulating the initiation of transcription of certain genes.

Regulation of eucaryotic genes

Although there are similarities, the regulation of genes in eukaryotic cells is more complex than in prokaryotic cells. One reason for this is that the typical eukaryotic genome is much larger than the typical prokaryotic genome. Another source of complexity is that many eukaryotic cells are part of a larger organism and do not serve the same purpose as other cells do within the same organism. For example, although some of the proteins are the same, a liver cell must produce a different overall set of proteins than a heart cell does.

Histones

Eukaryotic DNA has a group of proteins associated with it. These small, basic proteins are called *histones.* They are basic because approximately 25 percent of the amino acid residues present are either arginine or lysine. These are tightly bound to the DNA and total approximately half of the mass of a chromosome. A complex of the cell's DNA and associated protein is a *chromatin,* and there are five important histones present in chromatin: H1 — and four that associate with each other: H2A, H2B, H3, and H4.

A chromatin apparently consists of repeat units consisting of two copies each of H2A, H2B, H3, and H4, with a strand of DNA consisting of about 200 base pairs tightly wrapped around this histone octomer. Each of these repeating units is a nucleosome. The wrapping of the DNA to form a nucleosome yields a significant compaction of the DNA. Research indicates that about 145 of

the 200 base pairs are actually associated with the histone octomer, and the remaining base pairs are linker DNA that link one histone octomer to the next. Histone H1 usually binds to linker DNA.

The eight histones in a histone octomer are arranged into a tetramer with the composition $(H3)_2(H4)_2$ and two dimers each with the composition (H2A)(H2B). All the histone proteins have long tails rich in arginine and lysine residues that extend out of the core. Modification of these tails is important for gene regulation.

The structure of chromatin is a factor in eukaryotic gene regulation. For a gene to be available for transcription, the tightly packed chromatin structure must open. In addition, the structure regulates access to regulatory sites on DNA. Enhancers disturb this structure, explaining why enhancers can have an effect on the expression of a gene even though the enhancer site may be thousands of base pairs away from the gene. Certain enhancers only occur in specific types of cells. Thus, the genes they enhance are only expressed in these cells. For example, the gene to produce insulin is expressed only in pancreatic cells.

A modification of DNA can also inhibit gene expression. Approximately 70 percent of the 5'-CpG-3' sequences in mammals have the cytosine methylated. The distribution of the methylated cytosines (Figure 16-12) varies with cell type. Regions in chromatins necessary for gene expression in that cell are hypomethylated (have fewer methylated cytosines), relative to similar regions in cells where no expression of the gene occurs. The presence of the methyl group interferes with the binding of enhancers and promoters.

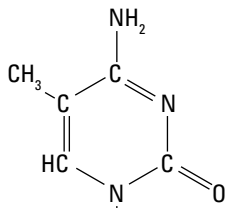


Figure 16-12: Structure of methylated cytosine.

Mediating transcription

Eukaryotic cells require a variety of transcription factors to initiate transcription — no factor can carry out the entire process on its own. This is called *combinatorial control,* and it is necessary in organisms with multiple cell types and helpful in other eukaryotic cells.

A number of nonpolar molecules, such as the steroid hormones, can easily pass through the hydrophobic cell membrane and bind to receptor proteins. They are very specific. Estrogen (Figure 16-13) is one example of a steroid hormone. Such molecules are known as *ligands.*

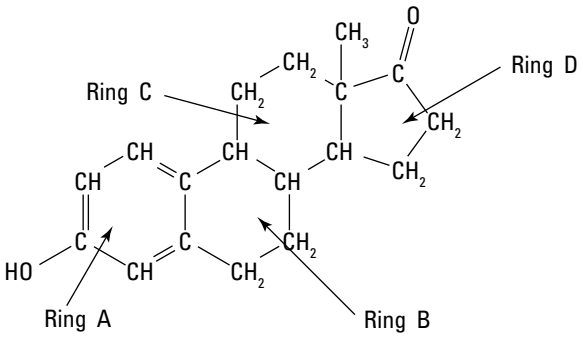


Figure 16-13: Structure of estrogen.

The ligand binds to a specific site — called, helpfully, the *ligand-binding site* — which is present near the end of a receptor protein. This portion of the protein contains many nonpolar residues, which have an affinity for hydrophobic molecules. Receptor proteins that bind hormones are called *nuclear hormone receptors.* There is a DNA binding site near the center of the protein that contains eight cysteine residues, which are necessary to bind zinc ions, four residues for each. The presence of the zinc ions stabilize structure and led to the name *zinc finger domains.* (There are other cysteine residues and zinc ions nearby.) The binding of a molecule to the ligand-binding site causes a significant structural rearrangement of the protein. This situation would seem to be similar to the lac repressor in prokaryotic cells; however, experiments indicate that there is no significant alteration in binding affinity

The next part of the puzzle involves a number of small proteins known as *coactivators.* Near the center of each of these are three regions with the pattern Leu-X-X-Leu-Leu. Each of these regions generates a short hydrophobic α-helix. These three helices bind to a hydrophobic region on the ligand-binding region. The presence of the ligand appears to enhance the binding of a coactivator. (A receptor protein may act as a repressor, especially in the presence of a corepressor.)

Just what are the roles of coactivators and corepressors? Their effectiveness appears to be linked to their ability to covalently bond to the tails of the histones. Histone acetyltransferases (HATs) catalyze this modification of the histone tails (a process that is reversed by histone deacetylase enzymes — see Figure 16-14). This process changes a very polar (positively charged lysine) to a much less polar (neutral) amide, resulting in a significant reduction in the affinity of the tail to the associated DNA. To a lesser degree, it reduces the affinity of the entire histone to the associated DNA. The reduction in the affinity allows access of a portion of the DNA to transcription.

The acetyated lysine residues also affect the acetyllysine-binding domain (the *bromodomain*) present in many of the eukaryotic transcription regulatory proteins.

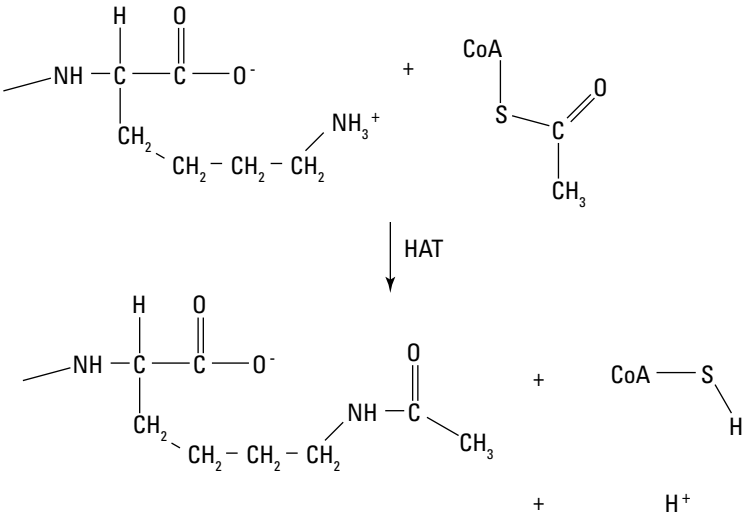


Figure 16-14: Reaction catalyzed by histone acetyltransferases (HATs).

There are two important bromodomain-containing proteins: One of these is a large complex of ten proteins that binds to the TATA-box-binding protein that is responsible for the transcription of many genes. The other proteins containing bromodomains are part of large complexes known as *chromatin-remodeling engines.* As the name implies, these proteins alter the structure of the chromatins, which changes the behavior of the chromatin.

All these factors alter the availability of portions of the DNA structure to transcription. Once the DNA becomes open, the procedures discussed earlier in the chapter come into play.

Chapter 17

Translation: Protein Synthesis

In This Chapter

- Talking about translation
- Looking at protein synthesis
- Examining eukaryotic cells
- Discussing the Human Genome Project

You are no doubt familiar with the process of translation — converting text from one language into another. The process of translation in biochemistry does exactly the same thing.

Hopefully Not Lost in Translation

Translation is the process where the four-letter alphabet of the nucleic acids becomes the twenty-letter alphabet of proteins. In doing so, genetic information is passed on. Translation occurs in the cell's ribosomes, which contain ribosomal RNA (rRNA). The information necessary for translation travels from the cell nucleus to the ribosomes via messenger RNA (mRNA). The messenger RNA binds to the smaller ribosomal body, and transfer RNA (tRNA) brings amino acids to it.

Why translation is necessary

The purpose of translation is to put together specific amino acids in a specific order to produce a specific protein. Messenger RNA provides the template or blueprint for this process. To utilize this template, something must bring the amino acids to the mRNA, and that thing is transfer RNA (tRNA). Transfer RNA has two important sites. One site is for the attachment of a specific amino acid. For example, only one specific type of tRNA will transfer the amino acid methionine. The other site is the recognition site, which contains an

anticodon. An *anticodon* is a sequence of three bases that match a codon on the mRNA. A codon sequence of AUG on the mRNA matches the UAC anticodon on a tRNA. All of this takes place in the ribosome, home of rRNA.

Home, home in the ribosome

The *ribosome* is the factory that produces proteins. Thousands of ribosomes are present in even the simplest of cells. They are complex units composed of RNA and protein. It is possible to dissociate a prokaryotic ribosome into two units. One unit is the 50S, or large unit, and the other is the 30S, or small unit. The large unit contains 34 different proteins, labeled L1 through L34, and two RNA molecules, labeled 23S and 5S. The small unit contains 21 different proteins, labeled S1 through S21, and an RNA molecule labeled 16S. The RNA molecules in the ribosomes act as transfer RNA in translation.

A prokaryotic ribosome contains three rRNA molecules (23S, 16S, and 5S), one copy of proteins S1 through S21, two copies of L7 and L12, and one each of the other L1-L34 proteins. L7 and L12 are identical except that L7 has an acetylated amino terminus. S20 and L26 are identical. Mixing the constituents in vitro leads to the two subunits reconstituting themselves. A version of the structure of the 16S form of ribosomal RNA appears in Figure 17-1.

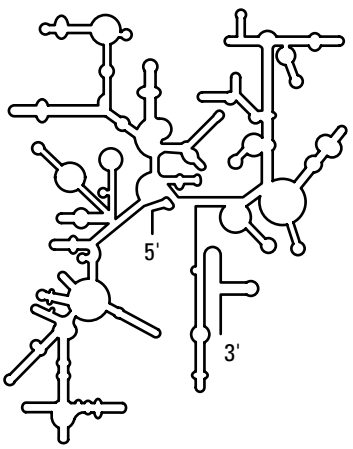


Figure 17-1: Simplified schematic of the structure of the 16S form of ribosomal RNA.

The Translation Team

A number of players, along with the rRNA, must interact in order to form a protein molecule. In addition, the structure of the ribosome is important to controlling protein synthesis. Both the rRNA and protein molecules control this structure. One possibly helpful analogy is the game of football.

The team captain: rRNA

RNA makes up approximately two-thirds of the mass of a ribosome. The three mRNA units play a key role in the shape and function of the ribosome (the proteins apparently fine-tune the shape and structure of the ribosome). The three mRNA form from the cleaving and processing of transcribed 30S RNA. A significant portion of each of the rRNA molecules have numerous duplex regions (short stretches of base-paired RNA).

The 30S and 50S rRNA subunits combine to form a 70S ribosome, which holds an mRNA in place during translation. There are also three sites for various tRNA molecules: the E, P, and A sites. The E site is the exit site. A tRNA occupies this position after delivery of its amino acid and just before exiting the ribosome. The P site is the peptidyl site, which holds the tRNA containing either the initial amino acid or the C-terminal amino acid of a protein chain. Finally, the A site is the aminoacyl site, which holds the tRNA attached to the next amino acid in sequence. When the 30S and 50S subunits join, they create A and E sites at the interface of the subunits. The P site of the 50S unit is the opening of a tunnel through which the growing protein chain passes out of the ribosome.

Here's the snap: mRNA

The base sequence of the mRNA is read in the 5' $\rightarrow$ 3' direction, and transcription occurs in this same direction. (Prokaryotic cells sometimes take advantage of this by beginning translation before transcription is over. This situation cannot occur in eukaryotic cells because the E sites of transcription and translation are physically separated.) The mRNA resulting from transcription gains a cap and a poly(A) tail before it ventures out of the nucleus on its trip to the ribosome.

Translation does not begin at the 5' terminus of the mRNA molecule. Just as there is a "stop" signal to terminate translation, there is a "start" signal. The 5' terminus base-pairs with the 3' terminus of the 16S rRNA. This region is normally about 30 nucleotides in length (a portion of this region, called the Shine-Dalgarno sequence, is purine-rich).

Shortly after this sequence is the start signal. In most cases, the start signal is AUG (methionine), though there are instances where the signal is GUG (valine). In *E. coli,* the first amino acid is formylmethionine instead of methionine. The formylmethionine is usually removed soon after translation begins. In prokaryotic cells, there may be more than one start and stop signal because many of the mRNA molecules are polygenic (polycistronic) — that is, they produce more than one protein. The structures of methionine and formylmethionine attached to tRNA are illustrated in Figure 17-2.

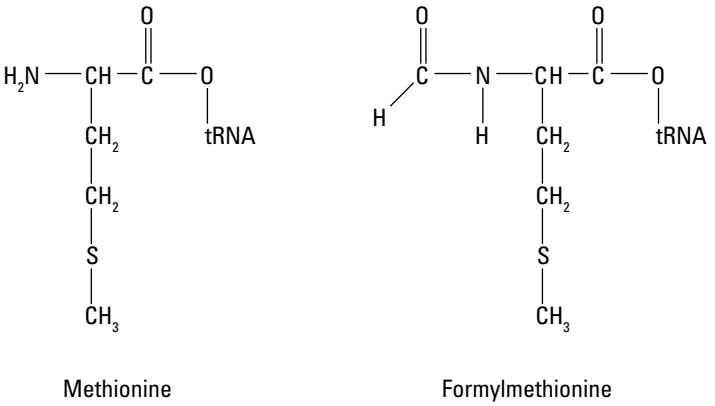


Figure 17-2: The structures of methionine and formylmethionine attached to tRNA.

Carrying the ball: tRNA

Several features are common to all forms of tRNA. Each form of tRNA is a single strand containing between 73 and 93 nucleotides. There are between seven and fifteen unusual bases (not one of the usual four, A, C, G, or U) in each molecule. Approximately half of the nucleotides present are base-paired. The activated amino acid is attached to the hydroxyl group at the 3'-end of the chain. The hydroxyl group is on the adenosine residue of a CCA segment. The other end, the 5'-end, is phosphorylated. The phosphorylation usually is a pG. The anticodon is contained in a loop near the center of the molecule.

Many of the unusual bases are methylated or dimethylated forms of A, C, G, or U, which are usually the result of post-transcription modification of the molecule. The presence of the methyl groups interferes with the formation of some base pairs, which prevents certain additional interactions. Methyl groups are nonpolar, so their presence makes regions of the tRNA hydrophobic, which affects their interaction with ribosomal proteins and syntheses. The unusual bases include dihydrouridine, dimethylguanosine, inosine, methylguanosine, methylinosine, pseudouridine, and ribothymidine. Inosine, shown in Figure 17-3, is part of the anticodon. Many of these are in or near the bends in structure of tRNA.

There are five regions, shown in Figure 17-4, that are not base-paired. (Note that the structure of tRNA shown in Figure 17-4 is not the actual three-dimensional structure of tRNA.) Starting at the 5'-end, the unpaired regions are, in order, the DHU loop, the anticodon loop, the extra arm, the TψC loop, and the 3-CCA terminus. (The name of the DHU loop derives from the presence of several dihydrouracil residues. The anticodon loop contains the segment that recognizes the codon on the mRNA, and the extra arm contains a variable number of

residues. The TψC loop derives its name from the presence of the sequence thymine-pseudouracil-cytosine.) These loops make each tRNA different, even though the overall structure is the same.

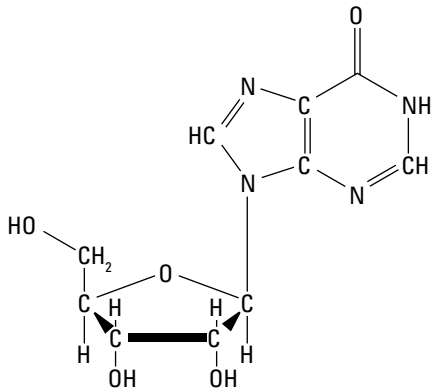


Figure 17-3: The structure of inosine.

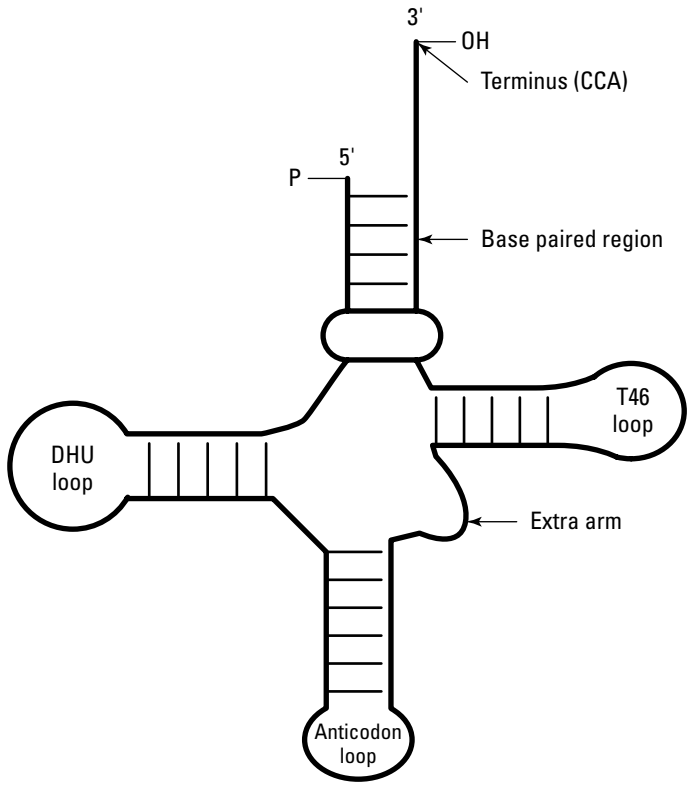


Figure 17-4: Some aspects of the structure of tRNA.

The anticodon is present in the 5' → 3' direction, and it base-pairs to a codon in the 3' → 5' direction. This matches the first base of the anticodon with the third base of the codon. (Don't forget the convention of writing base sequences in the 5' → 3' direction.)

Charging up the middle: Amino acid activation

It is imperative that the correct amino acid attaches to the tRNA because the presence of an incorrect amino acid or the absence of any amino acid would be devastating to translation. Connection of the amino acid to the tRNA activates the amino acid. Joining free amino acids is a nonspontaneous process, however, connecting the amino acid to the tRNA changes the free amino acid to a more reactive amino acid ester. The amino acid-tRNA combination is an aminoacyl-tRNA or a charged tRNA (Figure 17-5).

Specific aminoacyl-tRNA synthetases, called *activating enzymes,* catalyze the activation reaction. The process begins with an amino acid and an ATP forming an aminoacyl adenylate, which leads to the release of a pyrophosphate. Figure 17-6 shows an aminoacyl adenylate.

There is a separate aminoacyl-tRNA synthetase for each amino acid.

The two classes of aminoacyl-tRNA synthetases are denoted Class I (monomeric) and Class II (usually dimeric). Each class is responsible for ten amino acids. The CCA arm adopts different structures when interacting with members of the different classes, and ATP adopts a different conformation when interacting with members of different classes. Most Class II examples attach the amino acid as illustrated back in Figure 17-5, whereas Class I examples attach the amino acid to the alternate linking site. Some aspects of the structure of tRNA appear in the schematic structure shown in Figure 17-4.

The conversion of an aminoacyl adenylate, once formed, remains tightly bound to the synthetase until it can form an aminoacyl-tRNA.

In order to make sure that the aminoacyl-tRNA synthetase incorporates the correct amino acid, the enzyme must take advantage of specific properties of the amino acids. Examining the amino acids serine, valine, and threonine can give some insight into the selection process. These three amino acids appear in Figure 17-7, where they are drawn to emphasize similarities in the side-chain. (Recall that the threonine side-chain is chiral, but the others are not.) It is important to realize that there are size differences (–H for $–CH_3$) and

hydrogen bonding differences (–OH can, but $–CH_3$ cannot). The recognition site has the proper size and composition to take advantage of these specific properties. A significant species in this site is a zinc ion, which coordinates to the enzyme and the amino acid.

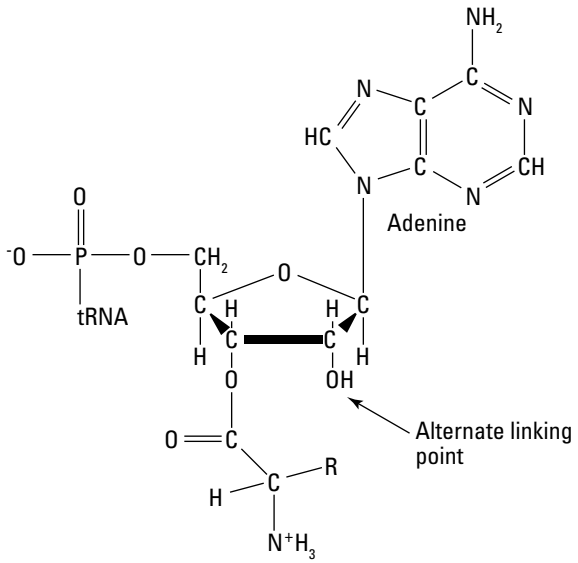


Figure 17-5: An example of an aminoacyl-tRNA.

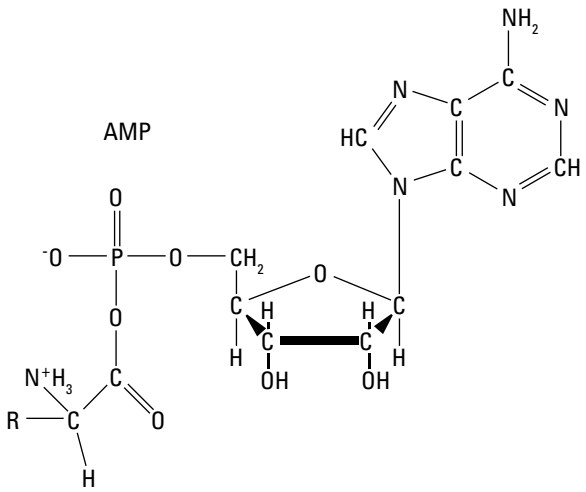


Figure 17-6: Structure of an aminoacyl adenylate.

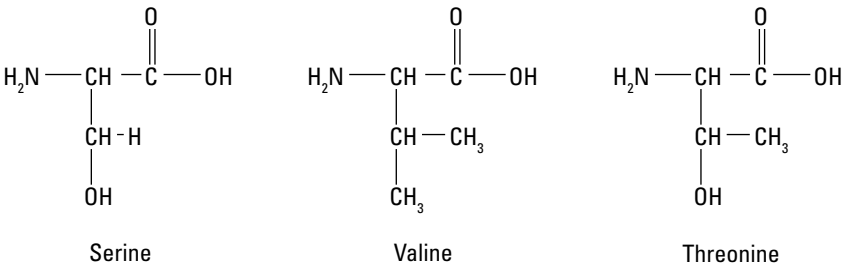


Figure 17-7: Structures of serine, valine, and threonine.

Even with these differences, serine sometimes replaces threonine. Fortunately, the enzyme includes an editing feature — the editing site is near the reaction site, but it is not the same. Similar editing occurs in other aminoacyl-tRNA synthetases. Amino acids, such as tryptophan, do not have closely similar analogues; thus, editing is far less important in these cases.

The aminoacyl-tRNA synthetases need to be able to recognize the anticodon present to make sure they interact with the appropriate tRNA, matching it to the correct amino acid. The enzymes may recognize other features of the tRNA structure. These features include the size of the extra arm and the hydrophobic character imparted by methylating some of the ribonucleotides.

Hooking Up: Protein Synthesis

The major steps in protein (polypeptide) synthesis are as follows:

- Activation
- Initiation
- Elongation
- Termination

These basics apply to all living organisms — there are no differences between human translation, fungi translation, or tulip translation. Synthesis proceeds from the amino to the carboxyl direction of the protein.

In this section we discuss these in greater detail. These steps involve tRNA, mRNA, and rRNA — along with a number of protein factors.

Activation

As mentioned earlier in this chapter, during *activation* an amino acid reacts with ATP to give aminoacyl adenylate. The aminoacyl adenylate then reacts

with a specific tRNA to give aminoacyl-tRNA plus AMP. This constitutes one of the players necessary for the translation game.

Initiation

During *initiation,* an mRNA attaches to a ribosome by interacting, through the Shine-Delgarno sequence, to the 30S rRNA subunit. Then the anticodon of the first tRNA attaches to the AUG (or GUG) codon on the mRNA. This occupies the P site of the 30S subunit. The amino acid extends into the P site of the 50S subunit. The 30S and 50S portions of the rRNA combine to produce the 70S ribosome. The combination of the two subunits allows the tRNA to interact with both parts.

In order to initiate translation it is necessary to bring the mRNA and the first tRNA to the ribosome. Three proteins, known as *initiation factors,* accomplish this task: IF1, IF2, and IF3. First, the 30S ribosome subunit, IF1 and IF3, form a complex. The two initiation factors bound to the 30S subunit interfere with a premature joining of the 30S and 50S subunit without the necessary mRNA. The remaining initiation factor, IF2, binds to GTP. The IF2-GTP combination binds to the initiator-tRNA, and the IF2-GTP-initiator-tRNA unit binds to the mRNA. Interaction of the Shine-Dalgarno sequence and the 16S rRNA manipulates the incoming group into the correct position.

Combining all these units with the 30S subunit gives the 30S initiation complex. Hydrolysis of the GTP as the 50S subunit approaches leads to expulsion of the initiation factors. With the initiation factors out of the way, the remaining moieties join to give the 70S initiation complex. (Wow, trying say that three times fast!) Once this complex forms, elongation can begin.

Elongation

During *elongation,* a second activated tRNA comes into the A site (which is adjacent to the P site) on the 30S subunit, where it binds to the appropriate codon. The activated tRNA is brought to the A site by a protein known as *elongation factor Tu* or *EF-Tu.* EF-Tu forms a complex with the activated tRNA (in the GTP form), and this complex protects the ester linkage holding the amino acid to the tRNA. In addition, the complex does not allow the activated tRNA to enter the A site if there is not a codon-anticodon match. EF-Tu interacts with all tRNAs except the initiator-tRNA. The energy needed for the EF-Tu to leave the tRNA in the ribosome comes from the hydrolysis of the GTP unit induced by the protein known as *elongation factor Ts.*

The two amino acids extend into the peptidyl transferase center of the ribosome. The amino group of the aminoacyl-tRNA from the A site is held in position to attack the ester linkage of the aminoacyl-tRNA in the P site. The catalyzed formation of the peptide bond occurs, accompanied by separation from the tRNA in the P site. The protein is now attached to the A site (30S).

With the loss of its amino acid, the tRNA no longer interacts in the same way with the ribosome. The tRNA moves to the E site of the 50S subunit as the next RNA, with its attached polypeptide, moves to the P (tunnel) site of the same subunit. The ribosome must now move over (the fancy way to say it is *translocate*) by one codon. For translocation to occur, it is necessary to employ the elongation factor G enzyme (EF-G or translocase is the protein that aids translocation). The hydrolysis of GTP to GDP supplies the energy for the move. This move places the polypeptide-tRNA into the P site of the 30S subunit. At the same time, the amino acid–stripped tRNA disengages from the mRNA and moves into the E site of the same subunit. Throughout this process, the polypeptide chain remains in the P site of the 50S subunit.

The first tRNA leaves the E site. Now the elongation cycle can begin again with the entry of another tRNA carrying the next amino acid. The process continuously cycles until a "stop" signal codon.

Termination

A "stop" signals *termination,* which results in the release of the protein, the last tRNA, and the mRNA.

Recall that the stop signals are UAA, UGA, and UAG.

Normal cells do not contain tRNAs with anticodons complementary to these codons. However, proteins known as *release factors* (RF) recognize these three codons. Release factor 1, RF1, recognizes UAA and UAG. Release factor 2, RF2, recognizes UAA and UGA. Release factor 3, RF3, is an intermediary between RF1, RF2, and the ribosome. The release factors carry a water molecule into the ribosome in place of an amino acid. The final reaction, the one that releases the newly formed protein, is the hydrolysis of the last ester linkage to a tRNA. The water brought in by the release factors is necessary for this hydrolysis.

The 70S ribosome remains together for a short time. Dissociation of the complex is mediated by a ribosome release factor and EF-G. GTP supplies the energy for this process.

The wobble hypothesis

Experimental studies have found that even pure tRNA molecules are capable of recognizing more than one codon. Biochemists developed the *wobble hypothesis* to explain this behavior, and subsequent work has firmly established this hypothesis.

The presence of the unusual base, inosine (shown back in Figure 17-3), in the anticodon loop is the key to understanding the wobble hypothesis. This base is capable of base pairing with adenine, cytosine, or uracil, allowing for some variation, or wobble, in the matching of codon to anticodon. The presence of inosine increases the number of different codons a particular tRNA can read. The first two bases in the codon pair to the corresponding bases in the anticodon. The third base is the wobble position.

Review the table of codons (Table 16-1 in Chapter 16) and see which amino acids depend only on the first two bases. Hint: Look at valine.

The base-pairing rules for the wobble hypothesis are shown in Table 17-1. The presence of an A or C as the first base allows the reading of only one codon. The presence of a G or U allows the reading of two codons, whereas an I allows the reading of three codons. Inosine is a useful base for allowing wobble; however, as Table 17-1 shows, it is only when the first anticodon base is an A or a C that there is no wobble. In general, the base in the wobble position forms weaker hydrogen bonds than normal because of the strain in the environment. The weaker hydrogen bonding aids in the loss of the tRNA after it delivers its amino acid.

Table 17-1 Base-pairing Rules for the Wobble Hypothesis

Base on Anticodon (1st Base)	*Bases Recognized on Codon (3rd Base)*
A	U
C	G
G	U, C
U	A, G
I	U, C, A

Four codons code for valine, comprising a four-codon family. If you examine three of the codons for valine: GUU, GUC, and GUA, they would all pair to the anticodon CAI instead of the anticodons CAA, CAG, and CAU. For this reason,

one CAI anticodon replaces three other anticodons. The remaining valine codon is GUG, which requires the synthesis of only two types of tRNA instead of four. Other four-codon families also work this way.

The only cases where the codons for a particular amino acid differ in the first two bases are the six-codon families, which are those of arginine, leucine, and serine. These families require three different tRNAs.

The presence of wobble reduces the number of necessary tRNAs in a cell from 61 to 31. However, cells usually have some number of tRNAs between these extremes. All the tRNAs coding for a specific amino acid require only one aminoacyl-tRNA synthetase.

Variation in Eukaryotic Cells

All cells follow the same basic pattern for translation. However, eukaryotic cells show some variations. More proteins are necessary to mediate translation, and the steps are, in general, more complicated.

Ribosomes

In eukaryotic cells, the ribosomes contain a 60S subunit and a 40S subunit, which combine to produce an 80S ribosome. The 40S subunit contains an 18S rRNA analogous to the 16S in the 30S subunit. There are three rRNA components in the 60S subunit: a 5S and a 23S, analogous to the 5S and the 23S of the prokaryotic 50S subunit, and a unique 5.8S rRNA.

Initiator tRNA

In eukaryotic cells, the initiator amino acid is methionine instead of formylmethionine. As in prokaryotic cells, a special tRNA is necessary for the first tRNA — a modification of the normal methionine-carrying tRNA.

Initiation

AUG is the only initiator codon in eukaryotic cells, and this is always the AUG nearest the 5' end of the mRNA. There is no purine-rich sequence immediately before this as in prokaryotic cells. The 40S ribosome subunit attaches to the mRNA cap and moves base by base in the 3' direction until it reaches

The Human Genome Project

The U.S. Human Genome Project was begun in 1990. It was originally scheduled to last for 15 years but because of rapid advances in the field of biotechnology it finished two years ahead of schedule in 2003. The U.S. Department of Energy and the National Institutes of Health coordinated the projects.

Goals

The Project had the following goals:

- Identify all the 20,000–25,000 genes in human DNA.
- Determine the sequences of the approximately 3 billion base pairs in human DNA.
- Store the information in databases.
- Improve data analysis tools.
- Transfer the developed technology to the private sector.
- Address the ethical, legal, and social issues associated with the project.

In addition to human DNA, researchers also studied the genetic blueprints of *E. coli*, a common bacterium found in humans as well as mice and fruit flies. The goal of transferring the technology to the private sector was included to develop the infant biotechnology industry and encourage the development of new medical applications.

Potential Benefits

Some potential benefits of the Human Genome Project include the following:

- Improved disease diagnosis
- Earlier detection of genetic predispositions to disease
- Drug design and gene therapy
- Creation of new biofuels
- More effective ways of detecting environmental pollutants
- Studying evolution through mutations in lineages
- Forensic identification of subjects through DNA analysis
- Establishing paternity
- Matching organ donors and patients
- Creation of insect- and disease-resistant crops
- Creation of biopesticides
- Increased productivity of crops and farm animals

Many of these potential benefits are showing up in our everyday life already.

Ethical, Legal, and Social Issues

One of the unique aspects of the Human Genome Project was that it was the first large scientific project that studied and addressed potential ethical, legal, and social implications that arose from the data generated from the study. Questions such as the following were addressed:

- Who should have access to personal genetic information?
- Who controls and owns genetic information?
- How reliable and useful is fetal genetic testing?
- How will genetic tests be checked for reliability and accuracy?
- Do parents have the right to test their children for adult-onset diseases?
- Do people's genes influence their behavior?
- Where is the line between medical treatment and enhancement?
- Are genetically modified foods safe for humans?

Many questions have been raised — but, as yet, few answers have resulted.

an AUG codon. The hydrolysis of ATP by helicases powers this process. Many more initiation factors are present in eukaryotic cells. A eukaryotic initiation factor has the symbol eIF instead of IF.

Elongation and termination

The EF-Tu and EF-Ts prokaryotic elongation factors have the eukaryotic counterparts EF1α and EF1$\beta\gamma$. Translocation is driven by eukaryotic EF2 with the aid of GTP. Only one release factor, eRF1, is present in eukaryotic cells instead of the two factors in prokaryotic cells. To prevent the reassembly of the two ribosome subunits, eIF3 functions like the IF3 protein in prokaryotic cells.

Part VI
The Part of Tens

The 5th Wave By Rich Tennant

"You can take that old jar for your science project, I'm sure I have some baking soda you can borrow, and let's see, where's that old particle accelerator of mine...here it is in the pantry."

In this part . . .

We wrap things up by zooming out a bit and looking at things we haven't covered yet. Here we compile two chapters' worth of short and sweet information about some of the lesser-known potential applications of biochemistry and some perhaps unexpected careers related to it.

Chapter 18

Ten Great Applications of Biochemistry

In This Chapter

- Examining tests
- Checking out genetically modified foods
- Considering cloning

In this chapter, we briefly look at some of the biochemical applications and tests that have changed our everyday lives. Although these are just a few of the hundreds of examples we could have chosen, we feel that all of these have made and continue to make a significant impact on society. And we hope you realize that more applications are being discovered almost daily.

Ames Test

The Ames test is a test that is used to determine whether a substance will affect (mutate) the structure of DNA. In this test salmonella bacteria is exposed to the chemical under question (food additives, for example), and changes in the way the bacteria grows are measured. Many substances that cause mutations in this bacteria also cause cancer in animals and humans. Indeed, this test is used today to screen chemicals for their potential ability to cause cancer in humans.

Pregnancy Testing

There are two types of pregnancy tests — one uses a urine sample and the other a blood sample. Both detect the presence of the hormone human chorionic gonadotropin (hCG). This hormone is produced by the placenta shortly after implantation of the embryo into the uterine walls and accumulates

rapidly in the body in the first few days after implantation. Home pregnancy tests, urine tests, are typically around 97 percent accurate if done two weeks after implantation. Blood tests, performed in a clinic, are more costly but can detect pregnancy as early as a week after implantation.

HIV Testing

Tests have been developed to screen for the presence of the human immunodeficiency virus. These tests may be done on urine, blood serum, or saliva and detect HIV antigens, antibodies, or nucleic acids (RNA). The nucleic acid tests (NAT) detect a 142-base sequence located on one of the HIV genes. Most blood banks use a combination of tests to ensure accuracy.

Breast Cancer Testing

Most breast cancer is not hereditary, but in 5–10 percent of cases, there *is* a heredity linkage. The vast majority of these cases is due to mutations in two genes: Breast Cancer-1 gene (BRCA1) and the Breast Cancer-2 gene (BRCA2), which were discovered in 1994 and 1995, respectively. Females who inherit a mutation in either one of these genes have a greatly increased chance of developing breast cancer over their lifetime. Postive tests for these mutations allow the individual to schedule increased screening tests at a more frequent rate than the general population.

Prenatal Genetic Testing

Prenatal genetic testing refers to testing the fetus for potential genetic defects. Tests commonly are performed on blood or tissue samples from the fetus. This may involve amniocentesis — collection of a sample of amniotic fluid that contains cells from the fetus — or collection of blood from the umbilical cord. Tests such as these are used to detect chromosomal abnormalities, such as Down syndrome or birth defects such as spina bifida.

PKU Screening

Phenylketonuria (PKU) is a metabolic disorder in which the individual is missing an enzyme called phenylalanine hydroxylase. Absence of this enzyme allows the buildup of phenylalanine, which can lead to mental retardation.

All states in the United States require PKU testing of all newborns. Infants who test positive are placed on a diet low in phenylalanine, allowing them to develop normally. Check out cans of soft drinks, and you will find a warning on many of them that they contain phenylalanine.

Genetically Modified Foods

Biochemists have developed the ability to transfer genes from one organism into other organisms, including plants and animals. This allows the creation of crops that are more pest and disease resistant and animals that are more disease resistant. Genetic modification can also be used to increase the yield of milk, eggs, or meat. In 1993, the first genetically modified food was given a license for human consumption from the U.S. Food and Drug Administration. It was a new tomato called Flavr Savr, which was resistant to rotting. However, the public has been slow to accept genetically altered foods and afraid of unforeseen effects on the population and environment.

Genetic Engineering

Genetic engineering involves taking a gene from one organism and placing it into another. The recipient may be a bacteria or a plant or an animal. One of the most well-known examples of genetic engineering involves the hormone insulin. Diabetes use to be treated with insulin derived from pigs or cows, and although very similar to human insulin, these animal-derived insulins were not identical and caused problems for some individuals. Biochemists solved this problem by inserting the gene for human insulin into bacteria. The bacteria, through the process of translation, created human insulin. (See Chapter 17 for much more on translation.)

Cloning

In 1996, Dolly the sheep was cloned — the first mammal ever cloned from adult animal cells. The cloned sheep was, of course, genetically identical to the original adult sheep. This clone was created by taking cells from the udder of a 6-year-old ewe and growing them in the lab. They then took unfertilized eggs and stripped the genetic material from them. Finally, they inserted the genetic material grown in the lab into these cells and implanted them into the uterus of another sheep. And Dolly was born. Since Dolly, many other animals have been successfully cloned. However, there is worldwide debate on the idea of cloning a human, which will surely continue for decades.

Gene-Replacement Therapy

In gene-replacement therapy, a modified or healthy gene is inserted into the organism to replace a disease-causing gene. Commonly a virus that has been altered to carry human DNA is used to deliver the healthy gene to the targeted cells of the patient. This process was first used successfully in 1990 on a 4-year-old patient who lacked an immune system due to a rare genetic disease called severe combined immunodeficiency (SCID). Individuals with SCID were prone to life-threatening infections. They lead isolated lives, avoiding people and commonly taking massive doses of antibiotics. Scientists removed white blood cells from the patient, grew them in the lab, and inserted the missing gene into the cells. They then inserted this genetically altered blood back into the patient. The process allowed the child to develop normally and even attend school, but the treatment must be repeated every few months.

Chapter 19

Ten Biochemistry Careers

In This Chapter

- Cruising careers
- Finding out about professions
- Juggling jobs

Because of recent advances in biochemistry and its related area biotechnology, many new professions have been created for the individual majoring in biochemistry. Those who stop at the B.S. degree often find themselves working as technicians in a variety of industries, but for those who go on for their M.S. or Ph.D. many more opportunities become possible.

Graduates at all levels find positions in a wide variety of career areas including forensics, industrial chemistry, molecular biology, pharmacology, technical sales, virology, horticulture, immunology, forestry, and so on. We have mentioned several careers throughout the book, so we are including here careers one might not normally associate with the field of biochemistry.

Research Assistant

A research assistant works in the area of biochemical research and development as part of a team. They may investigate new genetic tests, be involved in genetic engineering or cloning, or help with the development of new drugs or drug protocols. In addition to performing typical technical biochemical procedures, the research assistant analyzes data and prepares technical reports and summaries. Research assistants are often also involved in the search for inventions that can lead to patents. They may eventually head up their own research groups.

Plant Breeder

A plant breeder designs and implements plant breeding projects in conjunction with other research teams. They may be involved in the development of disease-resistant strains of crops or may search for ways to increase crop yields using biochemical and biotechnological techniques. They may also be involved in personnel management, public relations, and/or advising their company about future projects and plant-breeding goals.

Quality Control Analyst

The quality control analyst conducts analyses of raw materials and the finished products coming off the production line. They collect data concerning quality control test procedures and pinpoint sources of error. Along with quality control engineers, the analyst ensures that the quality of the product remains high. This is especially important, as you might imagine, when the product could be a genetically modified virus or a genetically altered food.

Clinical Research Associate

Clinical research associates design and implement clinical research projects such as a new drug protocol or the use of a new virus for gene therapy. They may travel to the various field sites where the clinical trials are being conducted to coordinate and/or supervise the trials. The clinical associate analyzes and evaluates data from the trials to ensure that clinical protocols were followed. A background in nursing or pharmacology is useful.

Technical Writer

Anyone who has ever read a poorly written set of directions or technical manual realizes the importance of a good technical writer. A technical writer in the biochemical world edits and writes operating procedures, laboratory manuals, clinical protocols, and so on. They ensure that these documents are written in a way that meets government regulations. They may develop professional development programs for staff members and write news releases. Part of their job is to take highly technical reports and edit them in such a way that they are understandable to the company's administration and the general public.

Biochemical Development Engineer

The biochemical development engineer takes the biochemical process developed in the laboratory and scales it up through the pilot plant stage to the full production plant. They help determine what instrumentation and equipment are needed and troubleshoot problems in the scale-up procedure. They work to develop more efficient manufacturing processes while maintaining a high degree of quality control. They may also be involved in technological advances from another area and apply them to their manufacturing process.

Market Research Analyst

Market research analysts analyze and research the company's market, the product mix, and the competition. They perform literature searches and make presentations on technical areas and new potential markets for the company. They predict future marketing trends based upon market research and may even be involved in the preparation of research proposals.

Patent Attorney

A patent attorney coordinates and prepares documentation for patent applications. They track the company's research studies and recommend the timing of patent filings. They collect supporting documentation and negotiate patent licenses and other legal agreements. They may become involved with interference and appeal hearings.

Pharmaceutical Sales

An individual with a degree in biochemistry becomes a natural for a career in pharmaceutical sales. These sales representatives spend much of their time on the road, talking to hospital personnel, physicians, pharmacists, and others. They are quite familiar with their company's products and try to be as persuasive as possible in touting their advantages over the competition. They have to be familiar with statistics and issues of concern in the medical community in order to successfully communicate with potential clients.

Biostatistician

Biostatisticians are statisticians who work in health-related fields. They design research studies and collect and analyze data on problems — such as how a disease progresses, how safe a new treatment or medication is, or the impact of certain risk factors associated with medical conditions. They may also design and analyze studies to determine health care costs and health care quality. They are instrumental in the designing stages of studies, providing expertise on experimental design, sample sizes, and other considerations.

Index

• Symbols & Numerics •

• A •

• B •

• C •

• D •

• E •

• F •

• G •

• H •

• I •

• J •

• K •

• L •

• M •

• N •

• O •

• P •

• Q •

• R •

• S •

• T •

• U •

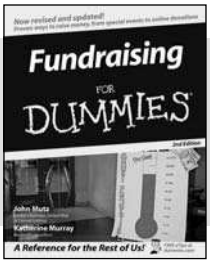
• V •

• W •

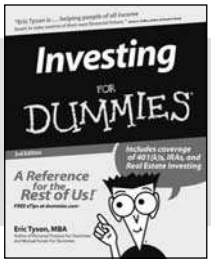
• X •

• Z •

BUSINESS, CAREERS & PERSONAL FINANCE



0-7645-9847-3

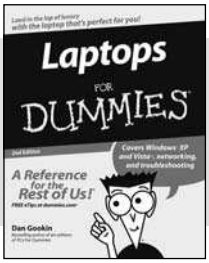


0-7645-2431-3

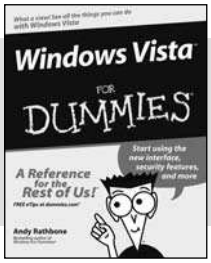
Also available:

- Business Plans Kit For Dummies 0-7645-9794-9
- Economics For Dummies 0-7645-5726-2
- Grant Writing For Dummies 0-7645-8416-2
- Home Buying For Dummies 0-7645-5331-3
- Managing For Dummies 0-7645-1771-6
- Marketing For Dummies 0-7645-5600-2
- Personal Finance For Dummies 0-7645-2590-5*
- Resumes For Dummies 0-7645-5471-9
- Selling For Dummies 0-7645-5363-1
- Six Sigma For Dummies 0-7645-6798-5
- Small Business Kit For Dummies 0-7645-5984-2
- Starting an eBay Business For Dummies 0-7645-6924-4
- Your Dream Career For Dummies 0-7645-9795-7

HOME & BUSINESS COMPUTER BASICS



0-470-05432-8

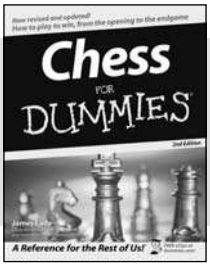


0-471-75421-8

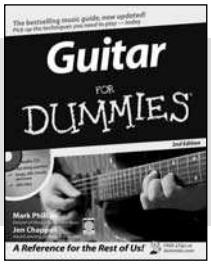
Also available:

- Cleaning Windows Vista For Dummies 0-471-78293-9
- Excel 2007 For Dummies 0-470-03737-7
- Mac OS X Tiger For Dummies 0-7645-7675-5
- MacBook For Dummies 0-470-04859-X
- Macs For Dummies 0-470-04849-2
- Office 2007 For Dummies 0-470-00923-3
- Outlook 2007 For Dummies 0-470-03830-6
- PCs For Dummies 0-7645-8958-X
- Salesforce.com For Dummies 0-470-04893-X
- Upgrading & Fixing Laptops For Dummies 0-7645-8959-8
- Word 2007 For Dummies 0-470-03658-3
- Quicken 2007 For Dummies 0-470-04600-7

FOOD, HOME, GARDEN, HOBBIES, MUSIC & PETS



0-7645-8404-9

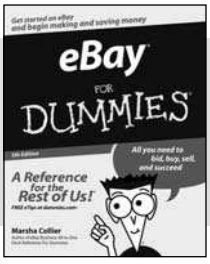


0-7645-9904-6

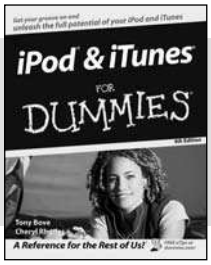
Also available:

- Candy Making For Dummies 0-7645-9734-5
- Card Games For Dummies 0-7645-9910-0
- Crocheting For Dummies 0-7645-4151-X
- Dog Training For Dummies 0-7645-8418-9
- Healthy Carb Cookbook For Dummies 0-7645-8476-6
- Home Maintenance For Dummies 0-7645-5215-5
- Horses For Dummies 0-7645-9797-3
- Jewelry Making & Beading For Dummies 0-7645-2571-9
- Orchids For Dummies 0-7645-6759-4
- Puppies For Dummies 0-7645-5255-4
- Rock Guitar For Dummies 0-7645-5356-9
- Sewing For Dummies 0-7645-6847-7
- Singing For Dummies 0-7645-2475-5

INTERNET & DIGITAL MEDIA



0-470-04529-9



0-470-04894-8

Also available:

- Blogging For Dummies 0-471-77084-1
- Digital Photography For Dummies 0-7645-9802-3
- Digital Photography All-in-One Desk Reference For Dummies 0-470-03743-1
- Digital SLR Cameras and Photography For Dummies 0-7645-9803-1
- eBay Business All-in-One Desk Reference For Dummies 0-7645-8438-3
- HDTV For Dummies 0-470-09673-X
- Home Entertainment PCs For Dummies 0-470-05523-5
- MySpace For Dummies 0-470-09529-6
- Search Engine Optimization For Dummies 0-471-97998-8
- Skype For Dummies 0-470-04891-3
- The Internet For Dummies 0-7645-8996-2
- Wiring Your Digital Home For Dummies 0-471-91830-X

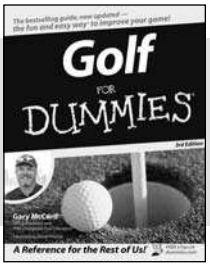
*** Separate Canadian edition also available**

† **Separate U.K. edition also available**

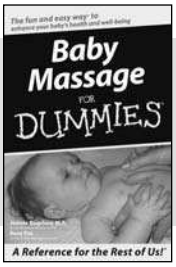
Available wherever books are sold. For more information or to order direct: U.S. customers visit www.dummies.com or call 1-877-762-2974. U.K. customers visit www.wileyeurope.com or call 0800 243407. Canadian customers visit www.wiley.ca or call 1-800-567-4797.

WILEY

SPORTS, FITNESS, PARENTING, RELIGION & SPIRITUALITY



0-471-76871-5

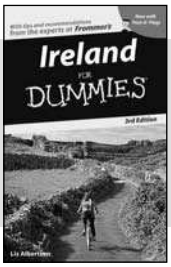


0-7645-7841-3

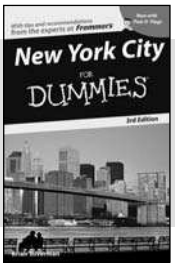
Also available:

- Catholicism For Dummies 0-7645-5391-7
- Exercise Balls For Dummies 0-7645-5623-1
- Fitness For Dummies 0-7645-7851-0
- Football For Dummies 0-7645-3936-1
- Judaism For Dummies 0-7645-5299-6
- Potty Training For Dummies 0-7645-5417-4
- Buddhism For Dummies 0-7645-5359-3
- Pregnancy For Dummies 0-7645-4483-7 †
- Ten Minute Tone-Ups For Dummies 0-7645-7207-5
- NASCAR For Dummies 0-7645-7681-X
- Religion For Dummies 0-7645-5264-3
- Soccer For Dummies 0-7645-5229-5
- Women in the Bible For Dummies 0-7645-8475-8

TRAVEL



0-7645-7749-2

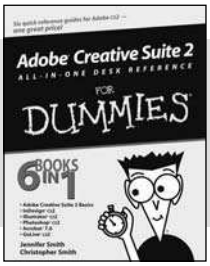


0-7645-6945-7

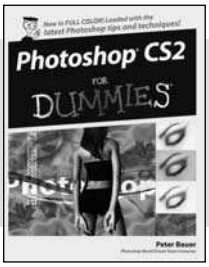
Also available:

- Alaska For Dummies 0-7645-7746-8
- Cruise Vacations For Dummies 0-7645-6941-4
- England For Dummies 0-7645-4276-1
- Europe For Dummies 0-7645-7529-5
- Germany For Dummies 0-7645-7823-5
- Hawaii For Dummies 0-7645-7402-7
- Italy For Dummies 0-7645-7386-1
- Las Vegas For Dummies 0-7645-7382-9
- London For Dummies 0-7645-4277-X
- Paris For Dummies 0-7645-7630-5
- RV Vacations For Dummies 0-7645-4442-X
- Walt Disney World & Orlando For Dummies 0-7645-9660-8

GRAPHICS, DESIGN & WEB DEVELOPMENT



0-7645-8815-X

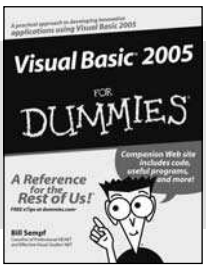


0-7645-9571-7

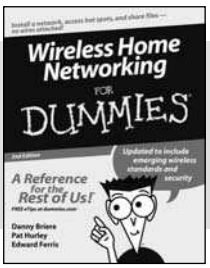
Also available:

- 3D Game Animation For Dummies 0-7645-8789-7
- AutoCAD 2006 For Dummies 0-7645-8925-3
- Building a Web Site For Dummies 0-7645-7144-3
- Creating Web Pages For Dummies 0-470-08030-2
- Creating Web Pages All-in-One Desk Reference For Dummies 0-7645-4345-8
- Dreamweaver 8 For Dummies 0-7645-9649-7
- InDesign CS2 For Dummies 0-7645-9572-5
- Macromedia Flash 8 For Dummies 0-7645-9691-8
- Photoshop CS2 and Digital Photography For Dummies 0-7645-9580-6
- Photoshop Elements 4 For Dummies 0-471-77483-9
- Syndicating Web Sites with RSS Feeds For Dummies 0-7645-8848-6
- Yahoo! SiteBuilder For Dummies 0-7645-9800-7

NETWORKING, SECURITY, PROGRAMMING & DATABASES



0-7645-7728-X



0-471-74940-0

Also available:

- Access 2007 For Dummies 0-470-04612-0
- ASP.NET 2 For Dummies 0-7645-7907-X
- C# 2005 For Dummies 0-7645-9704-3
- Hacking For Dummies 0-470-05235-X
- Hacking Wireless Networks For Dummies 0-7645-9730-2
- Java For Dummies 0-470-08716-1
- Microsoft SQL Server 2005 For Dummies 0-7645-7755-7
- Networking All-in-One Desk Reference For Dummies 0-7645-9939-9
- Preventing Identity Theft For Dummies 0-7645-7336-5
- Telecom For Dummies 0-471-77085-X
- Visual Studio 2005 All-in-One Desk Reference For Dummies 0-7645-9775-2
- XML For Dummies 0-7645-8845-1