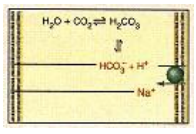
Chapter 25—
Gas Transport and pH Regulation

James Baggott



CLINICAL CORRELATION 25.1

Diaspirin Hemoglobin

Shock is a condition of inadequate tissue perfusion due, for example, to loss of blood. Hemorrhagic shock is a major cause of death following trauma. Rapid blood transfusion can be life-saving, but cross-matching must be done before transfusing blood, and transfusion is associated with a significant risk of disease. In addition, blood (or blood of the correct type) may be in short supply under certain circumstances. Hence there is considerable interest in developing a safe, effective blood substitute.

Hemoglobin in plasma has a very short lifetime. It rapidly dissociates into dimers, which bind to the plasma protein, haptoglobin, and are removed from circulation. Hemoglobin can be specifically cross-linked with bis(3,5-dibromosalicyl) fumarate at the Lys 99 of the α chains; the product is called diaspirin cross-linked hemoglobin (DCLHb). DCLHb has a longer lifetime in plasma than hemoglobin, and its lifetime can be extended still further by polymerizing the DCLHb. DCLHb has performed well as a blood replacement in experimental animals, and the possibility of using it in humans is being pursued.

25.1— Introduction to Gas Transport

Large organisms, especially terrestrial ones, require a relatively tough, impermeable outer covering to help shield them from dust, twigs, nonisotonic fluids like rain and seawater, and other elements in the environment that might be harmful to living cells. One of the consequences of being large and having an impermeable covering is that individual cells of the organism cannot exchange gases directly with the atmosphere. Instead there must exist a specialized exchange surface, such as a lung or a gill, and a system to circulate the gases (and other materials, such as nutrients and waste products) in a manner that will meet the needs of every living cell in the body.

The existence of a system for the transport of gases from the atmosphere to cells deep within the body is not merely necessary, it has definite advantages. Oxygen is a good oxidizing agent, and at its partial pressure in the atmosphere, about 160 mmHg or 21.3 kPa, it would oxidize and inactivate many components of the cells, such as essential sulfhydryl groups of enzymes. By the time O_2 gets through the transport system of the body its partial pressure is reduced to a much less damaging 20 mmHg (2.67 kPa) or less. In contrast, CO_2 is relatively concentrated in the body and becomes diluted in transit to the atmosphere. In the tissues, where it is produced, its partial pressure is 46 mmHg (6.13 kPa) or more. In the lungs it is 40 mmHg (5.33 kPa), and in the atmosphere only 0.2 mmHg (0.03 kPa), less abundant than the rare gas, argon. Its relatively high concentration in the body permits it to be used as one component of a physiologically important buffering system, a system that is particularly useful because, upon demand, the concentration of CO_2 in the extracellular fluid can be varied over a rather wide range. This is discussed in more detail later in the chapter.

Oxygen and CO_2 are carried between the lungs and the other tissues by the blood. In the blood some of each gas is present in simple physical solution, but mostly each is involved in some sort of interaction with hemoglobin, the major protein of the red blood cell. There is a reciprocal relation between hemoglobin's affinity for O_2 and CO_2, so that the relatively high level of O_2 in the lungs aids the release of CO_2, which is to be expired, and the high CO_2 level in other tissues aids the release of O_2 for their use. Thus a description of the physiological transport of O_2 and CO_2 is the story of the interaction of these two compounds with hemoglobin.

25.2— Need for a Carrier of Oxygen in Blood

An O_2 carrier is needed in blood because O_2 is not soluble enough in blood plasma to meet the body's needs. At 38°C, 1 L of plasma dissolves only 2.3 mL of O_2. Whole blood, because of its **hemoglobin,** has a much greater oxygen capacity (see Clin. Corr. 25.1). One liter of blood normally contains about 150 g of hemoglobin (contained within the erythrocytes), and each gram of hemoglobin can combine with 1.34 mL of O_2. Thus the hemoglobin in 1 L of blood can carry 200 mL of O_2, 87 times as much as plasma alone would carry. Without an O_2 carrier, the blood would have to circulate 87 times as fast to provide the same amount of O_2. As it is, the blood makes a complete circuit of the body in 60 s under resting conditions, and in the aorta it flows at the rate of about 18.6 m s^{-1}. An 87-fold faster flow would require a fabulous high-pressure pump, would produce tremendously turbulent flow and high shear forces in the plasma, would result in uncontrollable bleeding from wounds, and would not even allow the blood enough time in the lungs to take up O_2. The availability of a carrier not only permits us to avoid these impracticalities, but also gives us a way of controlling oxygen delivery, since the O_2 affinity of the carrier is responsive to changing physiological conditions.

Respiratory System Anatomy Affects Blood Gas Concentration

The respiratory system includes the trachea, in the neck, which bifurcates in the thorax into right and left bronchi, as shown schematically in Figure 25.1. The bronchi continue to bifurcate into smaller and smaller passages, ending with tiny bronchioles, which open into microscopic gas-filled sacs called alveoli. It is in the alveoli that gas exchange takes place with the alveolar capillary blood.

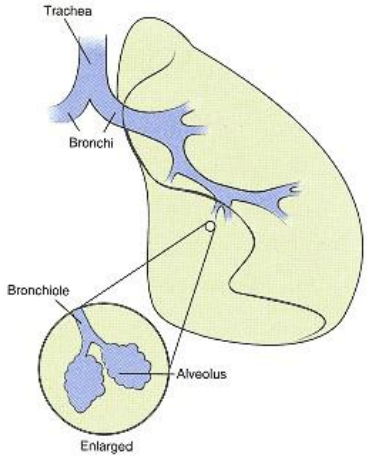


Figure 25.1
Diagram showing the respiratory tract.

As we inhale and exhale, the alveoli do not appreciably change in size. Rather, it is the airways that change in length and diameter as the air is pumped into and out of the lungs. Gas exchange between the airways and the alveoli then proceeds simply by diffusion. These anatomical and physiological facts have two important consequences. In the first place, since the alveoli are at the ends of long tubes that constitute a large dead space, and the gases in the alveoli are not completely replaced by fresh air with each breath, the gas composition of the alveolar air differs from that of the atmosphere, as shown in Table 25.1. Oxygen concentration is lower in the alveoli because it is removed by the blood. Carbon dioxide concentration is higher because it is added. Since we do not usually breathe air that is saturated with water vapor at 38°C, water vapor is generally added in the airways. The concentration of nitrogen is lower in the alveoli, not because it is taken up by the body, but simply because it is diluted by the CO_2 and water vapor.

A second consequence of the existence of alveoli of essentially constant size is that the blood that flows through the pulmonary capillaries during expiration, as well as the blood that flows through during inspiration, can exchange gases. This would not be possible if the alveoli collapsed during expiration and contained no gases, in which case the composition of the blood gases would fluctuate widely, depending on whether the blood passed through the lungs during an inspiratory or expiratory phase of the breathing cycle.

A Physiological Oxygen Carrier Must Have Unusual Properties

We have seen that an O_2 carrier is necessary. Clearly this carrier would have to be able to bind oxygen at an O_2 tension of about 100 mmHg (13.3 kPa), the partial pressure of oxygen in the alveoli. The earner must also be able to release O_2 to the extrapulmonary tissues. The O_2 tension in the capillary bed of an active muscle is about 20 mmHg (2.67 kPa). In resting muscle it is higher, but during extreme activity it is lower. These O_2 tensions represent the usual limits within which an oxygen carrier must work. An efficient carrier would be nearly fully saturated in the lungs but should be able to give up most of this to a working muscle.

Let us first see whether a carrier that binds O_2 in a simple equilibrium represented by

$$\text{Oxygen} + \text{carrier} \rightleftharpoons \text{oxygen} \cdot \text{carrier}$$

TABLE 25.1 Partial Pressures of Important Gases Given in Millimeters of Hg (kPa)

	In the Atmosphere		*In the Alveoli of the Lungs*	
Gas	***mmHg***	***kPa***	***mmHg***	***kPa***
O_2	159	21.2	100	13.3
N_2	601	80.1	573	76.4
CO_2	0.2	0.027	40	5.33
H_2O	0	0	47	6.27
Total	760	101	760	101

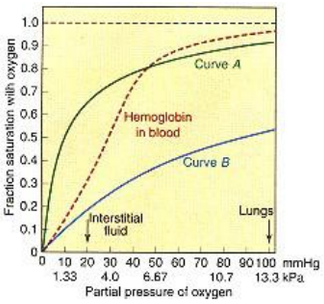


Figure 25.2
Oxygen saturation curves for two hypothetical oxygen carriers and for hemoglobin.
Curve *A*: Hypothetical carrier with hyperbolic saturation curve (a simple carrier), 90% saturated in the lungs and 66% saturated at the partial pressure found in interstitial fluid. Curve *B*: Hypothetical carrier with hyperbolic saturation curve (another simple carrier), 56% saturated in the lungs and 20% saturated at the partial pressure found in interstitial fluid. Dashed curve: Hemoglobin in whole blood.

CLINICAL CORRELATION 25.2

Cyanosis

Cyanosis is a condition in which a patient's skin or mucous membrane appears gray or (in severe cases) purple-magenta. It is due to an abnormally high concentration of deoxyhemoglobin below the surface, which is responsible for the observed color. The familiar blue of superficial veins is due to their deoxyhemoglobin content and is a normal manifestation of this color effect.

Cyanosis is most commonly caused by diseases of the cardiac or pulmonary systems, resulting in inadequate oxygenation of the blood. It can also be caused by certain hemoglobin abnormalities. Severely anemic individuals cannot become cyanotic; they do not have enough hemoglobin in their blood for the characteristic color of its deoxy form to be apparent.

Albert, R. K. Approach to the patient with cyanosis and/or hypoxemia. In: W. N. Kelley (Ed.), *Textbook of Internal Medicine*. Philadelphia: Lippincott, 1989, pp. 2041–2044.

would be satisfactory. For this type of carrier the dissociation constant would be given by the simple expression

$$K_d = \frac{[\text{oxygen}][\text{carrier}]}{[\text{oxygen} \cdot \text{carrier}]}$$

and the saturation curve would be a **rectangular hyperbola.** This model would be valid even for a carrier with several oxygen-binding sites per molecule, which we know is the case for hemoglobin, as long as each site were independent and not influenced by the presence or absence of O_2 at adjacent sites.

If such a carrier had a dissociation constant that permitted 90% saturation in the lungs, then, as shown in Figure 25.2, curve *A*, at a partial pressure of 20 mmHg (2.67 kPa) it would still be 66% saturated and would have delivered only 24% of its O_2 load. This would not be very efficient.

What about some other simple carrier, one that bound O_2 less tightly and therefore released most of it at low partial pressure, so that the carrier was, say, only 20% saturated at 20 mmHg (2.67 kPa)? Again, as shown in Figure 25.2, curve *B*, it would be relatively inefficient; in the lungs this carrier could fill only 56% of its maximum O_2 capacity and would deliver only 36% of what it could carry. It appears then that the mere fivefold change in O_2 tension between the lungs and the unloading site is not compatible with efficient operation of a simple carrier. Simple carriers are not sensitive enough to respond massively to a signal as small as a fivefold change.

Figure 25.2 also shows the oxygen-binding curve of hemoglobin in normal blood. The curve is **sigmoid,** not hyperbolic, and it cannot be described by a simple equilibrium expression. Hemoglobin, however, is a very good physiological O_2 carrier. It is 98% saturated in the lungs and only about 33% saturated in the working muscle. Under these conditions it delivers about 65% of the O_2 it can carry.

It can be seen in Figure 25.2 that hemoglobin is 50% saturated with O_2, at a partial pressure of 27 mmHg (3.60 kPa). The partial pressure corresponding to 50% saturation is called the $\boldsymbol{P_{50}}$. The term P_{50} is the most common way of expressing hemoglobin's O_2 affinity. By analogy with K_m for enzymes, a relatively high P_{50} corresponds to a relatively low O_2 affinity.

The Steep Part of the Curve Lies in the Physiological Range

Note that the steep part of hemoglobin's saturation curve lies in the range of O_2 tensions that prevail in the extrapulmonary tissues. This means that relatively small decreases in oxygen tension in these tissues will result in large increases in O_2 delivery, this effect becoming more pronounced as the partial pressure of O_2 diminishes within the physiological range. Furthermore, small shifts of the curve to the left or right will also strongly influence O_2 delivery. In Sections 25.3, 25.5, and 25.6 we see how physiological signals effect such shifts and result in enhanced delivery under conditions of increased O_2 demand. Small decreases of O_2 tension in the lungs, however, such as occur at moderately high altitudes, do not seriously compromise hemoglobin's ability to bind oxygen. This will be true as long as the alveolar partial pressure of O_2 remains in a range that corresponds to the relatively flat region of hemoglobin's O_2 dissociation curve (see Clin. Corr. 25.2).

Finally, we can see from Figure 25.2 that the binding of oxygen by hemoglobin is cooperative. At very low O_2 tension the hemoglobin curve tends to follow the hyperbolic curve, which represents relatively weak O_2 binding, but at higher tensions it actually rises above the hyperbolic curve that represents tight binding. Thus hemoglobin binds O_2 weakly at low oxygen tension and tightly at high tension. The binding of the first O_2 to each hemoglobin molecule enhances the binding of subsequent O_2 molecules.

Hemoglobin's ability to bind O_2 cooperatively is reflected in its **Hill coefficient,** which has a value of about 2.7. (The Hill equation is derived and interpreted on p. 119.) Since the maximum value of the Hill coefficient for a system at equilibrium is equal to the number of cooperating binding sites, a value of 2.7 means that hemoglobin, with its four oxygen-binding sites, is more cooperative than would be possible for a system with only two cooperating binding sites, but it is not as cooperative as it could be.

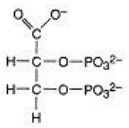


Figure 25.3
2,3-Bisphosphoglycerate (BPG).

25.3— Hemoglobin and Allosterism: Effect of 2,3-Bisphosphoglycerate

Hemoglobin's binding of O_2 was the original example of a **homotropic effect** (cooperativity and allosterism are discussed in Chapter 4), but hemoglobin also exhibits a **heterotropic effect** of great physiological significance. This involves its interaction with **2,3-bisphosphoglycerate** (BPG) (Figure 25.3), which is closely related to the glycolytic intermediate, 1,3-bisphosphoglycerate, from which it is biosynthesized.

It had been known for many years that hemoglobin in the red cell bound oxygen less tightly than purified hemoglobin could (Figure 25.4). It had also been known that the red cell contained high levels of BPG, nearly equimolar with hemoglobin. Finally, the appropriate experiment was done to demonstrate the relationships between these two facts. It was shown that the addition of BPG to purified hemoglobin produced a shift to the right of its oxygen-binding curve, bringing it into congruence with the curve observed for whole blood. Other organic polyphosphates, such as ATP and inositol pentaphosphate, also have this effect. Inositol pentaphosphate is the physiological effector in birds, where it replaces BPG, and ATP plays a similar role in some fish.

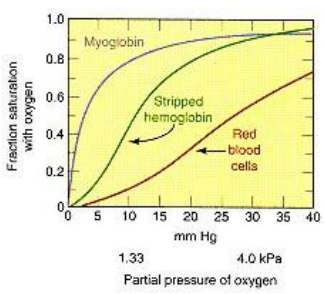


Figure 25.4
Oxygen dissociation curves for myoglobin, for hemoglobin that has been stripped of CO_2 and organic phosphates, and for whole red blood cells.
Data from Brenna, O., Luzzana, M., Pace, M., et al. *Adv. Exp. Biol. Med.* 28:19, 1972. Adapted from McGilvery, R. W. *Biochemistry: A Functional Approach,* 2nd ed. Philadelphia: Saunders, 1979, p. 236.

Monod's model of allosterism explains heterotropic interaction. Applying this model to hemoglobin, in the deoxy conformation (the **T state**) a cavity large enough to admit BPG exists between the β chains of hemoglobin. This cavity is lined with positively charged groups and firmly binds one molecule of the negatively charged BPG. In the oxy conformation (the **R state**) this cavity is smaller, and it no longer accommodates BPG as easily. The result is that the binding of BPG to oxyhemoglobin is much weaker. Since BPG binds preferentially to the T state, the presence of BPG shifts the R–T equilibrium in favor of the T state; the deoxyhemoglobin conformation is thus stabilized over the oxyhemoglobin conformation (Figure 25.5). For oxygen to overcome this and bind to hemoglobin, a higher concentration of oxygen is required. Oxygen tension in the lungs is sufficiently high under most conditions to saturate hemoglobin almost completely, even when BPG levels are high. The physiological effect of BPG can, therefore, be expected to be upon release of oxygen to the extrapulmonary tissues, where O_2 tensions are low.

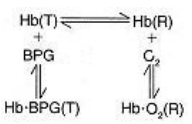


Figure 25.5
Schematic representation of equilibria among BPG, O_2, and the T and R states of hemoglobin.

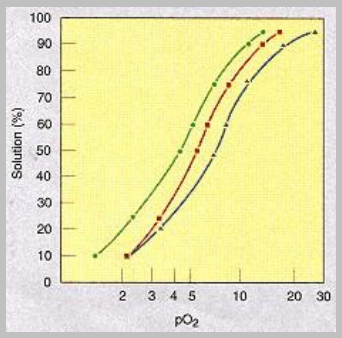
The significance of a high BPG concentration is that the efficiency of O_2 delivery is increased. Concentrations of BPG in the red cell rise in conditions associated with **tissue hypoxia,** such as various anemias, cardiopulmonary insufficiency, and high altitude. These high levels of BPG enhance the formation of deoxyhemoglobin at low partial pressures of oxygen; hemoglobin then delivers more of its O_2 to the tissues. This effect can result in a substantial increase in the amount of O_2 delivered because the venous blood returning to the heart of a normal individual is (at rest) at least 60% saturated with O_2. Much of this O_2 can dissociate in the peripheral tissues if the BPG concentration rises.

The BPG mechanism works very well as a compensation for tissue hypoxia as long as the partial pressure of oxygen in the lungs remains high enough that oxygen binding in the lungs is not compromised. Since, however, BPG shifts the oxygen-binding curve to the right, the mechanism will not compensate for tissue hypoxia when the partial pressure of O_2 in the lungs falls too low. Then

CLINICAL CORRELATION 25.3

Chemically Modified Hemoglobins: Methemoglobin and Sulfhemoglobin

Methemoglobin is a form of hemoglobin in which the iron is oxidized from the iron (II) state to the iron (III) state. A tendency for methemoglobin to be present in excess of its normal level of about 1% may be due to a hereditary defect of the globin chain or to exposure to oxidizing drugs or chemicals. Sulfhemoglobin is a species that forms when a sulfur atom is incorporated into the porphyrin ring of hemoglobin. Exposure to certain drugs or to soluble sulfides produces it. Sulfhemoglobin is green. Hemoglobin subunits containing these modified hemes do not bind oxygen, but they change the oxygen-binding characteristics of the normal subunits in hybrid hemoglobin molecules containing some normal subunits and one or more modified subunits. The accompanying figure shows the oxygen-binding curve of normal HbA, 15% methemoglobin and 12% sulfhemoglobin. The presence of methemoglobin shifts the curve to the left, impairing the delivery of the decreased amount of bound oxygen. In contrast, the sulfhemoglobin curve is shifted to the right, a BPG-like effect. As a result, oxygen delivery is enhanced, partially compensating for the inability of the sulfur-modified hemes to bind oxygen.



Oxygenation curves of unmodified hemoglobin A (squares) of a 15% oxidized hemolysate (circles) and of a hemolysate containing 12% sulfhemoglobin (triangles) in 0.1 M phosphate, pH 7.35, at 20°C. Data from Park, C. M., and Nagel, R. L., *N. Engl. J. Med.* 310:1579, 1984.

the increased efficiency of O_2 unloading to the tissues is counterbalanced by a decrease in the efficiency of loading in the lungs. This may be a factor in determining the maximum altitude at which people choose to establish permanent dwellings, which is about 18,000 ft (~5500 m). There is evidence that a better adaptation to extremely low ambient partial pressures of O_2 would be a shift of the curve to the left.

25.4— Other Hemoglobins

Although hemoglobin A is the major form of hemoglobin in adults and in children over seven months of age, accounting for about 90% of their total hemoglobin, it is not the only normal hemoglobin species. Normal adults also have 2–3% of **hemoglobin A_2,** which is composed of two α chains like those in hemoglobin A and two chains. It is represented as $_{2}{}_{2}$. The chains differ in amino acid sequence from the β chains and are under independent genetic control. Hemoglobin A_2 does not appear to be important in normal individuals.

Several species of modified hemoglobin A also occur normally. These are designated A_{1a1}, A_{1a2}, A_{1b}, and A_{1c}. They are adducts of hemoglobin with various sugars, such as glucose, glucose 6-phosphate, and fructose 1,6-bisphosphate. The quantitatively most significant is **hemoglobin A_{1c},** formed by covalent binding of a glucose residue to the N terminal of the β chain at a rate that depends on the concentration of glucose. As a result, hemoglobin A_{1c} forms more rapidly in uncontrolled diabetics and can comprise up to 12% of their total hemoglobin. Hemoglobin A_{1c} or total glycosylated hemoglobin levels are a useful measure of how well diabetes has been controlled during the days and weeks before the measurement is taken; measurement of blood glucose only indicates how well diabetes is under control when the blood sample is taken. Chemical modification of hemoglobin A can also occur from interaction with drugs or environmental pollutants (see Clin. Corr. 25.3).

Fetal hemoglobin, **hemoglobin F,** is the major hemoglobin in newborn infants. It contains two γ chains in place of the β chains and is represented as $_{2}{}_{2}$. Shortly before birth γ-chain synthesis diminishes and β-chain synthesis is initiated, and by the age of seven months well over 90% of the infant's hemoglobin is hemoglobin A.

Hemoglobin F is adapted to the environment of the fetus, who gets oxygen from maternal blood, a source that is far poorer than the atmosphere. To compete with the maternal hemoglobin for O_2, fetal hemoglobin must bind O_2 more tightly; its oxygen-binding curve is thus shifted to the left relative to hemoglobin A. This is accomplished through a difference in the influence of BPG upon the maternal and fetal hemoglobins. In hemoglobin F two of the groups that line the BPG-binding cavity have neutral side chains instead of the positively charged ones that occur in hemoglobin A. Consequently, hemoglobin F binds BPG less tightly and thus binds oxygen more tightly than hemoglobin A does. Also, about 15–20% of the hemoglobin F is acetylated at the N terminals; this is referred to as hemoglobin F_1. Hemoglobin F_1 does not bind BPG, and its affinity for oxygen is not affected at all by BPG. The postnatal change from hemoglobin F to hemoglobin A, combined with a rise in red cell BPG that peaks three months after birth, results in a gradual shift to the right of the infant's oxygen-binding curve (Figure 25.6). The result is greater delivery of oxygen to the tissues at this age than at birth, in spite of a 30% decrease in the infant's total hemoglobin concentration.

In many inherited anomalies of hemoglobin synthesis there is formation of a structurally abnormal hemoglobin; these are called **hemoglobinopathies.** They may involve the substitution of one amino acid in one type of polypeptide chain for some other amino acid or they may involve absence of one or more amino acid residues of a polypeptide chain. In some cases the change is clinically insignificant, but in others it causes serious disease (see Clin. Corr. 25.4).

25.5— Physical Factors That Affect Oxygen Binding

High Temperature Weakens Hemoglobin's Oxygen Affinity

Temperature has a significant effect on O_2 binding by hemoglobin (Figure 25.7). At below-normal temperatures the binding is tighter, resulting in a leftward shift of the curve; at higher temperatures the binding becomes weaker, and the curve is shifted to the right. The effect of elevated temperature is like that of high levels of BPG, in that both enhance unloading of oxygen. The temperature effect is physiologically useful, as it makes additional O_2 available to support the high metabolic rate found in fever or in exercising muscle with its elevated temperature. The relative insensitivity to temperature of O_2 binding at high partial pressure of oxygen minimizes compromise of O_2 uptake in the lungs under these conditions.

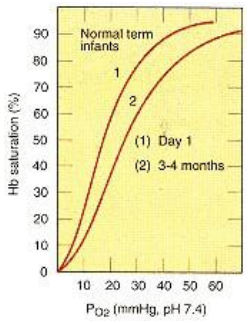


Figure 25.6
Oxygen dissociation curves after birth.
Adapted from Oski, F. A., and Delivoria-Papadopoulos, M. *J. Pediatr.* 77:941, 1970.

The tighter binding of O_2 that occurs in hypothermic conditions is not important in hypothermia induced for surgical purposes. Decreased O_2 utilization by the body and increased solubility of O_2 in plasma at lower temperatures, as well as the increased solubility of CO_2, which acidifies the blood, compensate for hemoglobin's diminished ability to release O_2.

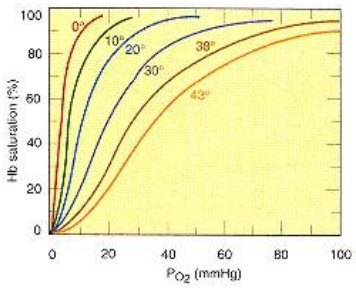


Figure 25.7
Oxygen dissociation curve for whole blood at various temperatures.
From Lambertson, C. J. In: P. Bard (Ed.), *Medical Physiology,* 11th ed. St. Louis, MO: Mosby, 1961, p. 596.

Low pH Weakens Hemoglobin's Oxygen Affinity

Hydrogen ion concentration influences hemoglobin's O_2 binding. As shown in Figure 25.8, low pH shifts the curve to the right, enhancing O_2 delivery, whereas high pH shifts the curve to the left. It is customary to express O_2 binding by hemoglobin as a function of plasma pH because it is this value, not the pH within the erythrocyte, that is usually measured. Erythrocyte cell sap pH is lower than the plasma pH, but these two fluids are in equilibrium, and changes in one reflect changes in the other.

The influence of pH upon O_2 binding is physiologically significant, since a decrease in pH is often associated with increased oxygen demand. Increased metabolic rate increases production of carbon dioxide and, as in muscular exercise and hypoxic tissue, lactic acid. These acids produced by metabolism help release oxygen to support that metabolism.

The increase in acidity of hemoglobin as it binds O_2 is known as the **Bohr effect;** an equivalent statement is that the Bohr effect is the increase in basicity of hemoglobin as it releases oxygen. The effect may be expressed by the equation

$$HHb + O_2 \leftrightharpoons HbO_2 + H^+$$

This equation gives the same information as Figure 25.8—that increases in hydrogen ion concentration favor formation of free oxygen from oxyhemoglobin, and conversely, that oxygenation of hemoglobin lowers the pH of the solution.

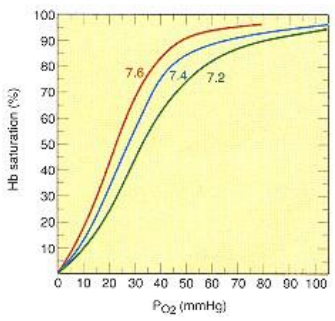


Figure 25.8
Oxygen dissociation curve for whole blood at various values of plasma pH.
Adapted from Lambertson, C. J. In: P. Bard (Ed.), *Medical Physiology,* 11th ed. St. Louis, MO: Mosby, 1961, p. 596.

25.6— Carbon Dioxide Transport

The carbon dioxide we produce is excreted by the lungs, to which it is transported by the blood. Carbon dioxide transport is closely tied to hemoglobin and to the problem of maintaining a constant pH in the blood, a problem that will be discussed subsequently.

Blood CO_2 Is Present in Three Major Forms

Carbon dioxide is present in the blood in three major forms, as dissolved CO_2, as HCO_{3-} (formed by ionization of H_2CO_3 produced when CO_2 reacts with H_2O), and as carbaminohemoglobin (formed when CO_2 reacts with amino groups of protein). Each of these is present both in arterial blood and in venous blood

(see the top three lines of Table 25.2). Net transport to the lungs for excretion is represented by the concentration difference between arterial and venous blood, shown in the last column. Note that for each form of carbon dioxide the arterial–venous difference is only a small fraction of the total amount present; venous blood contains only about 10% more **total carbon dioxide** (total CO_2 is the sum of HCO_{3-}, dissolved CO_2, and carbaminohemoglobin) than arterial blood.

After carbon dioxide enters the bloodstream for transport, it generates hydrogen ions. Most come from formation of bicarbonate ion, which occurs in the following manner.

Bicarbonate Formation

Carbon dioxide enters the blood and diffuses into erythrocytes, whose membranes, like most biological membranes, are freely permeable to dissolved CO_2. Within the erythrocytes most of the carbon dioxide is acted on by the intracellular enzyme, **carbonic anhydrase,** which catalyzes the reaction

$$CO_2 + H_2O \underset{\text{anhydrase}}{\overset{\text{carbonic}}{\rightleftharpoons}} H_2CO_3$$

This reaction proceeds in the absence of a catalyst, as is well known to all who drink carbonated beverages. Without the catalyst, however, it is too slow to meet the body's needs, taking over 100 s to reach equilibrium. Recall that at rest the blood makes a complete circuit of the body in 60 s. Carbonic anhydrase is a very active enzyme, having a turnover number of the order of 10_6, and inside the erythrocytes the reaction reaches equilibrium within 1 s, less than the time spent by the blood in the capillary bed. The enzyme contains zinc and accounts in part for our dietary requirement for this metal.

The ionization of carbonic acid, $H_2CO_3 \rightleftharpoons H^+ + HCO_3^-$, is a rapid, spontaneous reaction. It produces equivalent amounts of H^+ and HCO_{3-}. Since, as shown in the last column of line 2 in Table 25.2, 1.69 meq of bicarbonate was added to each liter of blood by this process, 1.69 meq of H^+ must also have been generated per liter of blood. Addition of this much acid, over 10_{-3} equiv of H^+, to 1 L of water would give a final pH below 3. Since the pH of venous plasma averages 7.37, most of the H^+ generated during HCO_{3-} production must be consumed by buffer action and/or other processes. This is discussed below.

Because of the compartmentalization of carbonic anhydrase, essentially all conversion of CO_2 to H_2CO_3, and ultimately to HCO_3^-, occurs inside the erythrocyte. Negligible amounts of CO_2 react nonenzymatically in the plasma. Thus virtually all of the increase in HCO_3^- in venous as compared to arterial blood is generated in erythrocytes. Most of this diffuses into the plasma, so that venous plasma HCO_3^- is higher than the arterial, but the erythrocyte was the site of its formation.

Carbaminohemoglobin Formation

It has been observed that in the presence of carbonic anhydrase inhibitors, such as acetazolamide or cyanide, blood will still take up a certain amount of carbon dioxide rapidly. This is due to the reaction of carbon dioxide with amino groups of proteins within erythrocytes to form **carbamino groups** (Figure 25.9). Hemoglobin is quantitatively the most important protein involved in this reaction. Deoxyhemoglobin forms **carbamino hemoglobin** more readily than oxyhemoglobin. Oxygenation causes release of CO_2 in carbaminohemoglobin.

Carbaminohemoglobin formation occurs only with uncharged aliphatic amino groups, not with the charged form, $R–NH_3^+$. The pH within erythrocytes is normally about 7.2, somewhat more acidic than the plasma. Since protein amino groups have pK values well to the alkaline side of 7.2, they will be mostly in the charged (undissociated acid) form. Removal of some of the un-

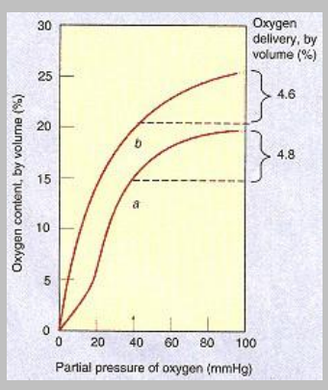
CLINICAL CORRELATION 25.4

Hemoglobins with Abnormal Oxygen Affinity

Some abnormal hemoglobins have an altered affinity for oxygen. If oxygen affinity is increased (P_{50} decreased), oxygen delivery to the tissues will be diminished unless some sort of compensation occurs. Typically, the body responds by producing more erythrocytes (polycythemia) and more hemoglobin. Hb Rainier is an abnormal hemoglobin in which the P_{50} is 12.9 mmHg, far below the normal value of 27 mmHg.

In the accompanying figure the oxygen content in volume percent (mL of O_2 per 100 mL of blood) is plotted versus partial pressure of oxygen, both for normal blood (curve *a*) and for the blood of a patient with Hb Rainier (curve *b*). Obviously, the patient's blood carries more oxygen; this is because it contains 19.5 g of Hb per 100 mL instead of the usual 15 g per 100 mL.

Since the partial pressure of oxygen in mixed venous blood is about 40 mmHg, the volume of oxygen the blood of each individual can deliver may be obtained from the graph by subtracting the oxygen content of the blood at 40 mmHg from its oxygen content at 100 mmHg. As shown in the figure, the blood of the patient with Hb Rainier delivers nearly as much oxygen as normal blood does, although Hb Rainier delivers a significantly smaller fraction of the total amount it carries. Evidently, polycythemia is an effective compensation for this condition, at least in the resting state.



Oxygen content plotted against partial pressure of oxygen.

(continued)

(Table continued from previous page)

Curve *a* shows the oxygen dissociation curve of normal blood with a hemoglobin of 15 g dL^{-1}, P_{50} 27 mmHg, n 2.8, at pH 7.4, 37°C. Curve *b* shows that of blood from a patient with Hb Rainier, having a hemoglobin of 19.5 g dL^{-1}, P_{50} 12.9 mmHg, n 1.2, at the same pH and temperature. (1 mmHg 133.3 Pa.) On the right is shown the oxygen delivery. The compensatory polycythemia and hyperbolic curve of Hb Rainier result in practically normal arterial and venous oxygen tensions. Arrow indicates normal mixed venous oxygen tension. *From Bellingham, A. J.* Br. Med. Bull. *32:234, 1976.*

charged form via carbamino group formation shifts the equilibrium, generating more uncharged amino groups and an equivalent amount of H^+, as shown in Figure 25.10. Carbamination, like HCO_{3-} formation, generates H+.

The N-terminal α-amino groups of proteins have p*K* values in the range of 7.6–8.4. The N terminals of hemoglobin's polypeptide chains are the principal sites of carbamination. If they are blocked chemically by reaction with cyanate, carbamino formation does not occur.

The N-terminal amino groups of the β-globin chains are part of the binding site for BPG. Since they cannot bind BPG and also form carbamino groups, a competition arises. Carbon dioxide diminishes the effect for BPG and, conversely, BPG diminishes the ability of hemoglobin to form carbaminohemoglobin. Ignorance of the latter interaction led to a major overestimation of the role of carbaminohemoglobin in carbon dioxide transport. Prior to the discovery of the BPG effect, careful measurements were made of the capacity of purified hemoglobin (no BPG present) to form carbaminohemoglobin. The results were assumed to be applicable to hemoglobin in the erythrocyte, leading to the erroneous conclusion that carbaminohemoglobin accounted for 25–30% or more of CO_2 transport. It now appears that 13–15% of CO_2 transport is via carbaminohemoglobin. Table 25.3 summarizes the contribution of each major form of blood carbon dioxide to overall CO_2 transport.

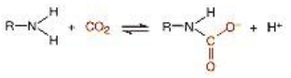


Figure 25.9
Carbamino formation from a free amino group and carbon dioxide.

Two Processes Regulate [H⁺] Derived from CO_2 Transport

Buffering

Hemoglobin, besides carrying O_2 and CO_2 in the covalently bound form of a carbamino group, also plays the major role in handling the H^+ produced in CO_2 transport. It does this by buffering and by the isohydric mechanism (discussed below). **Hemoglobin's buffering** power resides in its ionizable groups with p*K* values close to the intraerythrocyte pH. These include the four N-terminal amino groups and the imidazole side chains of the histidine residues. There are 38 histidines per hemoglobin tetramer; these provide most of hemoglobin's buffering ability.

Figure 25.10
Dissociation of an ammonium ion to yield a free amino group and H^+.

In whole blood, buffering takes up about 60% of the acid generated in normal carbon dioxide transport. Although hemoglobin is by far the most important nonbicarbonate buffer in blood, the organic phosphates in the eryth-

TABLE 25.2 Properties of Blood of Humans at Rest[a]

	Arterial			*Venous*			*A-V Difference*		
	Serum	***Cells***	***Blood***	***Serum***	***Cells***	***Blood***	***Serum***	***Cells***	***Blood***
Hb carbamino groups (meq L^{-1} of blood)		1.13	1.13		1.42	1.42		+0.29	+0.29
HCO_3^- (meq L^{-1} of blood)	13.83	5.73	19.56	14.84	6.41	21.25	+1.01	+0.68	+1.69
Dissolved CO_2 (meq L^{-1} of blood)	0.71	0.48	1.19	0.82	0.56	1.38	+0.11	+0.08	+0.19
Total CO_2 (meq L^{-1} of blood)	14.54	7.34	21.88	15.66	8.39	24.05	+1.12	+1.05	+2.17
Free O_2 (mmol L^{-1} of blood)			0.10			0.04			-0.06
Bound O_2 (mmol L^{-1} of blood)			8.60			6.01			-2.59
Total O_2 (mmol L^{-1} of blood)			8.70			6.05			-2.65
P_{O_2} (mmHg)			88.0			37.2			-50.8
P_{CO_2} (mmHg)			41.0			47.5			+6.5
pH	7.40	7.19		7.37	7.17		-0.03	-0.02	
Volume (cc L^{-1} of blood)	551.7	448.3	1000	548.9	451.1	1000	-2.8	+2.8	0.0
H_2O (cc L^{-1} of blood)	517.5	322.8	840.0	514.7	325.6	840.0	-2.8	+2.8	0.0
Cl^- (meq L^{-1} of blood)	57.71	24.30	82.01	56.84	25.17	82.01	-0.88	+0.88	0.0

Source: From Baggott, J. *Trends Biochem. Sci* 3:N207, 1978, with permission of the publisher.

[a] Hemoglobin, 9 mM; serum protein, 39.8 g L^{-1} of blood; respiratory quotient, 0.82.

TABLE 25.3 Major Forms of Carbon Dioxide Transport

Species	*Transport (%)*
HCO_3^-	78
CO_2 (dissolved)	9
Carbaminohemoglobin	13

TABLE 25.4 Processes occurring at the N Terminals of the α Chains and β Chains of Hemoglobin

	N Terminals	
Process	**α Chains**	**β Chains**
Carbamino formation	Yes	Yes
BPG binding	No	Yes
H^- binding in the Bohr effect	Yes	No

TABLE 25.5 Control of the Excess H^+ Generated During Normal Carbon Dioxide Transport

Buffering	
By hemoglobin	50%
By other buffers	10%
Isohydric mechanism (hemoglobin)	40%

rocytes, the plasma proteins, and so on also make a significant contribution. Buffering by these compounds accounts for about 10% of the H^+, leaving about 50% of acid control specifically attributable to buffering by hemoglobin. These buffer systems minimize the change in pH that occurs when acid or base is added but do not altogether prevent that change. A small difference in pH between arterial and venous blood is therefore observed.

Isohydric Mechanism

The remainder of the H^+ arising from carbon dioxide is taken up by hemoglobin, but not by buffering. Recall that when hemoglobin becomes oxygenated it becomes a stronger acid and releases H^+ (the Bohr effect). In the capillaries, where O_2 is released, the opposite occurs:

$$HbO_2 + H^+ \leftrightharpoons HHb + O_2$$

Simultaneously, CO_2 enters the capillaries and is hydrated:

$$CO_2 + H_2O \leftrightharpoons H^+ + HCO_3^-$$

Addition of these two equations gives

$$HbO_2 + CO_2 + H_2O \leftrightharpoons HHb + HCO_3^- + O_2$$

revealing that to some extent this system can take up H^+ arising from CO_2, and can do so without a change in H^+ concentration (i.e., with no change in pH). Hemoglobin's ability to do this, through the operation of the Bohr effect, is referred to as the **isohydric carriage of CO_2**. As already pointed out, there is a small A–V difference in plasma pH. This is because the isohydric mechanism cannot handle all the acid generated during normal CO_2 transport; if it could, no such difference would occur. Figure 25.11 is a schematic representation of

Figure 25.11
Schematic representation of oxygen transport and the isohydric carriage of CO_2 by hemoglobin.

In the lungs (left) O_2 from the atmosphere reacts with deoxyhemoglobin, forming oxyhemoglobin and H^+. The H^+ combines with the HCO_{3-} to form H_2O and CO_2. The CO_2 is exhaled. Oxyhemoglobin is carried to extrapulmonary tissues (right), where it dissociates in response to low P_{O_2}. The O_2 is used by metabolic processes, and CO_2 is produced. CO_2 combines with H_2O to give HCO_{3-} and H+. H+ can then react with deoxyhemoglobin to give HHb, which returns to the lungs, and the cycle repeats.

O_2 transport and the isohydric mechanism, showing what happens in the lungs and in the other tissues.

Estimates of the importance of the isohydric mechanism in handling normal respiratory acid production have changed upward and downward over the years. The older, erroneous estimates arose out of a lack of knowledge of the multiple interactions in which hemoglobin participates. The earliest experiments, titrations of purified oxyhemoglobin and purified deoxyhemoglobin, revealed that oxygenation of hemoglobin resulted in release of an average of 0.7 H^+ for every O_2 bound. This figure still appears in textbooks, and much is made of it. Authors point out that with a Bohr effect of this magnitude the isohydric mechanism alone could handle all of the acid produced by the metabolic oxidation of fat (RQ of fat is 0.7), and buffering would be unnecessary. Unfortunately, the experimental basis for this interpretation is physiologically unrealistic; the titrations were done in the total absence of carbon dioxide, which we now know binds to some of the Bohr groups, forming carbamino groups and diminishing the effect. When later experiments were carried out in the presence of physiological amounts of carbon dioxide, there was a drastic diminution of the Bohr effect, so much so that at pH 7.45 the isohydric mechanism was able to handle only the amount of acid arising from carbamino group formation. This work, however, was done prior to our appreciation of the competition between BPG and CO_2 for the same region of the hemoglobin molecule (see Table 25.4). Finally, in 1971, careful titrations of whole blood under presumably physiological conditions were carried out, yielding a value of 0.31 H^+ released per O_2 bound. This value is the basis of the present assertion that the isohydric mechanism accounts for about 40% of the H^+ generated during normal carbon dioxide transport. The quantitative contributions of various mechanisms to the handling of H^+ arising during carbon dioxide transport are summarized in Table 25.5. The major role of hemoglobin in handling this acid is obvious.

HCO_3^- Distribution between Plasma and Erythrocytes

We have seen that essentially all of HCO_3^- formation is intracellular, catalyzed by carbonic anhydrase, and that the vast bulk of the H^+ generated by CO_2 is handled within the erythrocyte. These two observations bear upon the final distribution of HCO_3^- between plasma and the erythrocyte.

Intracellular formation of HCO_3^- increases its intracellular concentration. Since HCO_3^- and Cl^- exchange freely across the erythrocyte membrane, HCO_3^- will diffuse out of the erythrocyte, increasing the plasma HCO_3^- concentration. Electrical neutrality must be maintained across the membrane as this happens. Maintenance of neutrality can be accomplished in principle either by having a positively charged ion accompany HCO_3^- out of the cell or by having some other negatively charged ion enter the cell in exchange for the HCO_3^-. Since the distribution of the major cations, Na^+ and K^+, is under strict control, it is the latter mechanism that is seen, and the ion that is exchanged for HCO_3^- is Cl^-. Thus as HCO_3^- is formed in red cells during their passage through the capillary bed, it moves out into the plasma and Cl^- comes in to replace it. The increase in intracellular Cl^- is shown in the last line of Table 25.2. In the lungs, all events that occur in the peripheral capillary beds are reversed; HCO_3^- enters the erythrocytes to be converted to CO_2 for exhalation, and Cl^- returns to the plasma. The exchange of Cl^- and HCO_3^- between the plasma and the erythrocyte is called the **chloride shift** (Figure 25.12).

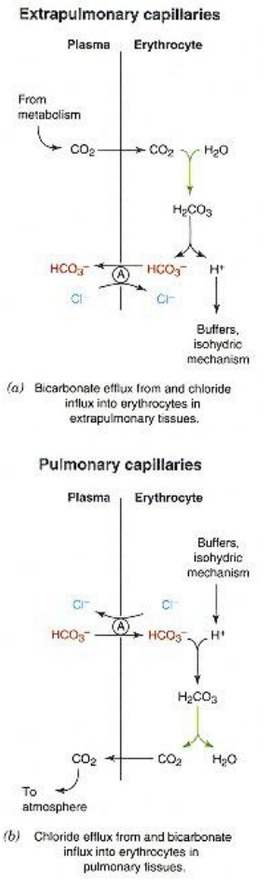


Figure 25.12
Schematic representation of the chloride shift.
(a) In the capillaries of the extrapulmonary tissues, CO_2 produced by tissue metabolism is converted to HCO_3^- in the erythrocytes. This HCO_3^- exits the erythrocytes in exchange for Cl^-.
(b) In the capillaries of the lungs, HCO_3^- enters the erythrocytes in exchange for Cl^-. Within the erythrocytes HCO_3^- is converted to CO_2. CO_2 subsequently diffuses out of the erythrocytes and is exhaled.

The intraerythrocytic buffering of H^+ from carbon dioxide causes these cells to swell, giving venous blood a slightly (0.6%) higher hematocrit than arterial blood. (Hematocrit is the volume percent of red cells in the blood.) This occurs because the charge on the hemoglobin molecule becomes more positive with every H^+ that binds to it. Each bound positive charge requires an accompanying negative charge to maintain neutrality. Thus as a result of buffering there is a net accumulation of HCO_{3-} or $Cl-$ inside the erythrocyte.

An increase in the osmotic pressure of the intracellular fluid results from this increase in concentration of particles. As a consequence, water enters the cells, causing them to swell slightly. Typically, an arterial hematocrit might be 44.8 and a venous hematocrit 45.1, as shown in Table 25.2 by the line labeled "volume (cc L^{-1} of blood)."

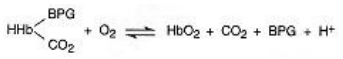


Figure 25.13
Interaction of H^+, BPG, CO_2, and O_2 with hemoglobin.
This is a schematic, intended to denote the direction of the equilibrium, not the stoichiometry of the reaction.

25.7—
Interrelationships among Hemoglobin, Oxygen, Carbon Dioxide, Hydrogen Ion, and 2,3-Bisphosphoglycerate

By now it should be clear that multiple interrelationships of physiological significance exist among the ligands of hemoglobin. These interrelationships are summarized schematically in Figure 25.13. This equation shows that changes in the concentration of H^+, BPG, or CO_2 have similar effects on O_2 binding. The equation will help you remember the effect of changes in any one of these variables upon hemoglobin's O_2 affinity.

BPG levels in the erythrocytes are controlled by product inhibition of its synthesis and by pH. Hypoxia results in increased levels of deoxyhemoglobin on a time-averaged basis. Since deoxyhemoglobin binds BPG more tightly, in hypoxia there is less free BPG to inhibit its own synthesis, and so BPG levels will rise due to increased synthesis. The effect of pH is that high pH increases BPG synthesis and low pH decreases BPG synthesis; this reflects the influence of pH on the activity of **BPG mutase,** the enzyme that catalyzes BPG formation. Since changes in BPG levels take many hours to become complete, this means that the immediate effect of a decrease in blood pH is to enhance oxygen delivery by the Bohr effect. If the acidosis is sustained (most causes of chronic metabolic acidosis are not associated with a need for enhanced oxygen delivery), diminished BPG synthesis leads to a decrease in intracellular BPG concentration, and hemoglobin's oxygen affinity returns toward normal (Figure 25.14). This system can respond appropriately to acute conditions, such as vigorous exercise, but when faced with a prolonged abnormality of pH, it readjusts to restore normal (and presumably optimal) oxygen delivery.

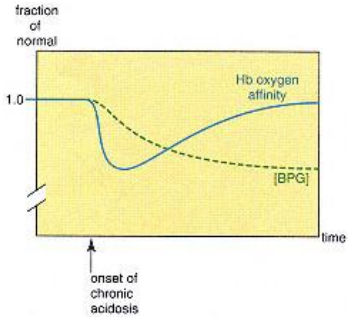


Figure 25.14
In chronic acidosis, BPG concentration decreases, returning hemoglobin's oxygen affinity toward normal.
This schematic diagram illustrates the rapid decrease in hemoglobin's oxygen affinity due to decreased pH. Lowering pH immediately lowers the activity of BPG mutase. In consequence, the concentration of BPG gradually diminishes as normal degradation proceeds. As BPG concentration diminishes, hemoglobin's oxygen affinity rises.

25.8—
Introduction to pH Regulation

We have noted the large amount of H^+ generated by carbon dioxide transport, and we considered the ways in which the blood pH is controlled. This is important because changes in blood pH will affect intracellular pH, which in turn may profoundly alter metabolism. Protein conformation is affected by pH, as is enzyme activity. In addition, the equilibria of important reactions that consume or generate hydrogen ions, such as any of the oxidation–reduction reactions involving pyridine nucleotides, are shifted by changes in pH.

Normal arterial plasma pH is 7.40 ± 0.05; the pH range compatible with life is about 6.8–7.8. Intracellular pH varies with cell type; that of the erythrocyte is nearly 7.2, but that of most other cells is lower, about 7.0. Values as low as 6.0 have been reported for skeletal muscle.

It is fortunate for both diagnosis and treatment of diseases that the acid–base status of intracellular fluid influences and is influenced by the acid–base status of the blood. Blood is readily available for analysis, and when alteration of body pH becomes necessary, intravenous administration of acidifying or alkalinizing agents is efficacious.

25.9—
Buffer Systems of Plasma, Interstitial Fluid, and Cells

Each body water compartment is defined spatially by one or more differentially permeable membranes. Each contains characteristic kinds and concentrations

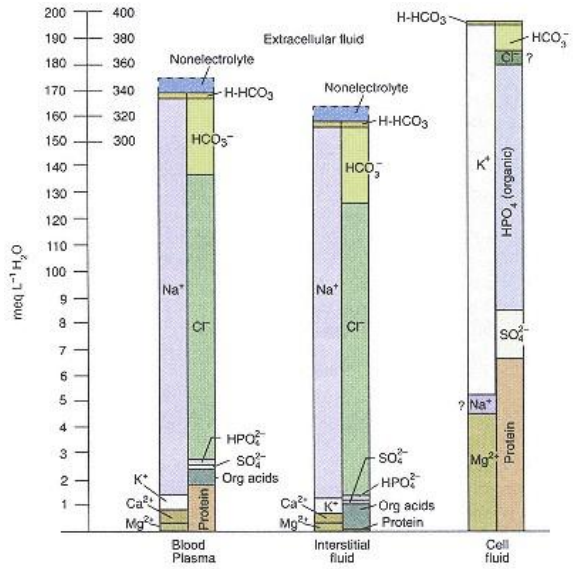


Figure 25.15
Chief chemical constituents of the three fluid compartments.
Height of left half of each column indicates total concentration of cations; that of right half, concentration of anions. Both are expressed in milliequivalents per liter (meq L^{-1}) of water. Note that chloride and sodium values in cell fluid are questioned. It is probable that, at least in muscle, the intracellular phase contains some sodium but no chloride.
Adapted from Gregersen, M. I. In: P. Bard (Ed.),
Medical Physiology, 11th ed. St. Louis, MO: Mosby, 1961, p. 307.

of solutes, some of which are buffers in the physiological range of pH. Although the solutes in each type of cell are different, most cells are similar enough to be considered together for purposes of acid–base balance. Thus there are, from this point of view, three major body water components: plasma, within the circulatory system; interstitial fluid, the fluid that bathes the cells; and intracellular fluid.

The compositions of these fluids are given in Figure 25.15. In plasma the major cation is Na^+; small amounts of K^+, Ca^{2+}, and Mg_{2+} are also present. The two dominant anions are HCO_{3-} and $Cl–$; smaller amounts of protein, phosphate, and SO_{42-} are also present, along with a mixture of organic anions (amino acids, etc.), each of which would be insignificant if taken separately. The sum of the anions equals, of course, the sum of the cations. It is apparent at a glance that the composition of interstitial fluid is very similar. The major difference is that interstitial fluid contains much less protein than plasma contains (capillary endothelium is not normally permeable to plasma proteins) and, correspondingly, a lower cation concentration. Plasma and interstitial fluid together comprise the extracellular fluid, and low molecular weight components equilibrate fairly rapidly between them. For example, $H+$ equilibrates between the plasma and interstitial fluid within about 1/2 h. The composition of intracellular fluid is strikingly different. The major cation is $K+$, while organic phosphates (ATP, BPG, glycolytic intermediates, etc.) and protein are the major anions.

TABLE 25.6 Acid Dissociation Constants of Major Physiological Buffers

Buffer System	*pK*
HCO_3^-/CO_2	6.1
Phosphate	
$HPO_4^{2-}/H_2PO_4^-$	6.7–7.2
Organic phosphate esters	6.5–7.6
Protein	
Histidine side chains	5.6–7.0
N-terminal amino groups	7.6–8.4

Because of these differences among the fluid compartments, each fluid makes a different contribution to buffering. The major buffer of extracellular fluid, for example, is the HCO_3^-/CO_2 system. Since its p*K* is 6.1 (Table 25.6 lists the major physiological buffers and their p*K* values), extracellular fluid at a pH of 7.4 is not very effective in resisting changes in pH arising from changes in P_{CO_2} changes. We have already seen the importance of buffering by hemoglobin and organic phosphates within erythrocytes. On the other hand, for reasons that will be explained

in Section 25.10, the bicarbonate buffer system is quite effective in controlling pH changes from causes other than changes in P_{CO_2}. Extracellular and intracellular fluids share almost equally in buffering strong organic or inorganic acids (see Table 25.7). Plasma is therefore an excellent indicator of the whole body's capacity to handle additional loads of these acids.

TABLE 25.7 Buffering of Metabolic Acids

Tissue	*Buffering (%)*
Extracellular fluids	42
Red cells	6
Tissue cells	52

Since acid–base imbalance arising from metabolic production of organic acids is common and potentially life-threatening, and since plasma is such a good indicator of the whole body's capacity to handle further metabolic acid loads, plasma composition is of major clinical concern. It is hydrogen ion concentration that must be kept within acceptable limits, but measuring pH alone is like walking on thin ice while observing merely whether or not you are still on the surface. Knowledge of $[HCO_3^-]$ tells you how close the ice is to the breaking point and how deep the water is underneath.

Because of the importance of the bicarbonate buffer system and its interaction with the other buffers of blood and other tissues, we will consider blood as a buffer in some detail. We will begin with a brief consideration of a model buffer.

Every buffer consists of a weak acid, HA, and its **conjugate base,** A^-. Examples of conjugate base/weak acid pairs are acetate$^-$/acetic acid, NH_3/NH_4^+, and $HPO_4^{2-}/H_2PO_4^-$. Note that the weak acid may be neutral, positively charged, or negatively charged, and that its conjugate base must (since a H^+ has been lost) have one less positive charge (or one more negative charge) than the weak acid. The degree of ionization of a weak acid depends on the concentration of free hydrogen ions. This may be expressed in the form of the **Henderson–Hasselbalch equation** (derived on p. 9) as follows:

$$\text{pH} = \text{p}K + \log \frac{[\text{conjugate base}]}{[\text{acid}]}$$

This is a mathematical rearrangement of the fundamental equilibrium equation. It states that there is a direct relationship between pH and the ratio [conjugate base]/[acid]. It is important to realize that this ratio, not the absolute concentration of any particular species, is the factor that is related to pH. Use of this equation will help you to understand the operation of and to predict the effects of various alterations upon acid–base balance in the body.

Blood plasma is a mixed buffer system; in the plasma the major buffers are HCO_3^-/CO_2, $HPO_4^{2-}/H_2PO_4^-$, and protein/Hprotein. The pH is the same throughout the plasma, so each of these buffer pairs distributes independently according to its own Henderson–Hasselbalch equation, shown in Figure 25.16. Because each pair has a different pK, the [conjugate base]/[acid] ratio is also different for each. Note, though, that if the ratio is known for any given buffer pair, information about the others can be calculated (assuming the pK values are known).

25.10— The Carbon Dioxide–Bicarbonate Buffer System

As we have seen, the major buffer of plasma and interstitial fluid is the **bicarbonate buffer system.** The bicarbonate system has two peculiar properties that make its operation unlike that of typical buffers. We will examine this important buffer in some detail, since a firm understanding of it is the key to a grasp of acid–base balance.

$$\begin{aligned}\text{pH} &= \text{p}K_1 + \log \frac{[HCO_3^-]}{[CO_2]} \\ &= \text{p}K_2 + \log \frac{[HPO_4^{2-}]}{[H_2PO_4^-]} \\ &= \text{p}K_3 + \log \frac{[\text{protein}^-]}{[\text{Hprotein}]}\end{aligned}$$

Figure 25.16
Some of the Henderson–Hasselbalch equations that are obeyed simultaneously in plasma.

The Chemistry of the System

The Equilibrium Expression Involves an Anhydride Instead of an Acid

In the first place, the component that we consider to be the acid in this buffer system is CO_2, which is an acid anhydride, not an acid. It reacts with water to

CLINICAL CORRELATION 25.5

The Case of the Variable Constant

In clinical laboratories plasma pH and P_{CO_2} are commonly measured with suitable electrodes, and plasma [HCO_3^-] is then calculated from the Henderson–Hasselbalch equation using pK = 6.1. Although this procedure is generally satisfactory, there have been several reports of severely erroneous results in patients whose acid–base status was changing rapidly.* Clinicians who are attuned to this phenomenon urge that direct measurements of all three variables be made in acutely ill patients.

The clinical literature discusses this problem in terms of departure of the value of pK from 6.1. Studies of model systems suggest that this interpretation is incorrect; pK does change with ionic strength, temperature, and so on, and so does α, but not enough to account for the magnitude of the clinical observations.

Astute commentators have speculated that the real basis of the phenomenon is disequilibrium. The detailed nature of the putative disequilibrium has not yet been established, but it is probably related to the difference in pH across the erythrocyte membrane. Normally, the pH of the erythrocyte is about 7.2, and the plasma pH is 7.4. If the plasma pH changes rapidly in an acute illness, the pH of the erythrocyte will also change, but the rate of change within the erythrocyte is not known. If the change within the erythrocyte lags sufficiently behind the change in the plasma, the system would indeed be in gross disequilibrium, and equilibrium calculations would not apply.

*See Hood, I., and Campbell, E. J. M. *N Engl. J. Med.* 306:864, 1982.

form carbonic acid, which is indeed a typical weak acid:

$$CO_2 + H_2O \leftrightharpoons H_2CO_3$$

Carbonic acid rapidly ionizes to give H^+ and HCO_{3-}:

$$H_2CO_3 \leftrightharpoons H^+ + HCO_3^-$$

If these two equations are added, H_2CO_3 cancels out, and the sum is

$$CO_2 + H_2O \leftrightharpoons H^+ + HCO_3^-$$

Elimination of H_2CO_3 from formal consideration is realistic, since not only does it simplify matters, but H_2CO_3 is, in fact, quantitatively insignificant. Because the equilibrium of the reaction,

$$CO_2 + H_2O \leftrightharpoons H_2CO_3$$

lies far to the left, H_2CO_3 is present only to the extent of 1/200 of the concentration of dissolved CO_2. Since the concentration of H_2O is virtually constant, it need not be included in the equilibrium expression for the reaction, and we may write:

$$K = \frac{[H^+][HCO_3^-]}{[CO_2]}$$

The value of K is 7.95×10^{-7}.

The concentration of a gas in solution is proportional to its partial pressure. Thus we measure partial pressure of CO_2(P_{CO_2} multiplied by a **conversion factor,** α, gives the millimolar concentration of dissolved CO_2.

$$\alpha P_{CO_2} = \text{meq } L^{-1}$$

α has a value of 0.03 meq L^{-1} $mmHg^{-1}$ (or 0.225 meq L^{-1} kPa^{-1}) at 37°C. The equilibrium expression thus becomes

$$K = \frac{[H^+][HCO_3^-]}{0.03 \cdot P_{CO_2}}$$

and the Henderson–Hasselbalch equation for this buffer system becomes

$$\text{pH} = 6.1 + \log \frac{[HCO_3^-]}{0.03 \cdot P_{CO_2}}$$

with [HCO_3^-] expressed in units of meq L^{-1} (see Clin. Corr. 25.5).

The Carbon Dioxide–Bicarbonate Buffer System Is an Open System

We said earlier that the bicarbonate buffer system, with a pK of 6.1, is not effective against carbonic acid in the pH range of 7.8–6.8 but is effective against noncarbonic acids. The usual rules of chemical equilibrium dictate that a buffer is not very useful in a pH range more than about one unit beyond its pK. Thus we need to explain how the bicarbonate system can be effective against noncarbonic acids; its failure to buffer carbonic acid is expected. The way it buffers noncarbonic acids in a pH range far from its pK is the second unusual property of this buffer system. Note that the explanation of this property in the following paragraph involves the flow of materials in a living system, and so departs from mere equilibrium considerations.

Consider first a typical buffer, consisting of a mixture of a weak acid and its conjugate base. When a strong acid is added, most of the added H^+ combines with the conjugate base. As a result, [weak acid] increases and simultaneously [conjugate base] diminishes. The ratio [conjugate base]/[weak acid] changes, and so does the pH, but much less than if there were no buffer present. Now imagine that the weak acid, as it is generated by reaction of added strong acid with conjugate base, is somehow removed so that while [conjugate base] diminishes, [weak acid] remains nearly constant. In this case the ratio of [conjugate base]/[weak acid] would change much less for a given addition of strong acid, and the pH would also change much less. This is exactly what happens with the body's bicarbonate buffer system. As strong acid is added, [HCO_{3-}]

diminishes and CO_2 is formed. But the excess CO_2 is exhaled, so that the ratio of P_{CO_2} changes strikingly, and the bicarbonate system would be relatively ineffective, in keeping with the prediction of chemical equilibrium.

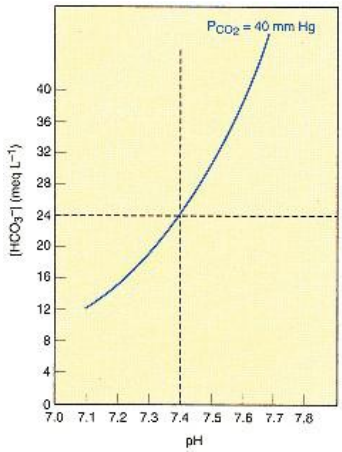


Figure 25.17
pH–Bicarbonate diagram including the 40-mmHg (5.33-kPa) CO_2 isobar, and showing the normal values of plasma pH and bicarbonate ion concentration.

Graphical Representation: The pH–Bicarbonate Diagram

A graphical representation of the Henderson–Hasselbalch equation for the bicarbonate buffer system assists in learning and understanding how this system reflects the body's acid–base status. A common representation is the **pH–bicarbonate diagram,** shown in Figure 25.17. [HCO_3^-] up to 40 meq L^{-1} is shown on the ordinate; enough to deal with most situations. Since plasma pH does not exceed 7.8 or (except transiently) fall below 7.0 in living patients, the abscissa is limited to 7.0–7.8. The normal plasma [HCO_3^-], 24 meq L^{-1}, and the normal plasma pH, 7.4, are indicated. The third variable, CO_2, can be shown on a two-dimensional graph by assigning a fixed value to P_{CO_2} is 40 mmHg (5.33-kPa), pH and [HCO_3^-] must be somewhere on that line.

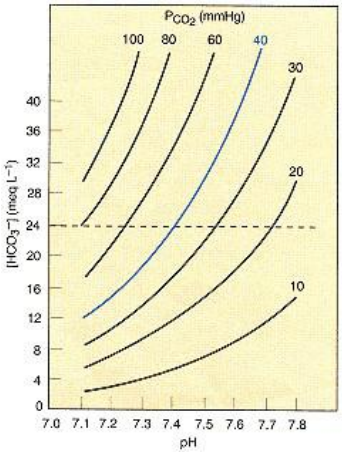


Figure 25.18
pH–Bicarbonate diagram showing CO_2 isobars from 10 to 100 mmHg.

Similarly, we can plot isobars for various abnormal values of P_{CO_2} (Figure 25.18). The range of values given covers those found in patients. Any point on the graph gives the values of the three variables of the Henderson–Hasselbalch equation for the bicarbonate system at that point. Since only two variables are needed to locate a point, the third can be read directly from the graph.

Let us now see how the bicarbonate buffer system behaves when it is in the presence of other buffers, as it is in whole blood. First, let us acidify the system by increasing the concentration of the acid-producing component, CO_2. For every CO_2 that reacts with water to produce a H^-, one HCO_{3-} forms. Most of the H^+, however, is buffered by protein and phosphate. As a result, [HCO_{3-}] rises much more than [H^+]. Similarly, if acid is removed from this system by decreasing P_{CO_2} is the only variable that is changed, the response of the system is confined to movements along this line.

The slope of the buffering line depends on the concentration of the nonbicarbonate buffers. If they were more concentrated, they would better resist changes in pH. An increase in P_{CO_2} to 80 mmHg (10.7 kPa) would then cause a smaller drop in pH, and since the more concentrated buffers would react with more hydrogen ions (produced by the ionization of carbonic acid), [HCO_3^-] would rise higher. Thus the slope of the buffering line would be steeper.

Hemoglobin is quantitatively the second most important blood buffer, exceeded only by the bicarbonate buffer system. Since hemoglobin concentration in the blood can fluctuate widely in various disease states, it is the most important physiological determinant of the slope of the blood buffer line. Figure 25.20 shows how this slope varies with hemoglobin concentration.

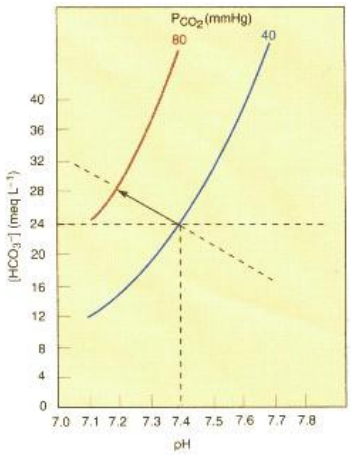


Figure 25.19
The buffering line of blood.
This pH–bicarbonate diagram shows the changes in pH that occur in whole blood in vitro when P_{CO_2} is changed. Note that the relationship between pH and $[HCO_3^-]$ is described by a straight line with a nonzero slope.

Having now seen how the bicarbonate buffer system in blood responds to changes in P_{CO_2} are represented by points confined to the CO_2 isobar.

The effects on blood of changing P_{CO_2} or of adding acid or alkali, as we have just described, are realistic qualitative models of what happens in certain disease states. We next see how these changes occur in the body and how the body compensates for them.

25.11— Acid–Base Balance and Its Maintenance

It should come as no surprise that mechanisms exist whereby the body normally rids itself of excess acid or alkali. The physiological implication is that if a patient is in a state of continuing **acidosis** (excess acid or deficiency of alkali in the body) or **alkalosis** (excess alkali or deficiency of acid in the body), there must be a continuing cause of the imbalance. In such a situation the body's first task is to somehow compensate so plasma pH does not exceed the limits compatible with life. Assistance from the physician is sometimes necessary. The body's second task is to eliminate the primary cause of the imbalance, that is, to cure the disease, so that a normal acid–base status can be reestablished. Again, intervention by the physician may be needed.

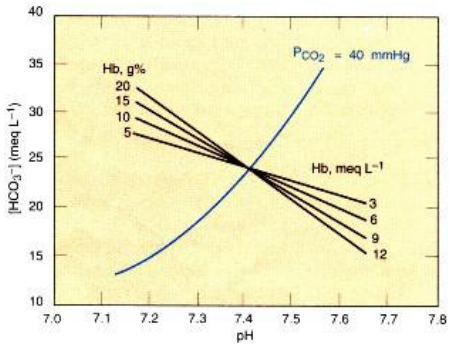


Figure 25.20
Slope of the buffering line of blood as it varies with hemoglobin concentration.
From Davenport, H. W. *The ABC of Acid–Base Chemistry*, 6th ed. revised. Chicago: University of Chicago Press, 1974, p. 55.

CLINICAL CORRELATION 25.6

The Role of Bone in Acid–Base Homeostasis

The average adult skeleton contains 50,000 meq of Ca^{2+} in the form of salts that are alkaline relative to the pH of plasma. In chronic acidosis this large reservoir of base is drawn upon to help control the plasma pH. Thus people with chronic kidney disease and severely impaired renal acid excretion do not experience a continuous decline in plasma pH and $[HCO_3^-]$. Rather, the pH and $[HCO_3^-]$ stabilize at some below-normal level. The resulting change in bone composition is not inconsequential, and clinical and roentgenologic evidence of rickets or osteomalacia often appear. Bone healing has been shown in these patients after prolonged administration of alkali in the form of sodium bicarbonate or citrate sufficient to restore plasma $[HCO_3^-]$.

Lemann, J. Jr. and Lennon, E. J. Kidney Int. *1:275, 1972.*

All individuals, in sickness or in health, produce large amounts of acids every day. The major acid is CO_2, the amount depending on the individual's caloric expenditure, and ranging between 12,500 and nearly 50,000 meq day^{-1}. In an average young adult male, about 22,000 meq of CO_2 are produced daily. This acid is volatile and is normally excreted by the lungs. Inability of the lungs to do this adequately leads to **respiratory acidosis** or alkalosis. Respiratory acidosis is the result of hypoventilation of the alveoli, so that CO_2 accumulates in the body. Alveolar hypoventilation occurs when the depth or rate of respiration diminishes. Airway obstruction, neuromuscular disorders, and diseases of the central nervous system are common causes of acute respiratory acidosis. Chronic respiratory acidosis is seen in patients with chronic obstructive lung disease, such as emphysema. Obviously, since the common element in all these conditions is increased alveolar P_{CO_2} would also cause respiratory acidosis.

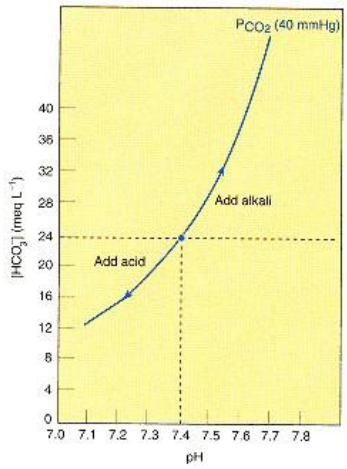


Figure 25.21
Effect of adding noncarbonic acid or alkali to whole blood with P_{CO_2} fixed at 40 mmHg.

Respiratory alkalosis, on the other hand, arises from decreased alveolar P_{CO_2} also falls, producing chronic respiratory alkalosis.

Nonvolatile acids are also produced by the body. The diet and physiological state of the individual determine the kinds and amounts of these acids. Oxidation of sulfur-containing amino acids produces H^+ and $SO_{4^{2-}}$, the equivalent of sulfuric acid. Hydrolysis of phosphate esters is equivalent to the formation of phosphoric acid. The contribution of these processes depends on the amount of acid precursors ingested; on an average American diet, net acid production is about 60 meq day^{-1}.

Metabolism normally produces lactic acid, acetoacetic acid, and β-hydroxybutyric acid. In some physiological or pathological states these are produced in excess, and accumulation of the excess causes acidosis. When an ammonium salt of a strong acid, such as ammonium chloride, or when arginine hydrochloride or lysine hydrochloride is administered, it is converted to urea, and the corresponding strong acid (HCl) is synthesized. Ingestion of salicylates, methyl alcohol, or ethylene glycol results in production of strong organic acids. Accumulation of any of these nonvolatile acids leads to **metabolic acidosis.**

While it is obvious that excess acid production can cause acidosis, the same net effect can arise from abnormal loss of base, as predicted from the Henderson-Hasselbalch equation for the bicarbonate buffer system. Renal tubular acidosis is a condition in which this occurs. Abnormal amounts of HCO_3^- escape from the blood into the urine, leaving the body acidotic (see Clin. Corr. 25.6). A more common cause of bicarbonate depletion is severe diarrhea. In this chapter it will be assumed that kidney function is normal.

Mammals do not synthesize alkaline compounds from neutral starting materials. **Metabolic alkalosis** therefore arises from intake of excess alkali or abnormal loss of acid. A commonly ingested alkali is sodium bicarbonate. A less obvious source of alkali is the salt of any metabolizable organic acid. Sodium lactate is often administered to combat acidosis; normal metabolism converts it to sodium bicarbonate. The net reaction is as follows:

$$Na^+ + CH_3CHOHCOO^- + 3O_2 \rightarrow Na^+ + HCO_3^- + 2CO_2 + 2H_2O$$

Most dietary fruits and vegetables have a net alkalinizing effect on the body for this reason. They contain a mixture of organic acids, which are metabolized to CO_2 and H_2O, and therefore have no long-term effect on acid–base balance, and salts of organic acids, which give rise to bicarbonate. Abnormal loss of acid, as occurs in prolonged vomiting or gastric lavage, causes alkalosis. Alkalosis may also be produced by rapid loss of body water, as in diuresis, which may temporarily increase $[HCO_{3^-}]$ in the plasma and extracellular fluid. Table 25.8 summarizes the causes of acid-base imbalances.

The Kidney Plays a Critical Role in Acid-Base Balance

Excess nonvolatile acid and excess bicarbonate are excreted by the kidney. As a result, urine pH varies as a function of the body's need to excrete these materials. For an individual on a typical American diet, urine pH is about 6, indicating a net acidification as compared to plasma. This is consistent with our knowledge that the typical diet results in a net production of acid. Urine pH can range from 4.4 to 8.0.

A typical daily urine volume is about 1.2 L. At the minimum urine pH of 4.4, $[H^+]$ is only 0.04 meq L^{-1}, and it would take 1250 L of urine to excrete 50 meq of acid as free hydrogen ions. Clearly, most of the acid we excrete must be in a form other than H^+. A form that can be excreted in a reasonable concentration, such as $H_2PO_4^-$ or NH_4^+, is needed.

TABLE 25.8 Causes of Acid-Base Imbalance Summarized

Acidosis
Respiratory
Alveolar hypoventilation
Metabolic
H^+ overproduction
HCO_3^- overexcretion
Alkalosis
Respiratory
Alveolar hyperventilation
Metabolic
Alkali ingestion
H^+ overexcretion

Urine Formation Occurs Primarily in the Nephron

Let us now see how the kidney accomplishes the excretion of acid or base. Figure 25.22 shows the fundamental functioning unit of the kidney, a nephron. Each human kidney contains at least a million, which first filter the blood and then modify the filtrate into urine.

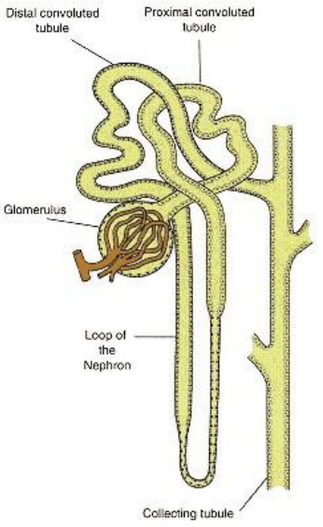


Figure 25.22
Essential features of a typical nephron in the human kidney.
Reprinted with permission from Smith, H. W. *The Physiology of the Kidney*. London: Oxford University Press, 1937, p. 6.

Filtration occurs in the glomerulus, a tuft of capillaries enclosed by an epithelial envelope called the glomerular capsule (formerly Bowman's capsule). Water and low molecular weight solutes, such as inorganic ions, urea, sugars, and amino acids (but not normally substances with molecular weights above 70,000, such as plasma proteins), pass from these capillaries into the capsular space. This ultrafiltrate of plasma then passes through the proximal convoluted tubule, where most of the water and solutes are reabsorbed. The tubule fluid continues through the loop of the nephron (loop of Henle) and through the distal convoluted tubule, where further reabsorption of some solutes or secretion of others occurs. The tubule fluid then passes into the collecting tubule, where additional concentration can occur if necessary. The fluid may now be called urine; it contains 1% or less of the water and solutes of the original glomerular filtrate.

The kidney regulates acid–base balance by controlling bicarbonate reabsorption and by secreting acid. Both processes depend on formation of H^+ and HCO_{3-} from CO_2 and H_2O within the tubule cells, shown in Figure 25.23*a*. The H^+ formed in this reaction is actively secreted into the tubule fluid in exchange for Na^+. Na^+ uptake by the tubule cell is partly passive, with Na^+ flowing down the electrochemical gradient, and partly active, via a Na^+, H^+-antiport system. At this point Na^+ has been reabsorbed in exchange for H^+, and sodium bicarbonate has been generated within the tubule cell. The sodium bicarbonate is then transported out of the cell into the interstitial fluid, which equilibrates with the plasma.

The Three Fates of Excreted H^+

The H^+ that has been secreted into the tubule fluid can now experience one of three fates. First, it can react with a HCO_{3-}, as shown in Figure 25.23*b*, to form CO_2 and H_2O. The overall net effect of this process is to move sodium bicarbonate from the tubule fluid back into the interstitial fluid. The name given to this is **reabsorption of sodium bicarbonate.**

As reabsorption of sodium bicarbonate proceeds, the tubule fluid becomes depleted of HCO_3^-, and the pH drops from its initial value, which was identical to the pH of the plasma from which it was derived. As HCO_3^- becomes less available and the pH comes closer to the p*K* of the $HPO_4^{2-}/H_2PO_4^-$ buffer system, more and more of the H^+ will be taken up by this buffer. **Buffering** is the second fate of H^+, represented in Figure 25.23*c*. $H_2PO_4^-$ is not readily reabsorbed by the kidney. It passes out in the urine, and its loss represents net excretion of H^+.

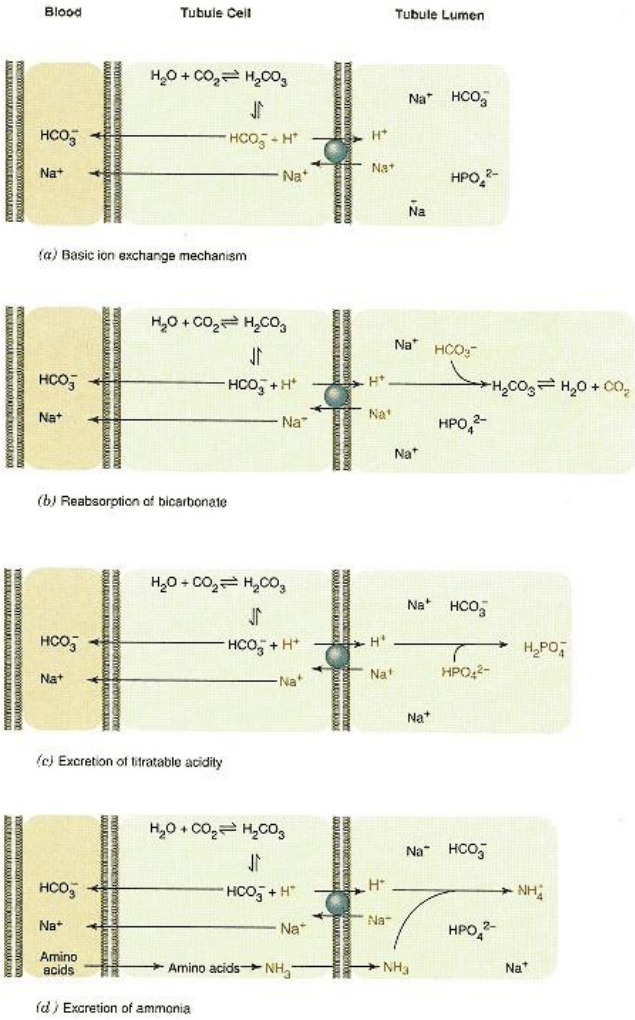


Figure 25.23
Role of the exchange of tubular cell H^+ ions in tubular fluid in renal regulation of acid–base balance.
(a) Basic ion exchange mechanism
(b) Reabsorption of bicarbonate.
(c) Excretion of titrable acid.
(d) Excretion of ammonia.
Adapted from Pitts, R. E. *N. Engl. J. Med.* 284:32, 1971, with permission of the publisher.

Although phosphate is normally the most important buffer in the urine, other ions can become significant. For example, in diabetic ketoacidosis, plasma levels of acetoacetate and β-hydroxybutyrate are elevated. These pass into the glomerular filtrate and appear in the tubule fluid. Since acetoacetic acid has a pK = 3.6 and β-hydroxybutyric acid has a pK = 4.7, as the urine pH approaches its minimum of 4,4, these begin to serve as buffers.

The effect of buffering is not only to excrete acid but to regenerate the bicarbonate that was lost when the acid was first neutralized. Let us consider a situation in which the metabolic defect of a diabetic patient has produced the elements of β-hydroxybutyric acid. The protons are neutralized by sodium

bicarbonate, leaving sodium β-hydroxybutyrate. In the kidney, then, β-hydroxybutyrate appears in the filtrate, it is converted to β-hydroxybutyric acid, which is excreted, and sodium bicarbonate returns to the extracellular fluid. Net acid excretion and bicarbonate regeneration occur no matter what anion in the tubule fluid acts as the H^+ acceptor.

The amount of acid excreted as the acid component of a urinary buffer is measured by titrating the urine back to the normal pH of the plasma, 7.4. The amount of base required is identical to the amount of acid excreted in this form and is called the **titratable acidity** of the urine.

The formation of titratable acidity accounts for about one-third to one-half of our normal daily acid excretion. It is thus an important mechanism for acid excretion and can put out as much as 250 meq of acid daily. There is, however, a limit to the amount of acid that can be excreted in this manner. Titratable acidity can be increased only by lowering the pH of the urine or by increasing the concentration of buffer in the urine, and neither of these processes can proceed indefinitely. The urine pH cannot go below about 4.4; evidently the Na^+/H^+ exchange mechanism is incapable of pumping H^+ out of the tubule cells against more than a 1000-fold concentration gradient. Buffer excretion is limited not only by the solubility of the buffer, but by limitations to the supply of the buffer ion and of the cations that are necessarily part of the important buffer systems. If a 600 meq day^{-1} of acid were excreted as NaH_2PO_4, the body would be totally depleted of sodium in less than one week.

The third fate that H^+ can experience in the tubule fluid is neutralization by NH_3. Tubule cells produce NH_4^+ from amino acids, particularly glutamine, as shown in Figure 25.23*d*. **Elimination of NH_4^+** in the urine contributes to net acid excretion.

NH_4^+ is normally a major urinary acid. Typically, one-half to two-thirds of our daily acid load is excreted as NH_4^+. For three reasons it becomes even more important in acidosis. In the first place, since the p*K* of NH_4^+ is 9.3, acid can be excreted in this form without lowering the pH of the urine, whereas formation of titratable acidity requires a decrease in urine pH. Second, enormous amounts of acid can be excreted in this form. Ammonia is readily available from amino acids, and in prolonged acidosis the NH_4^+ excretion system becomes activated. This activation, however, takes several days; it does not begin to adapt until after 2–3 days, and the process is not complete until 5–6 days after the onset of acidosis. Once complete, though, amounts of acid in excess of 500 meq can be excreted daily as NH_4^+. The third role of NH_4^+ in acidosis is that it spares the body's stores of Na^+ and K^+. Excretion of titratable acid, such as $H_2PO_4^-$, and of the anions of strong acids, such as acetoacetate, requires simultaneous excretion of a cation to maintain electrical neutrality. At the onset of acidosis this is Na^+, but as the body's Na^+ stores become depleted, K^+ excretion rises. If NH_4^+ were not available, even a moderate acidosis could quickly become fatal.

Total Acidity of the Urine

Total acid excretion, the **total acidity of the urine,** is the sum of titratable acidity and NH_4^+. Strictly speaking, we should subtract from this sum the urinary HCO_3^-, but this is seldom done in practice, since in severe metabolic acidosis, where the total acid excretion would be of greatest interest, the urine would be so acidic that $[HCO_3^-]$ would be nil.

In alkalosis the kidney's role is simply to allow HCO_3^- to escape. Metabolic alkalosis is therefore seldom long-lasting unless alkali is continuously administered or HCO_3^- elimination is somehow prevented. HCO_3^- elimination may be restricted if the kidney receives a strong signal to conserve Na^+ at a time when there is a deficiency of an easily reabsorbable anion, such as Cl^-, to be reabsorbed with it. Some diuretics cause this. The first renal response is to put out K^+ in exchange for Na^+ from the tubule fluid. When K^+ stores are depleted, H^+ is exchanged for Na^+. This results in the production of an acidic urine by

an alkalotic patient. If NaCl is administered, alkalosis associated with volume and Cl^- depletion may correct itself.

25.12— Compensatory Mechanisms

We have defined four primary types of acid–base imbalances and we have seen their chemical causes. Respiratory acidosis arises from an increased plasma P_{CO_2}. In metabolic acidosis addition of strong organic or inorganic acid (or loss of HCO_3^-) results in decreased plasma $[HCO_3^-]$. Conversely, in metabolic alkalosis loss of acid from the body or ingestion of alkali raises the plasma $[HCO_3^-]$. Recall that in an acute respiratory acid–base imbalance, as long as there is no attempt to compensate, pH will be abnormal, and $[HCO_3^-]$ will be somewhere on the buffer line. In an acute metabolic acid–base imbalance, if there is no attempt to compensate, pH will be abnormal and $[HCO_3^-]$ will be somewhere on the 40-mmHg (5.33-kPa) isobar.

Principles of Compensation

When the plasma pH deviates from the normal range, various compensatory mechanisms begin to operate. The general principle of compensation is that, since an abnormal condition has directly altered one term of the $[HCO_3^-]/[CO_2]$ ratio, plasma pH can be readjusted back toward normal by a compensatory alteration of the other term. For example, if a diabetic patient becomes acidotic due to excess production of ketone bodies, plasma $[HCO_3^-]$ will decrease. Compensation would involve decreasing plasma $[CO_2]$ so that the $[HCO_3^-]/[CO_2]$ ratio, and therefore the pH, is readjusted back toward normal. Note that compensation does not involve a return of $[HCO_3^-]$ and $[CO_2]$ toward normal. Rather, compensation is a secondary alteration in one of these that counteracts the primary alteration in the other. The result is that the plasma pH is readjusted toward normal. That this is necessarily so is evident from the Henderson–Hasselbalch equation.

$$pH = 6.1 + \log \frac{[HCO_3^-]}{0.03 \cdot P_{CO_2}}$$

If $[HCO_3^-]$ changes, the only way to restore the original $[HCO_3^-]/[CO_2]$ ratio is to change P_{CO_2}, the original ratio can be restored only by altering $[HCO_3^-]$ in the same direction.

The Three States of Compensation Defined

Although some compensatory mechanisms begin to operate rapidly and produce their effects rapidly, others are slower and show stages of compensation. First is the acute stage, before any significant degree of compensation could possibly occur. After the acid–base imbalance has been in effect for a period of time the patient may become **compensated.** This means the compensatory mechanisms have come into play in a normal manner, as expected on the basis of experience with other individuals with an acid–base imbalance of similar type and degree. The "compensated state" does not necessarily imply that the plasma pH is within the normal range. Alternatively, the patient may show no sign of compensation and may be in the **uncompensated** state; this occurs because compensation cannot occur due to some other abnormality. Finally, there is an intermediate state where compensation is occurring but is not yet as complete as it should be. This is the **partially compensated** state. Factors that limit the compensatory processes will be discussed at the end of this section.

Specific Compensatory Processes

Respiratory Acidosis

Let us now follow the course of acute onset of each type of acid–base imbalance and of the compensatory process. Each of these will be schematically illustrated in a pH–bicarbonate diagram. Imagine an individual in normal acid–base balance who goes into acute respiratory acidosis from breathing a gas mixture containing a high level of CO_2. As P_{CO_2} will occur. The abnormal condition has fixed this patient on an abnormally high CO_2 isobar. If the condition is returned to normal, he/she can drop back to the 40-mmHg (5.33-kPa) isobar and all will be well, but until that time all compensatory processes are confined to the higher CO_2 isobar. Compensation, of course, consists of renal excretion of H^+. Since this is a bicarbonate-producing process, $[HCO_3^-]$ should rise, even though it is already above normal. This could have been predicted from the pH–HCO_2^- diagram with no knowledge of the renal mechanism of compensation. Since it is assumed that the individual is fixed on the high CO_2 isobar by the abnormal condition, the only way the pH can possibly be adjusted toward normal is by sliding up the isobar to point *B* in Figure 25.24. This movement is necessarily linked to an increase in $[HCO_3^-]$. Thus the correct analysis of this compensation could be made either from an understanding of the nature of the compensatory mechanism or from an appreciation of the physical chemistry of the bicarbonate buffer system as expressed in the pH–HCO_3^- diagram.

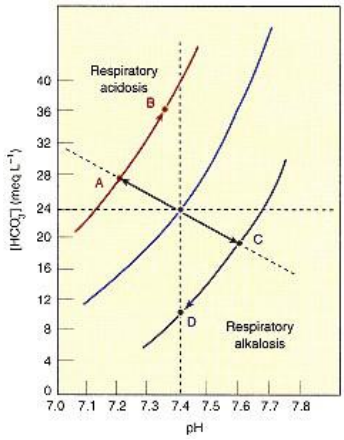


Figure 25.24
pH–Bicarbonate diagram showing compensation for respiratory acidosis (normal state to point B) and for respiratory alkalosis (normal state to point D).

Although the path we have described, up the buffer line to point *A* and then up the isobar to point *B*, is a real possibility, it is also possible that a respiratory acidosis would develop gradually, with compensation occurring simultaneously. The points describing this progress would fall on a curved line from the normal state to point *B*.

Respiratory Alkalosis

In sudden onset respiratory alkalosis P_{CO_2}), and the plasma pH

CLINICAL CORRELATION 25.7

Acute Respiratory Alkalosis

An anesthetized surgical patient with a urethral catheter in place was hyperventilated as an adjunct to the general anesthesia. Prior to hyperventilation normal values of plasma P_{CO_2} was 25 mmHg and the pH was 7.55. Plasma HCO_3^- was not directly measured, but interpolation from a pH–bicarbonate diagram (e.g., Figure 25.17) or calculation from the Henderson–Hasselbalch equation reveals that the plasma $[HCO_3^-]$ decreased to 21.2 meq L^{-1}. Analysis of the urine showed negligible loss of HCO_3^- through the kidneys. It can be concluded that the decrease in $[HCO_3^-]$ was due to titration of bicarbonate by the acid components of the body's buffer systems. The point representing the patient's new steady-state condition clearly must be on the buffering line that represents whole body buffering. (Since the buffers of the whole body are not identical in type or concentration to the blood buffers, the buffer line for the whole body will be analogous, but not identical, to the blood buffer line.)

Magarian, G. J. *Medicine*(*Baltimore*) 61:219, 1982.

CLINICAL CORRELATION 25.8

Chronic Respiratory Acidosis

H.W. was admitted to the hospital with marked dyspnea, cyanosis, and signs of mental confusion. As his acute problems were relieved by appropriate treatment, his symptoms disappeared except for a continuing dyspnea. Blood gas analysis performed eight days later yielded the following data: pH, 7.32; P_{CO_2}, 70 mmHg; [HCO_3^-], 34.9 meq L^{-1}. This is a typical compensation for this degree of chronic respiratory acidosis.

Another patient, C.Q., with chronic obstructive lung disease was found to have arterial plasma pH, 7.40; [HCO_3^-], 35.9 meq L^{-1}; and P_{CO_2} of 60 mmHg, a plasma pH of 7.4 lies outside the 95% probability range. Close questioning of the patient revealed that he had surreptitiously been taking a relative's thiazide diuretic, which superimposed a metabolic alkalosis upon respiratory acidosis.

Rastegar, A., and Thier, S. O. *Chest* 63:355, 1972.

decreases toward normal. This is described in Figure 25.24 by movement along the isobar from point *C* to point *D*. With a gradual onset of respiratory alkalosis, the bicarbonate buffer system would follow points along the curved line from the normal state to point *D*.

Metabolic Acidosis

In metabolic acidosis two mechanisms are usually available for dealing with the excess acid. One is that kidneys increase their H^+ excretion, but this is slow and inadequate to return [HCO_3^-] and pH to normal. The other, which begins to operate almost instantly, is respiratory compensation. Acidosis stimulates the respiratory system to hyperventilate, decreasing the P_{CO_2} but also a further small decrease in [HCO_3^-]. This is due to the same factor that causes the buffer line to have a slope: titration of nonbicarbonate buffers. The inevitability and magnitude of the further decrease in [HCO_3^-] can be seen clearly in the pH–bicarbonate diagram.

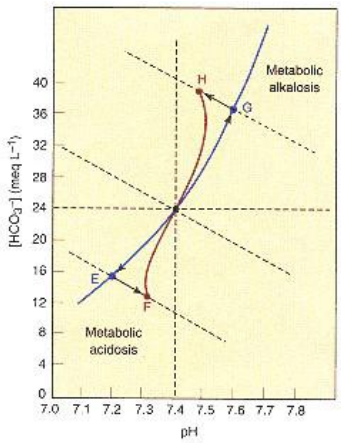


Figure 25.25
pH–Bicarbonate diagram showing compensation for metabolic acidosis (normal state to point F) and for metabolic alkalosis (normal state to point H).

Metabolic Alkalosis

The principles governing compensation for metabolic alkalosis are like those for metabolic acidosis, but operate in the opposite direction. In metabolic alkalosis the primary defect is an increase in plasma [HCO_3^-]; it rises from the normal state to point *G* in Figure 25.25. The immediate physiological response is hypoventilation, followed by increased renal excretion of HCO_3^-. As a result of hypoventilation P_{CO_2} increases along the line from *G* to *H*, and a further small rise in [HCO_3^-] occurs.

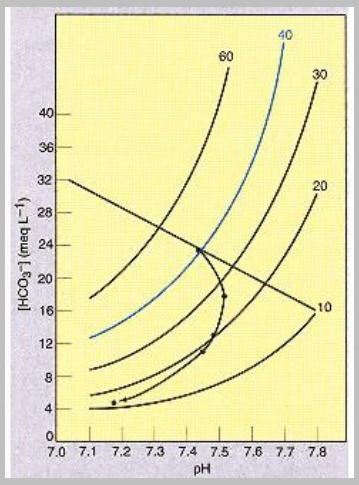
The respiratory response to metabolic acid–base imbalance is rapid, and the bicarbonate buffer system would in most cases be expected to follow points along the curved line from the normal state to the compensated state. An acute metabolic imbalance will not generally be seen outside the experimental laboratory. Indeed, if a physician sees a patient whose plasma pH, [HCO_3^-], and P_{CO_2} would be abnormal.

How complete can compensation be? Can the body totally compensate (bring the pH back to the normal range) for any imbalance? Generally, the answer is no. The compensatory organs, the lungs and kidneys, do not exist exclusively to deal with acid–base imbalance. There is a limit to how much one can hyperventilate; it is simply impossible to move air into and out of the lungs at an indefinitely high rate for an indefinitely long time. Also, one cannot suspend respiration merely to raise P_{CO_2}, rises above 70 mmHg (9.33 kPa) in respiratory acidosis, renal mechanisms for reabsorbing HCO_3^- fail to keep pace, and further increases in plasma [HCO_3^-] are only about what could be expected from titration of nonbicarbonate buffers (see Clin. Corr. 25.8). In respiratory alkalosis renal excretion of excess HCO_3^- can, with time, be sufficient to return plasma pH to within the normal range. Individuals who dwell at high altitude are typically

CLINICAL CORRELATION 25.9

Salicylate Poisoning

Salicylates are the most common cause of poisoning in children. A typical pathway of salicylate intoxication is plotted in the accompanying figure. The first effect of salicylate overdose is stimulation of the respiratory center, resulting in respiratory alkalosis. Renal compensation occurs, lowering the



A typical pathway of salicylate intoxication. Data replotted from Singer, R. B. *Medicine* (Baltimore) 33:1, 1954.

plasma [HCO_3^-]. A second, delayed effect of salicylate may then appear, metabolic acidosis. Since [HCO_3^-] had been lowered by the previous compensatory process, the victim is at a particular disadvantage in dealing with the metabolic acidosis. In addition, but not shown in the graph, respiratory stimulation sometimes persists after the acidosis has run its course. Rational management of salicylate intoxication requires knowledge of the plasma pH and the plasma [HCO_3^-] or its equivalent throughout the course of the condition.

in compensated respiratory alkalosis, with their plasma pH within the normal range. For the other types of acid–base imbalance, the exact degree of compensation expected of a patient with a given clinical picture is well worked out, but a detailed discussion is beyond the scope of this chapter. Suffice it to say that if a patient is compensating, but not as well as expected, this is taken to mean that the patient cannot compensate appropriately and must therefore have a mixed acid–base disturbance.

25.13— Alternative Measures of Acid–Base Imbalance

Modern clinical laboratories generally report plasma bicarbonate concentration, and the value is used by physicians just as we have used it here. Some laboratories, however, report **total plasma CO_2,** that is, the sum of bicarbonate and dissolved CO_2, and this is always slightly higher than [HCO_3^-]. At pH 7.4, for example, the ratio of [HCO_3^-] to [CO_2] is 20 : 1 (dissolved CO_2 is only 1 : 21 of the total CO_2); if [HCO_3^-] is 24 meq L^{-1}, [CO_2] is 1.2 meq L^{-1} and total CO_2 is 25.2 meq L^{-1}. At pH 7.1, HCO_3^- is still 10 times as concentrated as dissolved CO_2. Because the major contributor to total CO_2 is HCO_3^-, total CO_2 is often used in the same manner as bicarbonate to make clinical judgments. Strictly speaking, total CO_2 also includes that in carbamino proteins, but current clinical laboratory practice is to ignore this when making a blood gas and pH report. If it were included in a total CO_2 measurement, it would not change the interpretation of the measurement, since the CO_2 in carbamino proteins, like dissolved CO_2, represents only a small fraction of the total CO_2.

The clinical importance of bicarbonate as a gauge of the whole body's ability to buffer further loads of metabolic acid (see Clin. Corr. 25.9) has led to several ways of expressing what the [HCO_3^-] would be if there were no respiratory component or respiratory compensation involved in a patient's condition. **Base excess** is one of these expressions. It is defined as the amount of acid that would have to be added to blood to titrate it to pH 7.4 at a P_{CO_2}, only the metabolic contribution to acid–base imbalance (primary metabolic imbalance and nonrespiratory compensatory processes) would be measured. If a blood sample were acidic under the conditions of the titration, alkali would have to be added instead of acid, and the base excess would be negative.

The concept and the quantitation of base excess are most easily understood from the pH–bicarbonate diagram. In our discussion of the blood buffer line we saw how increasing the P_{CO_2} in blood, where other buffers are present, would result in a rise in [HCO_3^-] and a virtually identical decrease in the concentration of other buffer bases. This was because equivalent amounts of the other buffer bases were consumed as they buffered carbonic acid. Since virtually all the carbonic acid formed was buffered, for every HCO_3^- formed one conjugate base of some other system was consumed. In this situation the total base in the blood is not measurably changed; only the distribution of HCO_3^- and nonbicarbonate buffer conjugate base is changed. Thus, as long as one remains on the blood buffer line, [HCO_3^-] can change but total base will not. There will be no positive or negative base excess.

If, however, renal activity, diet, or some metabolic process adds or removes HCO_3^-, then a positive or negative base excess will occur. The patient's status will no longer be described by a point on the buffer line, and the base excess will be the difference between the observed plasma [HCO_3^-] and the [HCO_3^-] on the buffer line at the same pH (Figure 25.26). To calculate this difference, the position of the buffer line, which can be determined from knowledge of the slope and the point representing the normal state, must be known. In the

clinical laboratory it can be estimated from hemoglobin concentration and assuming that it is the major nonbicarbonate buffer.

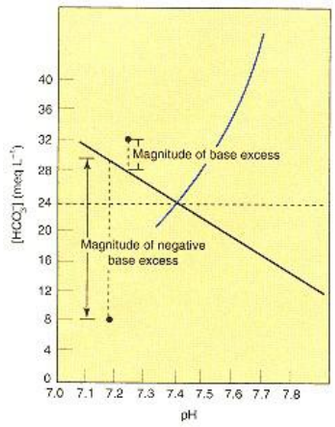


Figure 25.26
Calculation of base excess for a point above the blood buffer line, and calculation of negative base excess for a point below the blood buffer line.
Base excess is 32 – 28 = 4 meq L^{-1}. Negative base excess is 30 – 8 = 22 meq L^{-1}.

The buffer line, then, is the dividing line between positive and negative base excess. Any point above it is in the region of positive base excess, and any point below it is in the region of negative base excess. This gives rise to situations that may seem peculiar at first. In Figure 25.27 the $[HCO_3^-]$ at point *A* is normal, but the patient has a negative base excess. A positive or negative base excess occurs as a result of compensation for a respiratory acid–base imbalance or directly from a metabolic one. Respiratory compensation for a metabolic acid–base imbalance, since it involves movement along a line parallel to the buffer line (Figure 25.25), would cause no further change in the value of the base excess. Clinical Correlation 25.10 involves consideration of base excess.

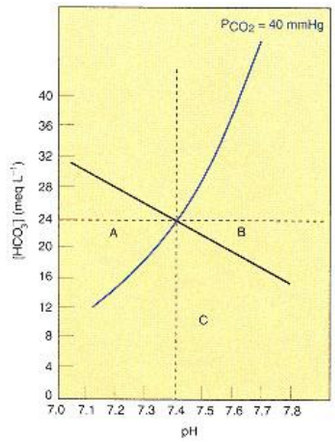


Figure 25.27
Examples showing the sign of the base excess at various points.
At points A and C there is a negative base excess. At point B the base excess is positive.

25.14— The Significance of Na^+ and Cl^- in Acid–Base Imbalance

An important concept in diagnosing certain acid–base disorders is the **anion gap.** Most clinical laboratories routinely measure plasma Na^+, K^+, Cl^-, and HCO_3^-. A glance at the graph in Figure 25.15 confirms that in the plasma of a normal individual the sum of the concentrations of Na^+ and K^+ is greater than the sum of the concentrations of Cl^- and HCO_3^-. This difference is called A, the anion gap; it represents the other plasma anions (Figure 25.15), which are not routinely measured. It is calculated as follows:

$$A = (Na^+ + K^+) - (Cl^- + HCO_3^-)$$

The normal value of A is in the range of 12–16 meq L^{-1}. In some clinical laboratories K^+ is not measured; then the normal value is 8–12 meq L^{-1}. The gap is changed only by conditions that change the sum of the cations or the sum of the anions, or by conditions that change both sums by different amounts. Thus administration or depletion of sodium bicarbonate would not change the anion gap because $[Na^+]$ and $[HCO_3^-]$ would be affected equally. Metabolic acidosis due to HCl or NH_4Cl administration would also leave the anion gap unaffected; here $[HCO_3^-]$ would decrease, but $[Cl^-]$ would increase by an equivalent amount, and the sum of $[HCO_3^-]$ plus $[Cl^-]$ would be unchanged. In contrast, diabetic ketoacidosis or methanol poisoning involves production of organic acids, which react with HCO_3^-, decreasing its concentration. But since the $[HCO_3^-]$ is replaced by some organic anion, the sum of $[HCO_3^-]$ plus $[Cl^-]$ decreases, and the anion gap increases.

The anion gap is most commonly used to establish a differential diagnosis for metabolic acidosis. In a metabolic acidosis with an increased anion gap, H^+ must have arisen in the body with some anion other than chloride. Metabolic acidosis without an increased anion gap must be due either to accumulation of H^+ with chloride or to a decrease in the concentration of sodium bicarbonate. Thus, on the basis of the anion gap, certain diseases can be ruled out, while others would have to be considered. This information can be especially important in dealing with patients who cannot give good histories due to language barriers, unconsciousness, and so on.

Electrolytes of body fluids interact in a multitude of ways. One important way involves the capacity of K^+ and H^+ to substitute for one another under certain circumstances. This can occur in cells, where K^+ is the major cation. In acidosis intracellular $[H^+]$ rises, and it replaces some of the intracellular K^+. The displaced K^+ appears in plasma and is excreted by the kidneys. This leaves the patient with normal plasma $[K^+]$ (normokalemia), but with seriously depleted body K^+ stores (hypokalia). Subsequent excessively rapid correction of the acidosis may then reverse events. As plasma pH rises, K^+ flows back into the cells, and plasma $[K^+]$ may decline to the point where muscular weakness sets in and respiratory insufficiency may become life-threatening.

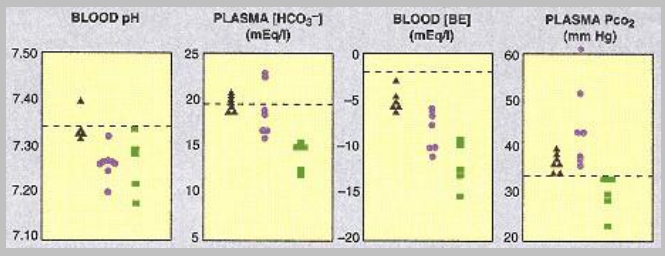
CLINICAL CORRELATION 25.10

Evaluation of Clinical Acid–Base Data

In a 1972 study of total parenteral nutrition of infants, it was found that infants who received amino acids in the form of a hydrolysate of the protein fibrin maintained normal acid–base balance. In contrast, infants receiving two different mixtures of synthetic amino acids, FreAmine and Neoaminosol, became acidotic. Both synthetic mixtures contained adequate amounts of all the essential amino acids, but neither contained aspartate or glutamate. The fibrin hydrolysate contained all of the common amino acids.

The accompanying figure shows the blood acid–base data from these infants. Note that the normal values for infants, given by the dashed lines, are not quite the same as normal values for adults. (A child is *not* a small adult.) The blood pH data show that the infants receiving synthetic mixtures were clearly acidotic. The low $[HCO_3^-]$ of the Neoaminosol group immediately suggests a metabolic acidosis, and the P_{CO_2} values indicate respiratory acidosis, but a simple respiratory acidosis should be associated with a slightly elevated $[HCO_3^-]$. The absence of this finding in most of the infants indicates that the acidosis must also have a metabolic component. This is confirmed by the observation that all the infants receiving FreAmine have a significant negative base excess.

The infants with mixed acid–base disturbances did, in fact, have pneumonia or respiratory distress syndrome. The metabolic acidosis, which all the infants receiving synthetic mixtures experienced, was due to synthesis of aspartic acid and glutamic acid from a neutral starting material (presumably glucose). Subsequent incorporation of these acids into body protein imposed a net acid load on the body. Addition of aspartate and/or glutamate to the synthetic mixtures was proposed as a solution of the problem.



Blood acid–base data of patients receiving fibrin hydrolysate (▲) and of those receiving synthetic L-amino acid mixtures, FreAmine (■).
Values are those observed at the time of the lowest blood base excess.
Dashed lines represent accepted normal values for infants.
Adapted from W. C. Heird, *N. Engl. J. Med.* 287:943, 1972.

In kidneys the reciprocal relationship between K^+ and H^+ results in an association between metabolic alkalosis and hypokalemia. If hypokalemia arises from long-term insufficiency of dietary potassium or long-term diuretic therapy, intracellular K^+ levels diminish, and intracellular $[H^+]$ will increase. This leads to increased acid excretion, acidic urine, and an alkaline arterial plasma pH. We have already seen how in an alkalotic individual a hormonal signal to absorb Na^+ can lead to K^+ loss and then to an exacerbation of the metabolic alkalosis (p. 1045). The opposite also occurs, with alkalosis leading to hypokalemia. In this case increased amounts of $Na^+ + HCO_{3-}$ are presented to the distal convoluted tubules, where all K+ secretion normally takes place (all filtered K+ is reabsorbed; K+ loss is due to distal tubular secretion). The distal tubules take up some Na+ but since HCO_{3-} does not readily follow across that membrane, the increased Na+ uptake is linked to increased K+ secretion. K+ excretion is complicated, being controlled by a variety of hormones and other

CLINICAL CORRELATION 25.11

Metabolic Alkalosis

Prolonged gastric lavage produces a metabolic alkalosis that is a good experimental model of the metabolic alkalosis that results from repeated vomiting. The following table gives plasma and urine acid–base and electrolyte data from a healthy volunteer on a low-sodium diet who, after a control period, was subjected to gastric lavage for two days. After a five day recovery period, he was placed on a low-potassium diet and given a sodium (130 meq day_{-1}) and chloride (121 meq day_{-1}) supplement. During the control period the data are within normal limits. After gastric lavage that selectively removed HCl (Na_+, K_+, and H_2O lost with the gastric juice were restored), an uncomplicated metabolic alkalosis developed. Note that the subject excreted an alkaline urine, containing a substantial amount of HCO_{3-}]. The Na_+ excretion increased, depleting the body's Na_+ stores. Plasma P_{CO_2} was not measured, but plotting the values of pH and [HCO_{3-}] on a pH-bicarbonate diagram (e.g., Figure 25.18) allows one to interpolate a value of about 47 mmHg. Clearly, respiratory compensation was occurring. Plasma [K_+] was decreased. Plasma [Cl_-] decreased, but no more than would be expected on the basis of the changes in [Na_+], [K_+], and [HCO_{3-}].

When the subject was placed on a low-potassium diet the alkalosis grew worse, and plasma [HCO_{3-}] rose. Additional compensatory hypoventilation evidently prevented a further rise in plasma pH. Note, though, that the urine became acidic, in spite of the increased severity of the alkalosis. The Na_+ was conserved, not in exchange for K_+, but in exchange for H_+. After several days of Na_+ and Cl_- administration, however, the subject was able to restore the depleted Cl_-, excrete the excess HCO_{3-}, and repair the acid–base imbalance with no other treatment.

	Control	*After Lavage*	*Low KCl*	*After NaCl*
Plasma				
pH	7.4	7.50	7.48	7.41
HCO_3^-	29.3	35.3	38.1	26.1
Na^+ (meq L^{-1})	138	134	141	144
K^+ (meq L^{-1})	4.2	3.2	2.9	3.2
Cl^- (meq L^{-1})	101	88	85	108
Urine				
pH	6.12	7.48	5.70	7.19
HCO_3^- (meq/day^{-1})	3	51	1	17
NH_4^+ (meq/day^{-1})	22	4	36	14
Titratable acidity (meq/day^{-1})	10	0	14	1
Total acidity (meq/day^{-1})	29	–49	49	–2
Na^+ (meq/day^{-1})	2	28	1	95

Source: Data from Kassirer, J. P., and Schwartz, W. B., *Am. J. Med*. 40:10, 1966.

factors. The end result, however, is that metabolic alkalosis and hypokalemia go hand in hand, so that the term "hypokalemic alkalosis" is often used synonymously with metabolic alkalosis. Clinical Correlation 25.11 discusses a case of experimental metabolic alkalosis in which this occurred.

Bibliography

Gas Transport

Bunn, H. F., and Forget, B. G. *Hemoglobin: Molecular, Genetic, and Clinical Aspects*. Philadelphia: Saunders, 1986.

Bunn, H. F., Gabbay, K. H., and Gallop, P. M. The glycosylation of hemoglobin: relevance to diabetes mellitus. *Science* 200:21, 1978.

Kilmartin, J. V. Interaction of haemoglobin with protons, CO_2 and 2,3-diphosphoglycerate. *Br. Med. Bull*. 32:209, 1976.

Perutz, M. F., and Lehmann, H. Molecular pathology of human haemoglobin. *Nature* 219:902, 1968.

Steffes, M. W., and Mauer, S. M. Toward a basic understanding of diabetic complications. *N. Engl. J. Med*. 325:883, 1991.

pH Regulation

Davenport, H. W. *The ABC of Acid–Base Chemistry*, 6th ed. Chicago: University of Chicago Press, 1974.

Gabow, P. A., Kaehny, W. D., Fennessey, P. V., et al. Diagnostic importance of an increased serum anion gap. *N. Engl. J. Med*. 303:854, 1980.

Gamble, J. L. Jr., and Bettice, J. A. Acid-base relationships in the different body compartments: the basis for a simplified diagnostic approach. *Johns Hopkins Med. J*. 140:213, 1977.

Masoro, E.J., and Siegel, P. D. *Acid–Base Regulation: Its Physiology, Patho-physiology and the Interpretation of Blood–Gas Analysis*, 2nd ed. Philadelphia: Saunders, 1977.

Siggaard-Andersen, O. *The Acid–Base Status of the Blood*, 4th ed. Baltimore: Williams & Wilkins, 1974.

Questions

J. Baggott and C. N. Angstadt

1. During a breathing cycle:

   A. the alveolar gases are completely exchanged for atmospheric gases.

   B. gas exchange between the alveoli and the capillary blood can occur at all times.

   C. gas exchange with the capillary blood occurs at the surface of all the airways.

   D. there is net uptake of nitrogen by the blood.

   E. atmospheric water vapor is taken up by the lungs.

2. From an oxygen saturation curve for normal blood we can determine that:

A. P_{50} is in the P_{CO_2} range found in extrapulmonary tissues.

B. oxygen binding is hyperbolic.

C. an oxygen carrier is necessary.

D. tighter oxygen binding occurs at lower P_{CO_2}.

E. shifts of the curve to the left or right would have little effect on oxygen delivery.

Refer to the following for Questions 3–5.

A. hemoglobin α-chains

B. hemoglobin β-chains

C. hemoglobin γ-chains

D. hemoglobin -chains

E. hemoglobin -chains

3. Found in HbA, HbA_2, and HbF.

4. Modified in HbA_{ic}.

5. Form the major binding sites for 2,3-bisphosphoglycerate (BPG).

6. At a P_{CO_2} of 30 mmHg hemoglobin's percent saturation will:

A. increase with increasing temperature.

B. increase with decreasing pH.

C. increase with increasing P_{CO_2}

D. increase with increasing 2,3-bisphosphoglycerate concentration.

E. none of the above.

7. Significant contributors to the total carbon dioxide of whole blood include all of the following EXCEPT:

A. bicarbonate ion.

B. dissolved carbon dioxide (CO_2).

C. carbaminohemoglobin.

D. carbonic acid (H_2CO_3).

8. 2,3-Bisphosphoglycerate (BPG):

A. is absent from the normal erythrocyte.

B. is a homotropic effector for hemoglobin.

C. binds more tightly to HbF than to HbA.

D. synthesis increases when hemoglobin's T R equilibrium is shifted in favor of the T state.

E. synthesis decreases when the erythrocyte pH rises.

9. Which of the following buffer systems is far less effective in controlling changes in physiological pH due to CO_2 than changes due to metabolic acids, like acetoacetic acid?

A. bicarbonate

B. inorganic phosphate

C. organic phosphate esters

D. intracellular protein

E. extracellular protein

10. The slope of the blood buffer line is most sensitive to pathological changes in the blood concentration of:

A. plasma bicarbonate.

B. plasma phosphate.

C. hemoglobin.

D. plasma proteins.

E. organic phosphates of the erythrocyte.

11. As P_{CO_2} is increased in a normal individual:

A. the plasma [CO_2] remains unchanged.

B. plasma bicarbonate increases.

C. the slope of the blood buffer line changes.

D. the base excess increases.

E. the base excess decreases.

12. All of the following produce H^+ EXCEPT:

A. formation of bicarbonate ion from CO_2 and water.

B. formation of carbaminohemoglobin from CO_2 and hemoglobin.

C. binding of oxygen by hemoglobin.

D. oxidation of sulfur-containing amino acids.

E. metabolism of sodium lactate.

13. A substantial fraction of the urinary titratable acidity of a normal individual consists of:

A. H_2CO_3.

B. NH_4^+.

C. acetoacetic acid.

D. H_2PO_{4-}.

E. HCO_3^-.

14. In a patient with diabetic ketoacidosis of long duration:

A. the major urinary acid is H_2PO_{4-}.

B. hemoglobin's oxygen dissociation curve would be shifted to the right.

C. the distribution of hemoglobin species would be the same as in a normal individual.

D. 1 mol of bicarbonate is regenerated for every mole of H_2PO_{4-} formed in the renal tubule.

E. hypoventilation would be expected.

15. The following laboratory data are obtained from a patient: P_{CO_2} = 60 mmHg, HCO_3^- = 27 meq L^{-1}, pH = 7.28. These values define a point on the patient's blood buffer line. We conclude:

A. The patient has an acute condition.

B. The condition would lead to production of an alkaline urine.

C. Of the blood buffers, the bicarbonate buffer system is the most important in resisting this pH change.

D. Increasing the alveolar P_{CO_2} could restore the plasma to normal.

E. Hyperventilation due to anxiety could cause this.

16. During compensation for a metabolic acid-base imbalance, which of the following would become increasingly abnormal?

A. plasma pH

B. blood P_{CO_2}

C. base excess

D. total hemoglobin

E. none of the above

17. In respiratory alkalosis:

A. the acute state is associated with an abnormally low plasma [HCO_3^-].

B. the mechanism of compensation causes an increase in the plasma HCO_3^-.

C. the plasma pH never returns to the normal range in the fully compensated state.

D. in the partially compensated state, there will be a negative base excess equal to the difference between 24 meq L^{-1} and the actual plasma HCO_{3-}.

E. compensation involves changing P_{CO_2}

18. Hypokalemia can be expected to:

A. occur if the plasma pH is rapidly raised.

B. lead to increased urine acidity.

C. be associated with a high plasma [HCO_3^-].

D. decrease the value of the anion gap slightly.

E. all of the above.

Answers

1. B A and C: The alveoli, where gas exchange with the blood occurs, are of constant size and exchange gases with the airways by diffusion. D and E: Water vapor and CO_2 are added to the alveolar gases by the lung tissue, diluting the nitrogen (p. 1027).

2. A P_{CO_2} (p. 1028, Figure 25.2). C: If O_2 were soluble enough in plasma, no carrier would be necessary (p. 1026). E: Shifts profoundly affect delivery (p. 1028).

3. A It is the non-α-chain that differs among these (p. 1030).

4. B The β-chains are nonenzymatically glycosylated in HbA_{1c} (p. 1030).

5. B BPG binds between the N terminals of the β-chains (p. 1029).

6. E All effects are opposite to those proposed in the question. A-C: High temperature, low pH (and therefore high P_{CO_2} favor dissociation; that is, decreased saturation (p. 1031). D: High BPG has the same effect (p. 1029).

7. D Carbonic acid is present in very small amounts; the equilibrium strongly favors CO_2 and H_2O (p. 1033, Table 25.2; see also p. 1039).

8. D A and B: BPG is a normal component of the red cell, where it serves as a heterotropic effector of HbA (p. 1029, Figure 25.4). C: It binds weakly or not at all to the HbF (p. 1030). D and E: BPG binds to the T state, relieving product inhibition of BPG synthesis; BPG synthesis is inhibited by low pH (p. 1036).

9. A The bicarbonate system is a major extracellular buffer; with a pK of 6.1 it is ineffective toward CO_2. The other buffers (phosphates and protein) are, effective (p. 1037, Table 25.6). All of these buffers, however, are effective against noncarbonic acids. The bicarbonate buffer system is included here because the response of the respiratory system to low pH, exhaling CO_2, compensates for the innate ineffectiveness of this system at a pH fairly distant from its pK (p. 1039).

10. C The slope of the blood buffer line is determined by the concentration of the nonbicarbonate buffers. Of these, hemoglobin is quantitatively the most important and is susceptible to change (i.e., anemias from any cause) (p. 1040).

11. B This is because the resulting H_2CO_3 is buffered by various nonbicarbonate buffers, producing HCO_{3-} (p. 1040).

12. E Metabolism of sodium lactate produces sodium bicarbonate and is used clinically to control acidosis (p. 1042). A and B are reactions whose products include H^+ (pp. 1032–1033). C is the Bohr effect (p. 1031). D is a major source of acid in the typical American diet (p. 1042).

13. D A: The level of H_2CO_3 is very low. B: NH_4^+ is an important urinary acid, but its pK is too high to be titrated at pH 7.4, the endpoint. C: Acetoacetic acid would appear only in some kinds of severe acidosis. E: HCO_3^- is physiologically a base; its dissociable H^+ has a pK that is far above the physiological range (pp. 1043 and 1045).

14. D See pp. 1043–1045. A: After adaptation to acidosis NH_4^+ excretion rises enormously, becoming the major urinary acid (p. 1045). B: True only in acidosis of short duration; decreasing BPG in prolonged acidosis tends to restore the normal position (p. 1036). C: Large amounts of HbA_{1c} would be expected (p. 1030). E: Hyperventilation, to expel CO_2, would be expected (p. 1048).

15. A High P_{CO_2}, low pH point on the blood buffer line define an acute respiratory acidosis. Buffering by nonbicarbonate buffer systems and excretion of acid in the urine would be the physiological responses (p. 1047).

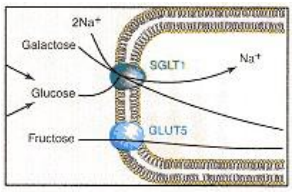
16. B P_{CO_2} would decrease during compensation for acidosis or rise during compensation for alkalosis. A: Plasma pH would be restored. C: Base excess would be unchanged. D: Hemoglobin would participate in buffering, but its total concentration would not be expected to change (pp. 1048–1050, Figures 25.25 and 25.26).

17. A A and B: See p. 1047, Figure 25.24. C: This is the only acid-base abnormality in which compensation is expected to restore the plasma pH to 7.4 (p. 1047). D: There is a negative base excess equivalent to the difference between the patient's [HCO_3^-] and the [HCO_3^-] of the point on the blood buffer line at the same pH, a point that will be less than 24 meq L^{-1} (p. 1050, Figure 25.26). E: This would either be a cure or an exacerbation, depending on the direction of the change; it would not be compensation.

18. E A, B, and C: See p. 1050. D: Decreasing K^+ would lower the anion gap by a small amount (p. 1050).

Chapter 26—
Digestion and Absorption of Basic Nutritional Constituents

Ulrich Hopfer



26.1— Overview

Secretion of digestive fluids and digestion of food were some of the earliest biochemical events to be investigated at the beginning of the era of modern science. Major milestones were the discovery of hydrochloric acid secretion by the stomach and enzymatic hydrolysis of protein and starch by gastric juice and saliva, respectively. The discovery of gastric HCl production goes back to the American physician William Beaumont (1785–1853). In 1822 he treated a patient with a stomach wound. The patient recovered from the wound, but retained a gastric fistula (abnormal opening through the skin). Beaumont seized the opportunity to obtain and study gastric juice at different times during and after meals. Chemical analysis revealed, to the surprise of chemists and biologists, the presence of the inorganic acid HCl. This discovery established the principle of unique secretions into the gastrointestinal tract, which are elaborated by specialized glands.

Soon thereafter, the principle of enzymatic breakdown of food was recognized. In 1836 Theodor Schwann, a German anatomist and physiologist (1810–1882), noticed that gastric juice degraded albumin in the presence of dilute acid. He recognized that a new principle was involved and coined the word pepsin from the Greek *pepsis*, meaning digestion. Today the process of secretion of digestive fluids, digestion of food, and absorption of nutrients and of electrolytes can be described in considerable detail.

The basic nutrients fall into the classes of proteins, carbohydrates, and fats. Many different types of food can satisfy the nutritional needs of humans, even though they differ in the ratios of proteins to carbohydrates and to fats and in the ratio of digestible to nondigestible materials. Unprocessed plant products are especially rich in **fibrous material** that can be neither digested by human enzymes nor easily degraded by intestinal bacteria. The fibers are mostly carbohydrates, such as **cellulose** (β-1,4-glucan) or **pectins** (mixtures of methyl esters of polygalacturonic acid, polygalactose, and polyarabinose). High-fiber diets enjoy a certain popularity nowadays because of a postulated preventive effect on development of colonic cancer.

26.1 describes average contributions of different food classes to the diet of North Americans. The intake of individuals may substantially deviate from the average, as food consumption depends mainly on availability and individual tastes. The ability to utilize a wide variety of food is possible because of the great adaptability and digestive reserve capacity of the gastrointestinal tract.

Knowledge of the nature of proteins and carbohydrates in the diet is important from a clinical point of view. Certain proteins and carbohydrates, although good nutrients for most humans, cannot be properly digested by some individuals and produce gastrointestinal ailments. Omission of the offending

TABLE 26.1 Contribution of Major Food Groups to Daily Nutrient Supplies in the United States

Type of Nutrient	*Total Daily Consumption (g)*	*Dairy Products, Except Butter (%)*	*Meat, Poultry, Fish (%)*	*Eggs (%)*	*Fruits, Nuts, Vegetables (%)*	*Flour, Cereal (%)*	*Sugar Sweeteners (%)*	*Fats, Oils (%)*
Protein	100	22	42	6	12	18	0	0
Carbohydrate	381	7	0.1	0.1	19	36	37	0
Fat	155	13	35	3	4	1	0	42

material and change to another diet can eliminate these gastrointestinal problems. Examples of food constituents that can be the cause of gastrointestinal disorders are **gluten,** one of the protein fractions of wheat, and **lactose,** the disaccharide in milk.

Gastrointestinal Organs Have Multiple Functions in Digestion

The bulk of ingested nutrients consists of large **polymers** that have to be broken down to **monomers** before they can be absorbed and made available to all cells of the body. The complete process from food intake to absorption of nutrients into the blood consists of a complicated sequence of events, which at the minimum includes (Figure 26.1):

1. Mechanical homogenization of food and mixing of ingested solids with fluids secreted by the glands of the gastrointestinal tract.

2. Secretion of digestive enzymes that hydrolyze macromolecules to oligomers, dimers, or monomers.

3. Secretion of electrolytes, acid, or base to provide an appropriate environment for optimal enzymatic digestion.

4. Secretion of bile acids as detergents to solubilize lipids and facilitate their absorption.

5. Hydrolysis of nutrient oligomers and dimers by enzymes on the intestinal surface.

6. Transport of nutrient molecules and of electrolytes from the intestinal lumen across the epithelial cells into blood or lymph.

To accomplish these functions, the gastrointestinal tract contains specialized glands and surface epithelia:

Organ	Major Function in Digestion and Absorption
Salivary glands	Elaboration of fluid and digestive enzymes
Stomach	Elaboration of HCl and proteases
Pancreas	Elaboration of $NaHCO_3$ and enzymes for intraluminal digestion
Liver	Elaboration of bile acids
Gallbladder	Storage and concentration of bile
Small intestine	Terminal digestion of food, absorption of nutrients and electrolytes
Large intestine	Absorption of electrolytes

The **pancreas** and **small intestine** are essential for digestion and absorption of all basic nutrients. Fortunately, both organs have large reserve capacities. For example, maldigestion due to pancreatic failure becomes a problem only when the pancreatic secretion rate of digestive enzymes drops below one-tenth of the normal rate. The secretion of the liver **(bile)** is important for efficient lipid absorption, which depends on the presence of bile acids. In contrast, gastric digestion of food is nonessential for adequate nutrition, and loss of this function can be compensated for by the pancreas and the small intestine. Yet normal gastric digestion greatly increases the smoothness and efficiency of the total digestive process. The stomach aids in the digestion through its reservoir function, its churning ability, and initiation of protein hydrolysis, which, although small, is important for stimulation of pancreatic and gallbladder output. Peptides and amino acids liberated in the stomach stimulate the coordinated release of pancreatic juice and bile into the lumen of the small intestine, thereby ensuring efficient digestion of food.

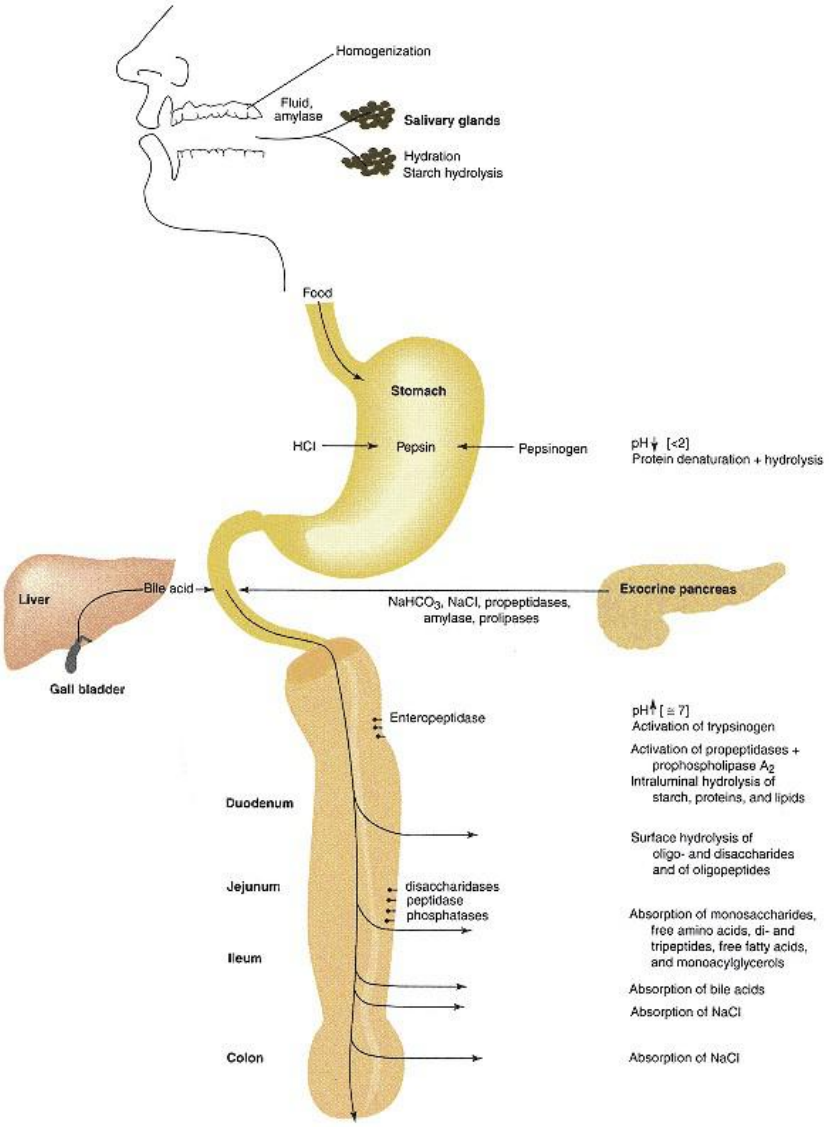


Figure 26.1
Gastrointestinal organs and their functions.

26.2— Digestion: General Considerations

Pancreas Supplies Enzymes for Intestinal Digestion

Most of the breakdown of food is catalyzed by soluble enzymes and occurs within the lumen of the stomach or small intestine. The **pancreas,** not the stomach, is the major organ that synthesizes and secretes the large amounts of enzymes needed for digestion. Secreted enzymes amount to at least 30 g of protein per day in a healthy adult. The **pancreatic enzymes** together with bile are poured into the lumen of the second (descending) part of the duodenum, so that the bulk of the **intraluminal digestion** occurs distal to this site in the small intestine. However, pancreatic enzymes cannot completely digest all nutrients to forms that can be absorbed. Even after exhaustive contact with pancreatic enzymes, a substantial portion of carbohydrates and amino acids are present as dimers and oligomers that depend for final digestion on enzymes present on the luminal surface or within the chief epithelial cells that line the lumen of the small intestine **(enterocytes).**

The luminal plasma membrane of enterocytes is enlarged by a regular array of projections, termed microvilli, which give it the appearance of a brush and have led to the name **brush border** for the luminal pole of enterocytes. This membrane contains many **di- and oligosaccharidases, amino- and dipeptidases,** as well as **esterases** (Table 26.2). Many of these enzymes protrude up to 100 Å into the intestinal lumen, attached to the plasma membrane by an anchoring polypeptide that itself has no role in the hydrolytic activity. The substrates for these enzymes are the oligomers and dimers that result from pancreatic digestion. The surface enzymes are glycoproteins that are relatively stable against digestion by pancreatic proteases or the effects of detergents.

A third site of digestion is the cytoplasm of enterocytes. **Intracellular digestion** is of some importance for the hydrolysis of di- and tripeptides, which can be absorbed across the luminal plasma membrane.

Digestive Enzymes Are Secreted as Proenzymes

Salivary glands, gastric mucosa, and pancreas contain specialized cells that synthesize and store digestive enzymes until the enzymes are needed during

TABLE 26.2 Digestive Enzymes of the Small Intestinal Surface

Enzyme (Common Name)	*Substrate*
Maltase	Maltose
Sucrase/isomaltase	Sucrose/α-limit dextrin
Glucoamylase	Amylose
Trehalase	Trehalose
β-Glucosidase	Glucosylceramide
Lactase	Lactose
Endopeptidase 24.11	Protein (cleavage at internal hydrophobic amino acids)
Aminopeptidase A	Oligopeptide with acidic NH_2 terminus
Aminopeptidase N	Oligopeptide with neutral NH_2 terminus
Dipeptidyl aminopeptidase IV	Oligopeptide with X-Pro or X-Ala at NH_2 terminus
Leucine aminopeptidase	Peptides with neutral amino acid at NH_2 terminus
γ-Glutamyltransferase	Glutathione + amino acid
Enteropeptidase (enterokinase)	Trypsinogen
Alkaline phosphatase	Organic phosphates

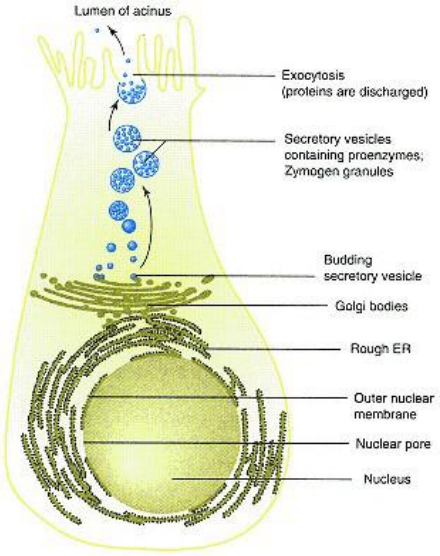


Figure 26.2
Exocrine secretion of digestive enzymes.
Redrawn with permission from Jamieson, J. D. Membrane and secretion. In: G. Weissmann and R. Claiborne (Eds.), *Cell Membranes: Biochemistry, Cell Biology and Pathology*. New York: HP Publishing Co., 1975. Figure by B. Tagawa.

a meal. The enzymes are then released into the lumen of the gastrointestinal tract (Figure 26.2). This secretion is termed **exocrine** because of its direction toward the lumen. Proteins destined for secretion are synthesized on the polysomes of the rough endoplasmic reticulum (see p. 739 for synthesis and glycosylation of membrane and secreted proteins) and transported via the Golgi complex to storage vesicles in the apical cytoplasm. The storage vesicles **(zymogen granules)** have a diameter of about 1 m. Most digestive enzymes are produced and stored as inactive **proenzymes (zymogens)** (see p. 101). The zymogen granules are bounded by a typical cellular membrane. When an appropriate stimulus for secretion is received by the cell, the granules move closer to the luminal plasma membrane, where their membranes fuse with the plasma membrane and release the contents into the lumen **(exocytosis).** Activation of proenzymes occurs only after they are released from the cells.

Regulation of Secretion Occurs through Secretagogues

The processes involved in the secretion of enzymes and electrolytes are regulated and coordinated. Elaboration of electrolytes and fluids simultaneously with that of enzymes is required to flush any discharged digestive enzymes out of the gland into the gastrointestinal lumen. The physiological regulation of secretion occurs through **secretagogues** that interact with receptors on the surface of the **exocrine cells** (Table 26.3). Neurotransmitters, hormones, pharmacological agents, and certain bacterial toxins can be secretagogues. Different exocrine cells, for example, in different glands, usually possess different sets of receptors. Binding of the secretagogues to receptors sets off a chain of signaling events that ends with fusion of zymogen granules with the plasma membrane. Two major signaling pathways have been identified (Figure 26.3): (1) activation of phosphatidylinositol-specific **phospholipase C** with liberation of **inositol 1,4,5-triphosphate** and **diacylglycerol** (see p. 862); in turn, triggering Ca^{2+} release into the cytosol and activation of protein kinase C, respec-

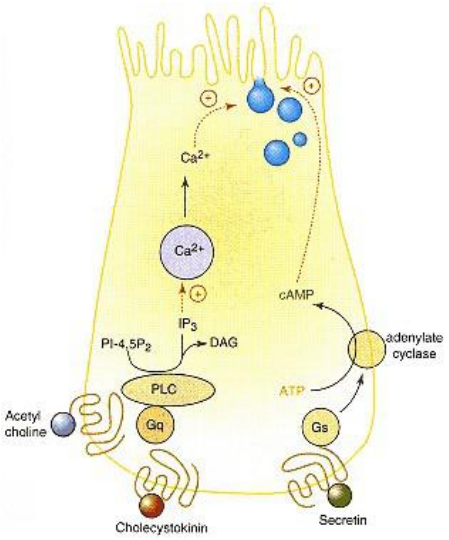


Figure 26.3
Cellular regulation of exocrine secretion in the pancreas.
Abbreviations: PI-4,5P_2, phosphatidylinositol-4,5-bisphosphate; DG, diacylglycerol; IP_3, inositol-1,4,5-triphosphate; PLC, phospolipase C. Adapted from Gardner, J. D. *Annu. Rev. Physiol.* 41:63, 1979.

tively; and (2) activation of **adenylate** or **guanylate cyclase,** resulting in elevated cAMP or cGMP levels, respectively (see p. 859). Secretion can be stimulated through either pathway.

Acetylcholine (Figure 26.4) elicits salivary, gastric, and pancreatic enzyme and electrolyte secretion. It is the major neurotransmitter for stimulating secretion, with input from the central nervous system in salivary and gastric glands, or via local reflexes in gastric glands and the pancreas. The acetylcholine receptor of exocrine cells is of the muscarinic type; that is, it can be blocked by atropine (Figure 26.5). Most people have experienced the effect of atropine because it is used by dentists to "dry up" the mouth for dental work.

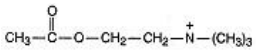


Figure 26.4
Acetylcholine.

Another class of secretagogues are the **biogenic amines,** consisting of **histamine** and **5-hydroxytryptamine. Histamine** (Figure 26.6) is a potent stimulator of HCl secretion. It interacts with a gastric-specific histamine receptor, also referred to as the H_2 receptor, on the contraluminal plasma membrane of parietal cells. Histamine is normally secreted by specialized regulatory cells in the stomach wall (**enterochromaffin-like cells,** ECC). Histamine analogs that are antagonists at the H_2 receptor are used medically to decrease HCl output during treatment for peptic ulcers. **5-Hydroxytryptamine (serotonin)** is pres-

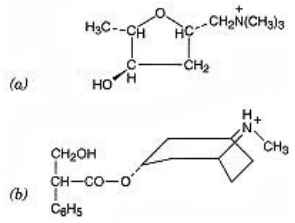


Figure 26.5
(a) L(+)-Muscarine and (b) atropine.

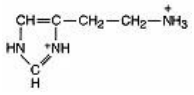


Figure 26.6
Histamine.

TABLE 26.3 Physiological Secretagogues

Organ	*Secretion*	*Secretagogue*
Salivary gland	NaCl, amylase	Acetylcholine, (catecholamines?)
Stomach	HCl, pepsinogen	Acetylcholine, histamine, gastrin
Pancreas—acini	NaCl, digestive enzymes	Acetylcholine, cholecystokinin (secretin)
Pancreas—duct	$NaHCO_3$, NaCl	Secretin
Small intestine	NaCl	Acetylcholine, serotonin, vasoactive intestinal peptide (VIP), guanylin

ent in relatively high amounts in the gastrointestinal tract (Figure 26.7). It stimulates secretion of NaCl by the small intestinal mucosa.

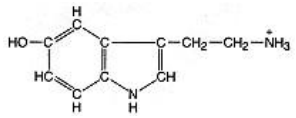


Figure 26.7
5-OH-Tryptamine (serotonin).

A third class of secretagogues consists of peptide-neurotransmitters and -hormones (Table 26.4). The intestinal nerve cells are rich in peptide-neurotransmitters that stimulate NaCl secretion. **Vasoactive intestinal peptide (VIP)** is a particularly potent one in this respect in the intestines and pancreas. Furthermore, the gastrointestinal tract contains many specialized epithelial cells that produce biologically active amines and peptides. The peptides are localized in granules, usually close to the contraluminal pole of these cells, and are released into the blood. Hence these cells are classified as **epithelial endocrine cells.** Of particular importance are the peptides gastrin, cholecystokinin (pancreozymin), and secretin. In contrast, a recently identified peptide—namely, **guanylin**—is released into the lumen and stimulates NaCl secretion by binding to a brush border receptor that activates guanylate cyclase and thus elevates cGMP levels.

Gastrin occurs as either a peptide of 34 amino acids (G-34) or one of 17 residues (G-17) from the COOH terminus of G-34. The functional portion of gastrin resides mainly in the last five amino acids of the COOH terminus. Thus pentagastrin, an artificial pentapeptide containing only the last five amino acids, can be used specifically to stimulate gastric HCl and pepsin secretion. Gastrin as well as cholecystokinin have an interesting chemical feature, a **sulfated tyrosine,** which considerably enhances the potency of both hormones.

Cholecystokinin and **pancreozymin** denote the same peptide. The different names allude to the different functions elicited by the peptide and had been coined before purification. The peptide stimulates gallbladder contraction (cholecystokinin) as well as secretion of pancreatic enzymes (pancreozymin). It is secreted by epithelial endocrine cells of the small intestine, particularly in the duodenum, and this secretion is stimulated by luminal amino acids and peptides, usually derived from gastric proteolysis, by fatty acids, and by an acid pH. Cholecystokinin and gastrin are thought to be related in an evolutionary sense, as both share an identical amino acid sequence at the COOH terminus.

TABLE 26.4 Structure of Human Intestinal Peptide Hormones

Vasoactive intestinal peptide (VIP)
His-Ser-Asp-Ala-Val-Phe-Thr-Asp-Asn-Tyr-Thr-Arg-Leu-Arg-Lys-Gln-Met-Ala-Val-Lys
Asn-Leu-Ile-Ser-Asn-Leu-Tyr-Lys

Secretin
His-Ser-Asp-Gly-Thr-Phe-Thr-Ser-Glu-Leu-Ser-Arg-Leu-Arg-Glu-Gly-Ala-Arg-Leu-Gln
[c]NH_2-Val-Leu-Gly-Gln-Leu-Leu-Arg

Guanylin
Pro-Gly-Thr-Cys-Glu-Ile-Cys-Ala-Tyr-Ala-Ala-Cys-Thr-Gly-Cys

Gastrin G-34-II[a] G-17-II
[b]Glp-Leu-Gly-Pro-Gln-Gly-Pro-Pro-His-Leu-Val-Ala-Asp-Pro-Ser-Lys-Lys-Gln
[c]NH_2-Phe-Asp-Met-Trp-Gly-Tyr(SO_3H)-Ala-$(Glu)_5$-Leu-Trp-Pro-Gly

Cholecystokinin
Lys-Ala-Pro-Ser-Gly-Arg-Met-Ser-Ile-Val-Lys-Asn-Leu-Gln-Asn-Leu-Asp-Pro
[c]NH_2-Phe-Asp-Met-Trp-Gly-Met-Tyr(SO_3H)-Asp-Arg-Asp-Ser-Ile-Arg-His-Ser

Source: Yanaihara, C. In: B. B. Rauner, G. M. Makhlouf, and S. G. Schultz (Eds.), *Handbook of Physiology. Section 6: Alimentary Canal. Vol. II: Neural and Endocrine Biology*. Bethesda, MD: American Physiological Society, 1989, pp. 45–62.

[a] Gastrin I is not sulfated.
[b] Glp = pyrrolidino carboxylic acid, derived from Glu through internal amide formation.
[c] NH_2 = amide of carboxy-terminal amino acid.

Secretin is a polypeptide of 27 amino acids. This peptide is secreted by yet other endocrine cells of the small intestine. Its secretion is stimulated particularly by luminal pH less than 5. The major biological activity of secretin is stimulation of secretion of pancreatic juice rich in $NaHCO_3$. Pancreatic $NaHCO_3$ is essential for neutralization of gastric HCl in the duodenum. Secretin also enhances pancreatic enzyme release, acting synergistically with cholecystokinin.

26.3— Epithelial Transport

Solute Transport May Be Transcellular or Paracellular

Solute movement across an epithelial cell layer is determined by the properties of epithelial cells, particularly their plasma membranes, and by the intercellular tight junctional complexes (Figure 26.8). The **tight junctions** extend in a belt-like manner around the perimeter of each epithelial cell and connect neighboring cells. Therefore the tight junctions constitute part of the barrier between the two extracellular spaces on either side of the epithelium, that is, the lumen of the gastrointestinal tract and the intercellular (interstitial) space on the other (blood or serosal) side. The tight junction marks the boundary between the luminal and contraluminal region of the plasma membrane of epithelial cells.

Two potentially parallel pathways for **solute transport** across epithelial cell layers can be distinguished: through the cells **(transcellular)** and through the tight junctions between cells **(paracellular)** (Figure 26.8). The transcellular route in turn consists mainly of two barriers in series, formed by the luminal and contraluminal plasma membranes. Because of this combination of different barriers in parallel (cellular and paracellular pathways) and in series (luminal and contraluminal plasma membranes), biochemical and biophysical information on all three barriers as well as their mutual influence is required for understanding the overall transport properties of the epithelium.

A major function of gastrointestinal epithelial cells is **active transport** of nutrients, electrolytes, and vitamins. The cellular basis for this vectorial solute movement lies in the different properties of the luminal and contraluminal regions of the plasma membrane. The small intestinal cells provide a prominent example of the differentiation and specialization of the two types of membrane. The luminal and contraluminal plasma membranes differ in morphological appearance, enzymatic composition, chemical composition, and transport functions (Table 26.5). The luminal membrane is in contact with the nutrients in the chyme (the semifluid mass of partially digested food) and is specialized for terminal digestion of nutrients through its digestive enzymes and for nutrient absorption through transport systems that accomplish concentrative uptake. Transport systems are present for monosaccharides, amino acids, peptides, and electrolytes. In contrast, the contraluminal plasma membrane, which is in contact with the intercellular fluid, capillaries, and lymph, has properties similar to the

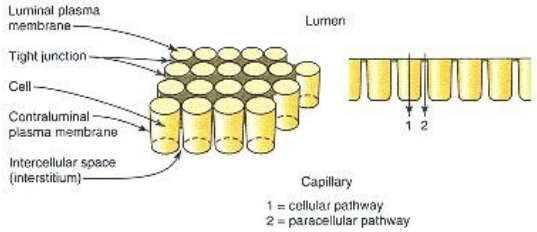


Figure 26.8
Pathways for transport across epithelia.

TABLE 26.5 Characteristic Differences Between Luminal and Contraluminal Plasma Membrane of Small Intestinal Epithelial Cells

Parameter	*Luminal*	*Contraluminal*
Morphological appearance	Microvilli in ordered arrangement (brush border)	Few microvilli
Enzymes	Di- and oligosaccharidases	Na^+,K^+–ATPase
	Aminopeptidase	Adenylyl cyclase
	Dipeptidases	
	γ-Glutamyltransferase	
	Alkaline phosphatase	
	Guanylate cyclase	
Transport systems	Na^+–monosaccharide cotransport (SGLT1)	Facilitated monosac- charide transport (GLUT2)
	Facilitated fructose transport (GLUT5)	Facilitated neutral amino acid transport
	Na^+–neutral amino acid cotransport	
	Na^+–bile acid cotransport	

plasma membrane of most cells. It possesses receptors for hormonal or neuronal regulation of cellular functions, a Na^+,K^+–ATPase for removal of Na^+ from the cell, and transport systems for the entry of nutrients for consumption by the cell. In addition, the contraluminal plasma membrane contains the transport systems necessary for exit of the nutrients derived from the lumen so that the digested food can become available to all cells of the body. Some of the transport systems in the contraluminal plasma membrane may fulfill both the function of catalyzing exit when the intracellular nutrient concentration is high after a meal and that of mediating their entry when the blood levels are higher than those within the cell.

NaCl Absorption Has Both Active and Passive Components

Transport of Na^+ plays a crucial role not only for epithelial NaCl absorption or secretion, but also in the energization of nutrient uptake. The **Na^+,K^+–ATPase** provides the dominant mechanism for transduction of chemical energy in the form of ATP into osmotic energy of a concentration (chemical) or a combined concentration and electrical (electrochemical) ion gradient across the plasma membrane. In epithelial cells this enzyme is located exclusively in the contraluminal plasma membrane (Figure 26.9). The stoichiometry of the Na^+,K^+–ATPase reaction is 1 mol of ATP coupled to the outward pumping of 3 mol of Na^+ and the simultaneous inward pumping of 2 mol of K^+. The Na^+,K^+–ATPase maintains the high K^+ and low Na^+ concentrations in the cytosol and is directly or indirectly responsible for an electrical potential of about –60 mV of the cytoplasm relative to the extracellular solution. The direct contribution comes from the charge movement when $3Na^+$ ions are replaced by $2K^+$; the indirect contribution is by way of the K^+ gradient, which becomes the dominant force for establishing the potential by the movement of K^+ through K^+ channels.

Transepithelial NaCl movements are produced by the combined actions of the Na^+,K^+–ATPase and additional "passive" transport systems in the plasma membrane, which allow the entry of Na^+ or Cl^- into the cell. NaCl absorption results from Na^+ entry into the cell across the luminal plasma membrane and its extrusion by the Na^+,K^+–ATPase across the contraluminal membrane. Epithelial cells of the lower portion of the large intestine possess a luminal **Na^+ channel (epithelial Na^+ channel** or **ENaC)** that allows the uncoupled entry of Na^+ down its electrochemical gradient (Figure 26.10). This Na^+ flux is **electrogenic;** that is, it is associated with an electrical current, and it can be inhibited by

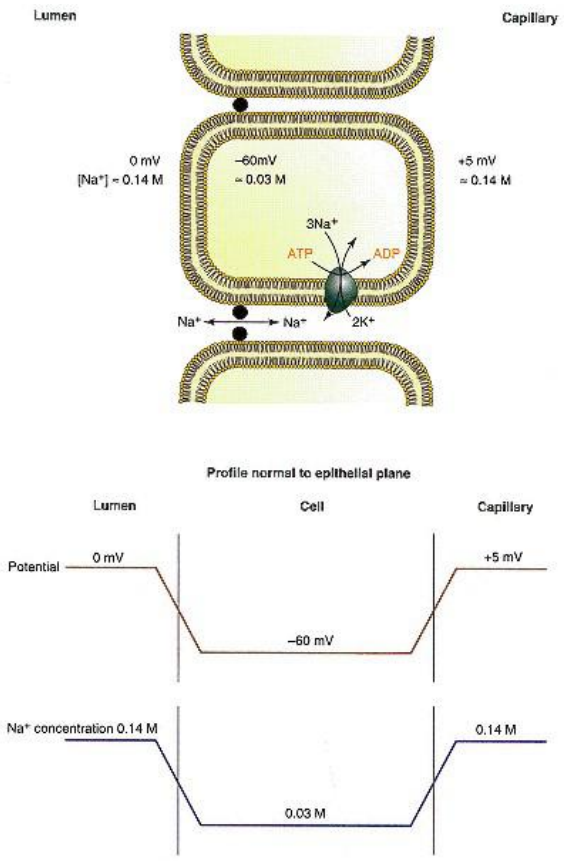


Figure 26.9
Na^+ concentrations and electrical potentials in enterocytes.

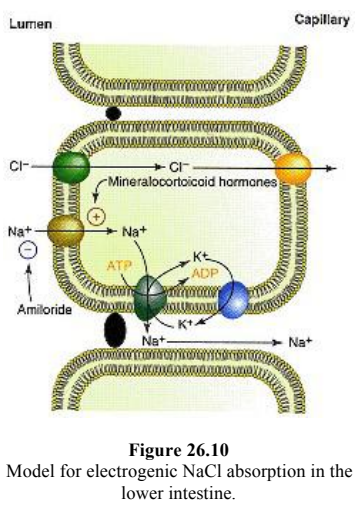


Figure 26.10
Model for electrogenic NaCl absorption in the lower intestine.

the diuretic drug amiloride at micromolar concentrations (Figure 26.11). The presence of this transport system, and hence NaCl absorption, is regulated by mineralocorticoid hormones of the adrenal cortex.

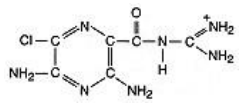


Figure 26.11
Amiloride.

Epithelial cells of the small intestine possess a transport system in the brush border membrane, which catalyzes an electrically neutral **Na^+/H^+ exchange (Na/H exchanger or NHE)** (Figure 26.12). The exchange is not affected by low concentrations of amiloride and not regulated by mineralocorticoids. The Na^+ absorption secondarily drives Cl^- absorption through a specific **Cl^-/HCO_{3-} exchanger (anion exchanger or AE)** in the luminal plasma membrane, as illustrated in Figure 26.12. The necessity for two types of NaCl absorption may arise from the different functions of upper and lower intestine, which require different regulation. The upper intestine reabsorbs the bulk of NaCl from the diet and from secretions of the exocrine glands after each meal, while the lower intestine participates in the fine regulation of NaCl retention, depending on the overall electrolyte balance of the body.

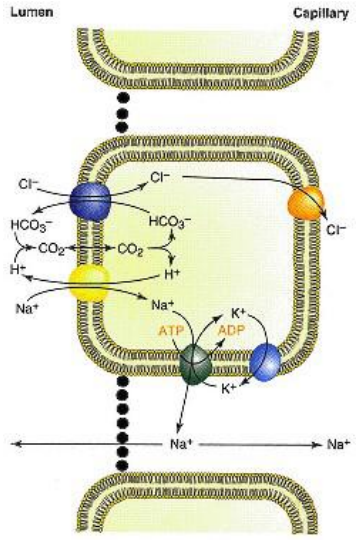


Figure 26.12
Model for electrically neutral NaCl absorption in the small intestine.

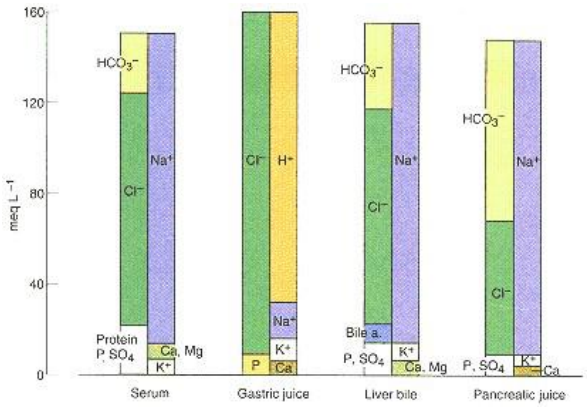


Figure 26.13
Ionic composition of secretions of the gastrointestinal tract.
Serum included for comparisons. Note the high H^+ concentration in gastric juice (pH + 1) and the high HCO_{3-} concentration in pancreatic juice. P, organic and inorganic phosphate; SO_4, inorganic and organic sulfate; Ca, calcium; Mg, magnesium; bile a., bile acids.
Adapted from *Biological Handbooks. Blood and Other Body Fluids*. Federation of American Societies for Experimental Biology, 1961.

NaCl Secretion Depends on Contraluminal Na^+,K^+–ATPase

The epithelial cells of most regions of the gastrointestinal tract have the potential for electrolyte and fluid secretions. The major secreted ions are Na^+ and Cl^-. Water follows passively because of the osmotic forces exerted by any secreted solute. Thus NaCl secretion secondarily results in fluid secretion. The fluid may be either hypertonic or isotonic, depending on the contact time of the secreted fluid with the epithelium and the tissue permeability to water. The longer the contact and the greater the water permeability, the closer the secreted fluid gets to osmotic equilibrium, that is, isotonicity. Ionic compositions of gastrointestinal secretions are presented in Figure 26.13.

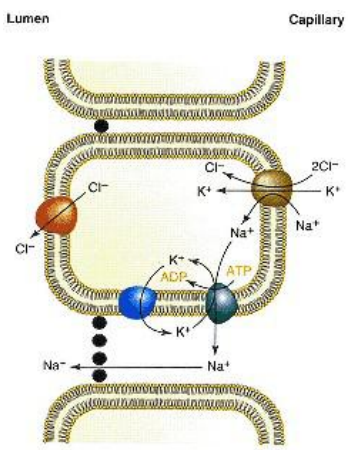


Figure 26.14
Model for epithelial NaCl secretion.

The cellular mechanisms for NaCl secretion involve the Na^+,K^+–ATPase located in the contraluminal plasma membrane of epithelial cells (Figure 26.14). The enzyme is implicated because cardiac glycosides, inhibitors of this enzyme, abolish salt secretion. However, the involvement of Na^+, K^+–ATPase does not provide a straightforward explanation for a NaCl movement from the capillary side to the lumen because the enzyme extrudes Na^+ from the cell toward the capillary side. Thus the active step of Na^+ transport across one of the plasma membranes has a direction opposite to that of overall transepithelial NaCl movements. The apparent paradox is resolved by an electrical coupling of Cl^- secretion across the luminal plasma membrane and Na^+ movements via the paracellular route, illustrated in Figure 26.14. The Cl^- secretion depends on coupled uptake of 2 Cl^- ions with Na^+ and K^+ via a specific cotransporter in the contraluminal plasma membrane and specific luminal **Cl^- channels**. The **Na^+,K^+,2 Cl^--cotransporter**, which can be identified by specific inhibitors such as the common diuretic **furosemide** (Figure 26.15), utilizes the energy of the Na^+ gradient to accumulate Cl^- within the cytoplasmic compartment above its electrochemical equilibrium concentration. Subsequent opening of luminal Cl^- channels allows efflux of Cl^- together with a negative charge (see Clin. Corr. 26.1 and 26.2).

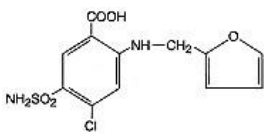


Figure 26.15
Furosemide.

In the pancreas a fluid rich in Na^+ and Cl^- is secreted by acinar cells. This fluid provides the vehicle for the movement of digestive enzymes from the acini, where they are released, to the lumen of the duodenum. The fluid is modified in the ducts by the additional secretion of $NaHCO_3$ (Figure 26.16). The HCO_{3-} concentration in the final pancreatic juice can reach concentrations of up to 120 mM.

The permeability of the tight junction to H_2O, Na^+ or other ions modifies

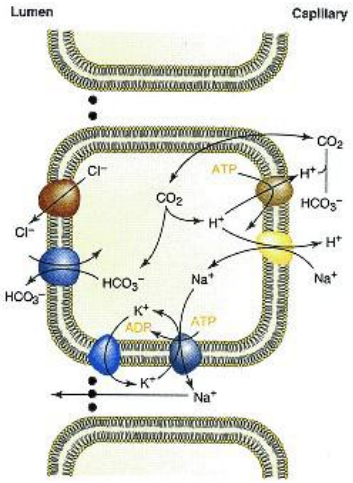


Figure 26.16
Model for epithelial $NaHCO_3$ secretion.
Note that two different mechanisms for H^+ secretion exist in the contraluminal plasma membrane: (1) Na^+/H^+ exchange and (2) H^+–ATPase.

CLINICAL CORRELATION 26.1

Cystic Fibrosis

Cystic fibrosis is an autosomal recessive inherited disease due to a mutation in the cystic fibrosis transmembrane regulatory (CFTR) protein. This protein contains 1480 amino acids organized into two membrane-spanning portions, which contain six transmembrane regions each, two ATP-binding domains, and a regulatory domain that undergoes phosphorylation by cAMP-dependent protein kinase. Some 400 mutations have been discovered since the gene was cloned in 1989.

The normal form of this protein is a Cl^- channel that is found in the luminal plasma membrane of epithelial cells in many tissues. The channel is normally closed but opens when phosphorylated by protein kinase A, thus providing regulated Cl^- and fluid secretion. The most common and severe mutation lacks one amino acid (F508 CFTR), which prevents the protein from properly maturing and reaching the plasma membrane. People who inherit this mutant CFTR from both parents lack Cl^- and fluid secretion in tissues that depend on CFTR for this function. Failure to secrete fluid, in turn, can lead to gross organ impairment due to partial or total blockage of passageways, for example, the ducts in the pancreas, the lumen of the intestine, or airways. (See Clin. Corr. 26.2 for activation of the CFTR Cl^- channel.)

active transepithelial solute movements. For example, a high permeability is necessary to allow Na^+ to equilibrate between extracellular solutions of the intercellular and luminal compartments during NaCl or $NaHCO_3$ secretion. Different regions of the gastrointestinal tract differ not only with respect to the transport systems that determine the passive entry (see above for amiloride-sensitive Na^+ channel and Na^+/H^+ exchange), but also with respect to the permeability characteristics of the tight junction. The distal portion (colon) is much tighter so as to prevent leakage of Na^+ from blood to lumen, in accordance with its function of scavenging of NaCl from the lumen.

Concentration Gradients or Electrical Potentials Drive Transport of Nutrients

Many solutes are absorbed across the intestinal epithelium against a concentration gradient. Energy for this "active" transport is directly derived from the Na^+ concentration gradient or the electrical potential across the luminal plasma membrane, rather than from the chemical energy of a covalent bond change, such as ATP hydrolysis. Glucose transport provides an example of uphill solute transport that is driven directly by the electrochemical Na^+ gradient and only indirectly by ATP (Figure 26.17).

Glucose is absorbed from the intestinal lumen into the blood against a concentration gradient. This vectorial transport is the combined result of several separate membrane events (Figure 26.18): (1) ATP-dependent Na^+ transport out of the cell at the contraluminal pole that establishes an electrochemical Na^+ gradient across the plasma membrane; (2) K^+ channels that convert a K^+ gradient into a membrane potential; (3) asymmetric insertion of two different transport systems for glucose into the luminal and contraluminal plasma membranes; and (4) coupling of Na^+ and glucose transport across the luminal membrane.

The luminal plasma membrane contains a transport system that facilitates a tightly coupled movement of Na^+ and D-glucose or structurally similar sugars

CLINICAL CORRELATION 26.2

Bacterial Toxigenic Diarrheas and Electrolyte Replacement Therapy

Voluminous, life-threatening intestinal electrolyte and fluid secretion (diarrhea) occurs in patients with cholera, an intestinal infection by *Vibrio cholerae*. Certain strains of *E. coli* also cause (traveler's!) diarrhea that can be serious in infants. The secretory state is a result of enterotoxins produced by the bacteria. The mechanisms of action of some of these enterotoxins are well understood at the biochemical level. Cholera toxin activates adenylyl cyclase by causing ADP-ribosylation of the G_{α_s}-protein, which stimulates the cyclase (see p. 859). Elevated cAMP levels in turn activate protein kinase A, which opens the luminal CFTR Cl^- channel and inhibits the Na^+/H^+ exchanger by protein phosphorylation. The net result is gross NaCl secretion. *Escherichia coli* produces a heat-stable toxin that binds to the receptor for the physiological peptide "guanylin," namely, the brush border guanylyl cyclase. When the receptor is occupied on the luminal side by either guanylin or the heat-stable *E. coli* toxin, the guanylyl cyclase domain of the protein on the cytosolic side is activated and cGMP levels rise. Elevated cGMP levels have the same effect on Cl^- secretion as elevated cAMP levels, except that a cGMP-activated protein kinase is involved in protein phosphorylation.

Modern, oral treatment of cholera takes advantage of the presence of Na^+-glucose cotransport in the intestine, which is not regulated by cAMP and remains fully active in this disease. In this case, the presence of glucose allows uptake of Na^+ to replenish body NaCl. Composition of solution for oral treatment of cholera patients is glucose 110 mM, Na^+ 99 mM, Cl^- 74 mM, HCO_{3-} 29 mM, and K+ 4 mM. The major advantages of this form of therapy are its low cost and ease of administration when compared with intravenous fluid therapy.

Carpenter, C. C. J. In: M. Field, J. S. Fordtran, and S. G. Schultz (Eds.), *Secretory Diarrhea*. Bethesda, MD: American Physiological Society, 1980, pp. 67–83.

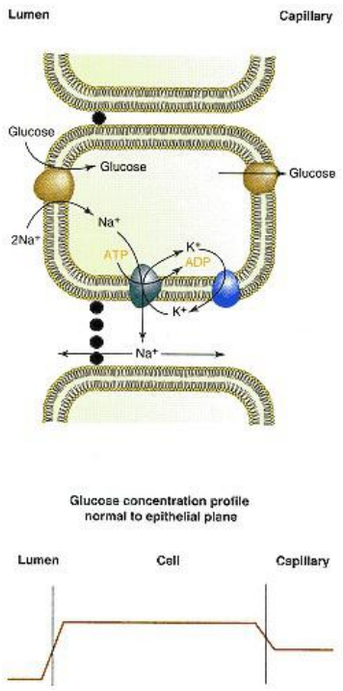


Figure 26.17
Model for epithelial glucose absorption.

$$2X \quad 3Na^+_{cell} + 2K^+_{interstitium} + ATP_{cell} \xrightarrow{Na^+,\,K^+\text{–ATPase}} 3Na^+_{interstitium} + 2K^+_{cell} + ADP_{cell} + P_{cell}$$

$$4X \quad K^+_{cell} \xrightleftharpoons{K^+\ channel} K^+_{interstitium}$$

$$3X \quad 2Na^+_{lumen} + Glc_{lumen} \xrightleftharpoons{SGLT1} 2Na^+_{cell} + Glc_{cell}$$

$$3X \quad Glc_{cell} \xrightleftharpoons{GLUT2} Glc_{interstitium}$$

$$6X \quad Na^+_{interstitium} \rightleftharpoons Na^+_{lumen}$$

$$Sum \quad 3Glc_{lumen} + 2ATP_{cell} \longrightarrow 3Glc_{interstitium} + 2ADP_{cell} + 2P_{cell}$$

Figure 26.18
Transepithelial glucose transport as translocation reactions across the plasma membranes and the tight junction.
SGLT1 (sodium glucose transporter 1) and GLUT2 (glucose transporter 2) are specific intestinal gene products mediating Na^+–glucose cotransport and facilitated glucose transport, respectively. Numbers in the left column indicate the minimal turnover of individual reactions to balance the overall reaction.

(Sodium GLucose Transporter or **SGLT).** The most common intestinal sodium-glucose cotransporter is **SGLT1** and it couples the movement of 2 Na^+ ions with that of 1 glucose molecule. It mediates glucose and Na^+ transport equally well in both directions. However, because of the higher Na^+ concentration in the lumen and the negative potential within the cell, the observed direction is from lumen to cell, even if the cellular glucose concentration is higher than the luminal one. In other words, downhill Na^+ movement normally supports concentrative glucose transport. Concentration ratios of up to 20-fold between intracellular and extracellular glucose have been observed *in vitro* under conditions of blocked efflux of cellular glucose. In some situations Na^+ uptake via this route is actually more important than glucose uptake (see Clin. Corr. 26.2).

The contraluminal plasma membrane contains a member of the **GLUcose Transporter** (or **GLUT**) family, which facilitates glucose exit and entry. The intestine contains the **GLUT2** transporter, which accepts many monosaccharides, including glucose. The direction of net flux is determined by the sugar concentration gradient. The two glucose transport systems SGLT1 and GLUT2 in the luminal and contraluminal plasma membranes, respectively, share glucose as substrate, but otherwise differ considerably in terms of amino acid sequence, secondary protein structure, Na^+ as cosubstrate, specificity for other sugars, sensitivity to inhibitors, or biological regulation. Since both SGLT and GLUT are not inherently directional, "active" transepithelial glucose transport can be maintained under steady-state conditions only if the Na^+,K^+–ATPase continues to move Na^+ out of the cell. Thus the active glucose transport is indirectly dependent on a supply of ATP and an active Na^+,K^+–ATPase.

The advantage of an electrochemical Na^+ gradient serving as intermediate is that the Na^+,K^+–ATPase can energize the transport of many different nutrients. The only requirement is presence of a transport system catalyzing cotransport of the nutrient with Na^+.

Gastric Parietal Cells Secrete HCl

The parietal (oxyntic) cells of gastric glands are capable of secreting HCl into the gastric lumen. Luminal H^+ concentrations of up to 0.14 M (pH 0.8) have been observed (see Figure 26.13). As the plasma pH = 7.4, the parietal cell transports protons against a concentration gradient of $10_{6.6}$. The free energy required for **HCl secretion** under these conditions is minimally 9.1 kcal mol^{-1} of HCl (= 38 J mol^{-1} of HCl), as calculated from

$$\Delta G' = RT\,2.3 \log 10^{6.6} \qquad RT = 0.6 \text{ kcal mol}^{-1} \text{ at } 37°C$$

A K^+-activated ATPase **(K^+,H^+–ATPase)** is intimately involved in the mechanism of active HCl secretion. This enzyme is unique to the parietal cell and is found only in the luminal region of the plasma membrane. It couples the hydrolysis of ATP to an electrically neutral obligatory exchange of K^+ for H^+, secreting H^+ and taking K^+ into the cell. The stoichiometry appears to be 1 mol of transported H^+ and K^+ for each mole of ATP.

$$ATP_{cell} + H^+_{cell} + K^+_{lumen} \rightleftharpoons ADP_{cell} + P_{i,cell} + H^+_{lumen} + K^+_{cell}$$

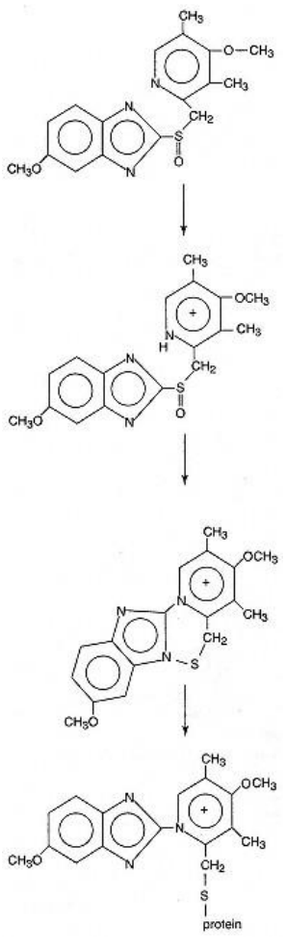


Figure 26.19
Omeprazole, an inhibitor of K^+,H^+–ATPase.
This drug accumulates in an acidic compartment (pK_a ~ 4) and is converted to a reactive sulfenamide, which reacts with cysteine SH groups.
From Sachs, G. The gastric H,K-ATPase. In. L. R. Johnson (Ed.), *Physiology of the Gastrointestinal Tract*. New York: Raven Press, 1994, p. 1133.

As the K^+,H^+–ATPase generates a very acidic solution, protein reagents that are activated by acid can become specific inhibitors of this enzyme. Figure 26.19 shows an example of such a reagent used to treat peptic ulcers. In the steady state, HCl can be elaborated by K^+, H^+–ATPase only if the luminal membrane is permeable to K^+ and Cl^- and the contraluminal plasma membrane catalyzes an exchange of Cl^- for HCO_{3-} (Figure 26.20). The exchange of Cl_- for HCO_{3-} is essential to resupply the cell with Cl_- and to prevent accumulation of base within the cell. Thus, under steady-state conditions, secretion of HCl into the gastric lumen is coupled to movement of HCO_{3-} into the plasma.

26.4— Digestion and Absorption of Proteins

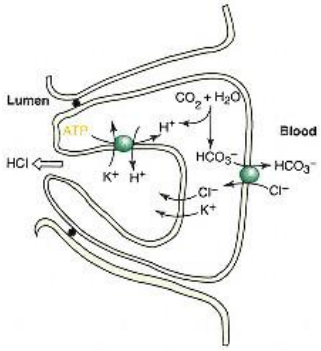


Figure 26.20
Model for secretion of hydrochloric acid.

Mixture of Peptidases Assures Efficient Protein Digestion

The total daily protein load to be digested consists of about 70–100 g of dietary proteins and 35–200 g of endogenous proteins from digestive enzymes and sloughed-off cells. Digestion and absorption of proteins are very efficient processes in healthy humans, since only about 1–2 g of nitrogen are lost through feces each day, which is equivalent to 6–12 g of protein.

Except for a short period after birth, oligo- and polypeptides (proteins) are not absorbed intact in appreciable quantities by the intestine. Proteins are broken down by hydrolases with specificity for the peptide bond, that is, by peptidases. This class of enzymes is divided into **endopeptidases** (proteases), which attack internal bonds and liberate large peptide fragments, and **exopeptidases,** which cleave off one amino acid at a time from either the COOH **(carboxypeptidases)** or the NH_2 terminus **(aminopeptidases).** Endopeptidases are important for an initial breakdown of long polypeptides into smaller products, which can then be attacked more efficiently by exopeptidases. The final products are free amino acids and di- and tripeptides, which are absorbed by epithelial cells (Figure 26.21).

The process of protein digestion can be divided into a gastric, a pancreatic, and an intestinal phase, depending on the source of peptidases.

Pepsins Catalyze Gastric Digestion of Protein

Gastric juice is characterized by the presence of HCl and therefore a low pH less than 2 as well as the presence of proteases of the pepsin family. The acid serves to kill off microorganisms and also to **denature proteins.** Denaturation makes proteins more susceptible to hydrolysis by proteases. **Pepsins** are unique in that they are acid stable; in fact, they are active at acid but not at neutral pH. The catalytic mechanism that is effective for peptide hydrolysis at the acid pH depends on two carboxylic groups at the active site of the enzymes. Pepsin A, the major gastric protease, prefers peptide bonds formed by the amino group of aromatic acids (Phe, Tyr) (Table 26.6).

Active pepsin is generated from the proenzyme **pepsinogen** by the removal of 44 amino acids from the NH_2 terminus (pig enzyme). Cleavage between residues 44 and 45 of pepsinogen occurs as either an intramolecular reaction **(autoactivation)** below pH 5 or by active pepsin (autocatalysis). The liberated peptide from the NH_2 terminus remains bound to pepsin and acts as "pepsin

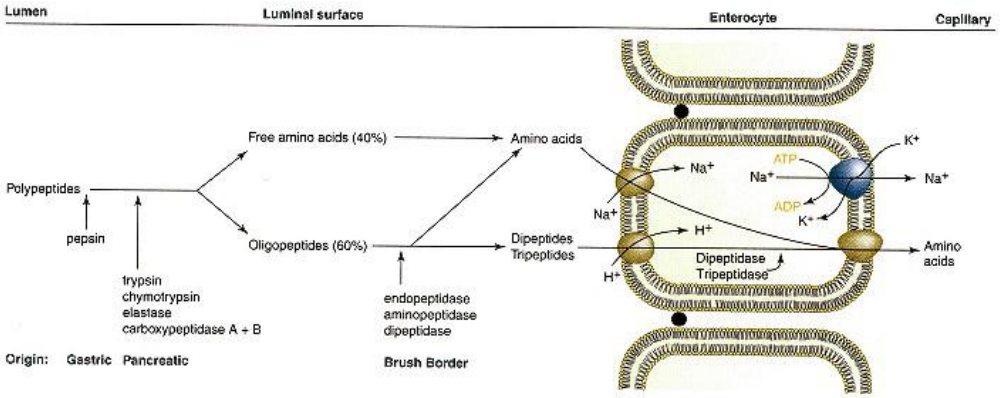
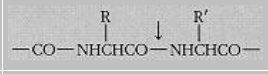
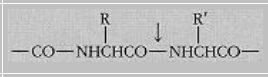
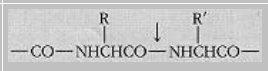
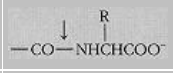
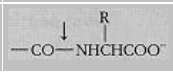


Figure 26.21
Digestion and absorption of proteins.

TABLE 26.6 Gastric and Pancreatic Peptidases

Enzyme	*Proenzyme*	*Activator*	*Cleavage Point*	*R*
CARBOXYL PROTEASES Pepsin A	Pepsinogen A	Autoactivation, pepsin		Tyr, Phe, Leu
SERINE PROTEASES Trypsin	Trypsinogen	Enteropeptidase, trypsin		Arg, Lys
Chymotrypsin	Chymotrypsinogen	Trypsin		Tyr, Trp, Phe, Met, Leu
Elastase	Proelastase	Trypsin		Ala, Gly, Ser
ZINC PEPTIDASES Carboxypeptidase A	Procarboxypeptidase A	Trypsin		Val, Leu, Ile, Ala
Carboxypeptidase B	Procarboxypeptidase B	Trypsin		Arg, Lys

inhibitor" above pH 2. This inhibition is released either by a drop of the pH below 2 or further degradation of the peptide by pepsin. Thus, once favorable conditions are reached, pepsinogen is converted to pepsin by autoactivation and subsequent autocatalysis at an exponential rate.

The major products of pepsin action are large peptide fragments and some free amino acids. The importance of gastric protein digestion does not lie so much in its contribution to the breakdown of ingested macromolecules, but rather in the generation of peptides and amino acids that act as stimulants for **cholecystokinin** release in the duodenum. The gastric peptides therefore are instrumental in the initiation of the pancreatic phase of protein digestion.

Pancreatic Zymogens Are Activated in Small Intestine

Pancreatic juice is rich in proenzymes of **endopeptidases** and **carboxypeptidases** (Figure 26.22), which are activated after they reach the lumen of the small intestine. **Enteropeptidase** (old name: enterokinase), a protease produced by duodenal epithelial cells, activates pancreatic **trypsinogen** to **trypsin** by scission of a hexapeptide from the NH_2 terminus. Trypsin in turn autocatalytically activates more trypsinogen to trypsin and also acts on the other proenzymes, thus liberating the endopeptidases chymotrypsin and elastase and the **carboxypeptidases A and B.** Since trypsin plays a pivotal role among pancreatic enzymes in the activation process, pancreatic juice normally contains a small-molecular-weight peptide that acts as a **trypsin inhibitor** and neutralizes any trypsin formed prematurely within the pancreatic cells or pancreatic ducts.

Trypsin, chymotrypsin, and elastase have different substrate specificity, as shown in Table 26.6. They are active only at neutral pH and depend on pancreatic $NaHCO_3$ for neutralization of gastric HCl. Their mechanism of catalysis involves an **essential serine residue** (see p. 97) and is thus similar to serine esterases, such as acetyl choline esterase. Reagents that interact with serine and modify it, inactivate serine esterases and peptidases. A prominent example of such a reagent is the highly toxic diisopropylphosphofluoridate, which was developed originally for chemical warfare (neurotoxic because of inhibition of acetyl choline esterase).

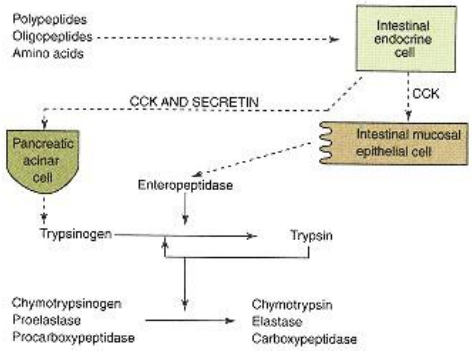


Figure 26.22
Secretion and activation of pancreatic enzymes.
Abbreviation: CCK, cholecystokinin.
Reproduced with permission from Freeman, H. J., and Kim, Y. S. *Annu. Rev. Med.* 29:102, 1978.

Polypeptides generated from ingested proteins are degraded within the small intestinal lumen by **carboxypeptidases A** and **B.** The **pancreatic carboxypeptidases** are $\mathbf{Zn^{2+}}$ metalloenzymes and possess a different type of catalytic mechanism than the carboxyl or serine peptidases. The combined action of pancreatic peptidases results in the formation of free amino acids and small peptides of 2–8 residues. Peptides account for about 60% of the amino nitrogen at this point.

Intestinal Peptidases Digest Small Peptides

Since pancreatic juice does not contain appreciable aminopeptidase activity, final digestion of di- and oligopeptides depends on small intestinal enzymes. The luminal surface of epithelial cells is particularly rich in endopeptidase and aminopeptidase activity, but also contains dipeptidases (Table 26.2). The end products of the cell surface digestion are free amino acids and di- and tripeptides, which are absorbed via specific **amino acid** or **peptide transport systems.** Transported di- and tripeptides are generally hydrolyzed within the cytoplasmic compartment before they leave the cell. The cytoplasmic dipeptidases explain why practically only free amino acids are found in the portal blood after a meal. The virtual absence of peptides had previously been taken as evidence that luminal protein digestion had to proceed all the way to free amino acids before absorption could occur. However, it is now established that a large portion of dietary amino nitrogen is absorbed in the form of small peptides with subsequent intracellular hydrolysis. However, di- and tripeptides containing proline and hydroxyproline or unusual amino acids, such as β-alanine as carnosine (β-alanylhistidine) or anserine (β-alanyl 1-methylhistidine), are absorbed without intracellular hydrolysis because they are not good substrates for the intestinal cytoplasmic dipeptidases. β-Alanine is present in chicken meat.

Free Amino Acids and Dipeptides Are Absorbed by Carrier-Mediated Transport

The small intestine has a high capacity to absorb free amino acids and small peptides. Most L-amino acids can be transported across the epithelium against a concentration gradient, although the need for concentrative transport *in vivo* is not obvious, since luminal concentrations are usually higher than the plasma levels of 0.1–0.2 mM. Amino acid and peptide transport in the small intestine has all the characteristics of carrier-mediated transport, such as discrimination between D- and L-amino acids and energy and temperature dependence. In addition, genetic defects are known to occur in humans (see Clin. Corr. 26.3).

CLINICAL CORRELATION 26.3

Neutral Amino Aciduria (Hartnup Disease)

Transport functions, like enzymatic functions, are subject to modification by mutations. An example of a genetic lesion in epithelial amino acid transport is Hartnup disease, named after the family in which the disease entity resulting from the defect was first recognized. The disease is characterized by the inability of renal and intestinal epithelial cells to absorb neutral amino acids from the lumen. In the kidney, in which plasma amino acids reach the lumen of the proximal tubule through the ultrafiltrate, the inability to reabsorb amino acids manifests itself as excretion of amino acids in the urine (amino aciduria). The intestinal defect results in malabsorption of free amino acids from the diet. Therefore the clinical symptoms of patients with this disease are mainly those due to essential amino acid and nicotinamide deficiencies. The pellagra-like features (see p. 1121) are explained by a deficiency of tryptophan, which serves as precursor for nicotinamide. Investigations of patients with Hartnup disease revealed the existence of intestinal transport systems for di- or tripeptides, which are different from the ones for free amino acids. The genetic lesion does not affect transport of peptides, which remains as a pathway for absorption of protein digestion products.

Silk, D. B. A. Disorders of nitrogen absorption. In: J. T. Harries (Ed.), *Clinics in Gastroenterology: Familial Inherited Abnormalities*, Vol. 11: London: Saunders, 1982, pp. 47–73.

On the basis of genetics, transport experiments, and expression cloning, at least seven **brush border specific transport systems** for the uptake of L-amino acids or small peptides in the luminal membrane can be distinguished: (1) for neutral amino acids with short or polar side chains (Ser, Thr, Ala); (2) for neutral amino acids with aromatic or hydrophobic side chains (Phe, Tyr, Met, Val, Leu, Ile); (3) for imino acids (Pro, Hyp); (4) for β-amino acids (β-Ala, taurine); (5) for basic amino acids and cystine (Lys, Arg, Cys-Cys); (6) for acidic amino acids (Asp, Glu); and (7) for dipeptides (Pept1) (Gly-sarcosine).

The concentration mechanisms for neutral L-amino acids appear to be similar to those discussed for D-glucose (see Figure 26.17). **Na^+-dependent transport systems** have been identified in the luminal (brush border) membrane and **Na^+-independent transporters** in the contraluminal plasma membrane of small intestinal epithelial cells. Similarly, as for active glucose transport, the energy for concentrative amino acid transport is derived directly from the electrochemical Na^+ gradient and only indirectly from ATP. Amino acids are not chemically modified during membrane transport, although they may be metabolized within the cytoplasmic compartment. The brush border transport for the other amino acids is energized in more complicated ways. For example, the **acidic amino acid transporter** mediates cotransport of the amino acid with 2 Na^+ ions and counter transport with 1 K^+ ion.

Neutral dipeptides are cotransported across the brush border membrane with a proton and thus are energized through the proton electrochemical gradient across this membrane. However, because of the Na^+/H^+ exchange, both gradients tend to be similar and interdependent. The dipeptide transporter also accepts β-lactam antibiotics (aminopenicillins) and is important for absorption of orally administered antibiotics of this class.

Fetus and Neonate Can Absorb Intact Proteins

The fetal and neonatal small intestines can absorb intact proteins. The uptake occurs by endocytosis, that is, the internalization of small vesicles of plasma membrane, which contain ingested macromolecules. The process is also termed **pinocytosis** because of the small size of vesicles. The small intestinal pinocytosis of protein is thought to be important for the transfer of maternal antibodies (γ-globulins) to the offspring, particularly in rodents. The pinocytotic uptake of proteins is not important for nutrition, and its magnitude usually declines after birth. Persistence of low levels of this process beyond the neonatal period may, however, be responsible for absorption of sufficient quantities of macro-molecules to induce antibody formation.

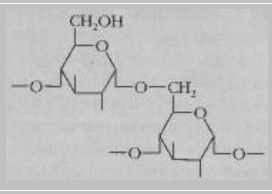
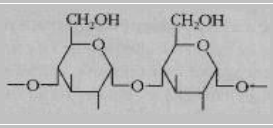
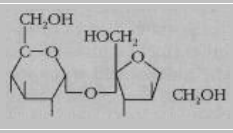
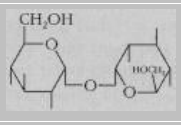
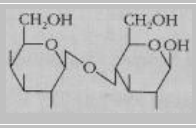
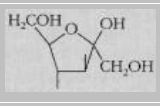
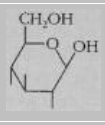
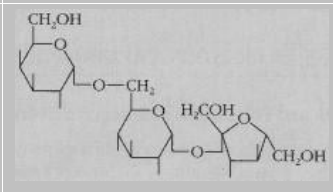
26.5— Digestion and Absorption of Carbohydrates

Di- and Polysaccharides Require Hydrolysis

Dietary carbohydrates provide a major portion of the daily caloric requirement. They consist of mono-, di-, and polysaccharides (Table 26.7). **Monosaccharides** need not be hydrolyzed for absorption. Disaccharides require the small intestinal surface enzymes for hydrolysis into monosaccharides, while polysaccharides depend on **pancreatic amylase** for degradation (Figure 26.23).

Starch, a major nutrient, is a plant polysaccharide with a molecular mass of more than 100 kDa. It consists of a mixture of linear chains of glucose molecules linked by α-1,4-glucosidic bonds **(amylose)** and of branched chains with branch points made up by α-1,6 linkages **(amylopectin).** The ratio of 1,4- to 1,6-glucosidic bonds is about 20 : 1. **Glycogen** is an animal polysaccharide similar in structure to amylopectin. The two compounds differ in terms of the number of branch points, which occur more frequently in glycogen.

TABLE 26.7 Dietary Carbohydrates

Carbohydrate	*Typical Source*		*Structure*
Amylopectin	Potatoes, rice, corn, bread	α-Glc(1 4)$_n$Glc with α-Glc(1 6) branches	
Amylose	Potatoes, rice, corn, bread	α-Glc(1 4)$_n$Glc	
Sucrose	Table sugar, desserts	α-Glc(1 2)β-Fru	
Trehalose	Young mushrooms	α-Glc(1 1)α-Glc	
Lactose	Milk, milk products	β-Gal(1 4)Glc	
Fructose	Fruit, honey	Fru	
Glucose	Fruit, honey, grape	Glc	
Raffinose	Leguminous seeds	α-Gal(1 6)α-Glc (1 2)β-Fru	

Hydrated starch and glycogen are attacked by the endosaccharidase **α-amylase** present in saliva and pancreatic juice (Figure 26.24). Hydration of the polysaccharides occurs during heating and is essential for efficient digestion. Amylase is specific for internal α-1,4-glucosidic bonds; α-1,6 bonds are not attacked, nor are α-1,4 bonds of glucose units that serve as branch points. The pancreatic isoenzyme is secreted in large excess relative to starch intake and

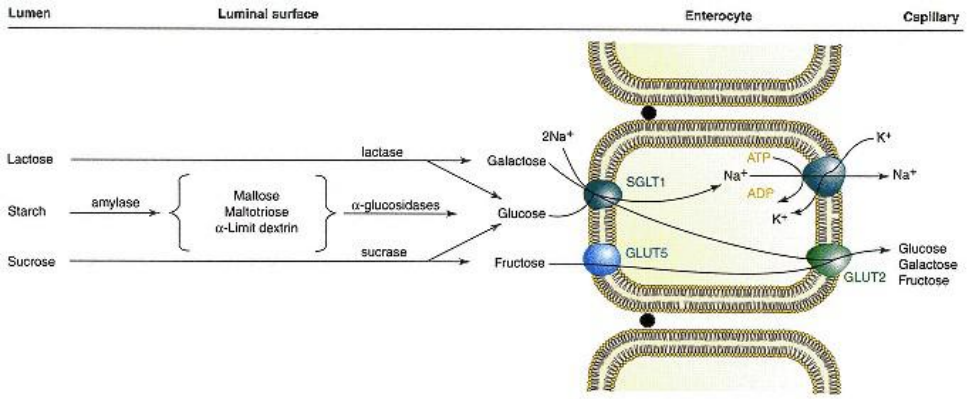


Figure 26.23
Digestion and absorption of carbohydrates.

CLINICAL CORRELATION 26.4

Disaccharidase Deficiency

Intestinal disaccharidase deficiencies are encountered relatively frequently in humans. Deficiency can be present in one enzyme or several enzymes for a variety of reasons (genetic defect, physiological decline with age, or the result of "injuries" to the mucosa). Of the disaccharidases, lactase is the most common enzyme with an absolute or relative deficiency, which is experienced as milk intolerance. The consequences of an inability to hydrolyze lactose in the upper small intestine are inability to absorb lactose and bacterial fermentation of ingested lactose in the lower small intestine. Bacterial fermentation results in the production of gas (distension of gut and flatulence) and osmotically active solutes that draw water into the intestinal lumen (diarrhea). The lactose in yogurt has already been partially hydrolyzed during the fermentation process of making yogurt. Thus individuals with lactase deficiency can often tolerate yogurt better than unfermented dairy products. The enzyme lactase is commercially available to pretreat milk so that the lactose is hydrolyzed.

Buller, H. A., and Grant, R. G. Lactose intolerance. *Annu. Rev. Med.* 41:141, 1990.

is more important than the salivary enzyme from a digestive point of view. The products of the digestion by α-amylase are mainly the **disaccharide maltose**, the **trisaccharide maltotriose**, and so-called **α-limit dextrins** containing on average eight glucose units with one or more α-1,6-glucosidic bonds.

Final hydrolysis of di- and oligosaccharides to monosaccharides is carried out by surface enzymes of the small intestinal epithelial cells (Table 26.8). Most of the surface oligosaccharidases are exoenzymes that cleave off one monosaccharide at a time from the nonreducing end. The capacity of the **α-glucosidases** is normally much greater than that needed for completion of the digestion of starch. Similarly, there is usually excess capacity for sucrose (table sugar) hydrolysis relative to dietary intake. In contrast, **β-galactosidase (lactase)** can be rate-limiting in humans for hydrolysis and utilization of lactose, the major milk carbohydrate (see Clin. Corr. 26.4).

Di-, oligo-, and polysaccharides that are not hydrolyzed by α-amylase and/ or intestinal surface enzymes cannot be absorbed; therefore they reach the lower tract of the intestine, which from the lower ileum on contains bacteria. Bacteria can utilize many of the remaining carbohydrates because they possess many more types of saccharidases than humans. Monosaccharides that are released as a result of **bacterial enzymes** are predominantly metabolized anaerobically by the bacteria themselves, resulting in degradation products such as short-chain fatty acids, lactate, hydrogen gas (H_2), methane (CH_4), and carbon

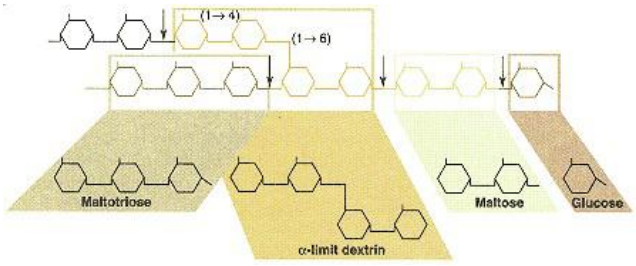


Figure 26.24
Digestion of amylopectin by salivary and pancreatic α-amylase.

TABLE 26.8 Saccharidases of the Surface Membrane of the Small Intestine

Enzyme	*Specificity*	*Natural Substrate*	*Product*
exo-1,4-α-Glucosidase (glucoamylase)	α-(1 4)Glucose	Amylose	Glucose
Oligo-1,6-glucosidase (isomaltase)	α-(1 6)Glucose	Isomaltose, α-dextrin	Glucose
α-Glucosidase (maltase)	α-(1 4)Glucose	Maltose, maltotriose	Glucose
Sucrose-α-Glucosidase (sucrase)	α-Glucose	Sucrose	Glucose, fructose
α,α-Trehalase	α-(1 1)Glucose	Trehalose	Glucose
β-Glucosidase	β-Glucose	Glucosylceramide	Glucose, ceramide
β-Galactosidase (lactase)	β-Galactose	Lactose	Glucose, galactose

dioxide (CO_2). These compounds can cause fluid secretion, increased intestinal motility, and cramps, either because of increased intraluminal osmotic pressure, and distension of the gut, or a direct irritant effect of the bacterial degradation products on the intestinal mucosa.

The well-known problem of flatulence after ingestion of leguminous seeds (beans, peas, and soya) is caused by oligosaccharides, which cannot be hydrolyzed by human intestinal enzymes. The leguminous seeds contain modified sucrose to which one or more galactose moieties are linked. The glycosidic bonds of galactose are in the α configuration, which can only be split by bacterial enzymes. The simplest sugar of this family is **raffinose** (see Table 26.7).

Trehalose, a disaccharide that occurs in young mushrooms, requires a special disaccharidase, **trehalase**.

Monosaccharides Are Absorbed by Carrier-Mediated Transport

The major monosaccharides that result from digestion of di- and polysaccharide are D-glucose, D-galactose, and D-fructose. Absorption of these and other minor monosaccharides are carrier-mediated processes that exhibit such features as substrate specificity, stereospecificity, saturation kinetics, and inhibition by specific inhibitors.

At least two types of monosaccharide transporters catalyze monosaccharide uptake from the lumen into the cell: (1) a **Na^+-monosaccharide cotransporter**, existing probably as a tetramer of 75-kDa peptides, has high specificity for D-glucose and D-galactose and catalyzes "active" sugar absorption (SGLT); and (2) a **Na^+-independent, facilitated-diffusion** type of monosaccharide transport system with specificity for D-fructose (GLUT5). In addition, a **Na^+-independent monosaccharide transporter** (GLUT2), consisting of 57-kDa peptide(s), which accepts all three monosaccharides, is present in the contraluminal plasma membrane. GLUT2 is also located in the liver and kidney, and other members of the GLUT family of glucose transporters are found in all cells. All GLUT transporters mediate uncoupled D-glucose flux down its concentration gradient. GLUT2 of gut, liver, and kidney moves D-glucose out of the cell into the blood under physiological conditions, while in other tissues GLUT1 (in erythrocytes and brain) or the insulin-sensitive GLUT4 (in fat and muscle tissue) are mainly involved in D-glucose uptake. Properties of intestinal SGLT1 and of GLUT2 are compared in Table 26.9, and their role in transepithelial glucose absorption is illustrated in Figure 26.18.

TABLE 26.9 Characteristics of Glucose Transport Systems in the Plasma Membranes of Enterocytes

Characteristic	*Luminal*	*Contraluminal*
Designation	SGLT1	GLUT2
Subunit molecular weight (kDa)	75	57
Effect of Na^+	Cotransport with Na^+	None
Good substrates	D-Glc, D-Gal, α-methyl-D-Glc	D-Glc, D-Gal, D-Man, 2-deoxy-D-Glc
Inhibitor	Phlorizin (Figure 26.25)	Cytochalasin B (Figure 26.26)

26.6— Digestion and Absorption of Lipids

Lipid Digestion Requires Overcoming the Limited Water Solubility of Lipids

An adult man ingests about 60–150 g of lipid per day. **Triacylglycerols** constitute more than 90% of the dietary fat. The rest is made up of phospholipids, cholesterol, cholesterol esters, and free fatty acids. In addition, 1–2 g of cholesterol and 7–22 g of phosphatidylcholine (lecithin) are secreted into the small intestine lumen as constituents of bile.

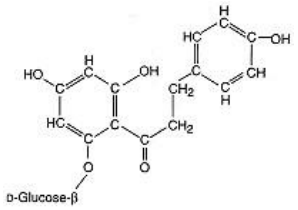


Figure 26.25
Phlorizin (phloretin-2 -glucoside).

Lipids are defined by their good solubility in organic solvents and their sparing or lack of solubility in aqueous solutions. The poor water solubility presents problems for digestion because the substrates are not easily accessible to the digestive enzymes in the aqueous phase. In addition, even if ingested lipids are hydrolyzed into simple constituents, the products tend to aggregate to larger complexes that make poor contact with the cell surface and therefore are not easily absorbed. These problems are overcome by (1) increases in the interfacial area between the aqueous and lipid phase and (2) "solubilization" of lipids with **detergents**. Thus changes in the physical state of lipids are intimately connected to chemical changes during digestion and absorption.

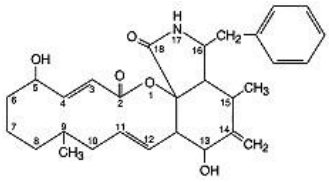


Figure 26.26
Cytochalasin B.

At least five different phases can be distinguished (Figure 26.27): (1) hydrolysis of triacylglycerols to free fatty acids and monoacylglycerols; (2) solubilization by detergents (bile acids) and transport from the intestinal lumen toward the cell surface; (3) uptake of free fatty acids and monoacylglycerols into the cell and resynthesis to triacylglycerols; (4) packaging of newly synthesized triacylglycerols into special lipid-rich globules, called chylomicrons, and (5) exocytosis of chylomicrons from cells and release into lymph.

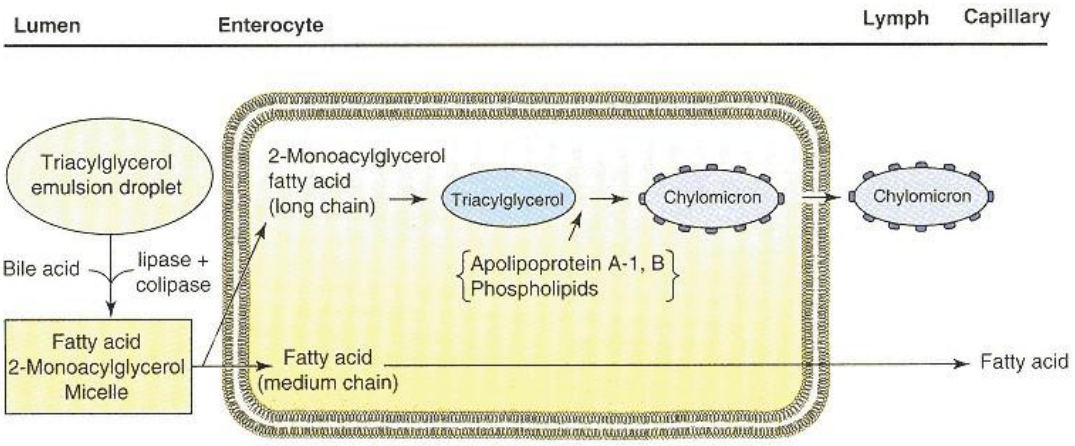


Figure 26.27
Digestion and absorption of lipids.

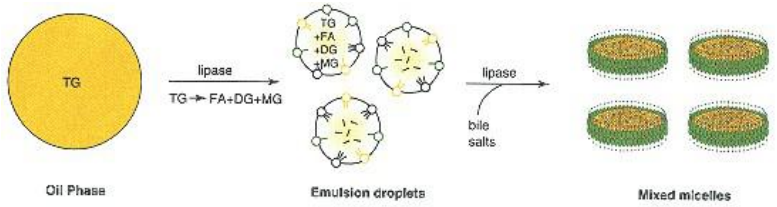


Figure 26.28
Changes in physical state during triacylglycerol digestion.
Abbreviations: TG, triacylglycerol; DG, diacylglycerol; MG, monoacylglycerol; FA, fatty acid.

Lipids Are Digested by Gastric and Pancreatic Lipases

Digestion of lipids is initiated in the stomach by an **acid-stable lipase,** most of which is thought to originate from glands at the back of the tongue. However, the rate of hydrolysis is slow because the ingested triacylglycerols form a separate lipid phase with a limited water–lipid interface. The lipase adsorbs to that interface and converts triacylglycerols into fatty acids and diacylglycerols (Figure 26.28). The importance of the initial hydrolysis is that some of the water-immiscible triacylglycerols are converted to products that possess both polar and nonpolar groups. Such surfactive products spontaneously adsorb to water–lipid interfaces and confer a hydrophilic surface to lipid droplets thereby providing a stable interface with the aqueous environment. At constant volume of the lipid phase, any increase in interfacial area produces dispersion of the lipid phase into smaller droplets (emulsification) and provides more sites for adsorption of more lipase molecules.

The major enzyme for triacylglycerol hydrolysis is the **pancreatic lipase** (Figure 26.29). This enzyme is specific for esters in the α-position of glycerol and prefers long-chain fatty acids with more than ten carbon atoms. Hydrolysis by the pancreatic enzyme also occurs at the water–lipid interface of emulsion droplets. The products are **free fatty acids** and **β-monoacylglycerols.** The purified form of the enzyme is strongly inhibited by the bile acids that normally are present in the small intestine during lipid digestion. The problem of inhibition is overcome by **colipase,** a small protein (12 kDa) that binds to both the water–lipid interface and to lipase, thereby anchoring and activating the enzyme. It is secreted by the pancreas as procolipase and depends on tryptic removal of a NH_2-terminal decapeptide for full activity.

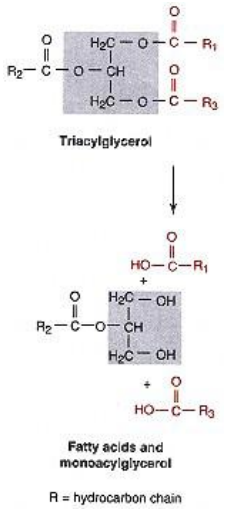


Figure 26.29
Mechanism of action of lipase.

Pancreatic juice also contains another less **specific lipid esterase,** which acts on cholesterol esters, monoglycerides, or other lipid esters, such as esters of vitamin A with carboxylic acids. In contrast to triacylglycerol lipase, this lipid esterase requires bile acids for activity.

Phospholipids are hydrolyzed by specific phospholipases. Pancreatic secretions are especially rich in the proenzyme for **phospholipase A_2** (Figure 26.30). As other pancreatic proenzymes, this one is also activated by trypsin. Phospholipase A_2 requires bile acids for activity.

Bile Acid Micelles Solubilize Lipids during Digestion

Bile acids are biological detergents that are synthesized by the liver and secreted as conjugates of **glycine** or **taurine** with the bile into the duodenum. At physiological pH values, they are present as anions, which have detergent

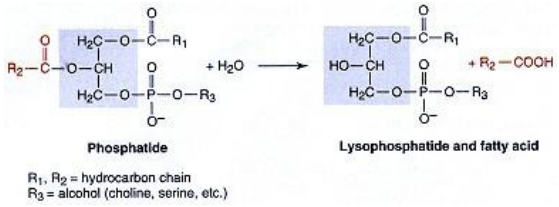


Figure 26.30
Mechanism of action of phospholipase A_2.

properties. Therefore the terms **bile acids** and **bile salts** are often used interchangeably (Figure 26.31). Bile acids at pH values above the p*K* (Table 26.10) reversibly form aggregates at concentrations above 2–5 mM. These aggregates are called **micelles**, and the minimal concentration necessary for micelle formation is the **critical micellar concentration** (Figure 26.32). The bile acids in micelles are in equilibrium with those free in solution. Thus micelles, in contrast to emulsified lipids, are equilibrium structures with well-defined sizes that are much smaller than emulsion droplets. Micelle sizes typically range between 40 and 600 m depending on bile acid concentration and the ratio of bile acids to lipids.

The arrangements of bile acids in micelles is such that the hydrophobic portions are removed from contact with water, while hydrophilic groups remain exposed to the water. The hydrophobic region of bile acids is formed by one surface of the fused ring system, while the carboxylate or sulfonate ion and the hydroxyl groups on the other side of the ring system are hydrophilic. Since the major driving forces for micelle formation are the removal of apolar, hydrophobic groups from and the interaction of polar groups with water molecules, the distribution of polar and apolar regions places some constraints on the stereochemical arrangements of bile acid molecules within a micelle. Four bile acid molecules are sufficient to form a very simple micelle as shown in Figure 26.33. Bile salt micelles can solubilize other lipids, such as phospholipids and fatty acids. These **mixed micelles** have disk-like shapes, whereby the phospholipids and fatty acids form a bilayer and the bile acids occupy the edge positions, rendering the edge of the disk hydrophilic (Figure 26.34). Within the mixed phospholipid–bile acid micelles, other water-insoluble lipids, such as cholesterol, can be accommodated and thereby "solubilized" (for potential problems see Clin. Corr. 26.5).

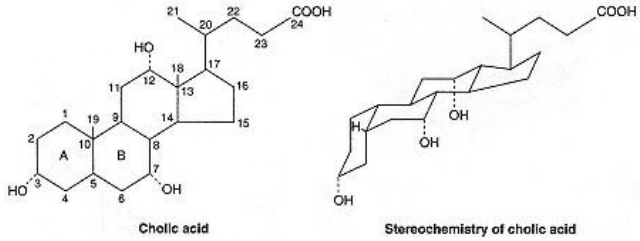


Figure 26.31
Cholic acid, a bile acid.

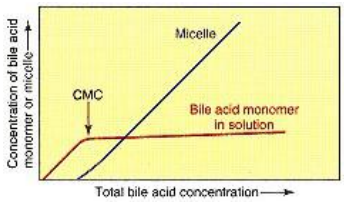


Figure 26.32
Solubility properties of bile acids in aqueous solutions. Abbreviation: CMC, critical micellar concentration.

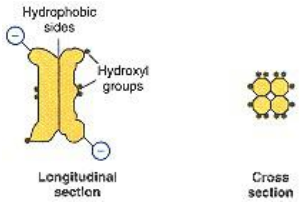
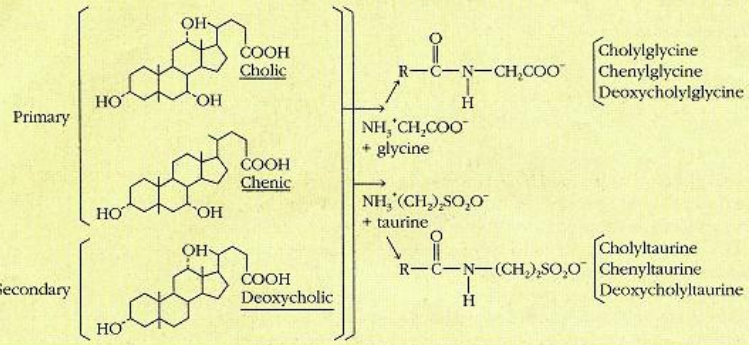


Figure 26.33
Diagrammatic representation of a Na^+ cholate micelle. Adapted from Small, D. M. *Biochim. Biophys. Acta* 176: 178, 1969.

TABLE 26.10 Effect of Conjugation on the Acidity of Cholic, Deoxycholic, and Chenodeoxycholic Acids

Bile Acid	*Ionized Group*	*pK_a*
Unconjugated bile acids	$-COO^-$ of cholestanoic acid	≃5
Glycoconjugates	$-COO^-$ of glycine	≃3.7
Tauroconjugates	$-SO_3^-$ of taurine	≃1.5



Source: Reproduced with permission from Hofmann, A. F. *Handbook of Physiology* 5:2508, 1968.

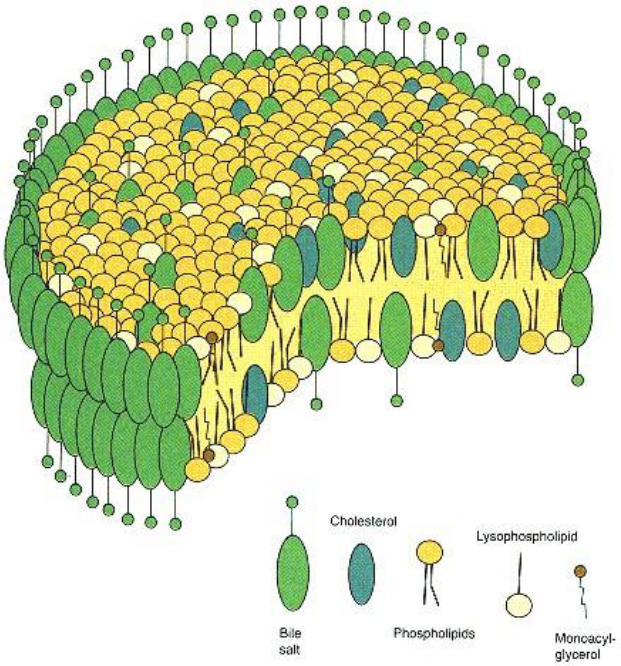


Figure 26.34
Proposed structure of the intestinal mixed micelle.
The bilayer disk has a band of bile salt at its periphery and other, more hydrophobic components (fatty acids, monoacylglycerol, phospholipids, and cholesterol) protected within its interior. Redrawn based on figure from Carey, M. C. In: A. M. Arias, H. Popper, D. Schachter, et al. (Eds.), *The Liver: Biology and Pathology*, New York: Raven Press, 1982.

CLINICAL CORRELATION 26.5

Cholesterol Stones

Liver secretes phospholipids and cholesterol together with bile acids into the bile. Because of the limited solubility of cholesterol, its secretion in bile can result in cholesterol stone formation in the gallbladder. Stone formation is a relatively frequent complication; up to 20% of North Americans will develop stones during their lifetime.

Cholesterol is practically insoluble in aqueous solutions. However, it can be incorporated into mixed phospholipid–bile acid micelles up to a mole ratio of 1:1 for cholesterol/phospholipids and thereby "solubilized" (see accompanying figure). The liver can produce supersaturated bile with a higher ratio than 1:1 of cholesterol/phospholipid. This excess cholesterol has a tendency to come out of solution and to crystallize. Such bile with excess cholesterol is considered lithogenic, that is, stoneforming. Crystal formation usually occurs in the gallbladder, rather than the hepatic bile ducts, because contact times between bile and any crystallization nuclei are greater in the gallbladder. In addition, the gallbladder concentrates bile by absorption of electrolytes and water. The bile salts chenodeoxycholate and ursodeoxycholate are now available for oral use to dissolve gallstones. Ingestion of these bile salts reduces cholesterol excretion into the bile and allows cholesterol in stones to be solubilized.

The tendency to secrete bile supersaturated with respect to cholesterol is inherited and found more frequently in females than in males, often associated with obesity. Supersaturation also appears to be a function of the size and nature of the bile acid pool as well as the secretion rate.

Schoenfield, L. J., and Lachin, J. M. Chenodiol (chenodeoxycholic acid) for dissolution of gallstones: The National Cooperative Gallstone Study. A controlled trial of safety and efficacy. *Ann. Intern. Med.* 95:257, 1981; and Carey, M. C., and Small, D. M. The physical chemistry of cholesterol solubility in bile. *J. Clin. Invest.* 61:998, 1978.

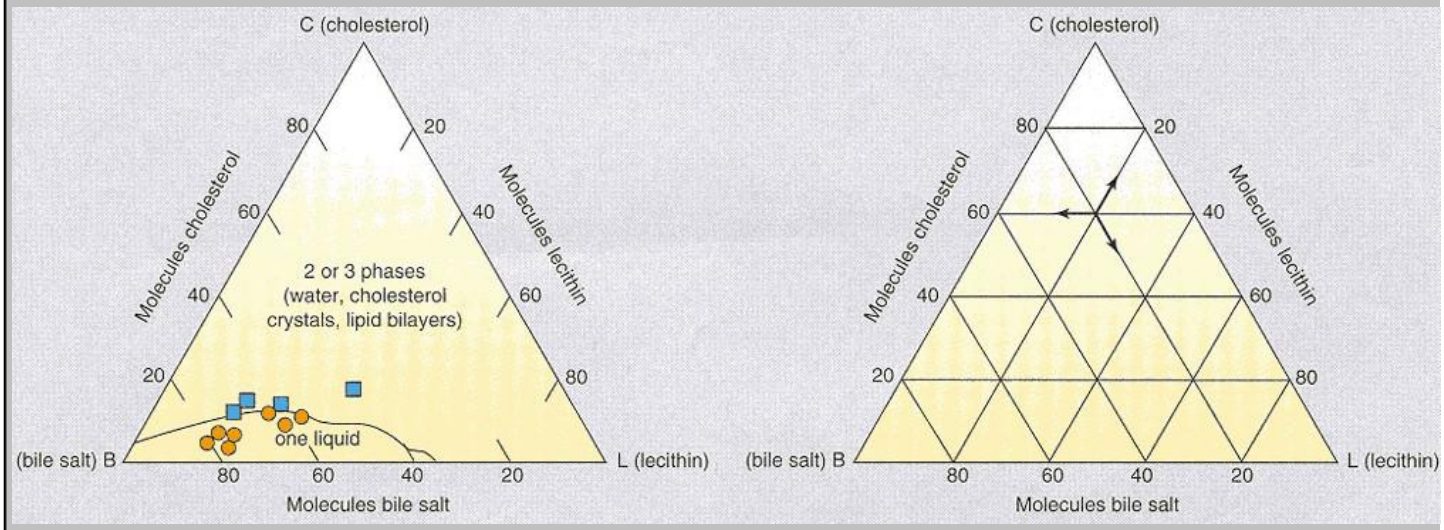


Diagram of the physical states of mixtures of 90% water and 10% lipid. The 10% lipid is made up of bile acids, lecithin, and cholesterol, and the triangle represents all possible ratios of the three lipid constituents. Each point within the triangle corresponds to a particular composition of the three components, which can be read off the graph as indicated; each point on one of the sides corresponds to a particular composition of just two components. The left triangle contains the composition of gallbladder bile samples from patients without stones (red ■). Lithogenic bile has a composition that falls outside the "one liquid" area in the lower left corner. Redrawn from Hofmann, A. F., and Small, D. M. *Annu. Rev. Med.* 18:362, 1967.

During triacylglycerol digestion, free fatty acids and monoacylglycerols are released at the surface of fat emulsion droplets. In contrast to triacylglycerols, which are water-insoluble, free fatty acids and monoacylglycerols are slightly water-soluble, and molecules at the surface equilibrate with those in solution. The latter in turn become incorporated into bile acid micelles. Thus the products of triacylglycerol hydrolysis are continuously transferred from emulsion droplets to the micelles (see Figure 26.27).

Micelles provide the major vehicle for moving lipids from the intestinal lumen to the cell surface where absorption occurs. Because the fluid layer next to the cell surface is poorly mixed, the major transport mechanism for solute

CLINICAL CORRELATION 26.6

A-β-Lipoproteinemia

A-β-lipoproteinemia is an autosomal recessive disorder characterized by the absence of all lipoproteins containing apo-β-lipoprotein, that is, chylomicrons, very low density lipoproteins (VLDLs), and low density lipoproteins (LDLs). Serum cholesterol is extremely low. This defect is associated with severe malabsorption of triacylglycerol and lipid-soluble vitamins (especially tocopherol and vitamin E) and accumulation of apo B in enterocytes and hepatocytes. The defect does not appear to involve the gene for apo B, but rather one of several proteins involved in processing of apo B in liver and intestinal mucosa, or in assembly and secretion of triacylglycerol-rich lipoproteins, that is, chylomicrons and VLDLs from these tissues, respectively.

Kane, J. P. Apolipoprotein B: structural and metabolic heterogeneity. *Annu. Rev. Physiol.* 45:673, 1983; and Kane, J. P., and Havel, R. J. In: C. R. Scriver, A. L. Beaudet, W. S. Sly, and D. Valle (Eds.), *The Metabolic and Molecular Bases of Inherited Disease*, Vol. 1, 7th ed. New York: McGraw-Hill, 1995, p. 1853.

flux across this "unstirred" fluid layer is diffusion down the concentration gradient. With this type of transport mechanism, the delivery rate of nutrients at the cell surface is proportional to their concentration difference between luminal bulk phase and cell surface. Obviously, the unstirred fluid layer presents problems for sparingly soluble or insoluble nutrients, in that reasonable delivery rates cannot be achieved. Bile acid micelles overcome this problem for lipids by increasing their effective concentration in the unstirred layer. The increase in transport rate is nearly proportional to the increase in effective concentration and can be 1000-fold over that of individually solubilized fatty acids, in accordance with the different solubility of fatty acids as micelles or as individual molecules. This relationship between flux and effective concentration holds because the diffusion constant, another parameter that determines the flux, is only slightly smaller for the mixed micelles as compared to lipid molecules free in solution. Thus efficient lipid absorption depends on the presence of sufficient bile acids to "solubilize" the ingested and hydrolyzed lipids in micelles. In the absence of bile acids, the absorption of triacylglycerols does not completely stop, although the efficiency is drastically reduced. The residual absorption depends on the slight water solubility of the free fatty acids and monoacylglycerols. Unabsorbed lipids reach the lower intestine where a small part can be metabolized by bacteria. The bulk of unabsorbed lipids, however, is excreted with the stool (this is called **steatorrhea**).

Micelles also transport cholesterol and the lipid-soluble vitamins A, D, E, and K through the unstirred fluid layers. Bile acid secretion is absolutely essential for their absorption.

Most Absorbed Lipids Are Incorporated into Chylomicrons

Uptake of lipids by the epithelial cells occurs by diffusion through the plasma membrane. Absorption is virtually complete for fatty acids and monoacylglycerols, which are slightly water-soluble. It is less efficient for water-insoluble lipids. For example, only 30–40% of the dietary cholesterol is absorbed.

Within the intestinal cells, the fate of absorbed fatty acids depends on chain length. **Fatty acids** of **medium chain length** (6–10 carbon atoms) pass through the cell into the portal blood without modification. Long-chain fatty acids (>12 carbon atoms) become bound to a cytosolic, specifically **intestinal fatty acid-binding protein** (I-FABP) and are transported to the endoplasmic reticulum, where they are resynthesized into triacylglycerols. Glycerol for this process is derived from the absorbed 2-monoacylglycerols and, to a minor degree, from glucose. The resynthesized triacylglycerols form lipid globules to which surface-active phospholipids and special proteins, termed **apolipoproteins**, adsorb. The lipid globules migrate within membrane-bounded vesicles through the Golgi to the basolateral plasma membrane. They are finally released into the intercellular space by fusion of the vesicles with the basolateral plasma membrane. Because the lipid globules can be several micrometers in diameter and because they leave the intestine via lymph vessels, they are called **chylomicrons** (chyle = milky lymph that is present in the intestinal lymph vessels, lacteals, and the thoracic duct after a lipid meal; chyle is derived from the Greek *chylos*, which means juice). The intestinal apolipoproteins are distinctly different from those of the liver and are designated A-1 and B. **Apolipoprotein B** is essential for chylomicron release from enterocytes (see Clin. Corr. 26.6).

While dietary medium-chain fatty acids reach the liver directly with the portal blood, the long-chain fatty acids bypass the liver by being released in the form of chylomicrons into the lymphatics. The intestinal lymph vessels drain into the large body veins via the thoracic duct. Blood from the large veins first reaches the lungs and then the capillaries of the peripheral tissues, including adipose tissue and muscle, before it comes into contact with the liver. Fat and

muscle cells in particular take up large amounts of dietary lipids for storage or metabolism. The bypass of the liver may have evolved to protect this organ from a lipid overload after a meal.

The differential handling of medium- and long-chain fatty acids by intestinal cells can be specifically exploited to provide the liver with high-caloric nutrients in the form of fatty acids. Short- and medium-chain fatty acids are not very palatable; however, triacylglycerols synthesized from these fatty acids are quite palatable and can be used as part of the diet.

26.7— Bile Acid Metabolism

All bile acids are synthesized within the liver from cholesterol but can be modified by bacterial enzymes in the intestinal lumen. **Primary bile acids** synthesized by the liver are **cholic** and **chenodeoxycholic** (or chenic) acid. The **secondary bile acids** are derived from the primary bile acids by bacterial dehydroxylation in position 7 of the ring structure, resulting in **deoxycholate** and **lithocholate,** respectively (Figure 26.35).

Primary and secondary bile acids are reabsorbed by the intestine into the portal blood, taken up by the liver, and then resecreted into bile. Within the liver, primary as well as secondary bile acids are linked to either glycine or

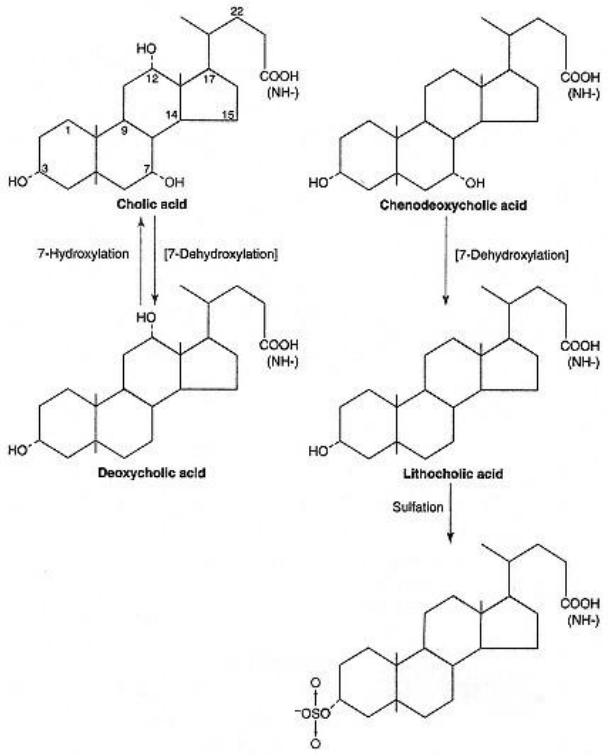


Figure 26.35
Bile acid metabolism in the rat.
Green and black arrows indicate reactions catalyzed by liver enzymes; red arrows indicate those of bacterial enzymes within the intestinal lumen. (NH—), glycine or taurine conjugate of the bile acids.

taurine via an isopeptide bond. These derivatives are called **glyco-** and **tauro-conjugates,** respectively, and constitute the forms that are secreted into bile. With the conjugation, the carboxyl group of the unconjugated acid is replaced by an even more polar group. The p*K* values of the carboxyl group of glycine and of the sulfonyl group of taurine are lower than that of unconjugated bile acids, so that conjugated bile acids remain ionized over a wider pH range (see Table 26.10). The conjugation is partially reversed within the intestinal lumen by hydrolysis of the isopeptide bond.

The total amount of conjugated and unconjugated bile acids secreted per day by the liver is 16–70 g for an adult. As the total body pool is only 3–4 g, bile acids have to recirculate 5–14 times each day between the intestinal lumen and the liver. Reabsorption of bile acids is important to conserve the pool. Most of the uptake is probably by passive diffusion along the entire small intestine. In addition, the lower ileum contains a specialized **Na^+-bile acid cotransport system** for concentrative reuptake. Thus during a meal, bile acids from the gallbladder and liver are released into the lumen of the upper small intestine, pass with the chyme down the small intestinal lumen, are reabsorbed by the epithelium of the lower small intestine into the portal blood, and are then extracted from the portal blood by the liver parenchymal cells. The process of secretion and reuptake is referred to as the **enterohepatic circulation** (Figure 26.36). Reabsorption of bile acids by the intestine is quite efficient as only about 0.5 g of bile acids escapes reuptake each day and is secreted with the feces. Serum levels of bile acids normally vary with the rate of reabsorption and therefore are highest during a meal.

Cholate, deoxycholate, chenodeoxycholate, and their conjugates continuously participate in the enterohepatic circulation. In contrast, most of the **lithocholic acid** that is produced by bacterial enzymes is sulfated during the next passage through the liver. The sulfate ester of lithocholic acid is not a substrate for the bile acid transport system in the ileum and therefore is excreted in the feces.

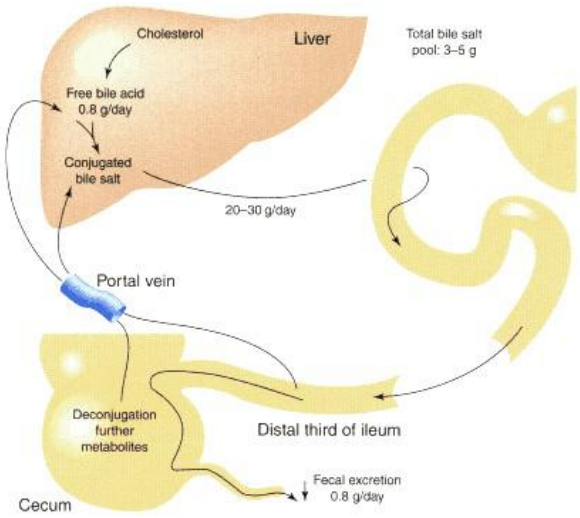


Figure 26.36
Enterohepatic circulation of bile acids.
Redrawn from Clark, M. L., and Harries, J. T. In: I. McColl and G. E. Sladen (Eds.), *Intestinal Absorption in Man*. New York: Academic Press, 1975, p. 195.

Bibliography

Cristofaro, E., Mottu, F., and Wuhrmann, J. J. Involvement of the raffinose family of oligosaccharides in flatulence. In: H. L. Sipple and K. W. McNutt (Eds.), *Sugars in Nutrition*. New York: Academic Press, 1974, p. 314.

Field, M., and Semrad, C. E. Toxigenic diarrheas, congenital diarrheas, and cystic fibrosis. *Annu. Rev. Physiol.* 55:631–655, 1993.

Hediger, M. A., and Rhoads, D. B. Molecular physiology of sodium-glucose cotransporters. *Physiol. Rev.* 74:993–1026, 1994.

Johnson, L. R. (ed.-in-chief). *Physiology of the Gastrointestinal Tract,* Vols. 1 and 2, 2nd and 3rd eds. New York: Raven Press, 1987, 1994.

Mathews, D. M. *Protein Absorption*. New York: Wiley-Liss, 1991.

Pandol, S., and Raybauld, H. E. Integrated response to a meal. The Undergraduate Teaching Project in Gastroenterology and Liver Disease. Unit #29. American Gastroenterological Assoc. Timonium, MD: Milner-Fenwick, Inc., 1995.

Porter, R., and Collins, G. M. *Brush Border Membranes*. Volume 95, Ciba Foundation Symposium. London: Pitman, 1983.

Schultz, S. G. (section ed.). *Handbook of Physiology. Section 6: The Gastrointestinal System. Vol. IV. Intestinal Absorption and Secretion* (M. Field and R. A. Frizzell, Eds.). Bethesda, MD: American Physiological Society, 1991

Sleisenger, M. H. Malabsorption and nutritional support. *Clin. Gastroenterol.* 12:323, 1983.

Thomson, A. B. R., Schoeller, C., Keelan, M., Smith, L., and Clandinin, M. T. Lipid absorption: passing through the unstirred layers near the brush border membrane, and beyond. *Can. J. Physiol. Pharmacol.* 71:531–555, 1993.

Questions

J. Baggott and C. N. Angstadt

Refer to the following for Questions 1–5:

A. liver

B. pancreas

C. spleen

D. stomach

E. none of the above

1. Has no role in digestion.

2. Synthesizes an essential emulsifier of lipids.

3. Participates in a nonessential manner in protein digestion.

4. Transports HCO_3^- from the cytoplasm across the contraluminal plasma membrane.

5. Site of chymotrypsinogen synthesis.

6. Active forms of enzymes that digest food may normally be found in all of the following EXCEPT:

   A. in soluble form in the lumen of the stomach.

   B. in the saliva.

   C. attached to the luminal surface of the plasma membrane of intestinal epithelial cells.

   D. dissolved in the cytoplasm of intestinal epithelial cells.

   E. in zymogen granules of pancreatic exocrine cells.

7. Histamine is a physiologically important secretagogue of:

   A. amylase by the salivary glands.

   B. HCl by the stomach.

   C. gastrin by the stomach.

   D. hydrolytic enzymes by the pancreas.

   E. $NaHCO_3$ by the pancreas.

8. The contraluminal membranes of small intestinal epithelial cells contain:

   A. aminopeptidases.

   B. Na^+,K^+–ATPase.

   C. disaccharidases.

   D. GLUT5.

   E. Na^+–monosaccharide transport (SGLT1).

9. Oral administration of large amounts of tyrosine could be expected to interfere with the intestinal absorption of:

   A. leucine.

   B. lysine.

   C. glycine.

   D. aspartate.

   E. none of the above.

10. Which of the following has two carboxyl groups essential for peptidase activity?

   A. carboxypeptidase

   B. chymotrypsin

   C. elastase

   D. pepsin

   E. trypsin

11. Starch digestion is more efficient after heating the starch with water because heating:

   A. hydrates the starch granules, making them more susceptible to pancreatic amylase.

   B. converts α-1,4 links to β-1,4 links, which are more susceptible to attack by mammalian amylases.

   C. partly hydrolyzes α-1,6 links.

   D. converts the linear amylose to branched amylopectin, which resembles glycogen.

   E. inactivates amylase inhibitors, which are common in the tissues of starchy plants.

12. In the cytoplasm of intestinal cells:

   A. all di- and tripeptides are hydrolyzed.

   B. aminopeptidases are especially active.

   C. during the neonatal period ingested proteins may be found.

   D. most disaccharides are hydrolyzed.

   E. raffinose and related sugars are degraded to yield hydrogen, methane, and carbon dioxide.

13. In the digestion and absorption of triacylglycerols:

   A. a pancreatic lipase initiates the process.

   B. an important colipase is activated by tryptic hydrolysis.

   C. hydrolysis occurs in the interior of the lipid droplets.

   D. most of the triacylglycerol hydrolysis is carried out by a lipase of gastric origin.

   E. efficiency is greatly increased if bile acids are absent.

14. Micelles:

A. are the same as emulsion droplets.

B. form from bile acids at all bile acid concentrations.

C. although they are formed during lipid digestion, do not significantly enhance utilization of dietary lipid.

D. always consist of only a single lipid species.

E. are essential for the absorption of vitamins A and K.

15. In the metabolism of bile acids:

A. the liver synthesizes the primary bile acids, cholic, and deoxycholic acids.

B. secondary bile acids are produced by conjugation of primary bile acids to glycine or taurine.

C. physiologically active bile acids are formed from primary bile acids by intestinal bacteria.

D. daily bile acid secretion by the liver is approximately equal to daily bile acid synthesis.

E. conjugation reduces the polarity of bile acids, enhancing their ability to interact with lipids.

Answers

1. C The spleen has no role in the digestion of food, though it does participate in other degradation processes.

2. A Bile acids are synthesized in the liver and are stored in the gallbladder (p. 1057).

3. D Loss of the stomach function can be compensated for by the intestinal processes (p. 1057).

4. D This occurs in the parietal (oxyntic) cells during HCl secretion (p. 1069).

5. B

6. E Zymogen granules contain inactive proenzymes or zymogens, which are not activated until after release from the cell (p. 1060).

7. B Stimulation of H_2 receptors of the stomach causes HCl secretion (p. 1061).

8. B. Only the contraluminal surface contains the Na^+,K^+–ATPase. All other activities are associated with the luminal surface (Table 26.5, p. 1064).

9. A Tyrosine shares a transport system with Val, Leu, Met, Phe, and Ile (p. 1073).

10. D The carboxylic acid groups are involved in the mechanism that depends on an acid pH (p. 1070).

11. A α-Amylase attacks hydrated starch more readily than unhydrated; heating hydrates the starch granules (p. 1074).

12. C They are taken up by pinocytosis (p. 1073).

13. B This colipase is required to overcome bile acid inhibition of pancreatic lipase, the major enzyme of lipid digestion. The colipase is secreted by the pancreas as a procolipase and must be activated by tryptic cleavage (p. 1078). A: Lipid digestion is initiated in the stomach by acid-stable lipase (p. 1078).

14. E The lipid-soluble vitamins must be dissolved in mixed micelles as a prerequisite for absorption (p. 1082). A: Micelles are of molecular dimensions and are highly ordered structures; emulsion droplets are much larger and are random (p. 1078, Figure 26.28; p. 1080, Figure 26.34). B: Micelle formation occurs only above the critical micellar concentration (CMC); below that concentration the components are in simple solution (p. 1079, Figure 26.32). C: See item 13. D: Micelles may consist of only one component, or they may be mixed (p. 1079.)

15. C The primary bile acids (cholic and chenodeoxycholic acids) are synthesized in the liver. In the intestine they may be dehydroxylated by bacteria to form the secondary bile acids—deoxycholate and lithocholate. Only a small fraction of the bile acid escapes reuptake; this must be replaced by synthesis. Both are reabsorbed and recirculated (enterohepatic circulation). Both are conjugated to glycine or taurine, increasing their polarity (p. 1078).