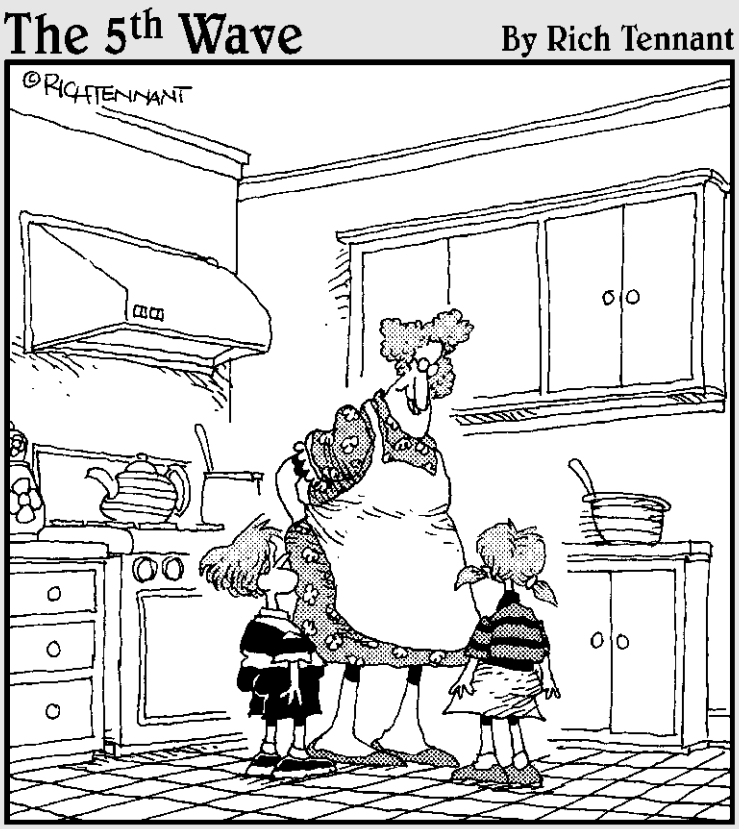
Part II

The Meat of Biochemistry: Proteins



"Who wants to help Grandma make her famous gingerbread man cookies? You kids get the flour, eggs, and sugar, and I'll get the amino acids and enzymes."

In this part . . .

We focus, not surprisingly, on proteins, starting with amino acids, protein's building blocks. After that we detail the processes of amino acid sequencing and the various kinds of protein structure. We finish up this part by discussing enzyme kinetics, covering catalysts (which speed up reactions) and inhibitors (which — can you guess? That's right — slow them down).

Chapter 4

Amino Acids: The Building Blocks of Protein

In This Chapter

- Looking at the structure and properties of amino acids
- Examining the common amino acids
- Finding out about the interactions of amino acids
- Seeing how amino acids combine

All cells contain thousands of types of proteins, and amino acids are the building blocks of these proteins. The sequential order, number, and chemical identity of the amino acids in the protein determine the structure of the protein as well as how the protein functions. That's why it's important to understand the chemical properties of amino acids before you can understand the behavior of proteins.

Amino acids are relatively simple molecules containing both an amine group and an acid group. The biologically important amino acids are the α-amino acids that have the amine and acid groups attached to the same carbon atom. There are more than 100 known natural amino acids; however, only 20 of them are used in protein synthesis. Francis Crick (who with James Watson determined the structure of DNA) labeled this set of amino acids the *magic 20*. Other amino acids are found in certain proteins, but in almost all cases these additional amino acids result from the modification of one of the magic 20 after the protein formed.

In this chapter, we examine the structure and properties of amino acids, especially the more common ones, and show how they interact and combine.

General Properties of Amino Acids

Like any organic compound, the properties of the molecules are largely determined by the functional groups present. In biological systems, the important properties of the amino acids include the following:

- **They can join to form proteins.** The average molecular weight of an amino acid is about 135. Proteins have molecular weights ranging from about 6,000 to several million. Thus, a large number of amino acids must be joined together to produce a protein.
- **They all have both an acid and a base.** The α-carbon (end carbon) not only has an amine group ($-NH_2$) and a carboxylic acid group (–COOH), but also two additional groups: a hydrogen atom and an R– group. The side chain, R group, identifies the amino acid.
- **They all have variations in what part of the structure is protonated depending on the pH of the solution and the structure of the rest of the molecule.**
- **They all, except glycine, have a chiral nature, influencing the reactions that the compound will undergo.**

Amino acids are positive and negative: The zwitterion formation

The presence of both an acid and a base (amine) in the same molecule leads to an interaction between the two. This interaction results in a transfer of a hydrogen ion from the acid portion to the base portion of the molecule. An amino acid with both positive and negative regions is a called a *zwitterion.* The net charge of the zwitterion is 0. This leaves the acid end of the amino acid with a negative charge ($-COO^-$) and a positive charge at the base end ($-NH_3^+$). The deprotonated portion (portion that has lost a hydrogen ion) is a carboxylate group, and the protonated group (group that has gained a hydrogen ion) is an ammonium group. The presence of a charge on the amino acid makes them water-soluble. Figure 4-1 shows zwitterion formation.

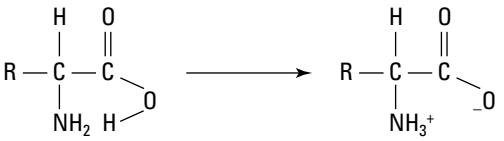


Figure 4-1: Zwitterion formation.

The unionized amino acid molecule shown in Figure 4-1 does not actually exist. However, many books and instructors draw the unionized form as a simplification, as if the ionization did not occur.

Protonated? pH and the isoelectric point

How amino acids react, because of their acid-base nature, is dependent on the pH of the solution in which they are found. Here we look at some of the implications of this pH dependency. The zwitterion is the predominant form at a particular pH, which is designated the *isoelectric point* (pI). The isoelectric point is midway between the two different pK_a values. Under most physiological conditions, isolated amino acids exist in their zwitterion form (Figure 4-2 (a)). Pure amino acids are also in the zwitterion form — and, for this reason, are *ionic solids.*

- **At a pH below the isoelectric point, some of the carboxylate groups will be protonated.** (See Figure 4-2 (b).) The pH required to cause this protonation of the carboxylate group depends on the K_a of the acid. For this reason the pK_a of the carboxylic acid group is important. Typical values are between 1 and 3. If, for example, the pK_a is 2.5, at a pH of 2.5, 50 percent of the carboxylate groups will be protonated. The net charge of the protonated form is +1.
- **At a pH above the isoelectric point, some of the ammonium groups will be deprotonated.** (See Figure 4-2 (c).) The pH required to cause this deprotonation of the ammonium group depends on the K_a of the ammonium group. For this reason, the pK_a of the ammonium group is important. Typical values are between 8 and 11. If, for example, the pK_a is 10, at a pH of 10, 50 percent of the ammonium groups will be deprotonated. The net charge of the protonated form is –1.

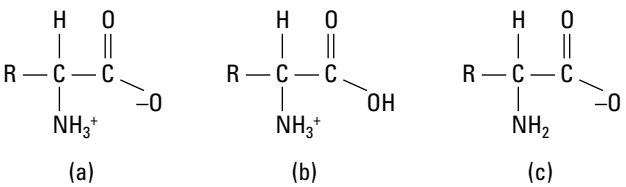


Figure 4-2: (a) Zwitterion form, (b) protonated form, and (c) deprotonated form.

Some of the side chains are also acidic or basic. In these cases, an additional pK_a becomes significant in the reactions of these molecules and will obviously complicate the pH behavior of the amino acid.

Asymmetry: Chiral amino acids

In a typical α-amino acid, four different groups are attached to the α-carbon (–COOH, $–NH_2$, –R, and –H). This makes the α-carbon asymmetric or *chiral*. The only exception is the amino acid glycine, where the R– group is a hydrogen atom. The presence of two hydrogen atoms on the α-carbon means that, in the case of glycine, the carbon atom is achiral. Chiral materials are optically active; the different forms affect light in different ways. (See Chapter 3 for more on what makes a molecule chiral.)

The arrangement of the groups around a chiral carbon atom is important. Just as your left hand only fits into your left glove, only certain arrangements of the groups will fit (because of what is called *handedness*).

There are two different forms of the chiral amino acids: the D- and the L- forms. Only the L- forms are constituents of proteins. The D- forms appear in some antibiotics and in the cell walls of certain bacteria. Fischer projections, as we explain in Chapter 3, are commonly used to represent the arrangement about the chiral carbon. Figure 4-3 illustrates some different ways to draw the Fischer projections of the structure of amino acids.

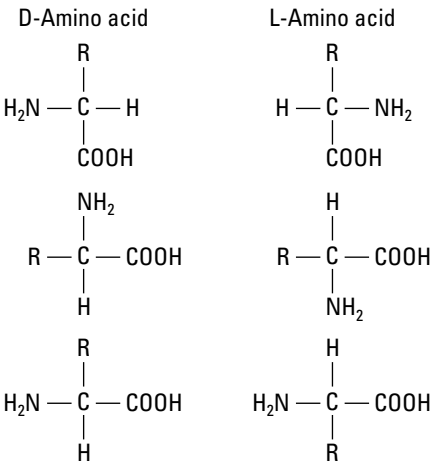


Figure 4-3: Different ways of drawing the Fischer projections of the amino acids.

A few amino acids contain two asymmetric carbon atoms. In these cases, there are four possible isomers. Biological activity is usually limited to only one of these four isomers.

The Magic 20 Amino Acids

Amino acids are subdivided into four subgroups based on the nature of the side chain (groups attached to the the α-carbon) and the general behavior of the amino acid:

- Nonpolar (hydrophobic) and uncharged
- Polar (hydrophilic) and uncharged
- Acidic (polar and charged)
- Basic (polar and charged)

The properties of the side chains are not only important to the behavior of the individual amino acids but also to the properties of the proteins resulting from the combination of certain amino acids.

In the following section we examine the structures of the individual amino acids. It is possible to represent each of the amino acids by either a three-letter or a one-letter abbreviation. Like the chemical symbols for the elements, these are fixed abbreviations. The three-letter abbreviations are easier to relate to the name of the specific amino acid. For example, we use glu for glutamine. The one-letter abbreviations are shorter, but not always related to the name. For example, we use Q for glutamine.

Nonpolar (hydrophobic) amino acids

The nonpolar amino acids are as follows:

- Alanine (ala, A)
- Valine (val, V)
- Leucine (leu, L)
- Isoleucine (ile, I)
- Proline (pro, P)
- Methionine (met, M)
- Phenylalanine (phe, F)
- Tryptophan (trp, W)

Figure 4-4 shows these amino acids.

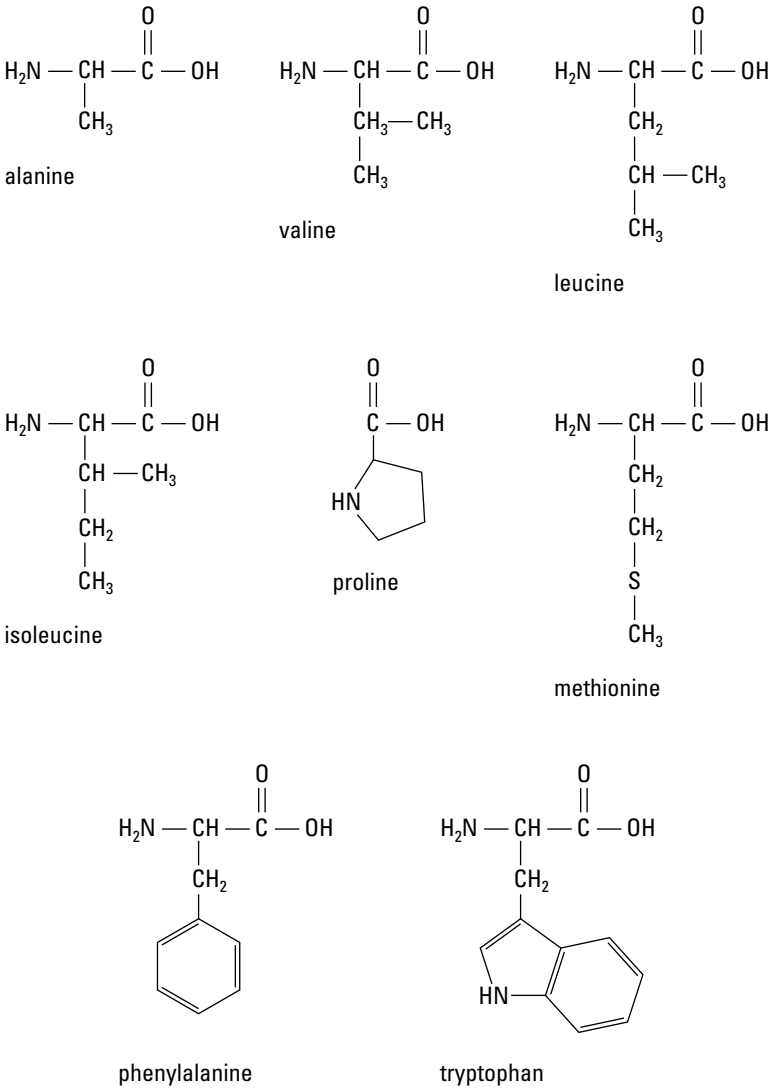


Figure 4-4: Nonpolar amino acids.

Proline has an unusual cyclic structure, which has a significant influence on protein structure. Tryptophan is a borderline case because the –NH from the ring system can interact with water to a limited extent.

Polar and uncharged (hydrophilic) amino acids

The polar and uncharged amino acids, other than glycine, can hydrogen bond to water. For this reason, they are usually more soluble than the nonpolar amino acids. The amino acids in this group are as follows:

- Glycine (gly, G)
- Serine (ser, S)
- Asparagine (asn, N)
- Glutamine (gln, Q)
- Threonine (thr, T)
- Tyrosine (tyr, Y)
- Cysteine (cys, C)

Glycine seems to be an unexpected member of this group. The small size of the R group in the case of glycine leads to the predominance of the amino and carboxylate functional groups, giving glycine's similarity to other amino acids in this group. The amide, alcohol, and sulfhydryl (–SH) groups of the remaining members of this group are very polar and neutral. At very high pH values, the phenolic group on tyrosine ionizes to yield a polar charged group. Figure 4-5 shows these amino acids.

Acidic amino acids

The acidic amino acids are as follows:

- Aspartic acid (asp, D)
- Glutamic acid (glu, E)

In both of these amino acids, the side group contains a carboxylic acid group. This secondary carboxylic acid group is a weaker acid (higher pK_a) than the primary carboxylic acid group. This additional carboxylate group leads to a net –1 charge at a pH where the "normal" zwitterion has a 0 net charge. The carboxylate side chain is important in the interaction of many proteins with metal ions, as *nucleophiles* (an electron-rich group replacing some group attached to a carbon) in many enzymes, and in ionic interactions. Figure 4-6 shows these amino acids.

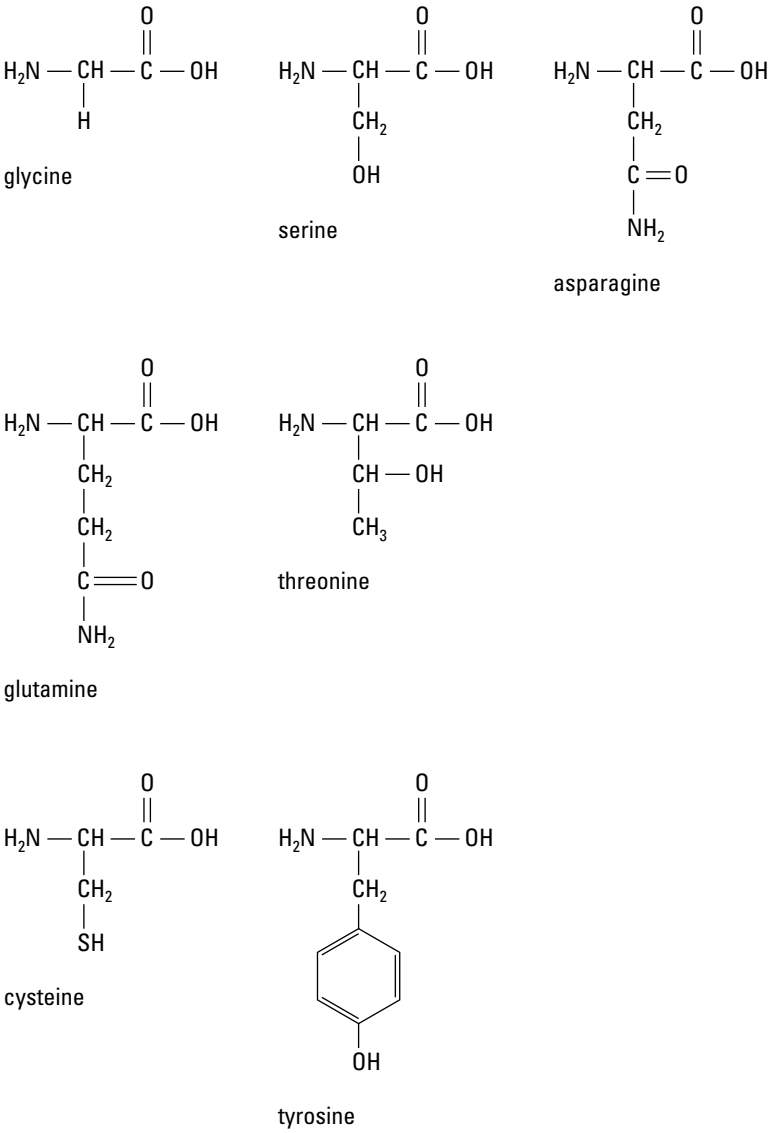


Figure 4-5: Polar and uncharged amino acids.

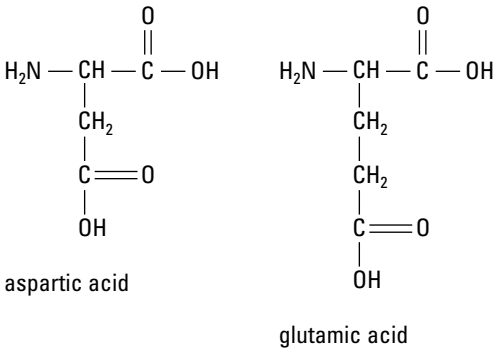


Figure 4-6: Acidic amino acids.

Basic amino acids

The basic amino acids are as follows:

- Histidine (his, H)
- Arginine (arg, R)
- Lysine (lys, K)

All of these are classified as basic amino acids, but dramatic changes in pH can affect their reactivities. This is especially true of histidine.

In all three of these amino acids, there is a basic group capable of accepting a hydrogen ion. In the case of lysine, this is a simple ammonium ion. Arginine forms the guanidinium group. Histidine forms an imidazolium group. As in the case of the acidic side chains, these side chains have a pK_a value. Both arginine and lysine are usually protonated at physiological pH values. In these cases, there is a net +1 charge present. In proteins, this net charge may be part of an ionic interaction. The pK_a of the side chain of histidine is lower than other basic groups. Protonation of histidine becomes significant at much lower pH values. In many proteins, histidine is not protonated, but is important in many enzymes in hydrogen ion transfer processes. Figure 4-7 shows these basic amino acids.

Lest We Forget: Rarer Amino Acids

In a few cases, an amino acid may undergo modification once it is incorporated into a protein. Collagen and gelatin, for example — proteins present in higher vertebrates — contain hydroxylysine and hydroxyproline. These two amino acids contain an additional –OH group on the side chain.

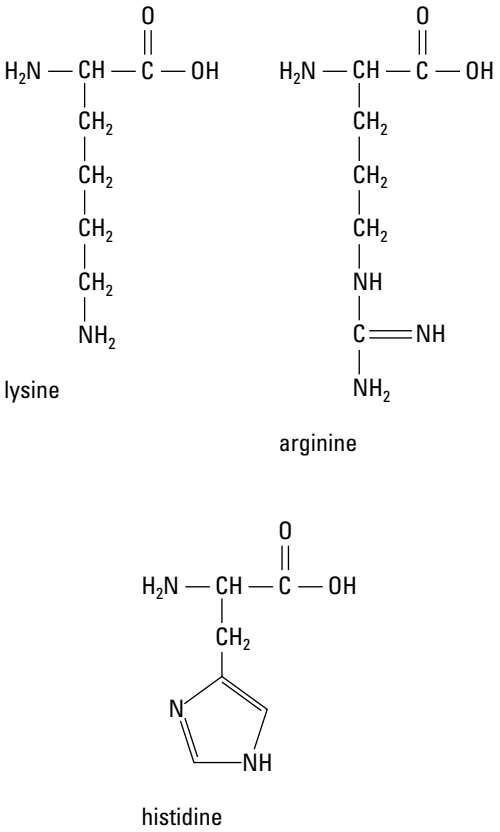


Figure 4-7: Basic amino acids.

Certain amino acids do not occur in proteins. The neurotransmitter γ-aminobutyric acid — GABA — is one example. Citrulline is the amino acid that serves as a precursor of arginine. Ornithine, homocysteine, and homoserine are important as metabolic intermediates. Figure 4-8 shows a couple of these amino acids.

Rudiments of Amino Acid Interactions

Amino acids are the ingredients used in the recipe in making a protein. Just as the individual ingredients in a recipe lead to distinct characteristics of what eventually shows up on the dinner table, the amino acids present contribute properties to proteins. And just as you cannot replace the flour in a recipe with pepper, you generally cannot replace one amino acid in a protein with another. In both cases, the final product will be different. In the next section, we show you some of the ways that amino acids interact. These interactions set the stage for our discussion of bonding among the amino acids.

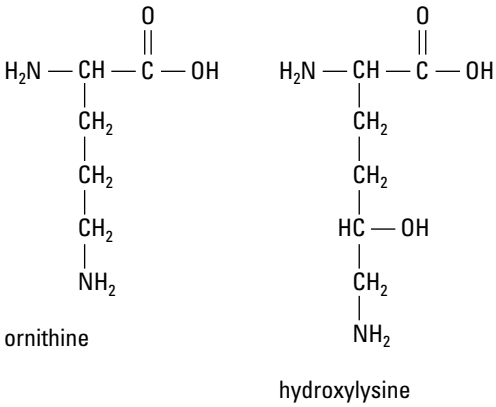


Figure 4-8: Two of the less common amino acids.

Intermolecular forces: How an amino acid reacts with other molecules

Amino acids can interact with other molecules — and we mean *any* other molecules, including fluids, other amino acids, and other biological molecules — in a variety of ways. We cover intermolecular forces in general in Chapter 3, but in this section we show you how they play out when amino acids are involved. The carboxylic acid and amine parts of the amino acids define much of the reactivity of the molecule, but the side chains can also interact with other molecules. There are three general ways in which they can interact.

- **Hydrophobic interactions:** The nonpolar side groups are hydrophobic and are attracted to each other through London dispersion forces. Nonpolar groups tend to clump together and exclude not only water but also all other types of side chains.
- **Hydrophilic reactions:** The polar and uncharged side groups are hydrophilic. The presence of a number of these groups increases the solubility of a protein. These groups hydrogen bond not only to water but also to each other. Polar groups tend to interact strongly and "push" the nonpolar groups out.
- **Ionic interactions:** The presence of acidic or basic side chains leads to ionic charges — opposite charges attract. A carboxylate group from one side chain is attracted to the ammonium ion of another side chain through an ionic interaction. This ionic bond is very strong.

The amino acid cysteine can interact with a second cysteine molecule through a different type of interaction (Figure 4-9). The mild oxidation of two cysteine sulfhydryl groups leads to the formation of cystine. A disulfide linkage joins the two amino acids with a covalent bond. Mild reduction can reverse this process.

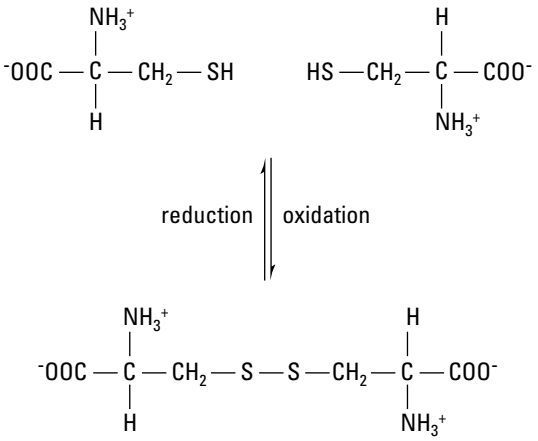


Figure 4-9: Joining two cysteine molecules to form cystine.

A hair perm utilizes an oxidation reduction reaction creating disulfide linkages. The greater the number of disulfide linkages, the curlier the hair!

Altering interactions by changing an amino acid's pH

As we discuss in Chapter 3, the function of many substances, especially biochemical ones, is dependent on pH. If you change the pH, you change some of the interactions. In this section we show how those changes affect interactions involving amino acids.

Just like any other molecule, an amino acid has two or three functional groups, depending on the amino acid. Those functional groups include those with oxygen and sulfur, those with nitrogen, and those with phosphorus. A change in pH affects one to three of those functional groups in terms of interactions. So if an amino acid has a functional group that changes from a dipole-dipole interaction to an ionic interaction.

One example of the dipole-dipole to ionic interaction change is the process of milk curdling. If you add an acid to milk, it coagulates. Casein has an isoelectric point at 4.6 pH, so that adding an acid causes the formation of ionic bonds among the molecules. This works against the dipole-dipole interactions with water, so that the protein precipitates.

The pK_a values for the various groups present in the different amino acids are shown in Table 4-1. If the pH of the solution matches one of these values, then half the species is in the protonated form and half is in the deprotonated

form. At a lower pH, more than half is protonated, whereas at a higher pH more than half is deprotonated.

The pH dependence of the protonation of amino acids aids in their separation and identification. Because the amino acids use the carboxylic acid and amine ends when they join to form a protein, only the pK_a values of the side chains are important in additional interactions and reactions.

Table 4-1 pK_a Values for the Amino Acids

Amino acid	*pK_a –COOH*	*pK_a –NH_3^+*	*pK_a R group*
Alanine	2.35	9.69	
Arginine	2.17	9.04	12.48
Asparagine	2.02	8.8	
Aspartic acid	2.09	9.82	3.86
Cysteine	1.71	10.78	8.33
Glutamic acid	2.19	9.67	4.25
Glutamine	2.17	9.13	
Glycine	2.34	9.6	
Histidine	1.82	9.17	6.0
Isoleucine	2.36	6.68	
Leucine	2.36	9.60	
Lysine	2.18	8.95	10.53
Methionine	2.28	9.21	
Phenylalanine	1.83	9.13	
Proline	1.99	10.60	
Serine	2.21	9.15	
Threonine	2.63	10.43	
Tryptophan	2.38	9.39	
Tyrosine	2.20	9.11	10.07
Valine	2.32	9.62	

Combining Amino Acids: How It Works

A *protein* is a string of at least 150 amino acids (residues) joined. We cover the fundamentals about protein creation in Chapter 5, but before you dive into that topic, this section gives you a solid understanding of how two amino acids join together in the first place, and how additional amino acids link onto the chain gang. The process is reversible (as in digestion).

When drawing the chemical structures of amino acids and their bonds, the standard convention is to first draw the structures from the ammonium group of the first amino acid (the N-terminal residue), starting at the left, and continuing the drawing to the right, ending with the carboxylate group (C-terminal residue) of the last amino acid.

The peptide bond and the dipeptide

One of the most important types of bonds in all of biochemistry is the *peptide bond.* As you will see, it is this type of bond that will be used in the synthesis of proteins. The interaction of two amino acids at the body's pH results in the formation of a peptide bond as illustrated in Figure 4-10.

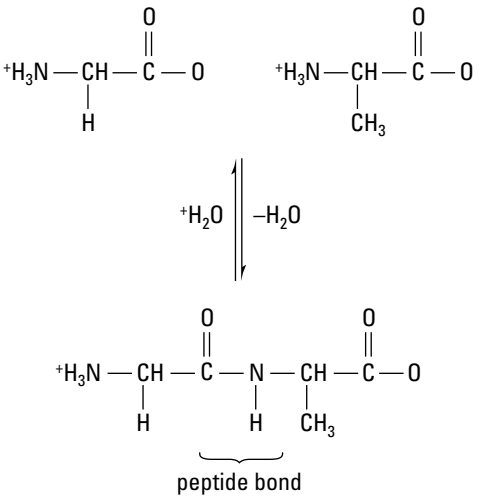


Figure 4-10: The formation of a peptide bond.

The two residues react to expel a water molecule, the same dehydration you used so much in organic chemistry. The reverse of this condensation reaction is hydrolysis. The resultant amide group is a peptide bond. The presence of two amino acid residues means the product is a dipeptide.

The peptide bond is a *flat* (planar) structure. It is stabilized by our old organic friend, *resonance.* Figure 4-11 illustrates the stabilization. The resonance increases the polarity of the nitrogen and oxygen. This increase in polarity leads to hydrogen bonds that are much stronger than most other hydrogen bonds. The double bond character between the carbon and the nitrogen restricts rotation about this bond. That's why the peptide bond is planar.

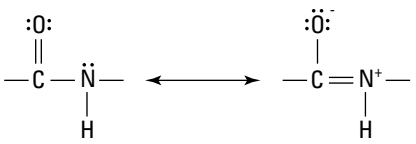


Figure 4-11: Resonance stabilization of a peptide bond.

Tripeptide: adding an amino acid to a dipeptide

A repetition of the process illustrated in Figure 4-10 joins a third amino acid to produce a *tripeptide.* For example, combining glycine, alanine, and serine yields the illustration in Figure 4-12. Notice that everything begins with the N-terminal residue and ends with the C-terminal residue. (You could designate this tripeptide as *gly-ala-ser* using the three letter abbreviations.)

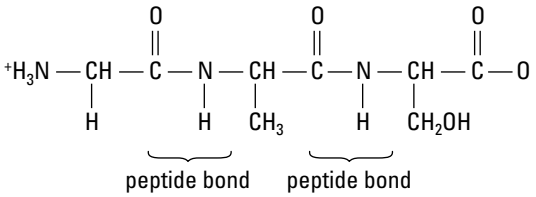


Figure 4-12: A tripeptide.

The repetition of the process of linking amino acids hundreds or thousands of times produces a protein. In the next chapter, we cover that topic in full.

Chapter 5

Protein Structure and Function

In This Chapter

- Finding out about the structure of proteins
- Understanding amino acid sequencing in proteins
- Learning about applications of protein sequencing

In Chapter 4, we show you how amino acids combine through the use of a peptide bond, and we mention there that if at least 150 or so amino acids join hands, they rise to the rank of a protein. However, distinguishing an amino acid chain as a protein isn't exactly simple — just as written English is an extremely diverse set of words made by combining letters from an alphabet of just 26 letters, proteins are an extremely diverse set of biochemicals made by combining 20 different amino acids.

In this chapter, we show you more about these proteins, including the four types of protein structure that determine a protein's function and the sequence of amino acids in a particular protein.

There are two general categories of proteins:

- **Fibrous proteins** are found only in animals. They usually serve as structural entities — for example, connective tissue, tendons, and muscle fiber. They are normally insoluble in water.
- **Globular proteins** usually do not serve a structural function — they act as transporters, like hemoglobin, and are often enzymes. They are usually water-soluble.

Proteins are utilized in living organisms in a number of ways, such as:

- **Structure:** Skin and bone contain collagen, a fibrous protein.
- **Catalysis:** These proteins, called enzymes, allow reactions to occur in the organism under mild conditions and with great specificity.
- **Movement:** Proteins make up a large protein of muscle fiber and help in the movement of various parts of our bodies.

- **Transport:** These proteins transport small molecules through the organism. Hemoglobin, the protein that transports oxygen to the cells, is a transport protein.
- **Hormones:** Proteins called hormones help regulate cell growth.
- **Protection:** Proteins called antibodies help rid the body of foreign proteins.
- **Storage:** These protein help store other substance in the organism. For example, iron is stored in the liver in a complex with the protein ferritin.
- **Regulation:** These proteins help mediate cell responses, such as the protein rhodopsin, found in the eye and involved in the vision process.

The function that a particular protein assumes is, in many cases, directly related to the structure of that protein. Proteins may have as many as four levels of structure (key word being *levels,* not different structures), each of which places the components into a position where these intermolecular forces can interact most advantageously. The levels are simply labeled primary, secondary, tertiary, and quaternary. Primary is the most fundamental level that all proteins have, and quaternary is the most specific level that only some proteins have. Intermolecular forces themselves are important to the function of a protein, of course, but the arrangement of the molecules is even more significant.

If present, the secondary, tertiary, and quaternary structures of a protein may be destroyed — in a number of ways:

- Heating (cooking) can break hydrogen bonds.
- Changing the pH can protonate or deprotonate the molecule and interrupt ionic interactions.
- Reducing agents can break disulfide linkages.

In some cases, the process may be reversible.

Primary Structure: The Structure Level All Proteins Have

The primary structure of a protein is simply the sequence of amino acids comprising the molecule. The primary structure of a protein is the amino acid sequence within the molecule. All proteins have a primary structure, because all proteins by definition consist of a sequence of amino acids. The primary structure serves as the foundation upon which all higher levels of protein structure build.

Next we take a look at how a protein is assembled from its building blocks, the amino acids.

Building a protein: Outlining the process

During the synthesis of a protein, the chain of amino acids is built one link at a time, roughly as follows:

1. **The transfer RNA (tRNA) molecule transfers specific amino acids to the mitochondria of the cell to connect to the growing chain.**
2. **Each amino acid joins to the chain through the formation of a peptide bond.** (See Chapter 4 for more on peptide bonds.)
3. **The first peptide bond joins two amino acids to form a dipeptide.**
4. **The second peptide bond joins three amino acids to produce a tripeptide.**
5. **This process continues hundreds, if not thousands, of times to produce a polypeptide — a protein.**

When two or more amino acids combine, a molecule of water is removed. What remains of each amino acid is called a *residue.* They lack a hydrogen atom on the amino group, or an –OH on the carboxyl group, or both.

The cell's DNA ultimately controls the sequence of amino acids. This information goes from the DNA to the messenger RNA (mRNA), which serves as the template for the creation of the primary structure of the protein. It is necessary to supply energy, as we will see later, to synthesize the protein.

Organizing the amino acids

One end of the primary structure has an amino group, and the other end has a carboxylate group. By convention, the end with the amino group is considered the "beginning" of the protein. Drawing, naming, numbering, and other treatments of the primary structure always begin with the amino end (called the *N-terminal*) and stop with the carboxylate end (the *C-terminal*). For example, in the hexapeptide Met-Thr-Ser-Val-Asp-Lys (see Chapter 4 for a list of the amino acids and their abbreviations), methionine (Met) is the N-terminal amino acid, and lysine (Lys) is the C-terminal amino acid. Note that reversing the sequence to Lys-Asp-Val-Ser-Thr-Met also gives a hexapeptide with the same composition but with different chemical properties because you initially started with a different amino acid. Therefore an amino acid that lost a hydrogen in one sequence will lose an –OH in the other.

The polypeptide chain has a backbone consisting of the same, rather simple, repeating unit. Variations take place in the form of side chains — the R groups of the amino acids. You can see this repeating sequence in Figure 5-1. Notice that the repeating unit (indicated by the brackets) is the amino-carbon-carbonyl sequence and that there can be different R groups attached to the carbon unit of this backbone.

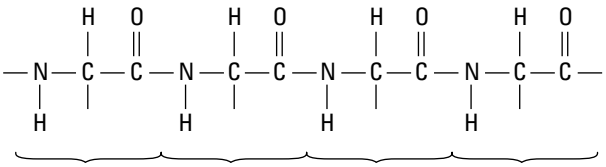


Figure 5-1: Repeating sequence of the protein backbone.

The protein backbone has many places where hydrogen bonds may form. Every residue — other than the amino acid proline — has an NH, which may serve as a *donor* to a hydrogen bond. And every residue has a carbonyl group, which can serve as the *acceptor* of a hydrogen bond. The presence of donors and acceptors leads to the possibility of forming numerous hydrogen bonds.

Each of the peptide bonds exhibits no free rotation about the carbon-nitrogen bond because of the contribution of the resonance form, which has a double bond. Thus, there is a planar unit of four atoms, and in almost all cases, the oxygen atom is trans to the hydrogen atom. The remainder of the backbone can rotate. The ability to rotate or not influences how the three-dimensional structure of the protein is established. There are restrictions to this rotation because the side-chains can "bump" into each other — called *steric* hindrance. The rigidity of the peptide bond and rotation restrictions lower the entropy of the three-dimensional structure of a protein relative to a random chain of amino acids. Lowering the entropy helps stabilize the structure.

Example: The primary structure of insulin

The first determination of the primary structure of a protein was that of bovine insulin, the structure of which appears in Figure 5-2. Since this landmark determination, the primary structures of more than 100,000 proteins have been determined. In all cases, the protein has a unique primary structure.

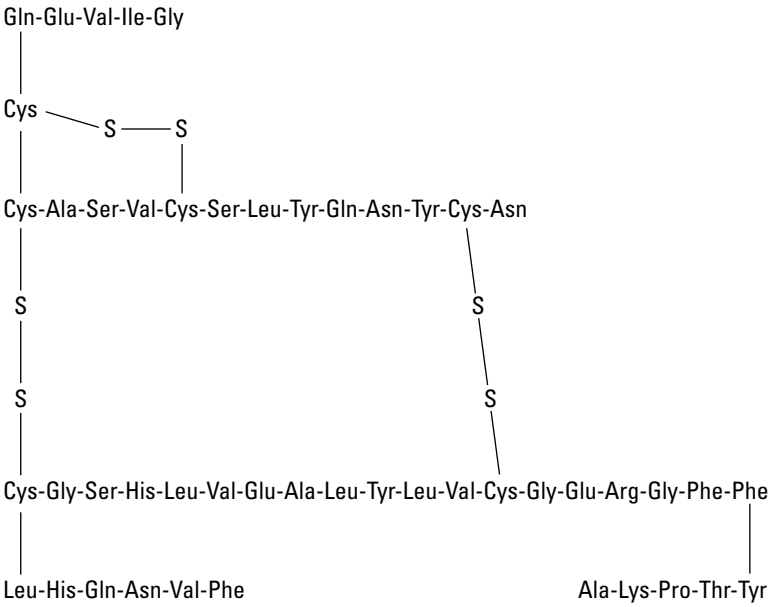


Figure 5-2: Structure of bovine insulin.

Secondary Structure: A Structure Level Most Proteins Have

It is possible for one peptide bond to form a hydrogen bond to another peptide bond. In general, the formation of these hydrogen bonds leads to the secondary structure of a protein. The secondary structure is the result of many hydrogen bonds, not just one. The hydrogen bonds are intramolecular, that is between segments of the same molecule, as shown in Figure 5–3:

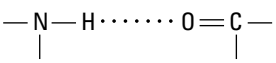


Figure 5-3: Hydrogen bonding between two peptide bonds.

The α-helix and β-pleated sheet are the secondary structures that result from this hydrogen bonding. Secondary structures may be only a small portion of the structure of a protein or can make up 75 percent or more.

The α-helix

In the α-helix, the primary structure twists into a tightly wound, spring or rod-like structure. Each turn consists of 3.6 amino acid residues. These turns allow hydrogen bonding between residues spaced four apart. Every peptide bond participates in two hydrogen bonds: one from an NH to a neighboring carbonyl, and one from a neighboring NH to the carbonyl (Figure 5-4).

Structurally, the helices may be either right-handed or left-handed (see Chapter 3 for more on handedness). Essentially all known polypeptides are right-handed. Slightly more steric hindrance is present in a left-handed helix, and the additional steric hindrance makes the structure less stable. Keratin — the protein of fur, hair, and nails — consists of three right-handed α-helices wrapped around each other in a left-handed coil.

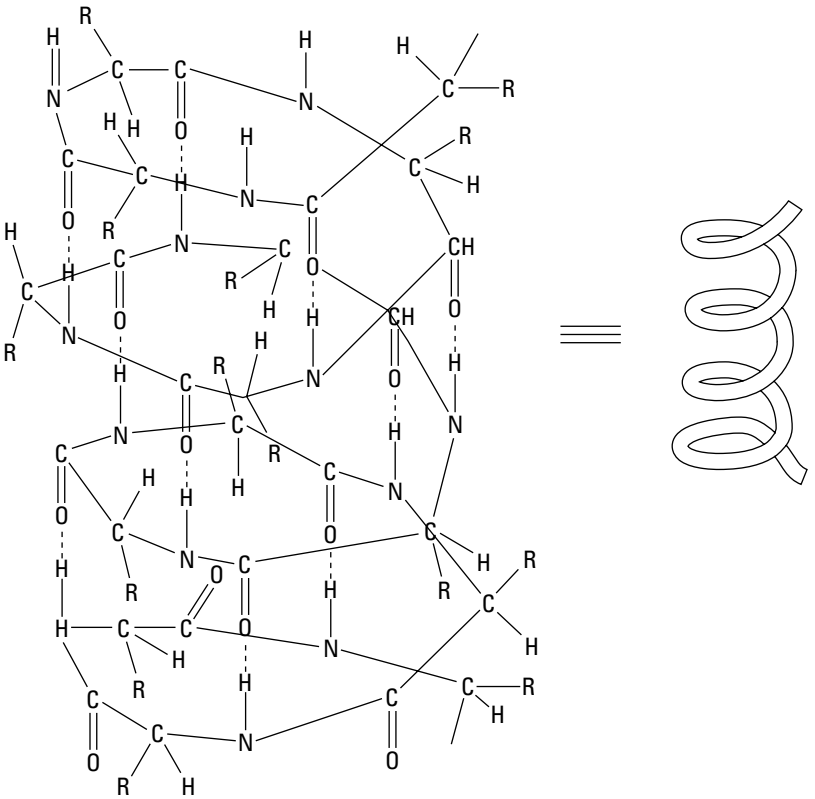


Figure 5-4: The α-helix.

Certain amino acids destabilize the α-helix. Proline, for example, creates bends or "kinks" in the primary structure, which inhibit the formation of a regular pattern of hydrogen bonds. A group of isoleucine residues disrupts the secondary structure because of the steric hindrance caused by their bulky R groups. The small R group of glycine, only an H, allows too much freedom of movement, which leads to a destabilization of the helix. A concentration of aspartic acid and/or glutamic acid residues also destabilizes the structure because the negative charges on the side chains repel each other. Other residues that destabilize the helix, for similar reasons, are lysine, arginine, serine, and threonine.

The β-pleated sheet

The β-pleated sheet, or simply the β sheet, is the other major secondary protein structure. Here, the primary structure is extended instead of tightly winding into a helix. There are two forms of this structure, known as the *parallel* β-pleated sheet and the *anti-parallel* β-pleated sheet. Again, hydrogen bonds are the source of these structures. A β-pleated sheet forms when two or more strands link by hydrogen bonds. The strands are different parts of the same primary structure.

In the parallel structure, the adjacent polypeptide strands align along the same direction from N-terminal end to C-terminal end. In the anti-parallel structure, the alignment is such that one strand goes from N-terminal end to C-terminal end, while the adjacent strand goes from C-terminal end to N-terminal end (Figure 5-5).

In the β-pleated sheet structures, the side chains of adjacent amino acids point in opposite directions. The hydrogen bonding pattern in the parallel structure is the more complicated. Here, the NH group of one residue links to a CO on the adjacent strand, whereas the CO of the first residue links to the NH on the adjacent strand that is two residues down the strand. In the anti-parallel structure, the NH and CO groups of one residue link to the respective CO and NH groups of one residue on the adjacent strand.

Schematically, broad arrows indicate the presence of β-pleated sheets. If the arrows point in the same direction, it is the parallel structure, and if they point in opposite directions, it is the anti-parallel structure. The sheets are typically 4 or 5 strands wide, but 10 or more strands are possible. The arrangements may be purely parallel, purely anti-parallel, or mixed (refer to Figure 5-5).

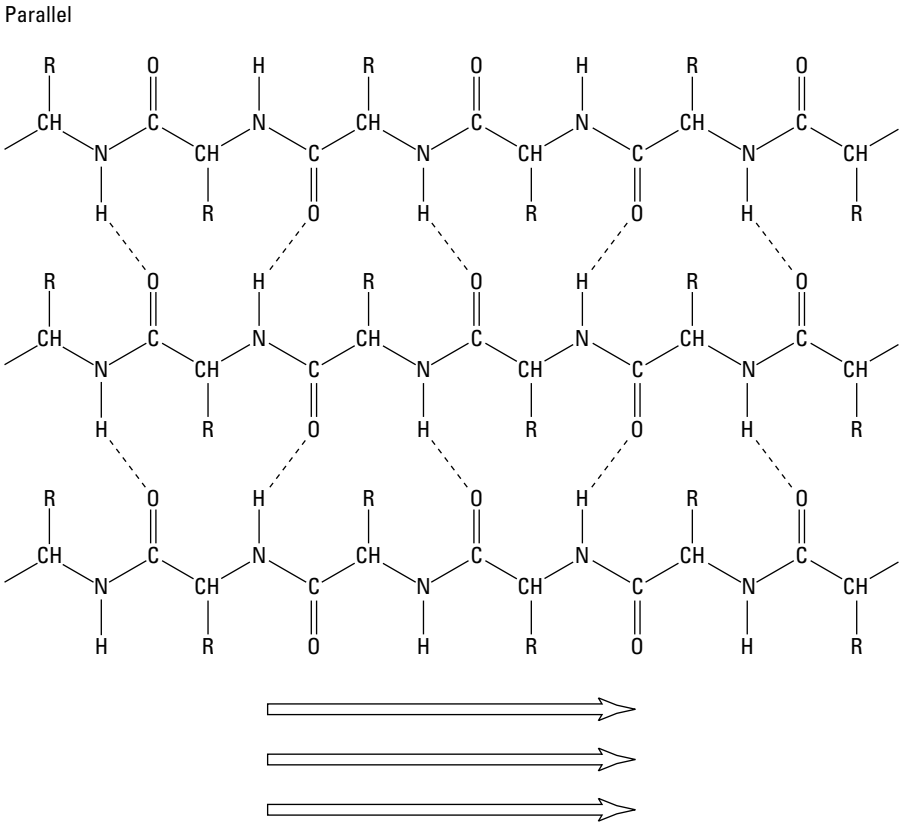


Figure 5-5: Parallel and anti-parallel β-pleated sheet structures.

β-turns and the Ω-loops

There are additional secondary structures involving hydrogen bonding between peptide bonds; these are much smaller units. The best known are the β-turn — or *hairpin bend* — and the Ω-loop. The hairpin bend is simply a bend in the primary structure held in place by a hydrogen bond. The Ω-loop gets its name because of the loose similarity of its shape to the Greek letter. Both are found on the exterior of proteins.

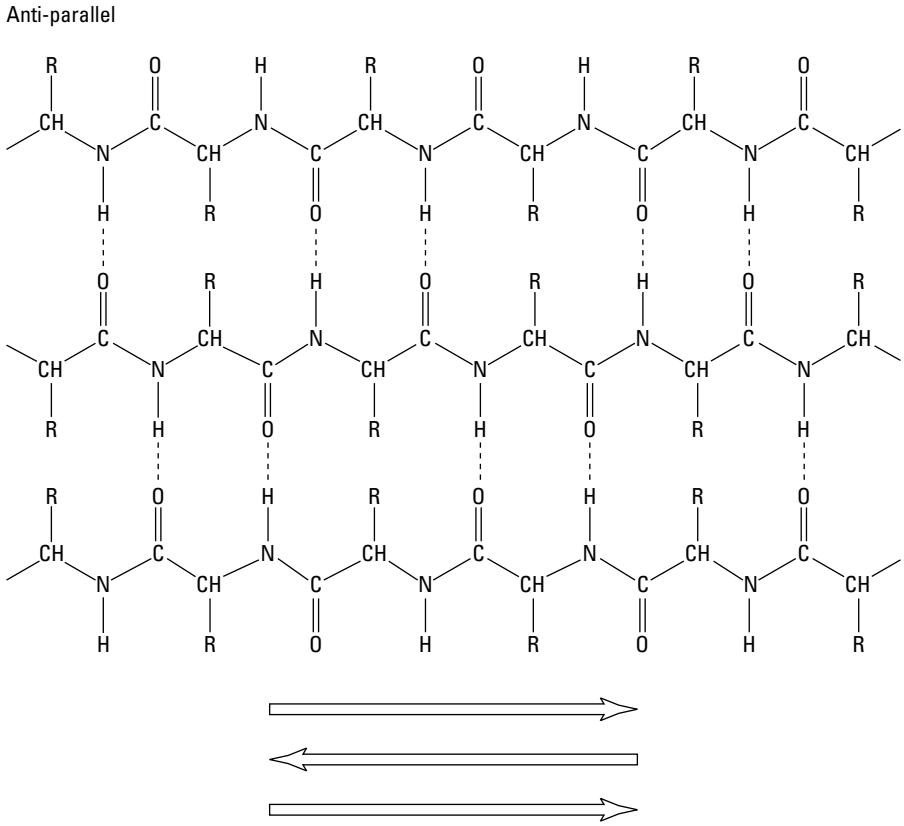


Figure 5-5: (continued)

Tertiary Structure: A Structure Level Many Proteins Have

The overall shape of a protein is determined by its primary and secondary structures along with interactions between the side chains. This gives rise to what is called the protein's *tertiary structure.* Nonpolar side chains are hydrophobic and, although repelled by water, are attracted to each other. Polar side chains attract other polar side chains through either dipole-dipole forces or hydrogen bonds.

For example, both aspartic acid and glutamic acid yield side chains with a negative charge that are strongly attracted to the positive charges in the side chains of lysine and arginine. Two cysteine residues can connect by forming a disulfide linkage — a covalent bond (Figure 5-6).

What induces a protein to adopt a very specific tertiary structure? Examination of the structures of many proteins shows a preponderance of nonpolar side chains in the interior with a large number of polar or ionic side chains on the exterior. In an aqueous environment, the hydrophobic (nonpolar) groups induce the protein to fold upon itself, burying the hydrophobic groups away from the water and leaving the hydrophilic groups adjacent to water. The result is similar in structure to a micelle.

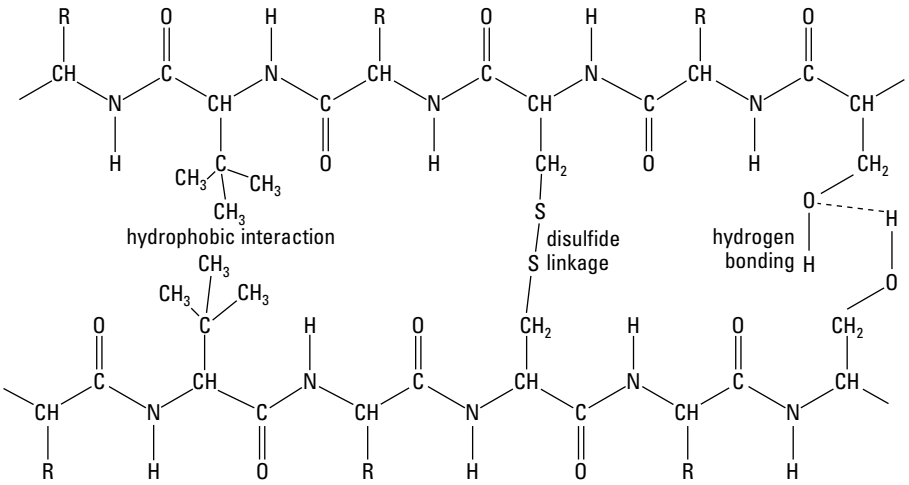


Figure 5-6: Some tertiary structures appearing in proteins.

Quaternary Structure: A Structure Level Some Proteins Have

The quaternary structure found in some proteins results from interactions between two or more polypeptide chains — interactions that are usually the same as those that give rise to the tertiary structure. These interactions include hydrogen bonding and disulfide bonds. This quaternary structure locks the complex of proteins into a specific geometry. An example is hemoglobin, which has four polypeptide chains. There are two identical α-chains and two identical β-chains. (The designations α and β simply refer to two different proteins and not to secondary structures.)

Dissecting a Protein for Study

The previous sections have discussed the different types of protein structure. Now it is time to see how a biochemist goes about determining the structure(s) of a particular protein.

Additional information about the structure of a protein comes from immunology. An animal generates an antibody in response to a foreign substance known as an *antigen*. *Antibodies* are proteins found in the blood serum. Exposure to diseases, certain chemicals, and allergies induce the formation of specific antibodies. These antigens collect on the surface of red blood cells. Every antigen has a specific antibody.

Antibodies are very specific and have a strong affinity for their specific antigens, recognizing specific amino acid sequences on the antigens. Animals have a large number of antibodies present in their bodies, based on their environmental history. One application of antibodies and antigens is in the analysis of bloods, specifically in the field of forensics investigations (see nearby "Forensics: Analysis of bloodstains" sidebar).

Separating proteins within a cell and purifying them

There are thousands of different proteins in each cell. In order to examine and study one of them, you need to separate it from all the others. The methods of separating proteins are, in general, applicable to all other types of biochemicals. Initially, simple filtration and solubility can remove gross impurities, but much more needs to be done before the sample is pure. The key separation and purification methods depend on two physical properties of the proteins: size and charge.

Separating proteins by size

Methods relying on separation by protein size and mass include ultrafiltration, ultracentrifugation, and size exclusion chromatography. *Ultrafiltration* is a modification of dialysis in which molecules smaller than a certain size diffuse through a semipermeable membrane, and larger ones don't. Ultrafiltration can separate smaller molecules from larger impurities or larger molecules from smaller impurities.

In *ultracentrifugation,* a powerful centrifuge causes heavier molecules to sink faster and, which allows their separation — much as the lighter water is separated from the heavier lettuce in a salad spinner. Ultracentrifugation also gives the molar mass of the protein.

In *size exclusion chromatography,* also known as *molecular sieve chromatography* or *gel filtration chromatography,* a solution passes through a chromatography column filled with porous beads. Molecules that are too large for the pores pass straight through. Molecules that may enter the pores are slowed. The molecules that may enter the pores undergo separation depending on how easily they can enter.

Forensics: Analysis of bloodstains

The study of proteins has many applications to forensics. One of them is the examination of bloodstains, blood being the most common form of evidence examined by a forensic serologist. The presence of blood can link a suspect to both a victim and a crime scene. Bloodstain patterns can also give evidence of how a violent attack took place. Criminals recognize the significance of this evidence and often try to conceal it.

Blood is mostly water, but it also contains a number of additional materials including cells, proteins, and enzymes. The fluid portion, or *plasma,* is mostly water. The *serum* is yellowish and contains platelets and white blood cells. The *platelets,* or red blood cells, outnumber the white blood cells by about 500 to 1. White blood cells are medically important, whereas red blood cells and, to a lesser extent, serum are important to the forensic serologist. Because blood quickly clots when exposed to air, serologists must separate the serum from the clotted material. The serum contains antibodies that have forensic applications, and red blood cells have substances such as antigens on their surfaces that also have forensic applications. Antibodies and antigens are the keys to forensic serology: Even identical twins with identical DNA have different antibodies. As you know from this chapter, antibodies, and some antigens, are proteins, and this is why methods of studying proteins are important to their analysis.

Analysis of bloodstains initially attempts to answer five questions.

- **Is this a blood sample?** To answer this question, the investigator can use a number of tests. The generic term for a test of this type is a *presumptive* test. The Kastle-Meyer test uses phenolphthalein, which, when it comes into contact with hemoglobin, or a few other substances, forms a bright pink color from the release of peroxidase enzymes. The luminol test is useful in detecting invisible bloodstains because, in contact with blood, or a few other chemicals, luminol emits light, which can be seen in a darkened room. The Wagenhaar, Takayama, and Teichman tests take advantage of the fact that long-dried blood will crystallize or can be induced to crystallize.
- **Is the blood from a human or an animal?** The forensic investigator answers this question (and the next one, if applicable) by means of an *antiserum* test. It is important to know whether the blood came from a human or an animal such as a pet. The standard test is the precipitin test. Injecting human blood into an animal results in the production of antibodies in the animal's bloodstream, and isolating these antibodies from the animal's blood yields an antiserum. If human antiserum creates clotting in a blood sample, the sample must be human.
- **If the sample is from an animal, what is the species?** It is possible to create animal antiserums in an analogous manner, and test for each type of animal.
- **If the blood is from a human, what is the blood type?** The procedure for answering this depends on the quantity and quality of the sample. If the quality is good, *direct* typing is done — otherwise, *indirect* typing is used. (Direct typing, to classify blood in the A-B-O system, is discussed in this chapter's other sidebar.) A dried bloodstain normally requires indirect typing. The most common indirect typing method is the absorption-elution test. Treatment of a sample with antiserum antibodies gives a solution which, upon addition to a known sample, causes coagulation.

- **Is it possible to determine the sex, race, and age of the source of the blood?** Here the answers become less precise. Clotting and crystallization indicate age. Testing for testosterone levels and chromosome testing can determine sex. And certain controversial, racial genetic markers based on protein and enzyme tests may indicate race.

  Other body fluids may contain the same antibodies and antigens found in blood. Therefore, similar tests work on these fluids as well.

Separating proteins by charge

Methods of separating proteins relying on the charge of the protein include solubility, ion exchange chromatography, and electrophoresis. Each of these methods is pH dependent.

Proteins are least soluble at their isoelectric point. (The *isoelectric point* is the pH where the net charge on the protein is 0.) At the isoelectric point, many proteins precipitate from solution. At a pH below the isoelectric point, the protein has a net positive charge, whereas a pH above the isoelectric point imparts a net negative charge. The magnitude of the charge depends on the pH and the identity of the protein. Therefore, two proteins coincidently having the same isoelectric point will not necessarily have the same net charge at a pH that is one unit lower than the isoelectric point.

Both ion exchange chromatography and electrophoresis take advantage of the net charge. In *ion exchange chromatography,* the greater the magnitude of the charge, the slower a protein moves through a column — this is similar to the ion-exchange process that occurs in water-softening units.

In *electrophoresis,* the sample solution is placed in an electrostatic field. Molecules with no net charge do not move, but species with a net positive charge move toward the negative end, and those with a net negative charge move toward the positive end. The magnitude of the net charge determines how fast the species moves. Other factors influence the rate of movement, but the charge is the key. There are numerous modifications of electrophoresis.

In protein analysis, rarely do biochemists use only one single technique. They commonly use several in order to confirm their findings.

Digging into the details: Uncovering a protein's amino acid sequence

Once a pure sample of protein is available, it is possible to begin determining its amino acid sequence, in order to identify the specific protein. The general

procedure for doing so, with slight modification, works for other biochemicals as well:

Step 1: Separating and purifying the polypeptide chains

If you determine that more than one polypeptide chain is present in the protein, you need to separate and purify the chains so you can sequence them individually. (Because many proteins only have one polypeptide chain, this step is not always necessary.) Denaturing the protein, disrupting its three-dimensional structure without breaking the peptide bonds, using pH extremes will normally suffice. If disulfide linkages are present between the chains, apply the procedure outlined in Step 2 to separate the chains for isolation.

Step 2: Slashing intrachain disulfide linkages

Step 2 requires breaking (cleaving) the disulfide linkages. A simple reduction accomplishes this. However, the linkages may reform later, so it is necessary to cleave the linkages and prevent their reformation via reductive cleavage followed by alkylation. Oxidative cleavage, where oxidation of the sulfur to $-SO_3^-$ occurs, also prevents a reversal of the process.

Step 3: Determining amino acid concentration of the chain

Step 3 is easily accomplished using an *amino acid analyzer*, an automated instrument that can determine the amino acid composition of a protein in less than an hour. The instrument requires less than a nanomole of protein. The analyzer's output is the percentages of each of the amino acids present.

Step 4: Identifying the terminal amino acids

Step 4 not only identifies the terminal amino acids but also indicates whether more than one chain is present. A polypeptide chain only has one N-terminal and one C-terminal amino acid. Therefore, if more than one N- or C-terminal amino acid is present, there must be more than one polypeptide chain.

It is possible to identify the N-terminal residue in a number of ways. In general, procedures begin by adding a reagent that reacts with the N-terminal amino acid and tags it. Subsequent hydrolysis destroys the polypeptide, allowing separation of the tagged residue and its identification. Such methods use Sanger's reagent, dansyl chloride, and leucine aminopeptidase. The method of choice nowadays is called the Edman degradation. This method, as do other methods, tags the N-terminal residue; however, only the terminal amino acid is cleaved from the chain, so the remainder of the chain is not destroyed as in other methods. It is possible to repeat the procedure on the shortened chain to determine the next residue. In principle, repetition of the Edman degradation can yield the entire sequence, but, in most cases, determination of the first 30 to 60 residues is the limit.

It is also possible to determine the C-terminal residue by tagging. The akabori reaction (hydrazinolysis) and reduction with lithium aluminum hydride tag the C-terminal residue. It is also possible to selectively cleave the C-terminal residue using the enzyme carboxypeptidase, a variety of which are available. Unfortunately, the enzyme doesn't stop with one cleavage — given sufficient time, it proceeds down the entire polypeptide chain.

Steps 5 and 6: Breaking the chain into smaller pieces

In Step 5, you cleave the polypeptide into smaller fragments and determine the amino acid composition and sequence of each fragment. Step 6 repeats Step 5 using a different cleavage procedure to give a different set of fragments. Steps 5 and 6 break the chain into smaller pieces to ease identification.

Most of the methods here employ enzymes; however, other less-specific methods are useful in some cases. Partial acid hydrolysis randomly cleaves the protein chain into a number of fragments. Trypsin, a digestive enzyme, specifically cleaves on the C-side of arginine or lysine. Using trypsin gives additional information that the total number of arginine and lysine residues present is one less than the number of fragments generated. The digestive enzyme chymotrypsin preferentially cleaves residues containing aromatic rings (tyrosine, phenylalanine, and tryptophan). It slowly cleaves other residues especially leucine. Clostripain cleaves positively charged amino acids, especially arginine. It cleaves lysine more slowly. Fragments with a C-terminal aspartic acid or glutamic acid form from the interaction of staphylococcal protease on a protein in a phosphate buffer. In the presence of bicarbonate or acetate buffer, only C-terminal glutamic acid fragments result. A number of less specific enzymes can complete the breakdown of the fragments, including elastase, subtilisin, thermolysin, pepsin, and papain.

Chemical methods of breaking up the fragments include treatment with cyanogen bromide, hydroxylamine, and heating an acidic solution. Cyanogen bromide specifically attacks methionine. Hydroxylamine specifically attacks asparagine-glycine bonds. If a solution at pH = 2.5 is heated to 40°C, selective cleavage of aspartic acid-proline bonds occurs.

It is possible to apply the Edman degradation on each of the fragments. This can simplify the determination of the sequence of a large protein.

Step 7: Combining information to get the total sequence

Step 7 is where the information from the various procedures comes together. For example, look at a simple octapeptide fragment from a protein. This fragment gave, upon complete hydrolysis, one molecule each of alanine (Ala), aspartic acid (Asp), glycine (Gly), lysine (Lys), phenylalanine (Phe), and valine (Val), and two molecules of cysteine (Cys). The following fragments were

Basics of blood typing

The determination of blood type in the A-B-O system, first begun in 1901, is based on antigen-antibody reactions. Over the years, additional reactions have been discovered. More than 256 antigens are known, leading to 23 different blood groups. Each blood group is defined by the antibodies present in the serum and the antigens present on the red blood cells.

In basic blood typing, one needs two antiserums, labeled *anti-A* and *anti-B*. Adding a drop of one of these to a blood sample causes coagulation if the appropriate antigens are present. Anti-A interacts with both A and AB blood. Anti-B interacts with both B and AB blood. Neither interacts with type O blood. The approximate distribution of the different blood types is: 43–45 percent type O; 40–42 percent type A; 10–12 percent B; and 3–5 percent AB. Subgrouping is also possible with designations such as O1 and O2. There are other very rare types as well.

The Rh factor provides an additional means of subdividing blood. The *Rh factor* (the name comes from the rhesus monkey) is an antigen on the surface of red blood cells. A person with a positive Rh factor contains a protein (antibody) that is also present in the bloodstream of the rhesus monkey. About 85 percent humans are Rh positive. A person lacking this protein is, naturally, Rh negative. Assigning a blood sample as Rh positive or Rh negative is a useful simplification. There are about 30 possible combinations of factors.

Additional factors can determine whether blood belongs to a specific individual: the identification of other proteins and enzymes present in the blood. A forensic serologist (see this chapter's other sidebar for more) does this level of testing in every case where the quality of the sample allows. One of the characteristics of proteins or enzymes in the blood is *polymorphism*, or the ability to be present as isoenzymes. Polymorphism means that the protein may exist in different forms or variants. One well-known example is the polymorphism of hemoglobin into the form causing sickle cell anemia. Some well-recognized polymorphisms are:

Adenyl kinase	AK
Adenosine deaminase	ADA
Erythrocyte acid phosphatase	EAP
Esterase D	EsD
Glucose-6-phosphate dehydrogenase	G-6-PD
Glutamic pyruvate transaminase	GPT
Phosphoglucomutase	PGM 2-1
6-phosphogluconate dehydrogenase	6-PGD
Transferrin	Tf

The distribution of each of these *polymorphs* in the population is well established. The determination of each of these additional factors narrows down the number of possible individuals.

isolated after partial hydrolysis: Gly-Cys, Phe-Val-Gly, Cys-Asp, Cys-Ala, Lys-Cys, and Cys-Asp-Lys. Now we match the fragments, deduce the amino acid sequence in the octapeptide, and write a primary structure for the peptide:

Cys-Asp Lys-Cys

Cys-Asp-Lys Cys-Ala

Gly-Cys

Phe-Val-Gly

Phe-Val-Gly-Cys-Asp-Lys-Cys-Ala

Step 8: Locating the disulfide linkages

Step 8 does not specifically deal with the primary structure of the protein, but it is related. If the disulfide linkages are left intact by skipping Step 2, different fragments result. x-ray diffraction analysis can locate each amino acid residue. This can be used to determine the overall shape of a protein. In some cases, more detailed structural information can be determined by sophisticated instrumental analysis techniques.

Chapter 6

Enzymes Kinetics: Getting There Faster

In This Chapter

- Understanding enzymes classification
- Examining kinetics
- Studying the Michaelis-Menten equation
- Comprehending enzyme inhibition and regulation

Enzymes are complex biological molecules, primarily or entirely protein, which behave as biological catalysts. As *catalysts,* they alter the rate of a chemical reaction without themselves being consumed in the reaction. Enzymes are normally very specific in their action, often targeting only one specific reacting species, known as the *substrate.*

This specificity includes *stereospecificity,* the arrangement of the substrate atoms in three-dimensional space. Stereospecificity is illustrated by the fact that if the D-glucose in your diet were replaced by its enantiomer, L-glucose, you would not be able to metabolize this otherwise identical enantiomer.

Enzymes occur in many forms. Some enzymes consist entirely of proteins, whereas others have non-protein portions known as *cofactors.* The cofactor may be a metal ion, such as magnesium, or an organic substance. We call an organic cofactor a *coenzyme* (there is no specific term for a metallic cofactor). An enzyme lacking its cofactor is an *apoenzyme,* and the combination of an apoenzyme and its cofactor is a *holoenzyme.* A metalloenzyme contains an apoenzyme and a metal ion cofactor. A tightly bound coenzyme is a prosthetic group. (Wow! We know that this is a lot of terminology, but hang in there. The key is the enzyme.)

One region on the enzyme, the *active site,* is directly responsible for interacting with the reacting molecule(s). When a reacting molecule, the substrate, binds to this active site, a reaction may occur. Other materials besides the enzyme and substrate, may be necessary for the reaction to occur.

In many cases, the cell initially produces the enzyme in an inactive form called a *proenzyme* or *zymogen,* which must undergo activation for it to function. The enzyme trypsin illustrates why it is sometimes necessary to generate an inactive form of an enzyme. Trypsin is one of the enzymes present in the stomach that is responsible for the digestion of proteins. Its production, as an inactive form, occurs in the cells of the stomach walls, and activation occurs after its release into the stomach. If trypsin were produced in the active form, it would immediately proceed to begin digesting the cell that produced it. Eating yourself is not a good thing.

The activation of the inactive form of an enzyme serves as one form of enzyme control. Inhibition is another method of enzyme control. The two general types of inhibition are competitive inhibition and noncompetitive inhibition. In *competitive* inhibition, another species competes with the substrate to interact with the active site on the enzyme. In *noncompetitive* inhibition, the other species binds to some site other than the active site. This binding alters the overall structure of the enzyme so that it no longer functions as a catalyst.

Enzyme Classification: The Best Catalyst for the Job

Ever wonder who gets to name chemicals? Well, the answer varies, but for enzymes it's the Enzyme Commission of the International Union of Biochemistry that's responsible. Common names for enzymes begin with some description of its action plus an *-ase* suffix. (Enzymes that were named before the implementation of the -ase system, such as trypsin, do not follow this convention.) The Enzyme Commission has also developed a numerical system for classifying enzymes. The names begin with EC, for Enzyme Commission, and end with four numbers, separated by decimal points, describing the enzyme. An example of this nomenclature is EC 2.7.4.4.

The first number in the EC name refers to the major enzyme *class*, and there are six major enzyme classes, summarized in Table 6-1. To continue with our example, the 2 in EC 2.7.4.4 designates the enzyme as a transferase. The second number, the 7, indicates what *group* the enzyme transfers. The third number, the first 4, indicates the *destination* of the transferred group. And the last number, the second 4, refines the information given by the third number.

Table 6-1	Six Basic Types of Enzymes
Class of Enzymes	***What They Catalyze***
Oxidoreductases	Redox reactions
Transferases	The transfer groups of atoms
Hydrolases	Hydrolysis
Lyases	Additions to a double bond, or the formation of a double bond
Isomerases	The isomerization of molecules
Ligases or synthetases	The joining of two molecules

Up one, down one: Oxidoreductases

Oxidoreductases catalyze a simultaneous oxidation and a reduction. An *oxidation* involves the increase in the oxidation state of an element, whereas a *reduction* involves the decrease in the oxidation state of an element. It is impossible to have one without the other. Examples of the types of reactions that qualify as oxidation and reduction reactions are in Table 6-2. In general, the substrate undergoes either oxidation or reduction, while the enzyme temporarily does the opposite but eventually returns to its original form.

Table 6-2	Some Possible Types of Oxidation and Reduction Reactions
Oxidation	***Reduction***
Loss of one or more electrons	Gain of one or more electrons
Addition of oxygen	Loss of oxygen
Loss of hydrogen	Gain of hydrogen

An example: Succinate dehydrogenase catalyzes the oxidation of the succinate ion. In this case, the oxidation involves the loss of two hydrogen atoms with the formation of a trans double bond. The enzyme alcohol dehydrogenase

removes two hydrogen atoms from an alcohol to produce an aldehyde. The general form, unbalanced, of these reactions is as follows:

$$^{-}OOCCH_2CH_2COO^{-} \xrightarrow{\text{succinate dehydrogenase}} {}^{-}OOCCH{=}CHCOO^{-}$$

$$CH_3CH_2OH \xrightarrow{\text{alcohol dehydrogenase}} CH_3CHO$$

You don't belong here: Transferases

The purpose of a transferase is to catalyze the transfer of a group from one molecule to another. Aminotransferase transfers an amino group, and phosphotransferase transfers a phosphoryl group. The general form, unbalanced, of these reactions appears in Figure 6-1.

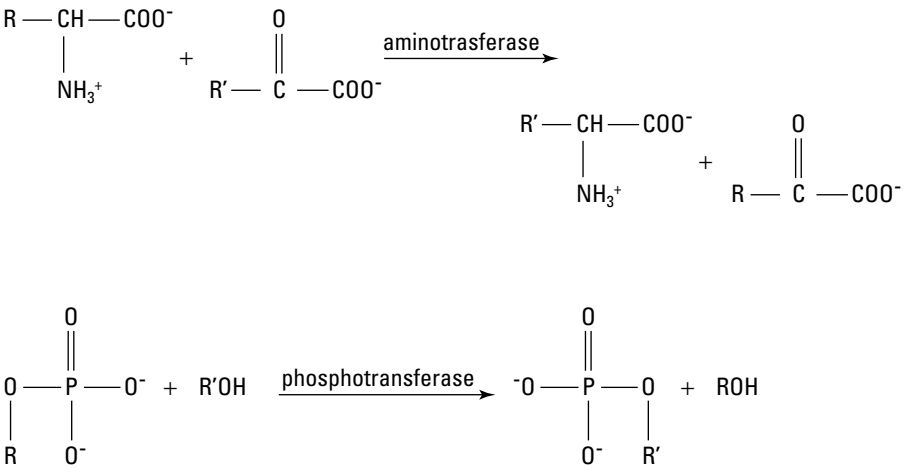


Figure 6-1: General form, unbalanced, of two transferase catalyzed reactions.

Water does it again: Hydrolases

Hydrolases catalyze the cleavage of a bond through the insertion of a water molecule (as an H and an OH). There may be a pH dependence, which results in the subsequent loss of a hydrogen ion. A *phosphatase* catalyzes the hydrolysis of a monophosphate ester, and a *peptidase* catalyzes the hydrolysis of a peptide bond. The general form of these reactions appears in Figure 6-2.

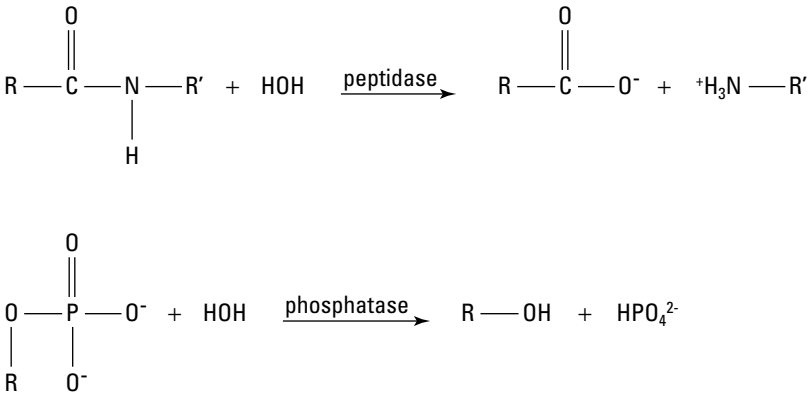
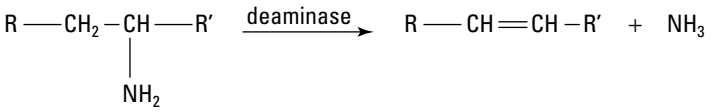


Figure 6-2: General form of two hydrolase catalyzed reactions.

Taking it apart: Lyases

Lyases catalyze the removal of a group. This process is accompanied by the formation of a double bond or the addition of a group to a double bond. A *deaminase* aids in the removal of ammonia, and a *decarboxylase* catalyzes the loss of CO_2. The general form of these reactions appears in Figure 6-3.



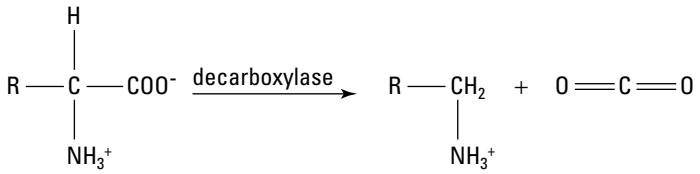


Figure 6-3: General form of two lyase catalyzed reactions.

Shuffling the deck: Isomerases

Racemase and epimerase are isomerases. *Isomerase* enzymes catalyze the conversion of one isomer to another. The *racemase* illustrated at the top of Figure 6-4 catalyzes the racemization of enantiomers. An *epimerase,* like the one at the bottom of Figure 6-4, catalizes the change of one epimer to another. Like all catalyzed reactions, these are equilibrium processes.

Putting it together: Ligases

Ligase enzymes catalyze reactions leading to the joining of two molecules in which a covalent bond forms between the two molecules. The process often utilizes high-energy bonds such as in ATP. Figure 6-5 illustrates the action of two ligases, pyruvate carboxylase and acetyl-CoA synthetase. *Pyruvate carboxylase* catalyzes the formation of a C-C bond. *Acetyl-CoA synthetase* catalyzes the formation of a C-S bond.

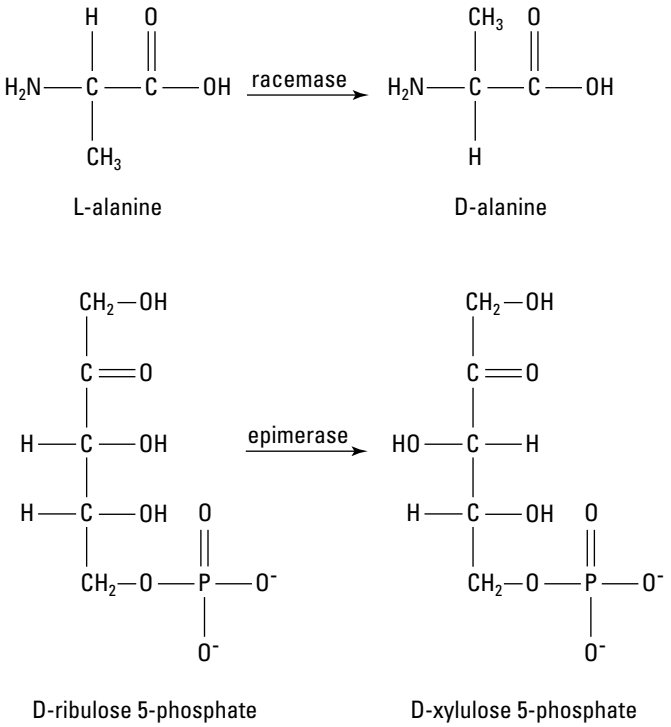


Figure 6-4: Examples of isomerase reactions catalyzed by a racemase and an epimerase.

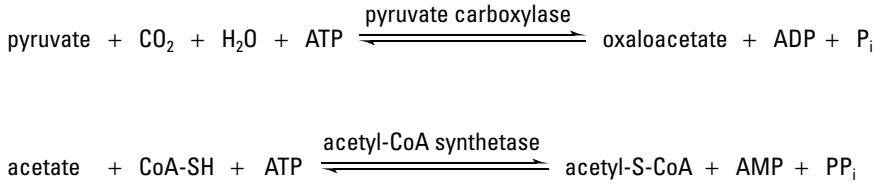


Figure 6-5: Reactions illustrating the action of the ligases pyruvate carboxylase and acetyl-CoA synthetase.

Enzymes as Catalysts: When Fast Is Not Fast Enough

The action of an enzyme begins with the formation of an enzyme-substrate complex. In this formation, the substrate in some way binds to the active site of the enzyme. The interaction between the enzyme and the substrate must, in some way, facilitate the reaction, and it opens a new reaction pathway.

The active site is typically a very small part of the overall enzyme structure. The amino acid residues comprising the active site may come from widely separated regions of the protein (primary structure), and it is only through interactions leading to higher structure levels that they are brought close together. Amino acid residues not in the active site serve many different functions that aid the function of the enzyme.

Models of catalysis: Lock and key versus induced-fit

The first attempt at explaining this process led to the *Lock and Key Model,* in which the substrate behaves as a key that fits into a lock, the enzyme (Figure 6-6). The Lock and Key Model, to a certain degree, explains the specificity of enzymes. Just as only the right key will fit into a lock, only the right substrate fits into the enzyme.

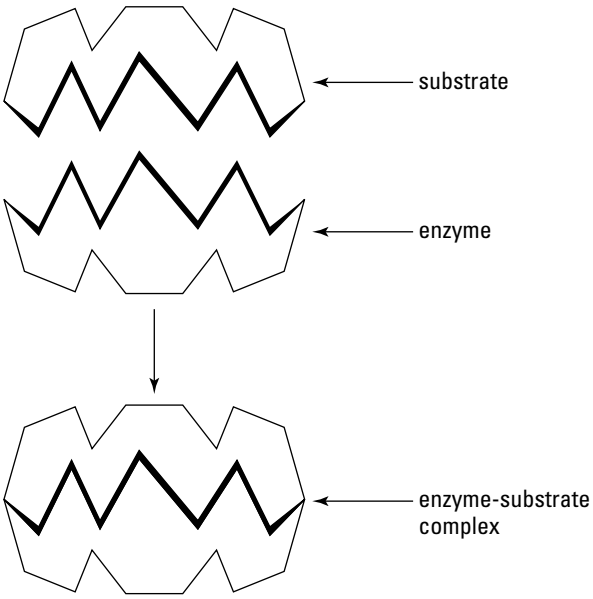


Figure 6-6: The Lock and Key Model of enzyme catalysis.

One limitation of the Lock and Key Model is that it does not explain why the reaction actually occurs, and another is that enzymes are flexible and not rigid as this theory implies.

The *Induced-Fit Model* overcomes some of the limitations of the Lock and Key Model. In this model, the substrate still needs to fit into the enzyme like a key, but instead of simply fitting into the "keyhole," some type of modification is induced in the substrate, enzyme, or both. The modification begins the process of the reaction. Figure 6-7 illustrates how the Induced-Fit Model applies to the formation of the same enzyme-substrate in Figure 6-6.

All About Kinetics

As you know, all reactions involve energy. The reactants begin with a certain level of energy, an additional quantity of energy is absorbed to reach the transition state (ΔG^*, where the asterisk indicates the transition state), and then energy is released to reach the products. The difference in the energy between the reactants and products is ΔG.

If the energy level of the products is greater than that of the reactants (energy is absorbed), the reaction is *endergonic,* and nonspontaneous. If the energy level of the products is less than the reactants (energy is released), the process is *exergonic,* and spontaneous.

But just because a reaction is spontaneous does not mean it will occur at an appreciable rate. The rate depends on the value of ΔG^*. The greater the value of ΔG^*, the slower the reaction is. An enzyme, like any catalyst, lowers the value of ΔG^* and consequently increases the rate of the reaction. The difference between the reactants and products remains unchanged, as does the equilibrium distribution of the reactants and products. The enzyme facilitates the formation of the transition state (Figure 6-8).

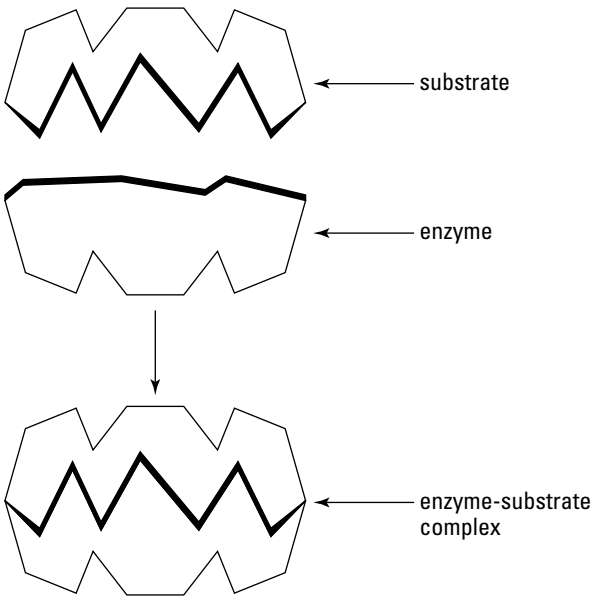


Figure 6-7: The Induced-Fit Model of enzyme catalysis.

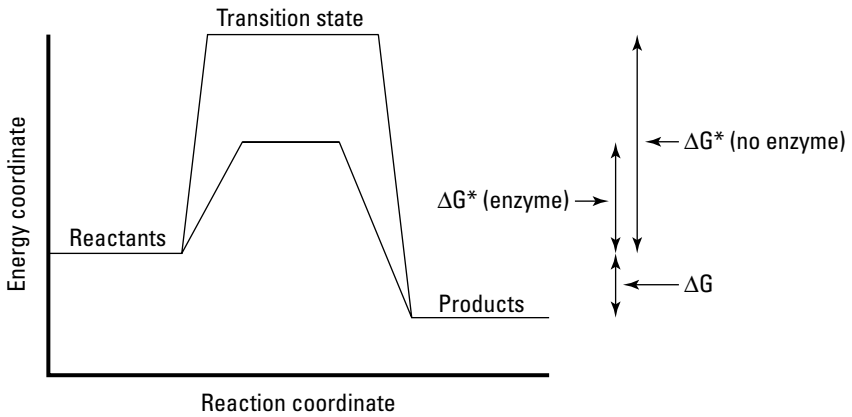


Figure 6-8: Effect of an enzyme on a reaction.

A species has two possible fates in the transition state: It may lose energy and return to the reactant form, or it may lose energy and move to the product form. These two fates lead to two equilibria. One of the equilibria involves the reactant (substrate) and the transition state, and the other involves the product(s) and the transition state. Rapid removal of the product(s) does not allow establishment of the reverse process that leads to the equilibrium. Removal of the product simplifies the analysis of the kinetic data.

Enzymes, like all catalysts, catalyze both the forward and the reverse reaction. The lowering of ΔG^* accelerates both reactions. The ultimate equilibrium concentrations of substrate and products will be the same whether an enzyme is present or not — the enzyme merely changes the amount of time necessary to reach this state.

Enzyme assays: Fixed time and kinetic

An *enzyme assay* is an experiment to determine the catalytic activity of an enzyme. It is possible to measure either the rate of disappearance of the substrate or the rate of appearance of a product. The experimental mode of detection depends on the particular chemical and physical properties of the substrate or the product, and the rate is the change in concentration per change in time. In *fixed time assay,* you simply measure the amount of reaction in a fixed amount of time. In *kinetic assay,* you monitor the progress of a reaction continuously. Once you determine the rate of change in concentration of any reactant or product, it is possible to determine the rate of change of for any other reactant or product of the reaction

It is important to control the conditions precisely. Minor changes in variables such as the temperature or the pH can drastically alter the catalytic activity of an enzyme. For example, the study of enzymes important to humans should be carried out at 37°C, because this is normal body temperature.

Rate determination: How fast is fast?

It is important to control kinetic experiments closely. Once you determine the basic conditions, you can run a series of experiments using a fixed enzyme concentration and varying concentrations of substrate. Up to a point, an increase in substrate concentration results in an increase in rate. The rate increases until the enzyme is saturated. This *saturation point* is where all the enzyme molecules are part of an enzyme-substrate complex. When this occurs, an increase in the substrate concentration yields no increase in the rate, because there are no enzymes available to interact with the

Enzymes in medical diagnosis and treatment

Enzyme levels may indicate medical problems, and that makes enzyme assays useful for both the diagnosis and treatment of medical problems. For examine, creatine kinese (CK) is an enzyme that aids in the synthesis and degradation of creatine phosphate.

CK exists as three different isoenzymes. Each is composed of two polypeptide chains. In the case of muscle CK, the chains are identical, and it's labeled CK-MM. CK found in the brain also has identical polypeptide chains, but they are different from the ones associated with muscle CK and are labeled CK-BB. Finally, the CK found in the heart is a hybrid of the two with one M chain and one B chain: CK-MB.

Normal blood serum contains a little CK-MM and almost no CK-BB and CK-MB. When tissue undergoes injury, though, some of the intracellular enzymes leak into the blood where they can be measured. Elevated levels of total CK (all three isoenzymes) may be indicative of sketal-muscle trauma or myocardial infarction (MI, or heart attack). Analysis of the individual isoenzymes may give additional clues.

For example, an individual falls off a ladder and suffers several broken bones. He is taken to the hospital, where his blood serum CK is measured. It is elevated as expected, but the physician also orders a CK-MB level determination. It turns out to also be highly elevated, indicating that the reason the man fell off the latter to begin with was that he was suffering a heart attack (CK-MB). This knowledge allows the doctor to start a regime of treatment that helps to minimize permanent heart damage.

additional substrate molecules. For most reactions, the rate of the reaction approaches the saturation level along a hyperbolic curve. Theoretically, the reaction rate will only reach saturation at infinite substrate concentration.

A plot of the reaction rate, V, versus the substrate concentration, [substrate], supplies several bits of useful data (see Figure 6-9). The experiment is at constant enzyme concentration. One piece of useful data is the maximum reaction rate, V_{max}. The rate approaches V_{max} asymptotically. At low substrate concentrations, the reaction approaches first-order kinetics, where the rate of reaction depends only on the concentration of one reactant. At high concentrations, the reaction approaches zero-order kinetics, where the rate of reaction is independent of reactant concentration. (Later in this chapter you will see that this graph varies with less simple enzyme-substrate interactions.) In the region between the zero-order region and the first-order region, the kinetics are mixed and difficult to interpret. Important values in the low-concentration region (first-order region) are $\frac{1}{2} V_{max}$ and K_M. The value $\frac{1}{2} V_{max}$ is one-half the V_{max} value. K_M is the Michaelis constant, which corresponds to the substrate concentration producing a rate of $\frac{1}{2} V_{max}$. The Michaelis constant, measured in terms of molarity, is a rough measure of the enzyme-substrate affinity. K_M values vary widely.

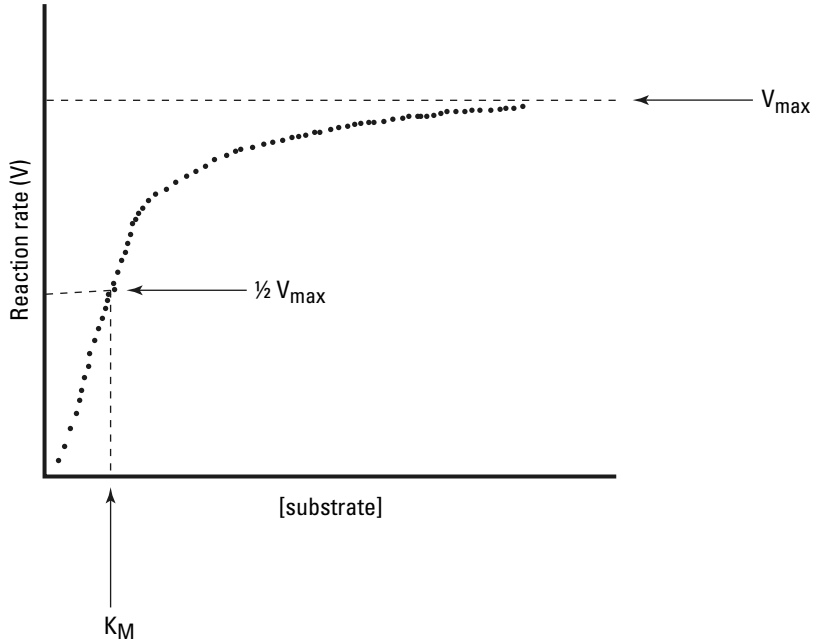


Figure 6-9: Plot of reaction rate, V, versus substrate concentration, [substrate].

At low substrate concentrations, there is an approximately linear relationship between [substrate] and V. At high substrate concentrations, though, V is nearly independent of [substrate]. The low substrate region is useful in the application of the Michaelis-Menten equation (see the next section).

In an uncatalyzed reaction, increasing the substrate concentration does not lead to a limiting V_{max}. The rate continues to increase with increasing substrate concentration. This indirect evidence leads to the conclusion that there is an *enzyme-substrate complex,* a tightly-bound grouping of the enzyme and the substrate. The limit occurs when all the enzyme molecules are part of a complex so that there are no free enzyme molecules available to accommodate the additional substrate molecules. Various x-ray and spectroscopic techniques provide direct evidence to confirm the formation of an enzyme-substrate complex.

Measuring Enzyme Behavior: The Michaelis-Menten Equation

One of the breakthroughs in the study of enzyme kinetics was the development of the Michaelis-Menten equation. It is possible to interpret the behavior of many enzymes by applying the equation to kinetic data. (There are exceptions, and they do not give a graph similar to the one appearing back in

Figure 6-9.) In general, the results of the kinetics experiments are for allosteric enzymes. The *Michaelis-Menten equation* is as follows:

$$V = \frac{V_{max}[S]}{[S] + K_M}$$

In this equation, V is the rate of the reaction, [S] is the substrate concentration, V_{max} is the maximum reaction rate, and K_M is the Michaelis constant. As seen in Figure 6-9, the rate of catalysis, V, increases linearly at low substrate concentration, but begins to level off at higher concentrations. Interpretation begins with examining the following general reaction pathway:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E + P$$

In this pathway, E refers to the enzyme, S is the substrate, ES is the enzyme-substrate complex, and P is the product. The various instances of k refer to the rate constants of the various steps — a negative rate constant is for the reverse process. In the first step, the separate enzyme and substrate combine to form the enzyme-substrate complex (transition state). The rate of formation of ES is k_1. After ES forms, it may break down to E and S (k_{-1}) or it may proceed to product (k_2). (Note: Some texts refer to k_2 as k_{cat}.)

Because the enzyme will catalyze the reverse process, E and P may combine to reform the complex (k_{-2}). Ignoring the reverse reaction (k_{-2}) simplifies the interpretation of the data. This is not an unreasonable assumption if data collection is near the beginning of the reaction, where the concentration of P is low. The assumption that k_{-2} is negligible leads to a simplification of the preceding equation to:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

Through this simplification, the chemists Leonor Michaelis and Maud Menten were able to propose a model that explains the kinetics of many different enzymes. Through their work, an expression relating the catalytic rate to the concentrations of the enzyme and substrate and to the individual rates was developed. The starting point for this expression is the relationship between the rate of the reaction and the concentration of the enzyme-substrate complex:

$$V = k_2[ES]$$

Similarly, the rate of formation of ES is $k_1[E]\,[S]$, and the rate for the breakdown of ES is $(k_{-1} + k_2)\,[ES]$. Throughout most of the reaction, the concentration of ES remains nearly constant. This is the steady-state assumption, which assumes that during a reaction the concentrations of any intermediates remain nearly constant. This assumption means that the rate of formation of ES must be equal to the rate of breakdown of ES, or:

$$k_1[E][S] = (k_{-1} + k_2)[ES]$$

This equation rearranges to:

$$\frac{[E][S]}{[ES]} = \frac{(k_{-1} + k_2)}{k_1} = K_M$$

The combination of the three rate constants yields a new constant: the Michaelis constant, K_M, which is independent of the enzyme and substrate concentrations and is an important characteristic of enzyme-substrate interactions. Using the Michaelis constant, the concentration of ES is:

$$[ES] = \frac{[E][S]}{K_M}$$

When the enzyme concentration is much lower than the substrate concentration, the value of [S] is very close to the total substrate concentration. The enzyme concentration, [E], is equal to the total enzyme concentration, $[E]_T$, minus the concentration of the enzyme-substrate complex, or $[E] = [E]_T - [ES]$. If we enter this relationship into the preceding equation, we get:

$$[ES] = \frac{([E]_T - [ES])[S]}{K_M}$$

Rearranging this equation gives:

$$[ES] = \frac{[E]_T / K_M}{1 + [S]/K_M} = \frac{[E]_T[S]}{[S] + K_M}$$

Substituting this relationship into $V = k_2[ES]$ or ($V = k_{cat}[ES]$) gives:

$$V = k_2[E]_T \frac{[S]}{[S] + K_M}$$

The maximum rate, V_{max}, occurs when all the enzyme molecules are associated with substrate. That is, $[ES] = [E]_T$. This changes $V = k_2[ES]$ to $V_{max} = k_2[ES]_T$. This relationship changes the preceding equation to the Michaelis-Menten equation:

$$V = \frac{V_{max}[S]}{[S] + K_M}$$

This equation accounts for the information depicted in Figure 6-9. At very low concentrations, $[S] << K_M$, we see $V = (V_{max} / K_M) / [S]$, and when [S] is greater than K_M (high [S]), $V = V_{max}$. When $[S] = K_M$ it leads to $V = V_{max} / 2$.

Ideal applications

The Michaelis-Menten equation explains the behavior of many enzymes. It is relatively easy to determine both the K_M and V_{max} values, and this is normally done graphically using computer programs that generate the best-fit curve.

The K_M values vary widely. The value depends on the identity of the substrate and on a variety of environmental factors such as temperature, ionic strength, and pH. Because K_M indicates the substrate concentration required

to fill half of the active sites on the enzyme, it gives an indication of the minimum substrate concentration for significant catalytic activity to occur. It is possible to determine the fraction of sites filled, f_{ES}, from the value of K_M:

$$f_{ES} = \frac{V}{V_{max}} = \frac{[S]}{[S] + K_M}$$

K_M also gives information about the rate constants for the reaction.

$$\frac{(k_{-1} + k_2)}{k_1} = K_M$$

In the special case where k_{-1} is significantly greater than k_2, $K_M = k_{-1} / k_1$, which relates to the equilibrium constant for the dissociation of the enzyme-substrate complex:

$$K_{ES} = \frac{[E][S]}{[ES]} = \frac{k_{-1}}{k_1}$$

Under these special conditions, K_M is a measure of the binding in the enzyme-substrate complex. A high K_M value indicates that the binding is weak, whereas a low value indicates that the binding is strong.

Don't forget: These conclusions only apply under the special conditions of $k_{-1} >> k_2$.

The value of V_{max} supplies the turnover number of the enzyme. The *turnover number* gives the number of substrate molecules transforming to products per unit of time for a fully saturated enzyme. You can determine k_2 from this value. (The constant k_2 is also known as the catalytic constant, k_{cat}.) If the concentration of active sites, $[E]_T$, is known, this relationship applies:

$$V_{max} = k_2[E]_T$$

And:

$$k_2 = V_{max} / [E]_T$$

Realistic applications

The ideas in the preceding section provide useful information about the behavior of many enzymes. In cells, however, the enzymes are seldom saturated with substrate. Under typical conditions [S] / K_M is usually between 1.0 and 0.01. If K_M is much greater than [S], the catalytic rate k_{cat} (or k_2) is significantly less than the ideal value because only a small portion of the active sites contain substrate. The ratio k_{cat} / K_M allows you to compare the substrate preferences of an enzyme.

The maximum rate of catalytic activity is limited by the rate of diffusion to bring the enzyme and substrate together. Some enzymes can exceed this limit by forming *assemblages.* In these groups, the product of one enzyme is the substrate for a closely associated enzyme. This allows a substrate to enter the group and pass from enzyme to enzyme as if it were in an assembly line.

Another complication is that many enzymes require more than one substrate. It is possible to utilize these multiple substrates through sequential displacement or through double displacement. In *sequential displacement,* all substrates must simultaneously bind to the enzyme before the release of the product. In this type of displacement, the order in which the substrates bind is unimportant. In *double displacement,* or *ping-pong,* situations, one or more products leave before all the substrates bind. Double displacement mechanisms temporarily modify the enzyme.

Here we go again: Lineweaver-Burk plots

Once upon a time, before the invention of computers, the determination of K_M and V_{max} was a tedious process. Today curve-fitting programs allow rapid analysis of the data to determine these values. However, a relatively simple method allows a relatively accurate determination of these two constants. This method is to construct a *Lineweaver-Burk plot,* also known as a *double-reciprocal plot.* The basis of a Lineweaver-Burk plot comes from the manipulation of the Michaelis-Menten equation to the form:

$$\frac{1}{V} = \frac{K_M}{V_{max}} \times \frac{1}{[S]} + \frac{1}{V_{max}}$$

This equation has the form y = mx + b, and describes a straight line. A plot of the reciprocal of the rate, 1 / V, versus the reciprocal of the substrate concentration, 1 / [S], gives a line with a y-intercept equal to 1 / V_{max} and an x-intercept of –1 / K_M. An example of this type of plot appears in Figure 6-10.

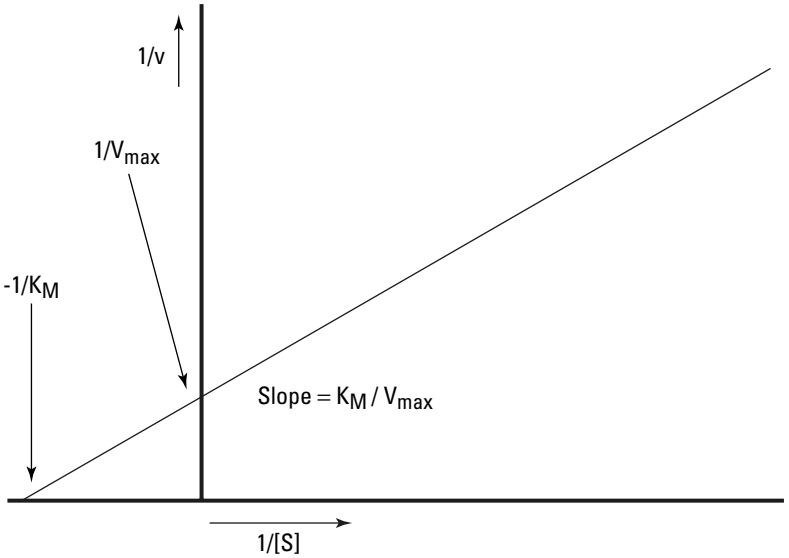


Figure 6-10: A Lineweaver-Burk plot.

The Lineweaver-Burk plot is the most widely used graphical technique for the determination of K_M and V_{max}. However, there are other methods. The Woolf plot, Figure 6-11, uses the equation:

$$\frac{[S]}{V} = \frac{1}{V_{max}} \times [S] + \frac{K_M}{V_{max}}$$

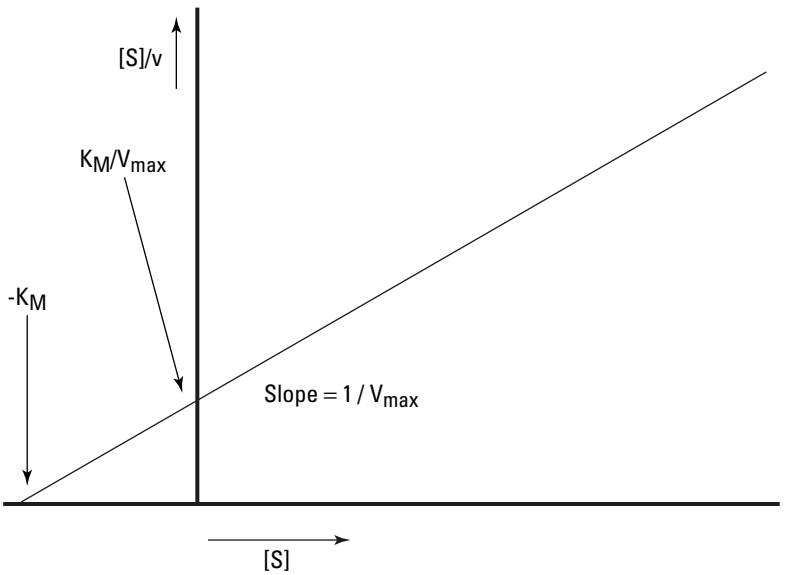


Figure 6-11: A Woolf plot.

Plotting [S] / V versus [S] gives a straight line. An Eadie-Hofstee plot, shown in Figure 6-12, uses the equation:

$$V = -K_M \times \frac{V}{[S]} + V_{max}$$

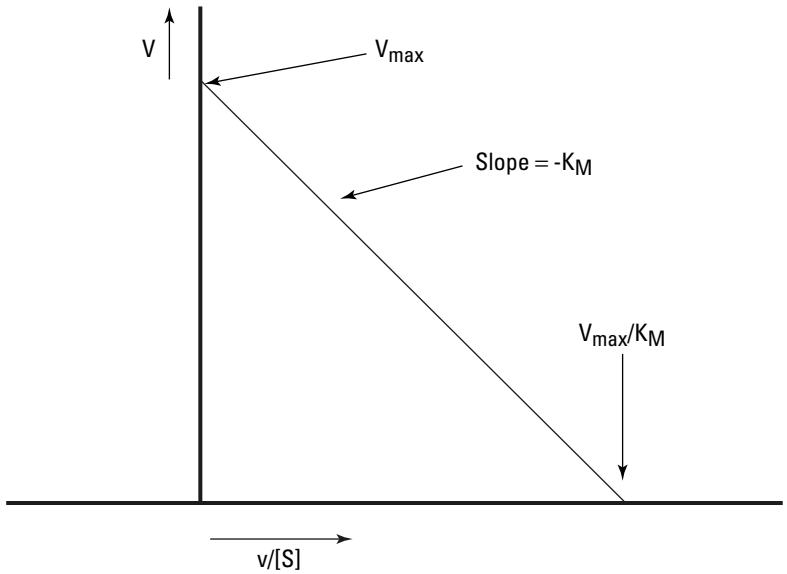


Figure 6-12: An Eadie-Hofstee plot.

Plotting V versus V / [S] gives a straight-line.

Enzyme Inhibition: Slowing It Down

Inibitors are substances that decrease the activity of an enzyme, and they come in two general classes: *competitive* inhibitors, which compete with the substrate, and *noncompetitive* inhibitors, which do not compete. (Mixed inhibition has characteristics of both competitive and noncompetitive inhibition.) In general, these processes are reversible, but there are also irreversible inhibiters that permanently alter the enzyme or bind very strongly to the enzyme. All inhibition may serve as a method of regulating enzymatic activity. There are also many medical applications of this form of inhibition. Examples include anti-epileptic and chemotherapy drugs, along with the ever-popular Viagra. The action of many poisons is also through inhibition.

Competititive inhibition

A competitive inhibitor enters the active site of an enzyme and, thus, prevents the substrate from entering. This prevention results in a decrease in the number of enzyme-substrate complexes that form, and, hence, a decrease in the rate of catalysis. In most cases, a portion of the inhibitor mimics a portion of the substrate. An increase in the substrate concentration overcomes this inhibition because of the increased probability of a substrate molecule entering the active site than an inhibitor molecule.

Noncompetitive inhibition

Noncompetitive inhibitors do not enter the active site but instead bind to some other region of the enzyme. These species usually do not mimic the substrate. This type of inhibitor reduces the turnover number of the enzyme. Unlike competitive inhibition, an increase in the substrate does not overcome noncompetitive inhibition. This type of inhibition takes many different forms, so there is no simple model.

Graphing inhibition

Lineweaver-Burk plots are useful in the study of enzyme inhibition. Figures 6-13 and 6-14 illustrate how the graph changes in the presence of a noncompetitive and a competitive inhibitor. The plot of enzyme inhibition allows us to quickly determine the type of inhibition. In noncompetitive enzyme inhibition, the value of K_M remains unchanged. In competitive inhibition, however, it is $1/V_{max}$ that remains unchanged.

Enzyme Regulation

In general, an increase in the concentration of a substrate, if unregulated, will induce an increase in the rate of reaction. An increase in the concentration of a product will, in general, have the reverse effect. Product regulation is a type of feedback control. In many cases, it is necessary to regulate the activity of enzymes more precisely. There are four general types of enzyme regulation:

- Allosteric control
- Multiple enzyme forms
- Covalent modification
- Proteolytic activation

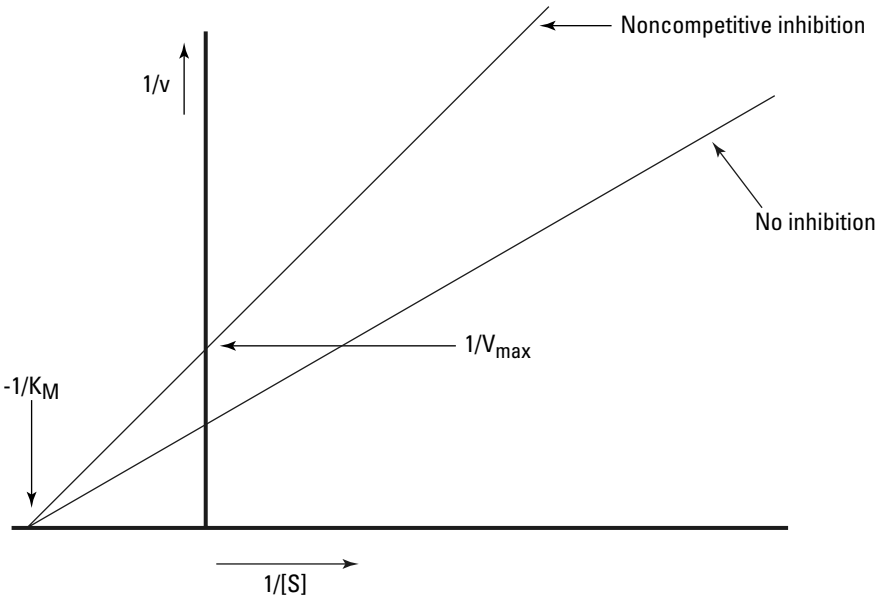


Figure 6-13: A Lineweaver-Burk plot indicating non-competitive inhibition.

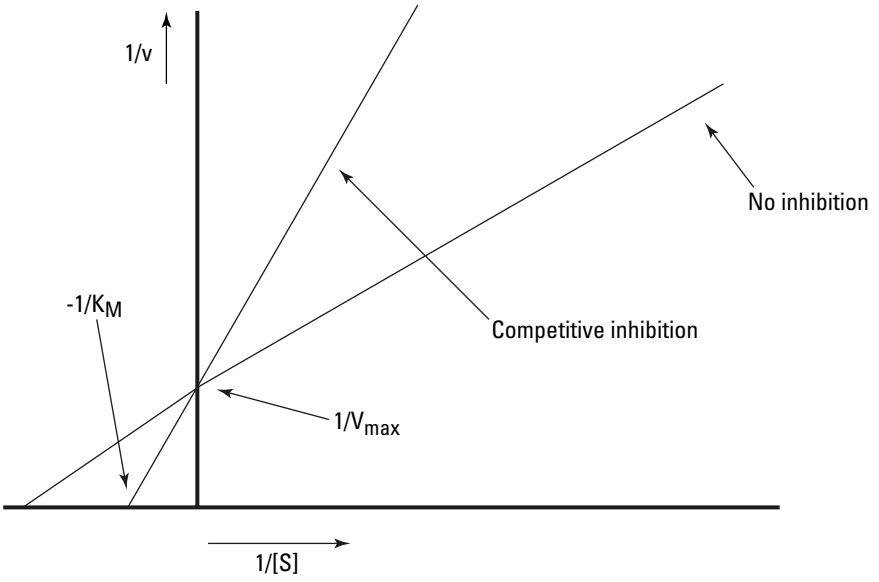


Figure 6-14: A Lineweaver-Burk plot indicating competitive inhibition.

Allosteric control

An allosterically regulated enzyme has a regulatory site. When a small molecule, called a *regulator,* binds to the regulatory site, it induces a conformational change in the enzyme, making it into its active form.

Multiple enzyme forms

Some enzymes have multiple forms known as *isozymes* or *isoenzymes.* There are slight differences in the structures of the forms. These differences lead to differences in the K_M and V_{max} values, and, therefore, in the general activity.

Covalent modification

In this form of regulation, the attachment of a group, often a phosphoryl group, alters the activity of the enzyme. This process is a reversible form of control. Protein kinases catalyze this type of activation, whereas other enzymes catalyze deactivation.

Proteolytic activation

In this form of regulation, an inactive form of an enzyme — a proenzyme or a zymogen — often undergoes irreversible conversion to the active form, often through the hydrolysis of one or more peptide bonds.

Where the money is: Enzymes and industry

The industrial implementation of enzymes originated from studies in the food, wine, and beer industries. Scientists, such as Louis Pasteur, laid much of the groundwork for these applications.

Many of the applications of enzymes to industry involve immobilized enzymes. An *immobilized enzyme* is covalently bonded to an insoluble matrix such as cellulose or glass beads. The immobilization of an enzyme stabilizes it and allows prolonged use. Some useful commercial enzymes are as follows:

Carbohydrases
Amylase: Digestive aid for precooked food
Amyloglucosidase: Converts starch to dextrose
Cellulase and hemicellulase: Conversion of sawdust to sugar and production of liquid coffee concentrates
Glucose isomerase: Production of fructose from cornstarch
Glucose oxidase: Removes glucose from egg solids
Invertase: Stabilizes sugars in soft-centered candy
Lactase: Prevents the crystallization of lactose in ice cream
Pectinase: Clarifies wine and fruit juice
Catalase
Removes H_2O_2 used in the "cold pasteurization" of milk
Proteases
Rennin: Used in cheesemaking
Ficin, Streptodornase, and Trypsin: Debridement of wounds
Pepsin: Digestive aid for precooked food
Papain: Meat tenderizer and beer stabilizer
Bromelain: Meat tenderizer
Alcalase: Additive to detergent for removal of protein stains
Lipoxygenase: Whitening of bread
Lipase: Produces flavor in cheese