SECTION I

Structures & Functions of Proteins & Enzymes

Amino Acids & Peptides

3

Victor W. Rodwell, PhD, & Peter J. Kennelly, PhD

BIOMEDICAL IMPORTANCE

In addition to providing the monomer units from which the long polypeptide chains of proteins are synthesized, the L-α-amino acids and their derivatives participate in cellular functions as diverse as nerve transmission and the biosynthesis of porphyrins, purines, pyrimidines, and urea. Short polymers of amino acids called peptides perform prominent roles in the neuroendocrine system as hormones, hormone-releasing factors, neuromodulators, or neurotransmitters. While proteins contain only L-α-amino acids, microorganisms elaborate peptides that contain both D- and L-α-amino acids. Several of these peptides are of therapeutic value, including the antibiotics bacitracin and gramicidin A and the antitumor agent bleomycin. Certain other microbial peptides are toxic. The cyanobacterial peptides microcystin and nodularin are lethal in large doses, while small quantities promote the formation of hepatic tumors. Neither humans nor any other higher animals can synthesize 10 of the 20 common L-α-amino acids in amounts adequate to support infant growth or to maintain health in adults. Consequently, the human diet must contain adequate quantities of these nutritionally essential amino acids.

PROPERTIES OF AMINO ACIDS

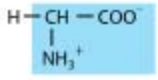
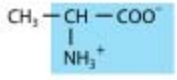
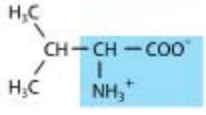
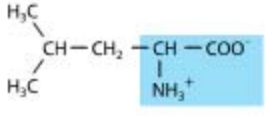
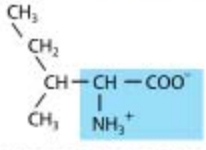
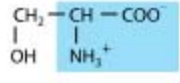
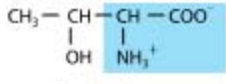
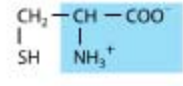
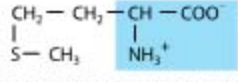
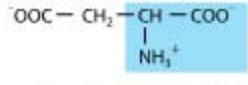
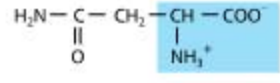
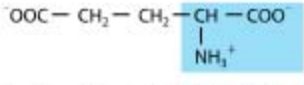
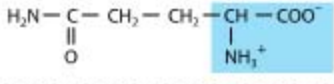
The Genetic Code Specifies 20 L-α-Amino Acids

Of the over 300 naturally occurring amino acids, 20 constitute the monomer units of proteins. While a nonredundant three-letter genetic code could accommodate more than 20 amino acids, its redundancy limits the available codons to the 20 L-α-amino acids listed in Table 3–1, classified according to the polarity of their R groups. Both one- and three-letter abbreviations for each amino acid can be used to represent the amino acids in peptides (Table 3–1). Some proteins contain additional amino acids that arise by modification of an amino acid already present in a peptide. Examples include conversion of peptidyl proline and lysine to 4-hydroxyproline and 5-hydroxylysine; the conversion of peptidyl glutamate to γ-carboxyglutamate; and the methylation, formylation, acetylation, prenylation, and phosphorylation of certain aminoacyl residues. These modifications extend the biologic diversity of proteins by altering their solubility, stability, and interaction with other proteins.

Only L-α-Amino Acids Occur in Proteins

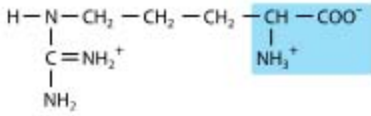
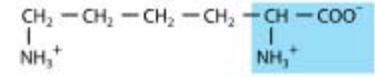
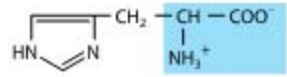
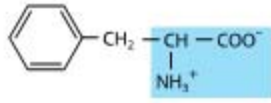
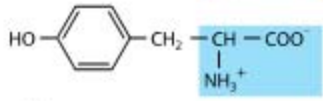
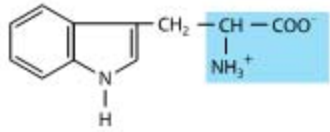
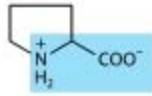
With the sole exception of glycine, the α-carbon of amino acids is chiral. Although some protein amino acids are dextrorotatory and some levorotatory, all share the absolute configuration of L-glyceraldehyde and thus are L-α-amino acids. Several free L-α-amino acids fulfill important roles in metabolic processes. Examples include ornithine, citrulline, and argininosuccinate that participate in urea synthesis; tyrosine in formation of thyroid hormones; and glutamate in neurotransmitter biosynthesis. D-Amino acids that occur naturally include free D-serine and D-aspartate in brain tissue, D-alanine and D-glutamate in the cell walls of gram-positive bacteria, and D-amino acids in some nonmammalian peptides and certain antibiotics.

Table 3–1. L- α-Amino acids present in proteins.

Name	Symbol	Structural Formula	pK_1 α-COOH	pK_2 α-NH_3^+	pK_3 R Group
With Aliphatic Side Chains					
Glycine	Gly [G]		2.4	9.8	
Alanine	Ala [A]		2.4	9.9	
Valine	Val [V]		2.2	9.7	
Leucine	Leu [L]		2.3	9.7	
Isoleucine	Ile [I]		2.3	9.8	
With Side Chains Containing Hydroxylic (OH) Groups					
Serine	Ser [S]		2.2	9.2	about 13
Threonine	Thr [T]		2.1	9.1	about 13
Tyrosine	Tyr [Y]	See below.			
With Side Chains Containing Sulfur Atoms					
Cysteine	Cys [C]		1.9	10.8	8.3
Methionine	Met [M]		2.1	9.3	
With Side Chains Containing Acidic Groups or Their Amides					
Aspartic acid	Asp [D]		2.0	9.9	3.9
Asparagine	Asn [N]		2.1	8.8	
Glutamic acid	Glu [E]		2.1	9.5	4.1
Glutamine	Gln [Q]		2.2	9.1	

(*continued*)

Table 3–1. L-α-Amino acids present in proteins. (continued)

Name	Symbol	Structural Formula	pK_1	pK_2	pK_3
With Side Chains Containing Basic Groups			α-COOH	α-NH_3^+	R Group
Arginine	Arg [R]		1.8	9.0	12.5
Lysine	Lys [K]		2.2	9.2	10.8
Histidine	His [H]		1.8	9.3	6.0
Containing Aromatic Rings					
Histidine	His [H]	See above.			
Phenylalanine	Phe [F]		2.2	9.2	
Tyrosine	Tyr [Y]		2.2	9.1	10.1
Tryptophan	Trp [W]		2.4	9.4	
Imino Acid					
Proline	Pro [P]		2.0	10.6	

Amino Acids May Have Positive, Negative, or Zero Net Charge

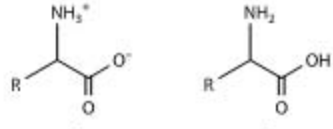
Charged and uncharged forms of the ionizable —COOH and —NH_3^+ weak acid groups exist in solution in protonic equilibrium:

$$R—COOH = R—COO^- + H^+$$

$$R—NH_3^+ = R—NH_2 + H^+$$

While both R—COOH and R—NH_3^+ are weak acids, R—COOH is a far stronger acid than R—NH_3^+. At physiologic pH (pH 7.4), carboxyl groups exist almost entirely as R—COO^- and amino groups predominantly as R—NH_3^+. Figure 3–1 illustrates the effect of pH on the charged state of aspartic acid.

Molecules that contain an equal number of ionizable groups of opposite charge and that therefore bear no net charge are termed **zwitterions.** Amino acids in blood and most tissues thus should be represented as in A, below.



A **B**

Structure B cannot exist in aqueous solution because at any pH low enough to protonate the carboxyl group the amino group would also be protonated. Similarly, at any pH sufficiently high for an uncharged amino

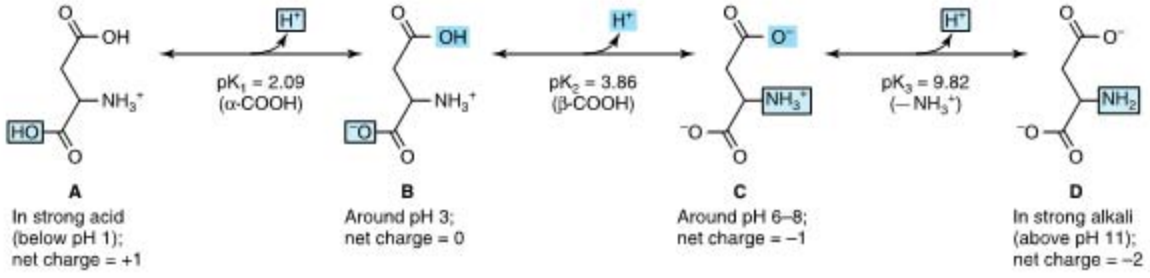


Figure 3–1. Protonic equilibria of aspartic acid.

group to predominate, a carboxyl group will be present as R—COO^-. The uncharged representation B (above) is, however, often used for reactions that do not involve protonic equilibria.

pK_a Values Express the Strengths of Weak Acids

The acid strengths of weak acids are expressed as their **pK_a** (Table 3–1). The imidazole group of histidine and the guanidino group of arginine exist as resonance hybrids with positive charge distributed between both nitrogens (histidine) or all three nitrogens (arginine) (Figure 3–2). The net charge on an amino acid—the algebraic sum of all the positively and negatively charged groups present—depends upon the pK_a values of its functional groups and on the pH of the surrounding medium. Altering the charge on amino acids and their derivatives by varying the pH facilitates the physical separation of amino acids, peptides, and proteins (see Chapter 4).

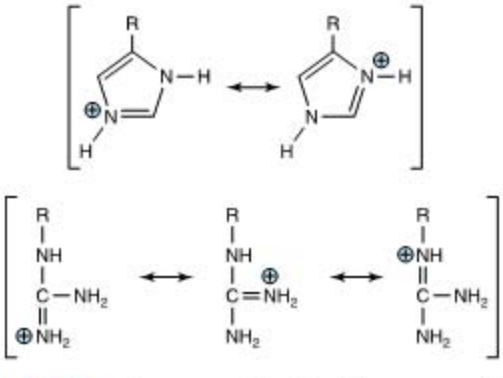


Figure 3–2. Resonance hybrids of the protonated forms of the R groups of histidine and arginine.

At Its Isoelectric pH (pI), an Amino Acid Bears No Net Charge

The **isoelectric** species is the form of a molecule that has an equal number of positive and negative charges and thus is electrically neutral. The isoelectric pH, also called the pI, is the pH midway between pK_a values on either side of the isoelectric species. For an amino acid such as alanine that has only two dissociating groups, there is no ambiguity. The first pK_a (R—COOH) is 2.35 and the second pK_a (R—NH_3^+) is 9.69. The isoelectric pH (pI) of alanine thus is

$$\text{pI} = \frac{pK_1 + pK_2}{2} = \frac{2.35 + 9.69}{2} = 6.02$$

For polyfunctional acids, pI is also the pH midway between the pK_a values on either side of the isoionic species. For example, the pI for aspartic acid is

$$\text{pI} = \frac{pK_1 + pK_2}{2} = \frac{2.09 + 3.96}{2} = 3.02$$

For lysine, pI is calculated from:

$$\text{pI} = \frac{pK_2 + pK_3}{2}$$

Similar considerations apply to all polyprotic acids (eg, proteins), regardless of the number of dissociating groups present. In the clinical laboratory, knowledge of the pI guides selection of conditions for electrophoretic separations. For example, electrophoresis at pH 7.0 will separate two molecules with pI values of 6.0 and 8.0 because at pH 8.0 the molecule with a pI of 6.0 will have a net positive charge, and that with pI of 8.0 a net negative charge. Similar considerations apply to understanding chromatographic separations on ionic supports such as DEAE cellulose (see Chapter 4).

pK_a Values Vary With the Environment

The environment of a dissociable group affects its pK_a. The pK_a values of the R groups of free amino acids in aqueous solution (Table 3–1) thus provide only an approximate guide to the pK_a values of the same amino acids when present in proteins. A polar environment favors the charged form (R—COO^- or R—NH_3^+), and a nonpolar environment favors the uncharged form (R—COOH or R—NH_2). A nonpolar environment thus *raises* the pK_a of a carboxyl group (making it a weaker acid) but *lowers* that of an amino group (making it a stronger acid). The presence of adjacent charged groups can reinforce or counteract solvent effects. The pK_a of a functional group thus will depend upon its location within a given protein. Variations in pK_a can encompass whole pH units (Table 3–2). pK_a values that diverge from those listed by as much as three pH units are common at the active sites of enzymes. An extreme example, a buried aspartic acid of thioredoxin, has a pK_a above 9—a shift of over six pH units!

The Solubility and Melting Points of Amino Acids Reflect Their Ionic Character

The charged functional groups of amino acids ensure that they are readily solvated by—and thus soluble in—polar solvents such as water and ethanol but insoluble in nonpolar solvents such as benzene, hexane, or ether. Similarly, the high amount of energy required to disrupt the ionic forces that stabilize the crystal lattice account for the high melting points of amino acids (> 200 °C).

Amino acids do not absorb visible light and thus are colorless. However, tyrosine, phenylalanine, and especially tryptophan absorb high-wavelength (250–290 nm) ultraviolet light. Tryptophan therefore makes the major contribution to the ability of most proteins to absorb light in the region of 280 nm.

Table 3–2. Typical range of pK_a values for ionizable groups in proteins.

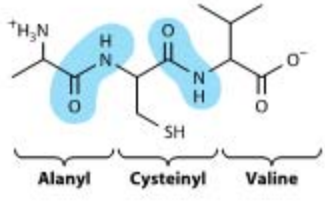
Dissociating Group	pK_a Range
α-Carboxyl	3.5–4.0
Non-α COOH of Asp or Glu	4.0–4.8
Imidazole of His	6.5–7.4
SH of Cys	8.5–9.0
OH of Tyr	9.5–10.5
α-Amino	8.0–9.0
ε-Amino of Lys	9.8–10.4
Guanidinium of Arg	~12.0

THE α-R GROUPS DETERMINE THE PROPERTIES OF AMINO ACIDS

Since glycine, the smallest amino acid, can be accommodated in places inaccessible to other amino acids, it often occurs where peptides bend sharply. The hydrophobic R groups of alanine, valine, leucine, and isoleucine and the aromatic R groups of phenylalanine, tyrosine, and tryptophan typically occur primarily in the interior of cytosolic proteins. The charged R groups of basic and acidic amino acids stabilize specific protein conformations via ionic interactions, or salt bonds. These bonds also function in "charge relay" systems during enzymatic catalysis and electron transport in respiring mitochondria. Histidine plays unique roles in enzymatic catalysis. The pK_a of its imidazole proton permits it to function at neutral pH as either a base or an acid catalyst. The primary alcohol group of serine and the primary thioalcohol (—SH) group of cysteine are excellent nucleophiles and can function as such during enzymatic catalysis. However, the secondary alcohol group of threonine, while a good nucleophile, does not fulfill an analogous role in catalysis. The —OH groups of serine, tyrosine, and threonine also participate in regulation of the activity of enzymes whose catalytic activity depends on the phosphorylation state of these residues.

FUNCTIONAL GROUPS DICTATE THE CHEMICAL REACTIONS OF AMINO ACIDS

Each functional group of an amino acid exhibits all of its characteristic chemical reactions. For carboxylic acid groups, these reactions include the formation of esters, amides, and acid anhydrides; for amino groups, acylation, amidation, and esterification; and for —OH and —SH groups, oxidation and esterification. The most important reaction of amino acids is the formation of a peptide bond (shaded blue).



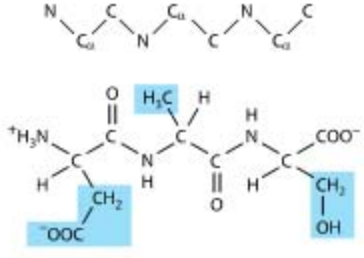
Amino Acid Sequence Determines Primary Structure

The number and order of all of the amino acid residues in a polypeptide constitute its primary structure. Amino acids present in peptides are called aminoacyl residues and are named by replacing the *-ate* or *-ine* suffixes of free amino acids with *-yl* (eg, alanyl, asparty*l*, ty-

rosyl). Peptides are then named as derivatives of the carboxyl terminal aminoacyl residue. For example, Lys-Leu-Tyr-Gln is called lysyl-leucyl-tyrosyl-glutamine. The *-ine* ending on glutamine indicates that its α-carboxyl group is *not* involved in peptide bond formation.

Peptide Structures Are Easy to Draw

Prefixes like *tri-* or *octa-* denote peptides with three or eight **residues,** respectively, not those with three or eight peptide **bonds.** By convention, peptides are written with the residue that bears the free α-amino group at the left. To draw a peptide, use a zigzag to represent the main chain or backbone. Add the main chain atoms, which occur in the repeating order: α-nitrogen, α-carbon, carbonyl carbon. Now add a hydrogen atom to each α-carbon and to each peptide nitrogen, and an oxygen to the carbonyl carbon. Finally, add the appropriate R groups (shaded) to each α-carbon atom.



Three-letter abbreviations linked by straight lines represent an unambiguous primary structure. Lines are omitted for single-letter abbreviations.

Glu - Ala - Lys - Gly - Tyr - Ala

E A K G Y A

Where there is uncertainty about the order of a portion of a polypeptide, the questionable residues are enclosed in brackets and separated by commas.

Glu - Lys - (Ala, Gly, Tyr) - His - Ala

Some Peptides Contain Unusual Amino Acids

In mammals, peptide hormones typically contain only the α-amino acids of proteins linked by standard peptide bonds. Other peptides may, however, contain nonprotein amino acids, derivatives of the protein amino acids, or amino acids linked by an atypical peptide bond. For example, the amino terminal glutamate of glutathione, which participates in protein folding and in the metabolism of xenobiotics (Chapter 53), is linked to cysteine by a non-α peptide bond (Figure 3–3). The amino terminal glutamate of thyrotropin-releasing hormone (TRH) is cyclized to pyroglutamic acid, and the carboxyl group of the carboxyl terminal prolyl residue is amidated. Peptides elaborated by fungi, bacteria, and lower animals can contain nonprotein amino acids. The antibiotics tyrocidin and gramicidin S are cyclic polypeptides that contain D-phenylalanine and ornithine. The heptapeptide opioids dermorphin and deltophorin in the skin of South American tree frogs contain D-tyrosine and D-alanine.

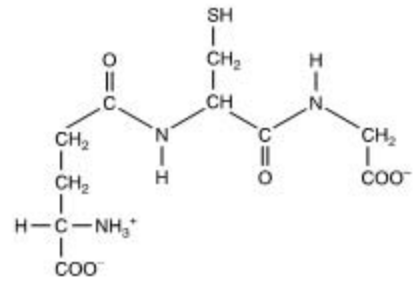


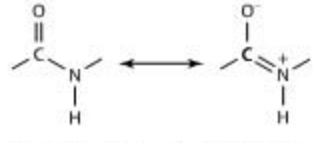
Figure 3–3. Glutathione (γ-glutamyl-cysteinyl-glycine). Note the non-α peptide bond that links Glu to Cys.

Peptides Are Polyelectrolytes

The peptide bond is uncharged at any pH of physiologic interest. Formation of peptides from amino acids is therefore accompanied by a net loss of one positive and one negative charge per peptide bond formed. Peptides nevertheless are charged at physiologic pH owing to their carboxyl and amino terminal groups and, where present, their acidic or basic R groups. As for amino acids, the net charge on a peptide depends on the pH of its environment and on the pK_a values of its dissociating groups.

The Peptide Bond Has Partial Double-Bond Character

Although peptides are written as if a single bond linked the α-carboxyl and α-nitrogen atoms, this bond in fact exhibits partial double-bond character:



There thus is no freedom of rotation about the bond that connects the carbonyl carbon and the nitrogen of a peptide bond. Consequently, all four of the colored atoms of Figure 3–4 are coplanar. The imposed semirigidity of the peptide bond has important conse-

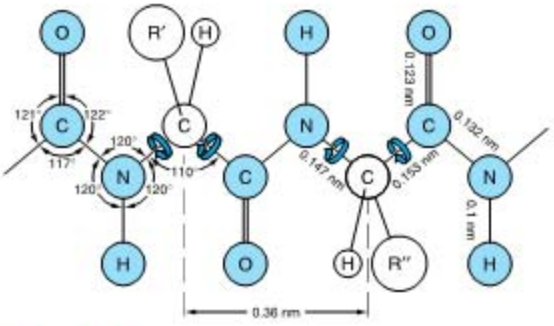


Figure 3–4. Dimensions of a fully extended polypeptide chain. The four atoms of the peptide bond (colored blue) are coplanar. The unshaded atoms are the α-carbon atom, the α-hydrogen atom, and the α-R group of the particular amino acid. Free rotation can occur about the bonds that connect the α-carbon with the α-nitrogen and with the α-carbonyl carbon (blue arrows). The extended polypeptide chain is thus a semirigid structure with two-thirds of the atoms of the backbone held in a fixed planar relationship one to another. The distance between adjacent α-carbon atoms is 0.36 nm (3.6 Å). The interatomic distances and bond angles, which are not equivalent, are also shown. (Redrawn and reproduced, with permission, from Pauling L, Corey LP, Branson HR: The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. Proc Natl Acad Sci U S A 1951;37:205.)

quences for higher orders of protein structure. Encircling arrows (Figure 3–4) indicate free rotation about the remaining bonds of the polypeptide backbone.

Noncovalent Forces Constrain Peptide Conformations

Folding of a peptide probably occurs coincident with its biosynthesis (see Chapter 38). The physiologically active conformation reflects the amino acid sequence, steric hindrance, and noncovalent interactions (eg, hydrogen bonding, hydrophobic interactions) between residues. Common conformations include α-helices and β-pleated sheets (see Chapter 5).

ANALYSIS OF THE AMINO ACID CONTENT OF BIOLOGIC MATERIALS

In order to determine the identity and quantity of each amino acid in a sample of biologic material, it is first necessary to hydrolyze the peptide bonds that link the amino acids together by treatment with hot HCl. The resulting mixture of free amino acids is then treated with 6-amino-*N*-hydroxysuccinimidyl carbamate, which reacts with their α-amino groups, forming fluorescent derivatives that are then separated and identified using high-pressure liquid chromatography (see Chapter 5). Ninhydrin, also widely used for detecting amino acids, forms a purple product with α-amino acids and a yellow adduct with the imine groups of proline and hydroxyproline.

SUMMARY

- Both D-amino acids and non-α-amino acids occur in nature, but only L-α-amino acids are present in proteins.
- All amino acids possess at least two weakly acidic functional groups, $R—NH_3^+$ and R—COOH. Many also possess additional weakly acidic functional groups such as —OH, —SH, guanidino, or imidazole groups.
- The pK_a values of all functional groups of an amino acid dictate its net charge at a given pH. pI is the pH at which an amino acid bears no net charge and thus does not move in a direct current electrical field.
- Of the biochemical reactions of amino acids, the most important is the formation of peptide bonds.
- The R groups of amino acids determine their unique biochemical functions. Amino acids are classified as basic, acidic, aromatic, aliphatic, or sulfur-containing based on the properties of their R groups.
- Peptides are named for the number of amino acid residues present, and as derivatives of the carboxyl terminal residue. The primary structure of a peptide is its amino acid sequence, starting from the amino-terminal residue.
- The partial double-bond character of the bond that links the carbonyl carbon and the nitrogen of a peptide renders four atoms of the peptide bond coplanar and restricts the number of possible peptide conformations.

REFERENCES

Doolittle RF: Reconstructing history with amino acid sequences. Protein Sci 1992;1:191.

Kreil G: D-Amino acids in animal peptides. Annu Rev Biochem 1997;66:337.

Nokihara K, Gerhardt J: Development of an improved automated gas-chromatographic chiral analysis system: application to non-natural amino acids and natural protein hydrolysates. Chirality 2001;13:431.

Sanger F: Sequences, sequences, and sequences. Annu Rev Biochem 1988;57:1.

Wilson NA et al: Aspartic acid 26 in reduced *Escherichia coli* thioredoxin has a pK_a greater than 9. Biochemistry 1995;34:8931.

Proteins: Determination of Primary Structure

4

Victor W. Rodwell, PhD, & Peter J. Kennelly, PhD

BIOMEDICAL IMPORTANCE

Proteins perform multiple critically important roles. An internal protein network, the cytoskeleton (Chapter 49), maintains cellular shape and physical integrity. Actin and myosin filaments form the contractile machinery of muscle (Chapter 49). Hemoglobin transports oxygen (Chapter 6), while circulating antibodies search out foreign invaders (Chapter 50). Enzymes catalyze reactions that generate energy, synthesize and degrade biomolecules, replicate and transcribe genes, process mRNAs, etc (Chapter 7). Receptors enable cells to sense and respond to hormones and other environmental cues (Chapters 42 and 43). An important goal of molecular medicine is the identification of proteins whose presence, absence, or deficiency is associated with specific physiologic states or diseases. The primary sequence of a protein provides both a molecular fingerprint for its identification and information that can be used to identify and clone the gene or genes that encode it.

PROTEINS & PEPTIDES MUST BE PURIFIED PRIOR TO ANALYSIS

Highly purified protein is essential for determination of its amino acid sequence. Cells contain thousands of different proteins, each in widely varying amounts. The isolation of a specific protein in quantities sufficient for analysis thus presents a formidable challenge that may require multiple successive purification techniques. Classic approaches exploit differences in relative solubility of individual proteins as a function of pH (isoelectric precipitation), polarity (precipitation with ethanol or acetone), or salt concentration (salting out with ammonium sulfate). Chromatographic separations partition molecules between two phases, one mobile and the other stationary. For separation of amino acids or sugars, the stationary phase, or matrix, may be a sheet of filter paper (paper chromatography) or a thin layer of cellulose, silica, or alumina (thin-layer chromatography; TLC).

Column Chromatography

Column chromatography of proteins employs as the stationary phase a column containing small spherical beads of modified cellulose, acrylamide, or silica whose surface typically has been coated with chemical functional groups. These stationary phase matrices interact with proteins based on their charge, hydrophobicity, and ligand-binding properties. A protein mixture is applied to the column and the liquid mobile phase is percolated through it. Small portions of the mobile phase or eluant are collected as they emerge (Figure 4–1).

Partition Chromatography

Column chromatographic separations depend on the relative affinity of different proteins for a given stationary phase and for the mobile phase. Association between each protein and the matrix is weak and transient. Proteins that interact more strongly with the stationary phase are retained longer. The length of time that a protein is associated with the stationary phase is a function of the composition of both the stationary and mobile phases. Optimal separation of the protein of interest from other proteins thus can be achieved by careful manipulation of the composition of the two phases.

Size Exclusion Chromatography

Size exclusion—or gel filtration—chromatography separates proteins based on their **Stokes radius,** the diameter of the sphere they occupy as they tumble in solution. The Stokes radius is a function of molecular mass and shape. A tumbling elongated protein occupies a larger volume than a spherical protein of the same mass. Size exclusion chromatography employs porous beads (Figure 4–2). The pores are analogous to indentations in a river bank. As objects move downstream, those that enter an indentation are retarded until they drift back into the main current. Similarly, proteins with Stokes radii too large to enter the pores (excluded proteins) remain in the flowing mobile phase and emerge before proteins that can enter the pores (included proteins).

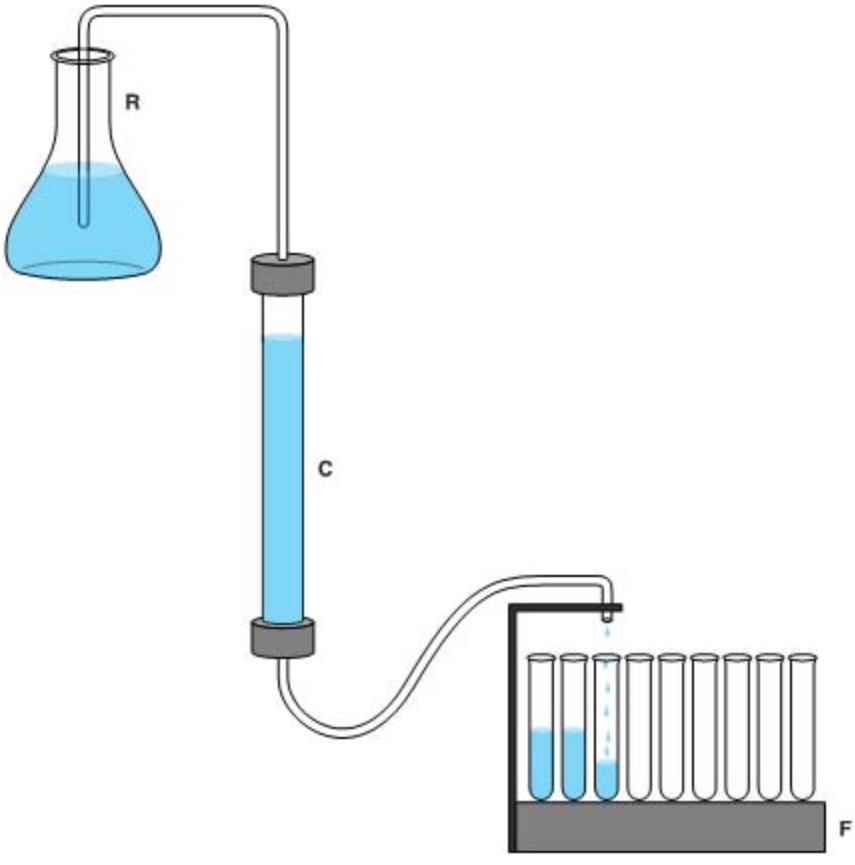


Figure 4–1. Components of a simple liquid chromatography apparatus. **R:** Reservoir of mobile phase liquid, delivered either by gravity or using a pump. **C:** Glass or plastic column containing stationary phase. **F:** Fraction collector for collecting portions, called fractions, of the eluant liquid in separate test tubes.

Proteins thus emerge from a gel filtration column in descending order of their Stokes radii.

Absorption Chromatography

For absorption chromatography, the protein mixture is applied to a column under conditions where the protein of interest associates with the stationary phase so tightly that its partition coefficient is essentially unity. Nonadhering molecules are first eluted and discarded. Proteins are then sequentially released by disrupting the forces that stabilize the protein-stationary phase complex, most often by using a gradient of increasing salt concentration. The composition of the mobile phase is altered gradually so that molecules are selectively released in descending order of their affinity for the stationary phase.

Ion Exchange Chromatography

In ion exchange chromatography, proteins interact with the stationary phase by charge-charge interactions. Proteins with a net positive charge at a given pH adhere to beads with negatively charged functional groups such as carboxylates or sulfates (cation exchangers). Similarly, proteins with a net negative charge adhere to beads with positively charged functional groups, typically tertiary or quaternary amines (anion exchangers). Proteins, which are polyanions, compete against monovalent ions for binding to the support—thus the term "ion exchange." For example, proteins bind to diethylaminoethyl (DEAE) cellulose by replacing the counter-ions (generally Cl^- or CH_3COO^-) that neutralize the protonated amine. Bound proteins are selectively displaced by gradually raising the concentration of monovalent ions in

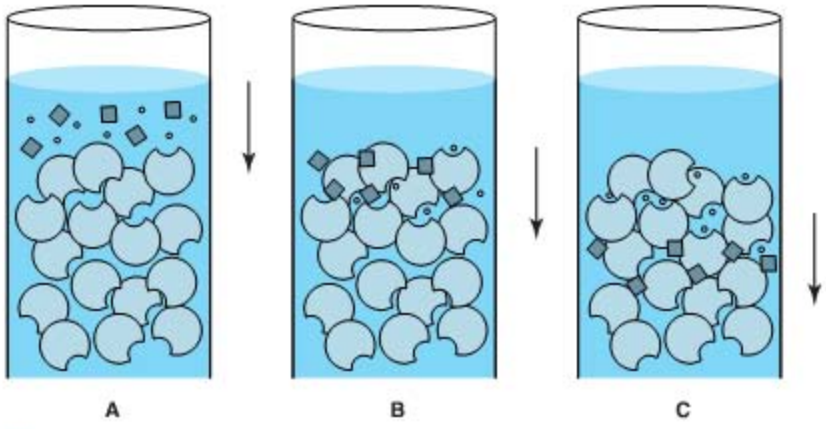


Figure 4–2. Size-exclusion chromatography. **A:** A mixture of large molecules (diamonds) and small molecules (circles) are applied to the top of a gel filtration column. **B:** Upon entering the column, the small molecules enter pores in the stationary phase matrix from which the large molecules are excluded. **C:** As the mobile phase flows down the column, the large, excluded molecules flow with it while the small molecules, which are temporarily sheltered from the flow when inside the pores, lag farther and farther behind.

the mobile phase. Proteins elute in inverse order of the strength of their interactions with the stationary phase.

Since the net charge on a protein is determined by the pH (see Chapter 3), sequential elution of proteins may be achieved by changing the pH of the mobile phase. Alternatively, a protein can be subjected to consecutive rounds of ion exchange chromatography, each at a different pH, such that proteins that co-elute at one pH elute at different salt concentrations at another pH.

Hydrophobic Interaction Chromatography

Hydrophobic interaction chromatography separates proteins based on their tendency to associate with a stationary phase matrix coated with hydrophobic groups (eg, phenyl Sepharose, octyl Sepharose). Proteins with exposed hydrophobic surfaces adhere to the matrix via hydrophobic interactions that are enhanced by a mobile phase of high ionic strength. Nonadherent proteins are first washed away. The polarity of the mobile phase is then decreased by gradually lowering the salt concentration. If the interaction between protein and stationary phase is particularly strong, ethanol or glycerol may be added to the mobile phase to decrease its polarity and further weaken hydrophobic interactions.

Affinity Chromatography

Affinity chromatography exploits the high selectivity of most proteins for their ligands. Enzymes may be purified by affinity chromatography using immobilized substrates, products, coenzymes, or inhibitors. In theory, only proteins that interact with the immobilized ligand adhere. Bound proteins are then eluted either by competition with soluble ligand or, less selectively, by disrupting protein-ligand interactions using urea, guanidine hydrochloride, mildly acidic pH, or high salt concentrations. Stationary phase matrices available commercially contain ligands such as NAD^+ or ATP analogs. Among the most powerful and widely applicable affinity matrices are those used for the purification of suitably modified recombinant proteins. These include a Ni^{2+} matrix that binds proteins with an attached polyhistidine "tag" and a glutathione matrix that binds a recombinant protein linked to glutathione *S*-transferase.

Peptides Are Purified by Reversed-Phase High-Pressure Chromatography

The stationary phase matrices used in classic column chromatography are spongy materials whose compressibility limits flow of the mobile phase. High-pressure liquid chromatography (HPLC) employs incompressible silica or alumina microbeads as the stationary phase and pressures of up to a few thousand psi. Incompressible matrices permit both high flow rates and enhanced resolution. HPLC can resolve complex mixtures of lipids or peptides whose properties differ only slightly. Reversed-phase HPLC exploits a hydrophobic stationary phase of

aliphatic polymers 3–18 carbon atoms in length. Peptide mixtures are eluted using a gradient of a water-miscible organic solvent such as acetonitrile or methanol.

Protein Purity Is Assessed by Polyacrylamide Gel Electrophoresis (PAGE)

The most widely used method for determining the purity of a protein is SDS-PAGE—polyacrylamide gel electrophoresis (PAGE) in the presence of the anionic detergent sodium dodecyl sulfate (SDS). Electrophoresis separates charged biomolecules based on the rates at which they migrate in an applied electrical field. For SDS-PAGE, acrylamide is polymerized and cross-linked to form a porous matrix. SDS denatures and binds to proteins at a ratio of one molecule of SDS per two peptide bonds. When used in conjunction with 2-mercaptoethanol or dithiothreitol to reduce and break disulfide bonds (Figure 4–3), SDS separates the component polypeptides of multimeric proteins. The large number of anionic SDS molecules, each bearing a charge of −1, on each polypeptide overwhelms the charge contributions of the amino acid functional groups. Since the charge-to-mass ratio of each SDS-polypeptide complex is approximately equal, the physical resistance each peptide encounters as it moves through the acrylamide matrix determines the rate of migration. Since large complexes encounter greater resistance, polypeptides separate based on their relative molecular mass (M_r). Individual polypeptides trapped in the acrylamide gel are visualized by staining with dyes such as Coomassie blue (Figure 4–4).

Isoelectric Focusing (IEF)

Ionic buffers called ampholytes and an applied electric field are used to generate a pH gradient within a polyacrylamide matrix. Applied proteins migrate until they reach the region of the matrix where the pH matches their isoelectric point (pI), the pH at which a peptide's net charge is zero. IEF is used in conjunction with SDS-PAGE for two-dimensional electrophoresis, which separates polypeptides based on pI in one dimension and based on M_r in the second (Figure 4–5). Two-dimensional electrophoresis is particularly well suited for separating the components of complex mixtures of proteins.

SANGER WAS THE FIRST TO DETERMINE THE SEQUENCE OF A POLYPEPTIDE

Mature insulin consists of the 21-residue A chain and the 30-residue B chain linked by disulfide bonds. Frederick Sanger reduced the disulfide bonds (Figure 4–3),

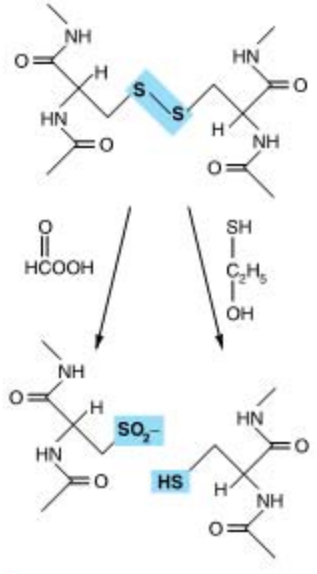


Figure 4–3. Oxidative cleavage of adjacent polypeptide chains linked by disulfide bonds (shaded) by performic acid **(left)** or reductive cleavage by β-mercaptoethanol **(right)** forms two peptides that contain cysteic acid residues or cysteinyl residues, respectively.

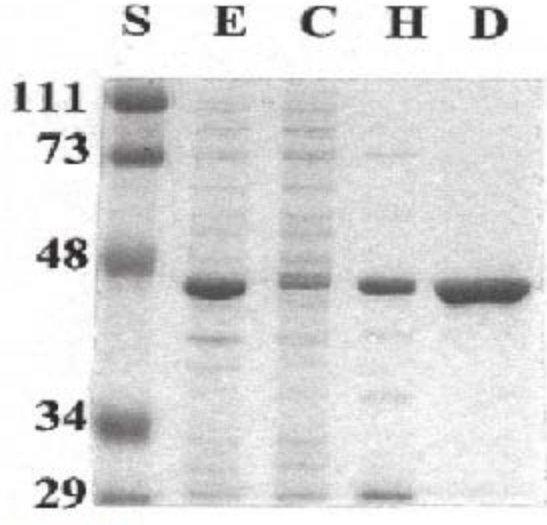


Figure 4–4. Use of SDS-PAGE to observe successive purification of a recombinant protein. The gel was stained with Coomassie blue. Shown are protein standards (lane **S**) of the indicated mass, crude cell extract **(E),** high-speed supernatant liquid **(H),** and the DEAE-Sepharose fraction **(D).** The recombinant protein has a mass of about 45 kDa.

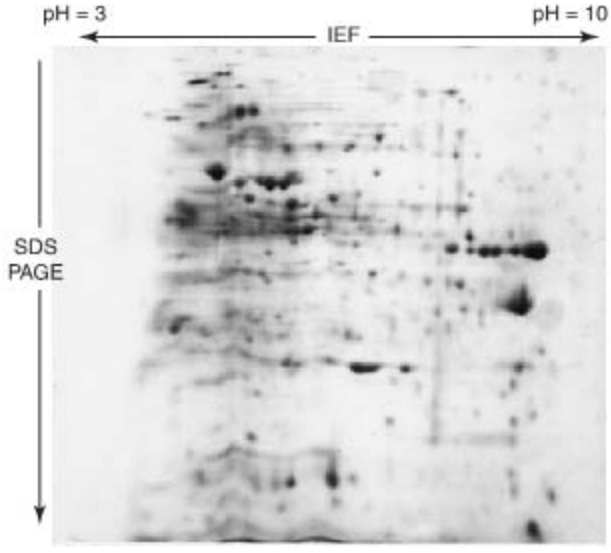


Figure 4–5. Two-dimensional IEF-SDS-PAGE. The gel was stained with Coomassie blue. A crude bacterial extract was first subjected to isoelectric focusing (IEF) in a pH 3–10 gradient. The IEF gel was then placed horizontally on the top of an SDS gel, and the proteins then further resolved by SDS-PAGE. Notice the greatly improved resolution of distinct polypeptides relative to ordinary SDS-PAGE gel (Figure 4–4).

separated the A and B chains, and cleaved each chain into smaller peptides using trypsin, chymotrypsin, and pepsin. The resulting peptides were then isolated and treated with acid to hydrolyze peptide bonds and generate peptides with as few as two or three amino acids. Each peptide was reacted with 1-fluoro-2,4-dinitrobenzene (Sanger's reagent), which derivatizes the exposed α-amino group of amino terminal residues. The amino acid content of each peptide was then determined. While the ε-amino group of lysine also reacts with Sanger's reagent, amino-terminal lysines can be distinguished from those at other positions because they react with 2 mol of Sanger's reagent. Working backwards to larger fragments enabled Sanger to determine the complete sequence of insulin, an accomplishment for which he received a Nobel Prize in 1958.

THE EDMAN REACTION ENABLES PEPTIDES & PROTEINS TO BE SEQUENCED

Pehr Edman introduced phenylisothiocyanate (Edman's reagent) to selectively label the amino-terminal residue of a peptide. In contrast to Sanger's reagent, the phenylthiohydantoin (PTH) derivative can be removed under mild conditions to generate a new amino terminal residue (Figure 4–6). Successive rounds of derivatization with Edman's reagent can therefore be used to sequence many residues of a single sample of peptide. Edman sequencing has been automated, using a thin film or solid matrix to immobilize the peptide and HPLC to identify PTH amino acids. Modern gas-phase sequencers can analyze as little as a few picomoles of peptide.

Large Polypeptides Are First Cleaved Into Smaller Segments

While the first 20–30 residues of a peptide can readily be determined by the Edman method, most polypeptides contain several hundred amino acids. Consequently, most polypeptides must first be cleaved into smaller peptides prior to Edman sequencing. Cleavage also may be necessary to circumvent posttranslational modifications that render a protein's α-amino group "blocked", or unreactive with the Edman reagent.

It usually is necessary to generate several peptides using more than one method of cleavage. This reflects both inconsistency in the spacing of chemically or enzymatically susceptible cleavage sites and the need for sets of peptides whose sequences overlap so one can infer the sequence of the polypeptide from which they derive (Figure 4–7). Reagents for the chemical or enzymatic cleavage of proteins include cyanogen bromide (CNBr), trypsin, and *Staphylococcus aureus* V8 protease (Table 4–1). Following cleavage, the resulting peptides are purified by reversed-phase HPLC—or occasionally by SDS-PAGE—and sequenced.

MOLECULAR BIOLOGY HAS REVOLUTIONIZED THE DETERMINATION OF PRIMARY STRUCTURE

Knowledge of DNA sequences permits deduction of the primary structures of polypeptides. DNA sequencing requires only minute amounts of DNA and can readily yield the sequence of hundreds of nucleotides. To clone and sequence the DNA that encodes a partic-

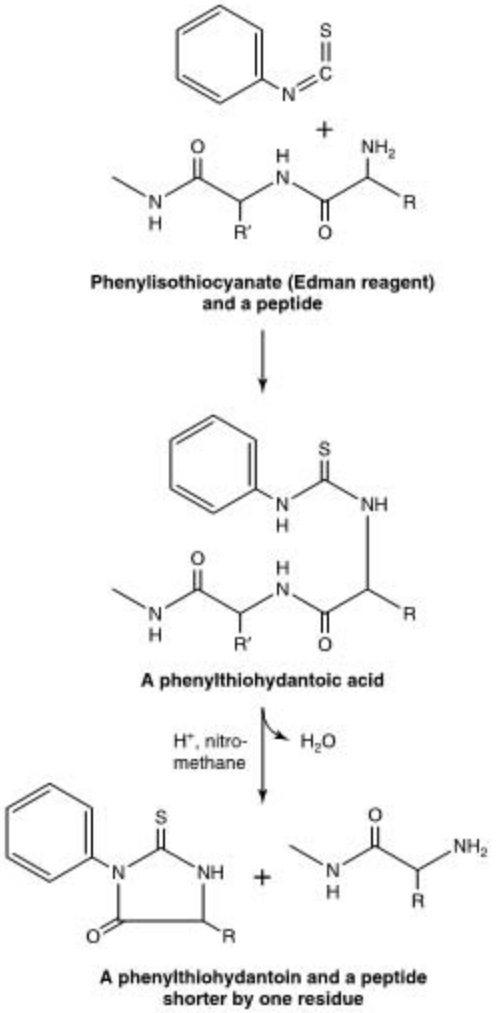


Figure 4–6. The Edman reaction. Phenylisothiocyanate derivatizes the amino-terminal residue of a peptide as a phenylthiohydantoic acid. Treatment with acid in a nonhydroxylic solvent releases a phenylthiohydantoin, which is subsequently identified by its chromatographic mobility, and a peptide one residue shorter. The process is then repeated.

ular protein, some means of identifying the correct clone—eg, knowledge of a portion of its nucleotide sequence—is essential. A hybrid approach thus has emerged. Edman sequencing is used to provide a partial amino acid sequence. Oligonucleotide primers modeled on this partial sequence can then be used to identify clones or to amplify the appropriate gene by the polymerase chain reaction (PCR) (see Chapter 40). Once an authentic DNA clone is obtained, its oligonucleotide

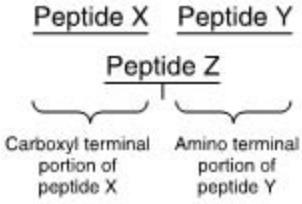


Figure 4–7. The overlapping peptide Z is used to deduce that peptides X and Y are present in the original protein in the order X → Y, not Y ← X.

sequence can be determined and the genetic code used to infer the primary structure of the encoded polypeptide.

The hybrid approach enhances the speed and efficiency of primary structure analysis and the range of proteins that can be sequenced. It also circumvents obstacles such as the presence of an amino-terminal blocking group or the lack of a key overlap peptide. Only a few segments of primary structure must be determined by Edman analysis.

DNA sequencing reveals the order in which amino acids are added to the nascent polypeptide chain as it is synthesized on the ribosomes. However, it provides no information about posttranslational modifications such as proteolytic processing, methylation, glycosylation, phosphorylation, hydroxylation of proline and lysine, and disulfide bond formation that accompany maturation. While Edman sequencing can detect the presence of most posttranslational events, technical limitations often prevent identification of a specific modification.

Table 4–1. Methods for cleaving polypeptides.

Method	Bond Cleaved
CNBr	Met-X
Trypsin	Lys-X and Arg-X
Chymotrypsin	Hydrophobic amino acid-X
Endoproteinase Lys-C	Lys-X
Endoproteinase Arg-C	Arg-X
Endoproteinase Asp-N	X-Asp
V8 protease	Glu-X, particularly where X is hydrophobic
Hydroxylamine	Asn-Gly
o-Iodosobenzene	Trp-X
Mild acid	Asp-Pro

MASS SPECTROMETRY DETECTS COVALENT MODIFICATIONS

Mass spectrometry, which discriminates molecules based solely on their mass, is ideal for detecting the phosphate, hydroxyl, and other groups on posttranslationally modified amino acids. Each adds a specific and readily identified increment of mass to the modified amino acid (Table 4–2). For analysis by mass spectrometry, a sample in a vacuum is vaporized under conditions where protonation can occur, imparting positive charge. An electrical field then propels the cations through a magnetic field which deflects them at a right angle to their original direction of flight and focuses them onto a detector (Figure 4–8). The magnetic force required to deflect the path of each ionic species onto the detector, measured as the current applied to the electromagnet, is recorded. For ions of identical net charge, this force is proportionate to their mass. In a time-of-flight mass spectrometer, a briefly applied electric field accelerates the ions towards a detector that records the time at which each ion arrives. For molecules of identical charge, the velocity to which they are accelerated—and hence the time required to reach the detector—will be inversely proportionate to their mass.

Conventional mass spectrometers generally are used to determine the masses of molecules of 1000 Da or less, whereas time-of-flight mass spectrometers are suited for determining the large masses of proteins. The analysis of peptides and proteins by mass spectometry initially was hindered by difficulties in volatilizing large organic molecules. However, matrix-assisted laser-desorption (MALDI) and electrospray dispersion (eg, nanospray) permit the masses of even large polypeptides (> 100,000 Da) to be determined with extraordinary accuracy (± 1 Da). Using electrospray dispersion, peptides eluting from a reversed-phase HPLC column are introduced directly into the mass spectrometer for immediate determination of their masses.

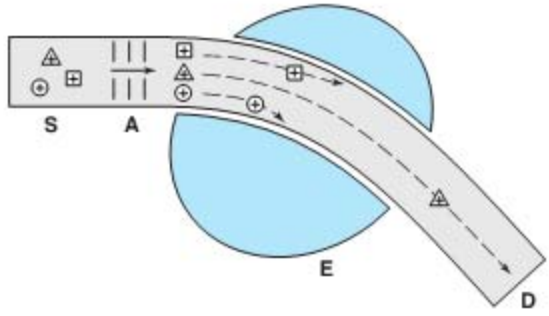


Figure 4–8. Basic components of a simple mass spectrometer. A mixture of molecules is vaporized in an ionized state in the sample chamber **S.** These molecules are then accelerated down the flight tube by an electrical potential applied to accelerator grid **A.** An adjustable electromagnet, **E,** applies a magnetic field that deflects the flight of the individual ions until they strike the detector, **D.** The greater the mass of the ion, the higher the magnetic field required to focus it onto the detector.

Peptides inside the mass spectrometer are broken down into smaller units by collisions with neutral helium atoms (collision-induced dissociation), and the masses of the individual fragments are determined. Since peptide bonds are much more labile than carbon-carbon bonds, the most abundant fragments will differ from one another by units equivalent to one or two amino acids. Since—with the exception of leucine and isoleucine—the molecular mass of each amino acid is unique, the sequence of the peptide can be reconstructed from the masses of its fragments.

Table 4–2. Mass increases resulting from common posttranslational modifications.

Modification	Mass Increase (Da)
Phosphorylation	80
Hydroxylation	16
Methylation	14
Acetylation	42
Myristylation	210
Palmitoylation	238
Glycosylation	162

Tandem Mass Spectrometry

Complex peptide mixtures can now be analyzed without prior purification by tandem mass spectrometry, which employs the equivalent of two mass spectrometers linked in series. The first spectrometer separates individual peptides based upon their differences in mass. By adjusting the field strength of the first magnet, a single peptide can be directed into the second mass spectrometer, where fragments are generated and their masses determined. As the sensitivity and versatility of mass spectrometry continue to increase, it is displacing Edman sequencers for the direct analysis of protein primary structure.

GENOMICS ENABLES PROTEINS TO BE IDENTIFIED FROM SMALL AMOUNTS OF SEQUENCE DATA

Primary structure analysis has been revolutionized by genomics, the application of automated oligonucleotide sequencing and computerized data retrieval and analysis to sequence an organism's entire genetic complement. The first genome sequenced was that of *Haemophilus influenzae,* in 1995. By mid 2001, the complete genome sequences for over 50 organisms had been determined. These include the human genome and those of several bacterial pathogens; the results and significance of the Human Genome Project are discussed in Chapter 54. Where genome sequence is known, the task of determining a protein's DNA-derived primary sequence is materially simplified. In essence, the second half of the hybrid approach has already been completed. All that remains is to acquire sufficient information to permit the open reading frame (ORF) that encodes the protein to be retrieved from an Internet-accessible genome database and identified. In some cases, a segment of amino acid sequence only four or five residues in length may be sufficient to identify the correct ORF.

Computerized search algorithms assist the identification of the gene encoding a given protein and clarify uncertainties that arise from Edman sequencing and mass spectrometry. By exploiting computers to solve complex puzzles, the spectrum of information suitable for identification of the ORF that encodes a particular polypeptide is greatly expanded. In peptide mass profiling, for example, a peptide digest is introduced into the mass spectrometer and the sizes of the peptides are determined. A computer is then used to find an ORF whose predicted protein product would, if broken down into peptides by the cleavage method selected, produce a set of peptides whose masses match those observed by mass spectrometry.

PROTEOMICS & THE PROTEOME

The Goal of Proteomics Is to Identify the Entire Complement of Proteins Elaborated by a Cell Under Diverse Conditions

While the sequence of the human genome is known, the picture provided by genomics alone is both static and incomplete. Proteomics aims to identify the entire complement of proteins elaborated by a cell under diverse conditions. As genes are switched on and off, proteins are synthesized in particular cell types at specific times of growth or differentiation and in response to external stimuli. Muscle cells express proteins not expressed by neural cells, and the type of subunits present in the hemoglobin tetramer undergo change pre- and postpartum. Many proteins undergo posttranslational modifications during maturation into functionally competent forms or as a means of regulating their properties. Knowledge of the human genome therefore represents only the beginning of the task of describing living organisms in molecular detail and understanding the dynamics of processes such as growth, aging, and disease. As the human body contains thousands of cell types, each containing thousands of proteins, the proteome—the set of all the proteins expressed by an individual cell at a particular time—represents a moving target of formidable dimensions.

Two-Dimensional Electrophoresis & Gene Array Chips Are Used to Survey Protein Expression

One goal of proteomics is the identification of proteins whose levels of expression correlate with medically significant events. The presumption is that proteins whose appearance or disappearance is associated with a specific physiologic condition or disease will provide insights into root causes and mechanisms. Determination of the proteomes characteristic of each cell type requires the utmost efficiency in the isolation and identification of individual proteins. The contemporary approach utilizes robotic automation to speed sample preparation and large two-dimensional gels to resolve cellular proteins. Individual polypeptides are then extracted and analyzed by Edman sequencing or mass spectroscopy. While only about 1000 proteins can be resolved on a single gel, two-dimensional electrophoresis has a major advantage in that it examines the proteins themselves. An alternative and complementary approach employs gene arrays, sometimes called DNA chips, to detect the expression of the mRNAs which encode proteins. While changes in the expression of the mRNA encoding a protein do not necessarily reflect comparable changes in the level of the corresponding protein, gene arrays are more sensitive probes than two-dimensional gels and thus can examine more gene products.

Bioinformatics Assists Identification of Protein Functions

The functions of a large proportion of the proteins encoded by the human genome are presently unknown. Recent advances in bioinformatics permit researchers to compare amino acid sequences to discover clues to potential properties, physiologic roles, and mechanisms of action of proteins. Algorithms exploit the tendency of nature to employ variations of a structural theme to perform similar functions in several proteins (eg, the Rossmann nucleotide binding fold to bind NAD(P)H,

nuclear targeting sequences, and EF hands to bind Ca^{2+}). These domains generally are detected in the primary structure by conservation of particular amino acids at key positions. Insights into the properties and physiologic role of a newly discovered protein thus may be inferred by comparing its primary structure with that of known proteins.

SUMMARY

- Long amino acid polymers or polypeptides constitute the basic structural unit of proteins, and the structure of a protein provides insight into how it fulfills its functions.
- The Edman reaction enabled amino acid sequence analysis to be automated. Mass spectrometry provides a sensitive and versatile tool for determining primary structure and for the identification of post-translational modifications.
- DNA cloning and molecular biology coupled with protein chemistry provide a hybrid approach that greatly increases the speed and efficiency for determination of primary structures of proteins.
- Genomics—the analysis of the entire oligonucleotide sequence of an organism's complete genetic material—has provided further enhancements.
- Computer algorithms facilitate identification of the open reading frames that encode a given protein by using partial sequences and peptide mass profiling to search sequence databases.
- Scientists are now trying to determine the primary sequence and functional role of every protein expressed in a living cell, known as its proteome.
- A major goal is the identification of proteins whose appearance or disappearance correlates with physiologic phenomena, aging, or specific diseases.

REFERENCES

Deutscher MP (editor): *Guide to Protein Purification.* Methods Enzymol 1990;182. (Entire volume.)

Geveart K, Vandekerckhove J: Protein identification methods in proteomics. Electrophoresis 2000;21:1145.

Helmuth L: Genome research: map of the human genome 3.0. Science 2001;293:583.

Khan J et al: DNA microarray technology: the anticipated impact on the study of human disease. Biochim Biophys Acta 1999;1423:M17.

McLafferty FW et al: Biomolecule mass spectrometry. Science 1999;284:1289.

Patnaik SK, Blumenfeld OO: Use of on-line tools and databases for routine sequence analyses. Anal Biochem 2001;289:1.

Schena M et al: Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 1995;270:467.

Semsarian C, Seidman CE: Molecular medicine in the 21st century. Intern Med J 2001;31:53.

Temple LK et al: Essays on science and society: defining disease in the genomics era. Science 2001;293:807.

Wilkins MR et al: High-throughput mass spectrometric discovery of protein post-translational modifications. J Mol Biol 1999;289:645.

Proteins: Higher Orders of Structure

5

Victor W. Rodwell, PhD, & Peter J. Kennelly, PhD

BIOMEDICAL IMPORTANCE

Proteins catalyze metabolic reactions, power cellular motion, and form macromolecular rods and cables that provide structural integrity to hair, bones, tendons, and teeth. In nature, form follows function. The structural variety of human proteins therefore reflects the sophistication and diversity of their biologic roles. Maturation of a newly synthesized polypeptide into a biologically functional protein requires that it be folded into a specific three-dimensional arrangement, or **conformation.** During maturation, **posttranslational modifications** may add new chemical groups or remove transiently needed peptide segments. Genetic or nutritional deficiencies that impede protein maturation are deleterious to health. Examples of the former include Creutzfeldt-Jakob disease, scrapie, Alzheimer's disease, and bovine spongiform encephalopathy (mad cow disease). Scurvy represents a nutritional deficiency that impairs protein maturation.

CONFORMATION VERSUS CONFIGURATION

The terms configuration and conformation are often confused. **Configuration** refers to the geometric relationship between a given set of atoms, for example, those that distinguish L- from D-amino acids. Interconversion of *configurational* alternatives requires breaking covalent bonds. **Conformation** refers to the spatial relationship of every atom in a molecule. Interconversion between conformers occurs without covalent bond rupture, with retention of configuration, and typically via rotation about single bonds.

PROTEINS WERE INITIALLY CLASSIFIED BY THEIR GROSS CHARACTERISTICS

Scientists initially approached structure-function relationships in proteins by separating them into classes based upon properties such as solubility, shape, or the presence of nonprotein groups. For example, the proteins that can be extracted from cells using solutions at physiologic pH and ionic strength are classified as **soluble.** Extraction of **integral membrane proteins** requires dissolution of the membrane with detergents. **Globular proteins** are compact, are roughly spherical or ovoid in shape, and have **axial ratios** (the ratio of their shortest to longest dimensions) of not over 3. Most enzymes are globular proteins, whose large internal volume provides ample space in which to construct cavities of the specific shape, charge, and hydrophobicity or hydrophilicity required to bind substrates and promote catalysis. By contrast, many structural proteins adopt highly extended conformations. These **fibrous proteins** possess axial ratios of 10 or more.

Lipoproteins and **glycoproteins** contain covalently bound lipid and carbohydrate, respectively. Myoglobin, hemoglobin, cytochromes, and many other proteins contain tightly associated metal ions and are termed **metalloproteins.** With the development and application of techniques for determining the amino acid sequences of proteins (Chapter 4), more precise classification schemes have emerged based upon similarity, or **homology,** in amino acid sequence and structure. However, many early classification terms remain in common use.

PROTEINS ARE CONSTRUCTED USING MODULAR PRINCIPLES

Proteins perform complex physical and catalytic functions by positioning specific chemical groups in a precise three-dimensional arrangement. The polypeptide scaffold containing these groups must adopt a conformation that is both functionally efficient and physically strong. At first glance, the biosynthesis of polypeptides comprised of tens of thousands of individual atoms would appear to be extremely challenging. When one considers that a typical polypeptide can adopt $\geq 10^{50}$ distinct conformations, folding into the conformation appropriate to their biologic function would appear to be even more difficult. As described in Chapters 3 and 4, synthesis of the polypeptide backbones of proteins employs a small set of common building blocks or modules, the amino acids, joined by a common linkage, the peptide bond. A stepwise modular pathway simplifies the folding and processing of newly synthesized polypeptides into mature proteins.

THE FOUR ORDERS OF PROTEIN STRUCTURE

The modular nature of protein synthesis and folding are embodied in the concept of orders of protein structure: **primary structure,** the sequence of the amino acids in a polypeptide chain; **secondary structure,** the folding of short (3- to 30-residue), contiguous segments of polypeptide into geometrically ordered units; **tertiary structure,** the three-dimensional assembly of secondary structural units to form larger functional units such as the mature polypeptide and its component domains; and **quaternary structure,** the number and types of polypeptide units of oligomeric proteins and their spatial arrangement.

SECONDARY STRUCTURE

Peptide Bonds Restrict Possible Secondary Conformations

Free rotation is possible about only two of the three covalent bonds of the polypeptide backbone: the α-carbon (Cα) to the carbonyl carbon (Co) bond and the Cα to nitrogen bond (Figure 3–4). The partial double-bond character of the peptide bond that links Co to the α-nitrogen requires that the carbonyl carbon, carbonyl oxygen, and α-nitrogen remain coplanar, thus preventing rotation. The angle about the Cα—N bond is termed the phi (Φ) angle, and that about the Co—Cα bond the psi (Ψ) angle. For amino acids other than glycine, most combinations of phi and psi angles are disallowed because of steric hindrance (Figure 5–1). The conformations of proline are even more restricted due to the absence of free rotation of the N—Cα bond.

Regions of ordered secondary structure arise when a series of aminoacyl residues adopt similar phi and psi angles. Extended segments of polypeptide (eg, loops) can possess a variety of such angles. The angles that define the two most common types of secondary structure, the **α helix** and the **β sheet,** fall within the lower and upper left-hand quadrants of a Ramachandran plot, respectively (Figure 5–1).

The Alpha Helix

The polypeptide backbone of an α helix is twisted by an equal amount about each α-carbon with a phi angle of approximately −57 degrees and a psi angle of approximately −47 degrees. A complete turn of the helix contains an average of 3.6 aminoacyl residues, and the distance it rises per turn (its pitch) is 0.54 nm (Figure 5–2). The R groups of each aminoacyl residue in an α helix face outward (Figure 5–3). Proteins contain only L-amino acids, for which a right-handed α helix is by far the more stable, and only right-handed α helices occur in nature. Schematic diagrams of proteins represent α helices as cylinders.

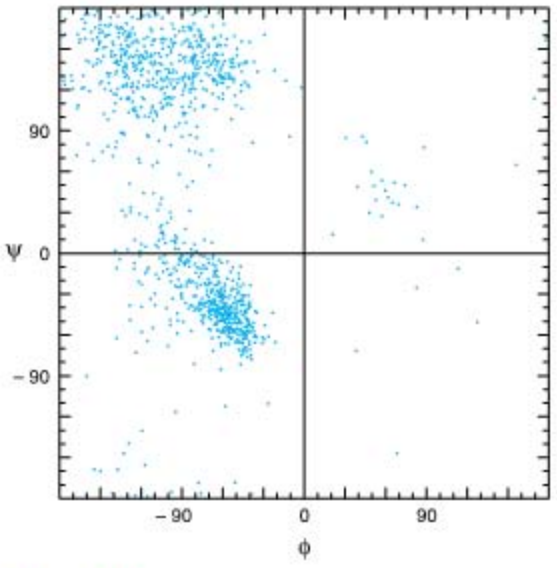


Figure 5–1. Ramachandran plot of the main chain phi (Φ) and psi (Ψ) angles for approximately 1000 nonglycine residues in eight proteins whose structures were solved at high resolution. The dots represent allowable combinations and the spaces prohibited combinations of phi and psi angles. (Reproduced, with permission, from Richardson JS: The anatomy and taxonomy of protein structures. Adv Protein Chem 1981;34:167.)

The stability of an α helix arises primarily from hydrogen bonds formed between the oxygen of the peptide bond carbonyl and the hydrogen atom of the peptide bond nitrogen of the fourth residue down the polypeptide chain (Figure 5–4). The ability to form the maximum number of hydrogen bonds, supplemented by van der Waals interactions in the core of this tightly packed structure, provides the thermodynamic driving force for the formation of an α helix. Since the peptide bond nitrogen of proline lacks a hydrogen atom to contribute to a hydrogen bond, proline can only be stably accommodated within the first turn of an α helix. When present elsewhere, proline disrupts the conformation of the helix, producing a bend. Because of its small size, glycine also often induces bends in α helices.

Many α helices have predominantly hydrophobic R groups on one side of the axis of the helix and predominantly hydrophilic ones on the other. These **amphipathic helices** are well adapted to the formation of interfaces between polar and nonpolar regions such as the hydrophobic interior of a protein and its aqueous envi-

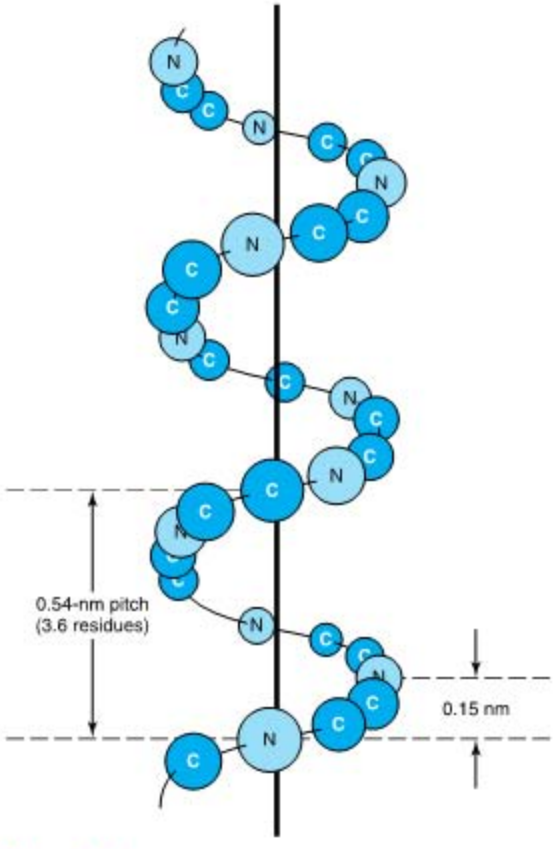


Figure 5–2. Orientation of the main chain atoms of a peptide about the axis of an α helix.

ronment. Clusters of amphipathic helices can create a channel, or pore, that permits specific polar molecules to pass through hydrophobic cell membranes.

The Beta Sheet

The second (hence "beta") recognizable regular secondary structure in proteins is the β sheet. The amino acid residues of a β sheet, when viewed edge-on, form a zigzag or pleated pattern in which the R groups of adjacent residues point in opposite directions. Unlike the compact backbone of the α helix, the peptide backbone of the β sheet is highly extended. But like the α helix, β sheets derive much of their stability from hydrogen bonds between the carbonyl oxygens and amide hydrogens of peptide bonds. However, in contrast to the α helix, these bonds are formed with adjacent segments of β sheet (Figure 5–5).

Interacting β sheets can be arranged either to form a **parallel** β sheet, in which the adjacent segments of the

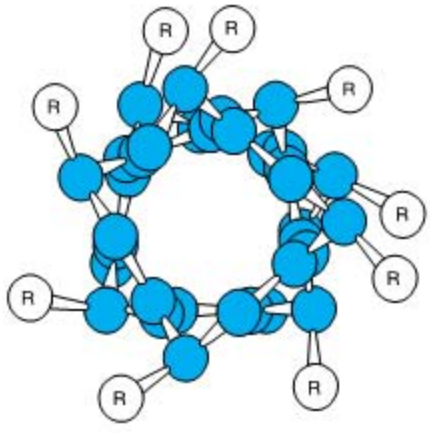


Figure 5–3. View down the axis of an α helix. The side chains (R) are on the outside of the helix. The van der Waals radii of the atoms are larger than shown here; hence, there is almost no free space inside the helix. (Slightly modified and reproduced, with permission, from Stryer L: *Biochemistry*, 3rd ed. Freeman, 1995. Copyright © 1995 by W.H. Freeman and Co.)

polypeptide chain proceed in the same direction amino to carboxyl, or an **antiparallel** sheet, in which they proceed in opposite directions (Figure 5–5). Either configuration permits the maximum number of hydrogen bonds between segments, or strands, of the sheet. Most β sheets are not perfectly flat but tend to have a right-handed twist. Clusters of twisted strands of β sheet form the core of many globular proteins (Figure 5–6). Schematic diagrams represent β sheets as arrows that point in the amino to carboxyl terminal direction.

Loops & Bends

Roughly half of the residues in a "typical" globular protein reside in α helices and β sheets and half in loops, turns, bends, and other extended conformational features. Turns and bends refer to short segments of amino acids that join two units of secondary structure, such as two adjacent strands of an antiparallel β sheet. A β turn involves four aminoacyl residues, in which the first residue is hydrogen-bonded to the fourth, resulting in a tight 180-degree turn (Figure 5–7). Proline and glycine often are present in β turns.

Loops are regions that contain residues beyond the minimum number necessary to connect adjacent regions of secondary structure. Irregular in conformation, loops nevertheless serve key biologic roles. For many enzymes, the loops that bridge domains responsible for binding substrates often contain aminoacyl residues

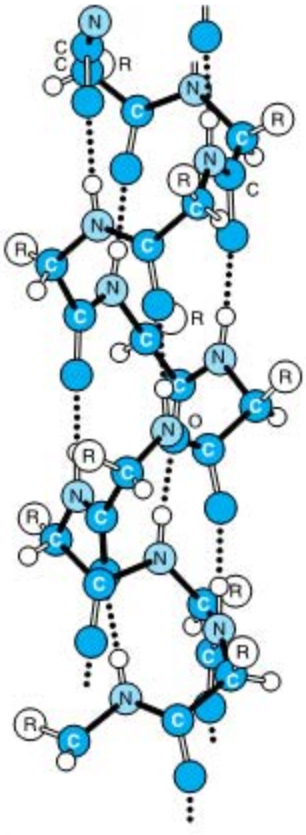


Figure 5–4. Hydrogen bonds (dotted lines) formed between H and O atoms stabilize a polypeptide in an α-helical conformation. (Reprinted, with permission, from Haggis GH et al: *Introduction to Molecular Biology.* Wiley, 1964.)

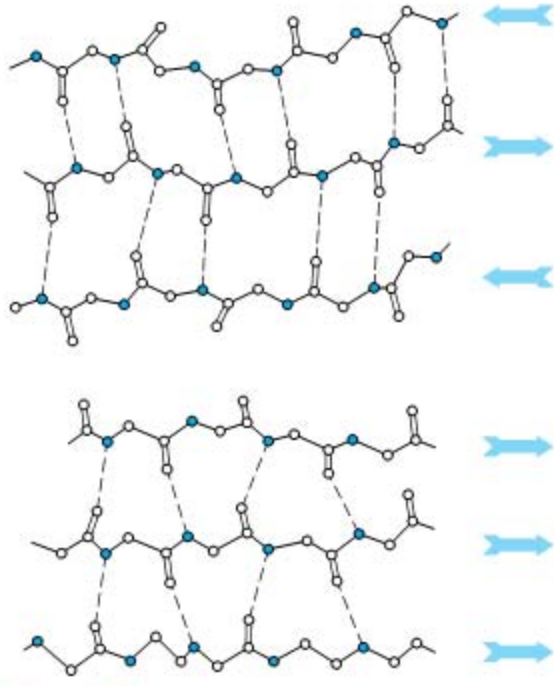


Figure 5–5. Spacing and bond angles of the hydrogen bonds of antiparallel and parallel pleated β sheets. Arrows indicate the direction of each strand. The hydrogen-donating α-nitrogen atoms are shown as blue circles. Hydrogen bonds are indicated by dotted lines. For clarity in presentation, R groups and hydrogens are omitted. **Top:** Antiparallel β sheet. Pairs of hydrogen bonds alternate between being close together and wide apart and are oriented approximately perpendicular to the polypeptide backbone. **Bottom:** Parallel β sheet. The hydrogen bonds are evenly spaced but slant in alternate directions.

that participate in catalysis. **Helix-loop-helix motifs** provide the oligonucleotide-binding portion of DNA-binding proteins such as repressors and transcription factors. Structural motifs such as the helix-loop-helix motif that are intermediate between secondary and tertiary structures are often termed **supersecondary structures.** Since many loops and bends reside on the surface of proteins and are thus exposed to solvent, they constitute readily accessible sites, or **epitopes,** for recognition and binding of antibodies.

While loops lack apparent structural regularity, they exist in a specific conformation stabilized through hydrogen bonding, salt bridges, and hydrophobic interactions with other portions of the protein. However, not all portions of proteins are necessarily ordered. Proteins may contain "disordered" regions, often at the extreme amino or carboxyl terminal, characterized by high conformational flexibility. In many instances, these disordered regions assume an ordered conformation upon binding of a ligand. This structural flexibility enables such regions to act as ligand-controlled switches that affect protein structure and function.

Tertiary & Quaternary Structure

The term "tertiary structure" refers to the entire three-dimensional conformation of a polypeptide. It indicates, in three-dimensional space, how secondary structural features—helices, sheets, bends, turns, and loops—assemble to form **domains** and how these domains relate spatially to one another. A domain is a section of protein structure sufficient to perform a particular chemical or physical task such as binding of a substrate

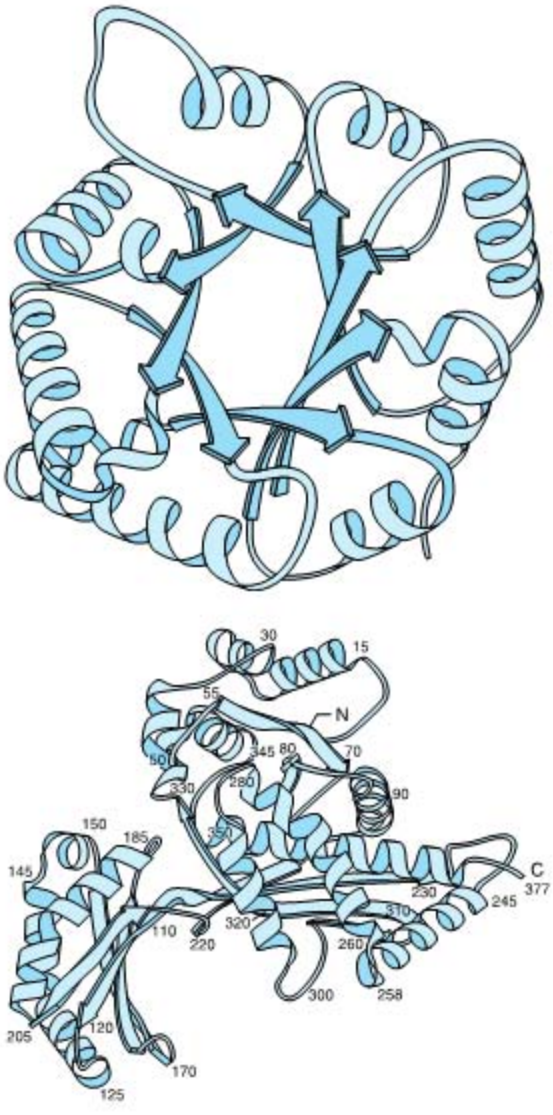


Figure 5–6. Examples of tertiary structure of proteins. **Top:** The enzyme triose phosphate isomerase. Note the elegant and symmetrical arrangement of alternating β sheets and α helices. (Courtesy of J Richardson.) **Bottom:** Two-domain structure of the subunit of a homodimeric enzyme, a bacterial class II HMG-CoA reductase. As indicated by the numbered residues, the single polypeptide begins in the large domain, enters the small domain, and ends in the large domain. (Courtesy of C Lawrence, V Rodwell, and C Stauffacher, Purdue University.)

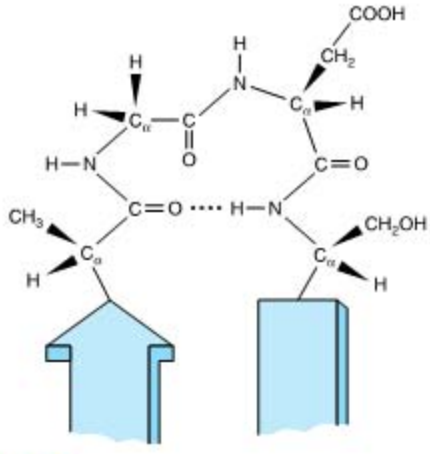


Figure 5–7. A β-turn that links two segments of antiparallel β sheet. The dotted line indicates the hydrogen bond between the first and fourth amino acids of the four-residue segment Ala-Gly-Asp-Ser.

or other ligand. Other domains may anchor a protein to a membrane or interact with a regulatory molecule that modulates its function. A small polypeptide such as triose phosphate isomerase (Figure 5–6) or myoglobin (Chapter 6) may consist of a single domain. By contrast, protein kinases contain two domains. Protein kinases catalyze the transfer of a phosphoryl group from ATP to a peptide or protein. The amino terminal portion of the polypeptide, which is rich in β sheet, binds ATP, while the carboxyl terminal domain, which is rich in α helix, binds the peptide or protein substrate (Figure 5–8). The groups that catalyze phosphoryl transfer reside in a loop positioned at the interface of the two domains.

In some cases, proteins are assembled from more than one polypeptide, or protomer. Quaternary structure defines the polypeptide composition of a protein and, for an oligomeric protein, the spatial relationships between its subunits or protomers. **Monomeric** proteins consist of a single polypeptide chain. **Dimeric** proteins contain two polypeptide chains. Homodimers contain two copies of the same polypeptide chain, while in a heterodimer the polypeptides differ. Greek letters (α, β, γ etc) are used to distinguish different subunits of a heterooligomeric protein, and subscripts indicate the number of each subunit type. For example, α_4 designates a homotetrameric protein, and $\alpha_2\beta_2\gamma$ a protein with five subunits of three different types.

Since even small proteins contain many thousands of atoms, depictions of protein structure that indicate the position of every atom are generally too complex to be readily interpreted. Simplified schematic diagrams thus are used to depict key features of a protein's ter-

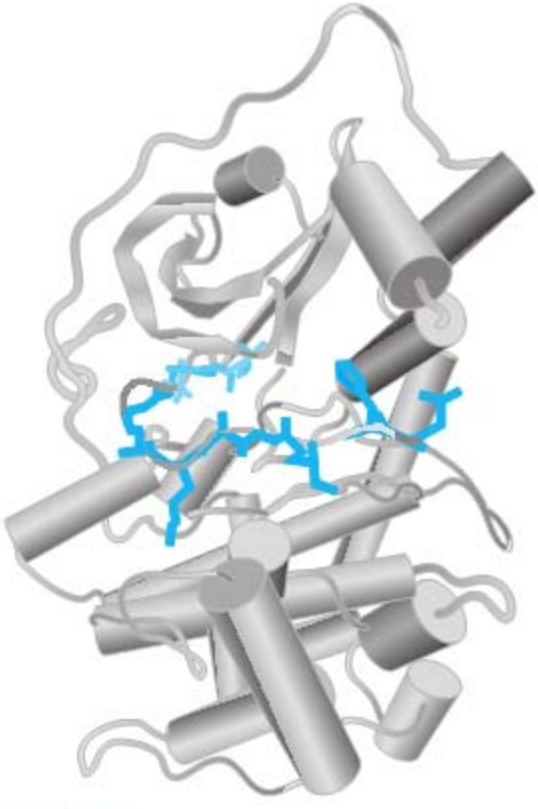


Figure 5–8. Domain structure. Protein kinases contain two domains. The upper, amino terminal domain binds the phosphoryl donor ATP (light blue). The lower, carboxyl terminal domain is shown binding a synthetic peptide substrate (dark blue).

tiary and quaternary structure. Ribbon diagrams (Figures 5–6 and 5–8) trace the conformation of the polypeptide backbone, with cylinders and arrows indicating regions of α helix and β sheet, respectively. In an even simpler representation, line segments that link the α carbons indicate the path of the polypeptide backbone. These schematic diagrams often include the side chains of selected amino acids that emphasize specific structure-function relationships.

MULTIPLE FACTORS STABILIZE TERTIARY & QUATERNARY STRUCTURE

Higher orders of protein structure are stabilized primarily—and often exclusively—by noncovalent interactions. Principal among these are hydrophobic interactions that drive most hydrophobic amino acid side chains into the interior of the protein, shielding them from water. Other significant contributors include hydrogen bonds and salt bridges between the carboxylates of aspartic and glutamic acid and the oppositely charged side chains of protonated lysyl, argininyl, and histidyl residues. While individually weak relative to a typical covalent bond of 80–120 kcal/mol, collectively these numerous interactions confer a high degree of stability to the biologically functional conformation of a protein, just as a Velcro fastener harnesses the cumulative strength of multiple plastic loops and hooks.

Some proteins contain covalent disulfide (S—S) bonds that link the sulfhydryl groups of cysteinyl residues. Formation of disulfide bonds involves oxidation of the cysteinyl sulfhydryl groups and requires oxygen. Intrapolypeptide disulfide bonds further enhance the stability of the folded conformation of a peptide, while interpolypeptide disulfide bonds stabilize the quaternary structure of certain oligomeric proteins.

THREE-DIMENSIONAL STRUCTURE IS DETERMINED BY X-RAY CRYSTALLOGRAPHY OR BY NMR SPECTROSCOPY

X-Ray Crystallography

Since the determination of the three-dimensional structure of myoglobin over 40 years ago, the three-dimensional structures of thousands of proteins have been determined by x-ray crystallography. The key to x-ray crystallography is the precipitation of a protein under conditions in which it forms ordered crystals that diffract x-rays. This is generally accomplished by exposing small drops of the protein solution to various combinations of pH and precipitating agents such as salts and organic solutes such as polyethylene glycol. A detailed three-dimensional structure of a protein can be constructed from its primary structure using the pattern by which it diffracts a beam of monochromatic x-rays. While the development of increasingly capable computer-based tools has rendered the analysis of complex x-ray diffraction patterns increasingly facile, a major stumbling block remains the requirement of inducing highly purified samples of the protein of interest to crystallize. Several lines of evidence, including the ability of some crystallized enzymes to catalyze chemical reactions, indicate that the vast majority of the structures determined by crystallography faithfully represent the structures of proteins in free solution.

Nuclear Magnetic Resonance Spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy, a powerful complement to x-ray crystallography, mea-

sures the absorbance of radio frequency electromagnetic energy by certain atomic nuclei. "NMR-active" isotopes of biologically relevant atoms include ^{1}H, ^{13}C, ^{15}N, and ^{31}P. The frequency, or chemical shift, at which a particular nucleus absorbs energy is a function of both the functional group within which it resides and the proximity of other NMR-active nuclei. Two-dimensional NMR spectroscopy permits a three-dimensional representation of a protein to be constructed by determining the proximity of these nuclei to one another. NMR spectroscopy analyzes proteins in aqueous solution, obviating the need to form crystals. It thus is possible to observe changes in conformation that accompany ligand binding or catalysis using NMR spectroscopy. However, only the spectra of relatively small proteins, ≤ 20 kDa in size, can be analyzed with current technology.

Molecular Modeling

An increasingly useful adjunct to the empirical determination of the three-dimensional structure of proteins is the use of computer technology for molecular modeling. The types of models created take two forms. In the first, the known three-dimensional structure of a protein is used as a template to build a model of the probable structure of a homologous protein. In the second, computer software is used to manipulate the static model provided by crystallography to explore how a protein's conformation might change when ligands are bound or when temperature, pH, or ionic strength is altered. Scientists also are examining the library of available protein structures in an attempt to devise computer programs that can predict the three-dimensional conformation of a protein directly from its primary sequence.

PROTEIN FOLDING

The Native Conformation of a Protein Is Thermodynamically Favored

The number of distinct combinations of phi and psi angles specifying potential conformations of even a relatively small—15-kDa—polypeptide is unbelievably vast. Proteins are guided through this vast labyrinth of possibilities by thermodynamics. Since the biologically relevant—or native—conformation of a protein generally is that which is most energetically favored, knowledge of the native conformation is specified in the primary sequence. However, if one were to wait for a polypeptide to find its native conformation by random exploration of all possible conformations, the process would require billions of years to complete. Clearly, protein folding in cells takes place in a more orderly and guided fashion.

Folding Is Modular

Protein folding generally occurs via a stepwise process. In the first stage, the newly synthesized polypeptide emerges from ribosomes, and short segments fold into secondary structural units that provide local regions of organized structure. Folding is now reduced to the selection of an appropriate arrangement of this relatively small number of secondary structural elements. In the second stage, the forces that drive hydrophobic regions into the interior of the protein away from solvent drive the partially folded polypeptide into a "molten globule" in which the modules of secondary structure rearrange to arrive at the mature conformation of the protein. This process is orderly but not rigid. Considerable flexibility exists in the ways and in the order in which elements of secondary structure can be rearranged. In general, each element of secondary or supersecondary structure facilitates proper folding by directing the folding process toward the native conformation and away from unproductive alternatives. For oligomeric proteins, individual protomers tend to fold before they associate with other subunits.

Auxiliary Proteins Assist Folding

Under appropriate conditions, many proteins will spontaneously refold after being previously **denatured** (ie, unfolded) by treatment with acid or base, chaotropic agents, or detergents. However, unlike the folding process in vivo, refolding under laboratory conditions is a far slower process. Moreover, some proteins fail to spontaneously refold in vitro, often forming insoluble **aggregates,** disordered complexes of unfolded or partially folded polypeptides held together by hydrophobic interactions. Aggregates represent unproductive dead ends in the folding process. Cells employ auxiliary proteins to speed the process of folding and to guide it toward a productive conclusion.

Chaperones

Chaperone proteins participate in the folding of over half of mammalian proteins. The hsp70 (70-kDa heat shock protein) family of chaperones binds short sequences of hydrophobic amino acids in newly synthesized polypeptides, shielding them from solvent. Chaperones prevent aggregation, thus providing an opportunity for the formation of appropriate secondary structural elements and their subsequent coalescence into a molten globule. The hsp60 family of chaperones, sometimes called **chaperonins,** differ in sequence and structure from hsp70 and its homologs. Hsp60 acts later in the folding process, often together with an hsp70 chaperone. The central cavity of the donut-

shaped hsp60 chaperone provides a sheltered environment in which a polypeptide can fold until all hydrophobic regions are buried in its interior, eliminating aggregation. Chaperone proteins can also "rescue" proteins that have become thermodynamically trapped in a misfolded dead end by unfolding hydrophobic regions and providing a second chance to fold productively.

Protein Disulfide Isomerase

Disulfide bonds between and within polypeptides stabilize tertiary and quaternary structure. However, disulfide bond formation is nonspecific. Under oxidizing conditions, a given cysteine can form a disulfide bond with the —SH of any accessible cysteinyl residue. By catalyzing disulfide exchange, the rupture of an S—S bond and its reformation with a different partner cysteine, protein disulfide isomerase facilitates the formation of disulfide bonds that stabilize their native conformation.

Proline-*cis,trans*-Isomerase

All X-Pro peptide bonds—where X represents any residue—are synthesized in the trans configuration. However, of the X-Pro bonds of mature proteins, approximately 6% are cis. The cis configuration is particularly common in β-turns. Isomerization from trans to cis is catalyzed by the enzyme proline-*cis,trans*-isomerase (Figure 5–9).

SEVERAL NEUROLOGIC DISEASES RESULT FROM ALTERED PROTEIN CONFORMATION

Prions

The transmissible spongiform encephalopathies, or **prion diseases,** are fatal neurodegenerative diseases characterized by spongiform changes, astrocytic gliomas, and neuronal loss resulting from the deposition of insoluble protein aggregates in neural cells. They include Creutzfeldt-Jakob disease in humans, scrapie in sheep, and bovine spongiform encephalopathy (mad cow disease) in cattle. Prion diseases may manifest themselves as infectious, genetic, or sporadic disorders. Because no viral or bacterial gene encoding the pathologic prion protein could be identified, the source and mechanism of transmission of prion disease long remained elusive. Today it is believed that prion diseases are protein conformation diseases transmitted by altering the conformation, and hence the physical properties, of proteins endogenous to the host. Human prion-related protein, PrP, a glycoprotein encoded on the short arm of chromosome 20, normally is monomeric and rich in α helix. Pathologic prion proteins serve as the templates for the conformational transformation of normal PrP, known as PrPc, into PrPsc. PrPsc is rich in β sheet with many hydrophobic aminoacyl side chains exposed to solvent. PrPsc molecules therefore associate strongly with one other, forming insoluble protease-resistant aggregates. Since one pathologic prion or prion-related protein can serve as template for the conformational transformation of many times its number of PrPc molecules, prion diseases can be transmitted by the protein alone without involvement of DNA or RNA.

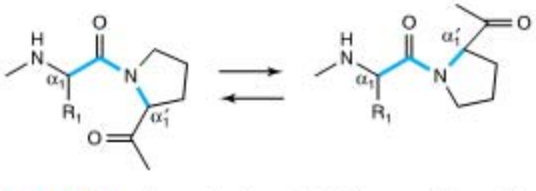


Figure 5–9. Isomerization of the N-α_1 prolyl peptide bond from a cis to a trans configuration relative to the backbone of the polypeptide.

Alzheimer's Disease

Refolding or misfolding of another protein endogenous to human brain tissue, β-amyloid, is also a prominent feature of Alzheimer's disease. While the root cause of Alzheimer's disease remains elusive, the characteristic senile plaques and neurofibrillary bundles contain aggregates of the protein β-amyloid, a 4.3-kDa polypeptide produced by proteolytic cleavage of a larger protein known as amyloid precursor protein. In Alzheimer's disease patients, levels of β-amyloid become elevated, and this protein undergoes a conformational transformation from a soluble α helix–rich state to a state rich in β sheet and prone to self-aggregation. Apolipoprotein E has been implicated as a potential mediator of this conformational transformation.

COLLAGEN ILLUSTRATES THE ROLE OF POSTTRANSLATIONAL PROCESSING IN PROTEIN MATURATION

Protein Maturation Often Involves Making & Breaking Covalent Bonds

The maturation of proteins into their final structural state often involves the cleavage or formation (or both) of covalent bonds, a process termed **posttranslational modification.** Many polypeptides are initially synthesized as larger precursors, called **proproteins.** The "extra" polypeptide segments in these proproteins often serve as leader sequences that target a polypeptide

to a particular organelle or facilitate its passage through a membrane. Others ensure that the potentially harmful activity of a protein such as the proteases trypsin and chymotrypsin remains inhibited until these proteins reach their final destination. However, once these transient requirements are fulfilled, the now superfluous peptide regions are removed by selective proteolysis. Other covalent modifications may take place that add new chemical functionalities to a protein. The maturation of collagen illustrates both of these processes.

Collagen Is a Fibrous Protein

Collagen is the most abundant of the fibrous proteins that constitute more than 25% of the protein mass in the human body. Other prominent fibrous proteins include keratin and myosin. These proteins represent a primary source of structural strength for cells (ie, the cytoskeleton) and tissues. Skin derives its strength and flexibility from a crisscrossed mesh of collagen and keratin fibers, while bones and teeth are buttressed by an underlying network of collagen fibers analogous to the steel strands in reinforced concrete. Collagen also is present in connective tissues such as ligaments and tendons. The high degree of tensile strength required to fulfill these structural roles requires elongated proteins characterized by repetitive amino acid sequences and a regular secondary structure.

Collagen Forms a Unique Triple Helix

Tropocollagen consists of three fibers, each containing about 1000 amino acids, bundled together in a unique conformation, the collagen triple helix (Figure 5–10). A mature collagen fiber forms an elongated rod with an axial ratio of about 200. Three intertwined polypeptide strands, which twist to the left, wrap around one another in a right-handed fashion to form the collagen triple helix. The opposing handedness of this superhelix and its component polypeptides makes the collagen triple helix highly resistant to unwinding—the same principle used in the steel cables of suspension bridges. A collagen triple helix has 3.3 residues per turn and a rise per residue nearly twice that of an α helix. The R groups of each polypeptide strand of the triple helix pack so closely that in order to fit, one must be glycine. Thus, every third amino acid residue in collagen is a glycine residue. Staggering of the three strands provides appropriate positioning of the requisite glycines throughout the helix. Collagen is also rich in proline and hydroxyproline, yielding a repetitive Gly-X-Y pattern (Figure 5–10) in which Y generally is proline or hydroxyproline.

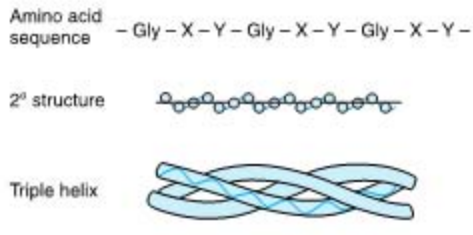


Figure 5–10. Primary, secondary, and tertiary structures of collagen.

Collagen triple helices are stabilized by hydrogen bonds between residues in *different* polypeptide chains. The hydroxyl groups of hydroxyprolyl residues also participate in interchain hydrogen bonding. Additional stability is provided by covalent cross-links formed between modified lysyl residues both within and between polypeptide chains.

Collagen Is Synthesized as a Larger Precursor

Collagen is initially synthesized as a larger precursor polypeptide, procollagen. Numerous prolyl and lysyl residues of procollagen are hydroxylated by prolyl hydroxylase and lysyl hydroxylase, enzymes that require ascorbic acid (vitamin C). Hydroxyprolyl and hydroxylysyl residues provide additional hydrogen bonding capability that stabilizes the mature protein. In addition, glucosyl and galactosyl transferases attach glucosyl or galactosyl residues to the hydroxyl groups of specific hydroxylysyl residues.

The central portion of the precursor polypeptide then associates with other molecules to form the characteristic triple helix. This process is accompanied by the removal of the globular amino terminal and carboxyl terminal extensions of the precursor polypeptide by selective proteolysis. Certain lysyl residues are modified by lysyl oxidase, a copper-containing protein that converts ε-amino groups to aldehydes. The aldehydes can either undergo an aldol condensation to form a C═C double bond or to form a Schiff base (eneimine) with the ε-amino group of an unmodified lysyl residue, which is subsequently reduced to form a C—N single bond. These covalent bonds cross-link the individual polypeptides and imbue the fiber with exceptional strength and rigidity.

Nutritional & Genetic Disorders Can Impair Collagen Maturation

The complex series of events in collagen maturation provide a model that illustrates the biologic consequences of incomplete polypeptide maturation. The best-known defect in collagen biosynthesis is scurvy, a result of a dietary deficiency of vitamin C required by

prolyl and lysyl hydroxylases. The resulting deficit in the number of hydroxyproline and hydroxylysine residues undermines the conformational stability of collagen fibers, leading to bleeding gums, swelling joints, poor wound healing, and ultimately to death. Menkes' syndrome, characterized by kinky hair and growth retardation, reflects a dietary deficiency of the copper required by lysyl oxidase, which catalyzes a key step in formation of the covalent cross-links that strengthen collagen fibers.

Genetic disorders of collagen biosynthesis include several forms of osteogenesis imperfecta, characterized by fragile bones. In Ehlers-Dahlos syndrome, a group of connective tissue disorders that involve impaired integrity of supporting structures, defects in the genes that encode α collagen-1, procollagen *N*-peptidase, or lysyl hydroxylase result in mobile joints and skin abnormalities.

SUMMARY

- Proteins may be classified on the basis of the solubility, shape, or function or of the presence of a prosthetic group such as heme. Proteins perform complex physical and catalytic functions by positioning specific chemical groups in a precise three-dimensional arrangement that is both functionally efficient and physically strong.
- The gene-encoded primary structure of a polypeptide is the sequence of its amino acids. Its secondary structure results from folding of polypeptides into hydrogen-bonded motifs such as the α helix, the β-pleated sheet, β bends, and loops. Combinations of these motifs can form supersecondary motifs.
- Tertiary structure concerns the relationships between secondary structural domains. Quaternary structure of proteins with two or more polypeptides (oligomeric proteins) is a feature based on the spatial relationships between various types of polypeptides.
- Primary structures are stabilized by covalent peptide bonds. Higher orders of structure are stabilized by weak forces—multiple hydrogen bonds, salt (electrostatic) bonds, and association of hydrophobic R groups.
- The phi (Φ) angle of a polypeptide is the angle about the C_α—N bond; the psi (Ψ) angle is that about the C_α-C_o bond. Most combinations of phi-psi angles are disallowed due to steric hindrance. The phi-psi angles that form the α helix and the β sheet fall within the lower and upper left-hand quadrants of a Ramachandran plot, respectively.
- Protein folding is a poorly understood process. Broadly speaking, short segments of newly synthesized polypeptide fold into secondary structural units. Forces that bury hydrophobic regions from solvent then drive the partially folded polypeptide into a "molten globule" in which the modules of secondary structure are rearranged to give the native conformation of the protein.
- Proteins that assist folding include protein disulfide isomerase, proline-*cis,trans*,-isomerase, and the chaperones that participate in the folding of over half of mammalian proteins. Chaperones shield newly synthesized polypeptides from solvent and provide an environment for elements of secondary structure to emerge and coalesce into molten globules.
- Techniques for study of higher orders of protein structure include x-ray crystallography, NMR spectroscopy, analytical ultracentrifugation, gel filtration, and gel electrophoresis.
- Silk fibroin and collagen illustrate the close linkage of protein structure and biologic function. Diseases of collagen maturation include Ehlers-Danlos syndrome and the vitamin C deficiency disease scurvy.
- Prions—protein particles that lack nucleic acid—cause fatal transmissible spongiform encephalopathies such as Creutzfeldt-Jakob disease, scrapie, and bovine spongiform encephalopathy. Prion diseases involve an altered secondary-tertiary structure of a naturally occurring protein, PrPc. When PrPc interacts with its pathologic isoform PrPSc, its conformation is transformed from a predominantly α-helical structure to the β-sheet structure characteristic of PrPSc.

REFERENCES

Branden C, Tooze J: *Introduction to Protein Structure.* Garland, 1991.

Burkhard P, Stetefeld J, Strelkov SV: Coiled coils: A highly versatile protein folding motif. Trends Cell Biol 2001;11:82.

Collinge J: Prion diseases of humans and animals: Their causes and molecular basis. Annu Rev Neurosci 2001;24:519.

Frydman J: Folding of newly translated proteins in vivo: The role of molecular chaperones. Annu Rev Biochem 2001;70:603.

Radord S: Protein folding: Progress made and promises ahead. Trends Biochem Sci 2000;25:611.

Schmid FX: Proly isomerase: Enzymatic catalysis of slow protein folding reactions. Ann Rev Biophys Biomol Struct 1993;22: 123.

Segrest MP et al: The amphipathic alpha-helix: A multifunctional structural motif in plasma lipoproteins. Adv Protein Chem 1995;45:1.

Soto C: Alzheimer's and prion disease as disorders of protein conformation: Implications for the design of novel therapeutic approaches. J Mol Med 1999;77:412.

Proteins: Myoglobin & Hemoglobin

6

Victor W. Rodwell, PhD, & Peter J. Kennelly, PhD

BIOMEDICAL IMPORTANCE

The heme proteins myoglobin and hemoglobin maintain a supply of oxygen essential for oxidative metabolism. Myoglobin, a monomeric protein of red muscle, stores oxygen as a reserve against oxygen deprivation. Hemoglobin, a tetrameric protein of erythrocytes, transports O_2 to the tissues and returns CO_2 and protons to the lungs. Cyanide and carbon monoxide kill because they disrupt the physiologic function of the heme proteins cytochrome oxidase and hemoglobin, respectively. The secondary-tertiary structure of the subunits of hemoglobin resembles myoglobin. However, the tetrameric structure of hemoglobin permits cooperative interactions that are central to its function. For example, 2,3-bisphosphoglycerate (BPG) promotes the efficient release of O_2 by stabilizing the quaternary structure of deoxyhemoglobin. Hemoglobin and myoglobin illustrate both protein structure-function relationships and the molecular basis of genetic diseases such as sickle cell disease and the thalassemias.

HEME & FERROUS IRON CONFER THE ABILITY TO STORE & TO TRANSPORT OXYGEN

Myoglobin and hemoglobin contain **heme,** a cyclic tetrapyrrole consisting of four molecules of pyrrole linked by α-methylene bridges. This planar network of conjugated double bonds absorbs visible light and colors heme deep red. The substituents at the β-positions of heme are methyl (M), vinyl (V), and propionate (Pr) groups arranged in the order M, V, M, V, M, Pr, Pr, M (Figure 6–1). One atom of ferrous iron (Fe_2^+) resides at the center of the planar tetrapyrrole. Other proteins with metal-containing tetrapyrrole prosthetic groups include the cytochromes (Fe and Cu) and chlorophyll (Mg) (see Chapter 12). Oxidation and reduction of the Fe and Cu atoms of cytochromes is essential to their biologic function as carriers of electrons. By contrast, oxidation of the Fe_2^+ of myoglobin or hemoglobin to Fe_3^+ destroys their biologic activity.

Myoglobin Is Rich in α Helix

Oxygen stored in red muscle myoglobin is released during O_2 deprivation (eg, severe exercise) for use in muscle mitochondria for aerobic synthesis of ATP (see Chapter 12). A 153-aminoacyl residue polypeptide (MW 17,000), myoglobin folds into a compact shape that measures 4.5 × 3.5 × 2.5 nm (Figure 6–2). Unusually high proportions, about 75%, of the residues are present in eight right-handed, 7–20 residue α helices. Starting at the amino terminal, these are termed helices A–H. Typical of globular proteins, the surface of myoglobin is polar, while—with only two exceptions—the interior contains only nonpolar residues such as Leu, Val, Phe, and Met. The exceptions are His E7 and His F8, the seventh and eighth residues in helices E and F, which lie close to the heme iron where they function in O_2 binding.

Histidines F8 & E7 Perform Unique Roles in Oxygen Binding

The heme of myoglobin lies in a crevice between helices E and F oriented with its polar propionate groups facing the surface of the globin (Figure 6–2). The remainder resides in the nonpolar interior. The fifth coordination position of the iron is linked to a ring nitrogen of the **proximal histidine,** His F8. The **distal histidine,** His E7, lies on the side of the heme ring opposite to His F8.

The Iron Moves Toward the Plane of the Heme When Oxygen Is Bound

The iron of unoxygenated myoglobin lies 0.03 nm (0.3 Å) outside the plane of the heme ring, toward His F8. The heme therefore "puckers" slightly. When O_2 occupies the sixth coordination position, the iron moves to within 0.01 nm (0.1 Å) of the plane of the heme ring. Oxygenation of myoglobin thus is accompanied by motion of the iron, of His F8, and of residues linked to His F8.

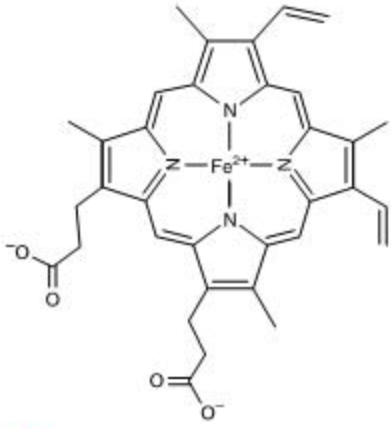


Figure 6–1. Heme. The pyrrole rings and methylene bridge carbons are coplanar, and the iron atom (Fe_2^+) resides in almost the same plane. The fifth and sixth coordination positions of Fe_2^+ are directed perpendicular to—and directly above and below—the plane of the heme ring. Observe the nature of the substituent groups on the β carbons of the pyrrole rings, the central iron atom, and the location of the polar side of the heme ring (at about 7 o'clock) that faces the surface of the myoglobin molecule.

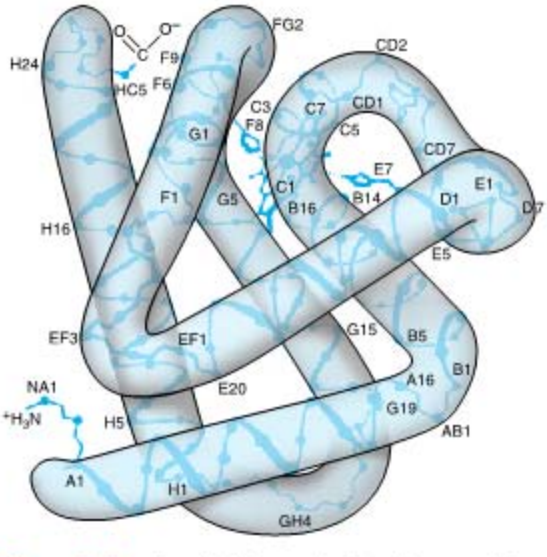


Figure 6–2. A model of myoglobin at low resolution. Only the α-carbon atoms are shown. The α-helical regions are named A through H. (Based on Dickerson RE in: The Proteins, 2nd ed. Vol 2. Neurath H [editor]. Academic Press, 1964. Reproduced with permission.)

Apomyoglobin Provides a Hindered Environment for Heme Iron

When O_2 binds to myoglobin, the bond between the first oxygen atom and the Fe_2^+ is perpendicular to the plane of the heme ring. The bond linking the first and second oxygen atoms lies at an angle of 121 degrees to the plane of the heme, orienting the second oxygen away from the distal histidine (Figure 6–3, left). Isolated heme binds carbon monoxide (CO) 25,000 times more strongly than oxygen. Since CO is present in small quantities in the atmosphere and arises in cells from the catabolism of heme, why is it that CO does not completely displace O_2 from heme iron? The accepted explanation is that the apoproteins of myoglobin and hemoglobin create a **hindered environment.** While CO can bind to isolated heme in its preferred orientation, ie, with all three atoms (Fe, C, and O) perpendicular to the plane of the heme, in myoglobin and hemoglobin the distal histidine sterically precludes this orientation. Binding at a less favored angle reduces the strength of the heme-CO bond to about 200 times that of the heme-O_2 bond (Figure 6–3, right) at which level the great excess of O_2 over CO normally present dominates. Nevertheless, about 1% of myoglobin typically is present combined with carbon monoxide.

THE OXYGEN DISSOCIATION CURVES FOR MYOGLOBIN & HEMOGLOBIN SUIT THEIR PHYSIOLOGIC ROLES

Why is myoglobin unsuitable as an O_2 transport protein but well suited for O_2 storage? The relationship between the concentration, or partial pressure, of O_2 (P_{O_2}) and the quantity of O_2 bound is expressed as an O_2 saturation isotherm (Figure 6–4). The oxygen-

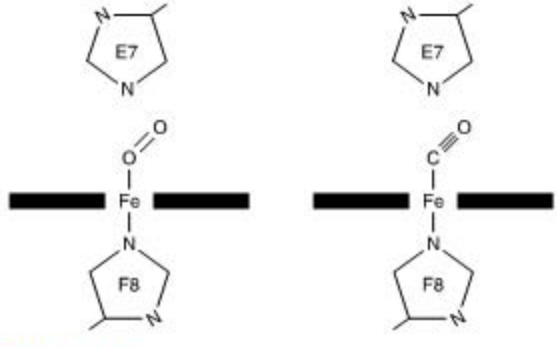


Figure 6–3. Angles for bonding of oxygen and carbon monoxide to the heme iron of myoglobin. The distal E7 histidine hinders bonding of CO at the preferred (180 degree) angle to the plane of the heme ring.

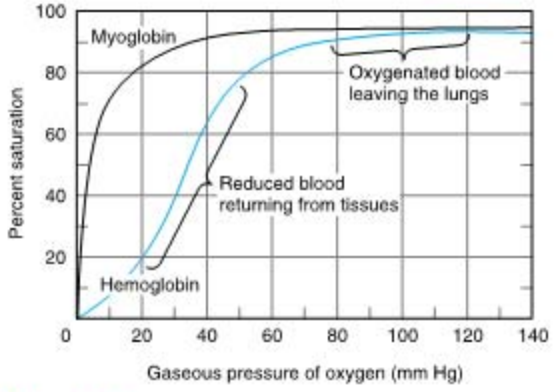


Figure 6–4. Oxygen-binding curves of both hemoglobin and myoglobin. Arterial oxygen tension is about 100 mm Hg; mixed venous oxygen tension is about 40 mm Hg; capillary (active muscle) oxygen tension is about 20 mm Hg; and the minimum oxygen tension required for cytochrome oxidase is about 5 mm Hg. Association of chains into a tetrameric structure (hemoglobin) results in much greater oxygen delivery than would be possible with single chains. (Modified, with permission, from Scriver CR et al [editors]: The Molecular and Metabolic Bases of Inherited Disease, 7th ed. McGraw-Hill, 1995.)

binding curve for myoglobin is hyperbolic. Myoglobin therefore loads O_2 readily at the P_{O_2} of the lung capillary bed (100 mm Hg). However, since myoglobin releases only a small fraction of its bound O_2 at the P_{O_2} values typically encountered in active muscle (20 mm Hg) or other tissues (40 mm Hg), it represents an ineffective vehicle for delivery of O_2. However, when strenuous exercise lowers the P_{O_2} of muscle tissue to about 5 mm Hg, myoglobin releases O_2 for mitochondrial synthesis of ATP, permitting continued muscular activity.

THE ALLOSTERIC PROPERTIES OF HEMOGLOBINS RESULT FROM THEIR QUATERNARY STRUCTURES

The properties of individual hemoglobins are consequences of their quaternary as well as of their secondary and tertiary structures. The quaternary structure of hemoglobin confers striking additional properties, absent from monomeric myoglobin, which adapts it to its unique biologic roles. The **allosteric** (Gk *allos* "other," *steros* "space") properties of hemoglobin provide, in addition, a model for understanding other allosteric proteins (see Chapter 11).

Hemoglobin Is Tetrameric

Hemoglobins are tetramers comprised of pairs of two different polypeptide subunits. Greek letters are used to designate each subunit type. The subunit composition of the principal hemoglobins are $\alpha_2\beta_2$ (HbA; normal adult hemoglobin), $\alpha_2\gamma_2$ (HbF; fetal hemoglobin), α_2S_2 (HbS; sickle cell hemoglobin), and $\alpha_2\delta_2$ (HbA_2; a minor adult hemoglobin). The primary structures of the β, γ, and δ chains of human hemoglobin are highly conserved.

Myoglobin & the β Subunits of Hemoglobin Share Almost Identical Secondary and Tertiary Structures

Despite differences in the kind and number of amino acids present, myoglobin and the β polypeptide of hemoglobin A have almost identical secondary and tertiary structures. Similarities include the location of the heme and the eight helical regions and the presence of amino acids with similar properties at comparable locations. Although it possesses seven rather than eight helical regions, the α polypeptide of hemoglobin also closely resembles myoglobin.

Oxygenation of Hemoglobin Triggers Conformational Changes in the Apoprotein

Hemoglobins bind four molecules of O_2 per tetramer, one per heme. A molecule of O_2 binds to a hemoglobin tetramer more readily if other O_2 molecules are already bound (Figure 6–4). Termed **cooperative binding,** this phenomenon permits hemoglobin to maximize both the quantity of O_2 loaded at the P_{O_2} of the lungs and the quantity of O_2 released at the P_{O_2} of the peripheral tissues. Cooperative interactions, an exclusive property of multimeric proteins, are critically important to aerobic life.

P_{50} Expresses the Relative Affinities of Different Hemoglobins for Oxygen

The quantity P_{50}, a measure of O_2 concentration, is the partial pressure of O_2 that half-saturates a given hemoglobin. Depending on the organism, P_{50} can vary widely, but in all instances it will exceed the P_{O_2} of the peripheral tissues. For example, values of P_{50} for HbA and fetal HbF are 26 and 20 mm Hg, respectively. In the placenta, this difference enables HbF to extract oxygen from the HbA in the mother's blood. However, HbF is suboptimal postpartum since its high affinity for O_2 dictates that it can deliver less O_2 to the tissues.

The subunit composition of hemoglobin tetramers undergoes complex changes during development. The

human fetus initially synthesizes a $\zeta_2\varepsilon_2$ tetramer. By the end of the first trimester, ζ and γ subunits have been replaced by α and ε subunits, forming HbF ($\alpha_2\gamma_2$), the hemoglobin of late fetal life. While synthesis of β subunits begins in the third trimester, β subunits do not completely replace γ subunits to yield adult HbA ($\alpha_2\beta_2$) until some weeks postpartum (Figure 6–5).

Oxygenation of Hemoglobin Is Accompanied by Large Conformational Changes

The binding of the first O_2 molecule to deoxyHb shifts the heme iron towards the plane of the heme ring from a position about 0.6 nm beyond it (Figure 6–6). This motion is transmitted to the proximal (F8) histidine and to the residues attached thereto, which in turn causes the rupture of salt bridges between the carboxyl terminal residues of all four subunits. As a consequence, one pair of α/β subunits rotates 15 degrees with respect to the other, compacting the tetramer (Figure 6–7). Profound changes in secondary, tertiary, and quaternary structure accompany the high-affinity O_2-induced transition of hemoglobin from the low-affinity **T (taut) state** to the **R (relaxed) state.** These changes significantly increase the affinity of the remaining unoxygenated hemes for O_2, as subsequent binding events require the rupture of fewer salt bridges (Figure 6–8). The terms T and R also are used to refer to the low-affinity and high-affinity conformations of allosteric enzymes, respectively.

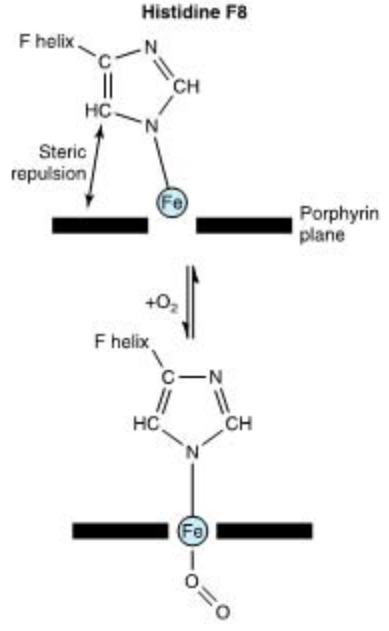


Figure 6–6. The iron atom moves into the plane of the heme on oxygenation. Histidine F8 and its associated residues are pulled along with the iron atom. (Slightly modified and reproduced, with permission, from Stryer L: *Biochemistry,* 4th ed. Freeman, 1995.)

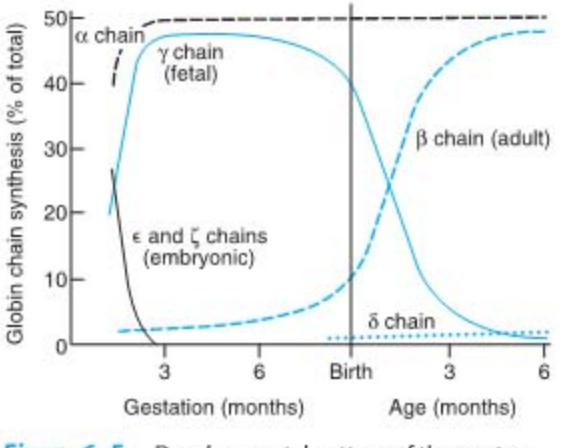


Figure 6–5. Developmental pattern of the quaternary structure of fetal and newborn hemoglobins. (Reproduced, with permission, from Ganong WF: *Review of Medical Physiology,* 20th ed. McGraw-Hill, 2001.)

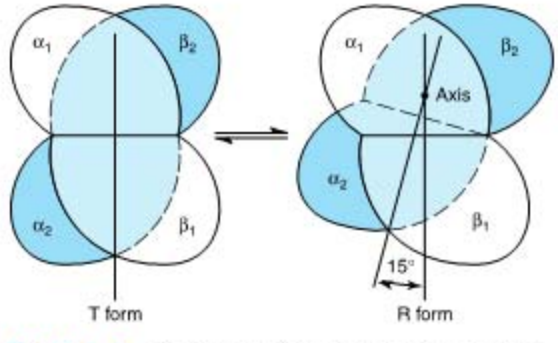


Figure 6–7. During transition of the T form to the R form of hemoglobin, one pair of subunits (α_2/β_2) rotates through 15 degrees relative to the other pair (α_1/β_1). The axis of rotation is eccentric, and the α_2/β_2 pair also shifts toward the axis somewhat. In the diagram, the unshaded α_1/β_1 pair is shown fixed while the colored α_2/β_2 pair both shifts and rotates.

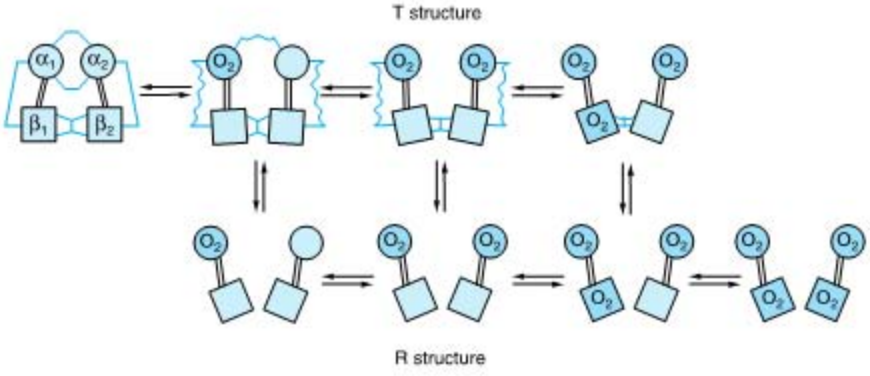
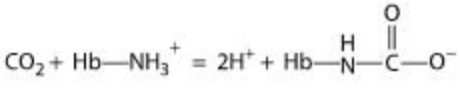


Figure 6–8. Transition from the T structure to the R structure. In this model, salt bridges (thin lines) linking the subunits in the T structure break progressively as oxygen is added, and even those salt bridges that have not yet ruptured are progressively weakened (wavy lines). The transition from T to R does not take place after a fixed number of oxygen molecules have been bound but becomes more probable as each successive oxygen binds. The transition between the two structures is influenced by protons, carbon dioxide, chloride, and BPG; the higher their concentration, the more oxygen must be bound to trigger the transition. Fully oxygenated molecules in the T structure and fully deoxygenated molecules in the R structure are not shown because they are unstable. (Modified and redrawn, with permission, from Perutz MF: Hemoglobin structure and respiratory transport. Sci Am [Dec] 1978;239:92.)

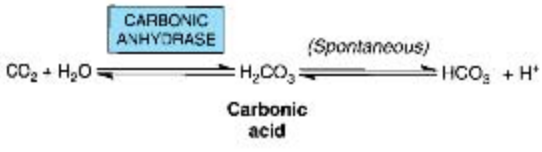
After Releasing O_2 at the Tissues, Hemoglobin Transports CO_2 & Protons to the Lungs

In addition to transporting O_2 from the lungs to peripheral tissues, hemoglobin transports CO_2, the byproduct of respiration, and protons from peripheral tissues to the lungs. Hemoglobin carries CO_2 as carbamates formed with the amino terminal nitrogens of the polypeptide chains.



Carbamates change the charge on amino terminals from positive to negative, favoring salt bond formation between the α and β chains.

Hemoglobin carbamates account for about 15% of the CO_2 in venous blood. Much of the remaining CO_2 is carried as bicarbonate, which is formed in erythrocytes by the hydration of CO_2 to carbonic acid (H_2CO_3), a process catalyzed by carbonic anhydrase. At the pH of venous blood, H_2CO_3 dissociates into bicarbonate and a proton.



Deoxyhemoglobin binds one proton for every two O_2 molecules released, contributing significantly to the buffering capacity of blood. The somewhat lower pH of peripheral tissues, aided by carbamation, stabilizes the T state and thus enhances the delivery of O_2. In the lungs, the process reverses. As O_2 binds to deoxyhemoglobin, protons are released and combine with bicarbonate to form carbonic acid. Dehydration of H_2CO_3, catalyzed by carbonic anhydrase, forms CO_2, which is exhaled. Binding of oxygen thus drives the exhalation of CO_2 (Figure 6–9). This reciprocal coupling of proton and O_2 binding is termed the **Bohr effect.** The Bohr effect is dependent upon **cooperative interactions between the hemes of the hemoglobin tetramer.** Myoglobin, a monomer, exhibits no Bohr effect.

Protons Arise From Rupture of Salt Bonds When O_2 Binds

Protons responsible for the Bohr effect arise from rupture of salt bridges during the binding of O_2 to T state

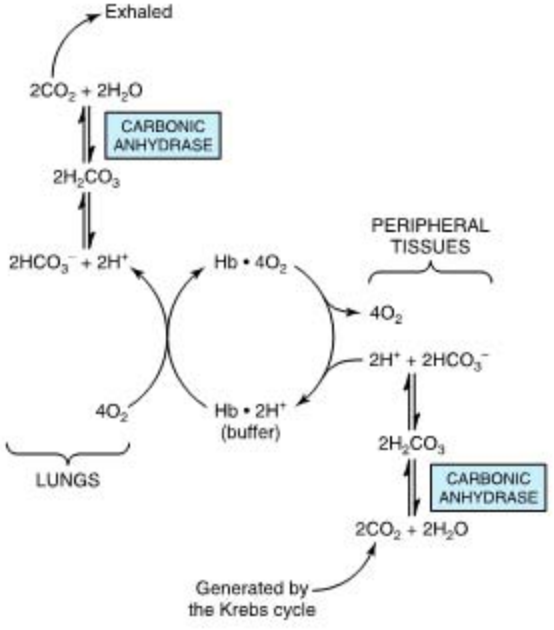
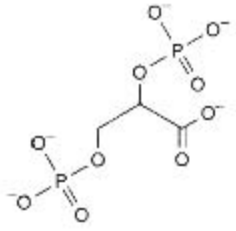


Figure 6–9. The Bohr effect. Carbon dioxide generated in peripheral tissues combines with water to form carbonic acid, which dissociates into protons and bicarbonate ions. Deoxyhemoglobin acts as a buffer by binding protons and delivering them to the lungs. In the lungs, the uptake of oxygen by hemoglobin releases protons that combine with bicarbonate ion, forming carbonic acid, which when dehydrated by carbonic anhydrase becomes carbon dioxide, which then is exhaled.

hemoglobin. Conversion to the oxygenated R state breaks salt bridges involving β-chain residue His 146. The subsequent dissociation of protons from His 146 drives the conversion of bicarbonate to carbonic acid (Figure 6–9). Upon the release of O_2, the T structure and its salt bridges re-form. This conformational change increases the pK_a of the β-chain His 146 residues, which bind protons. By facilitating the re-formation of salt bridges, an increase in proton concentration enhances the release of O_2 from oxygenated (R state) hemoglobin. Conversely, an increase in P_{O_2} promotes proton release.

2,3-Bisphosphoglycerate (BPG) Stabilizes the T Structure of Hemoglobin

A low P_{O_2} in peripheral tissues promotes the synthesis in erythrocytes of 2,3-bisphosphoglycerate (BPG) from the glycolytic intermediate 1,3-bisphosphoglycerate.



The hemoglobin tetramer binds one molecule of BPG in the central cavity formed by its four subunits. However, the space between the H helices of the β chains lining the cavity is sufficiently wide to accommodate BPG only when hemoglobin is in the T state. BPG forms salt bridges with the terminal amino groups of both β chains via Val NA1 and with Lys EF6 and His H21 (Figure 6–10). BPG therefore stabilizes deoxygenated (T state) hemoglobin by forming additional salt bridges that must be broken prior to conversion to the R state.

Residue H21 of the γ subunit of fetal hemoglobin (HbF) is Ser rather than His. Since Ser cannot form a salt bridge, BPG binds more weakly to HbF than to HbA. The lower stabilization afforded to the T state by BPG accounts for HbF having a higher affinity for O_2 than HbA.

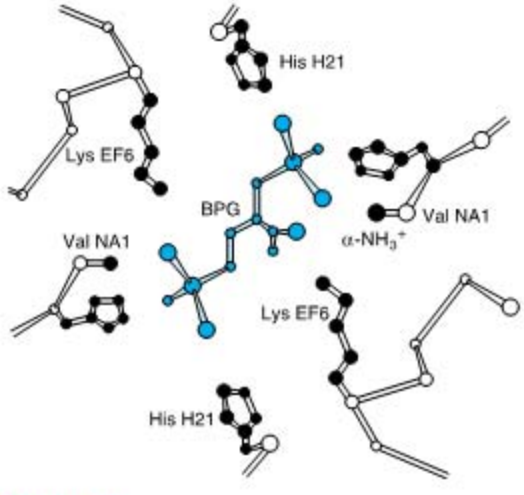


Figure 6–10. Mode of binding of 2,3-bisphosphoglycerate to human deoxyhemoglobin. BPG interacts with three positively charged groups on each β chain. (Based on Amone A: X-ray diffraction study of binding of 2,3-diphosphoglycerate to human deoxyhemoglobin. Nature 1972;237:146. Reproduced with permission.)

Adaptation to High Altitude

Physiologic changes that accompany prolonged exposure to high altitude include an increase in the number of erythrocytes and in their concentrations of hemoglobin and of BPG. Elevated BPG lowers the affinity of HbA for O_2 (decreases P_{50}), which enhances release of O_2 at the tissues.

NUMEROUS MUTANT HUMAN HEMOGLOBINS HAVE BEEN IDENTIFIED

Mutations in the genes that encode the α or β subunits of hemoglobin potentially can affect its biologic function. However, almost all of the over 800 known mutant human hemoglobins are both extremely rare and benign, presenting no clinical abnormalities. When a mutation does compromise biologic function, the condition is termed a **hemoglobinopathy.** The URL http://globin.cse.psu.edu/ (Globin Gene Server) provides information about—and links for—normal and mutant hemoglobins.

Methemoglobin & Hemoglobin M

In methemoglobinemia, the heme iron is ferric rather than ferrous. Methemoglobin thus can neither bind nor transport O_2. Normally, the enzyme methemoglobin reductase reduces the Fe_3^+ of methemoglobin to Fe_2^+. Methemoglobin can arise by oxidation of Fe_2^+ to Fe_3^+ as a side effect of agents such as sulfonamides, from hereditary hemoglobin M, or consequent to reduced activity of the enzyme methemoglobin reductase.

In hemoglobin M, histidine F8 (His F8) has been replaced by tyrosine. The iron of HbM forms a tight ionic complex with the phenolate anion of tyrosine that stabilizes the Fe_3^+ form. In α-chain hemoglobin M variants, the R-T equilibrium favors the T state. Oxygen affinity is reduced, and the Bohr effect is absent. β-Chain hemoglobin M variants exhibit R-T switching, and the Bohr effect is therefore present.

Mutations (eg, hemoglobin Chesapeake) that favor the R state increase O_2 affinity. These hemoglobins therefore fail to deliver adequate O_2 to peripheral tissues. The resulting tissue hypoxia leads to **polycythemia,** an increased concentration of erythrocytes.

Hemoglobin S

In HbS, the nonpolar amino acid valine has replaced the polar surface residue Glu6 of the β subunit, generating a hydrophobic **"sticky patch"** on the surface of the β subunit of both oxyHbS and deoxyHbS. Both HbA and HbS contain a complementary sticky patch on their surfaces that is exposed only in the deoxygenated, R state. Thus, at low P_{O_2}, deoxyHbS can polymerize to form long, insoluble fibers. Binding of deoxyHbA terminates fiber polymerization, since HbA lacks the second sticky patch necessary to bind another Hb molecule (Figure 6–11). These twisted helical fibers distort the erythrocyte into a characteristic sickle shape, rendering it vulnerable to lysis in the interstices of the splenic sinusoids. They also cause multiple secondary clinical effects. A low P_{O_2} such as that at high altitudes exacerbates the tendency to polymerize.

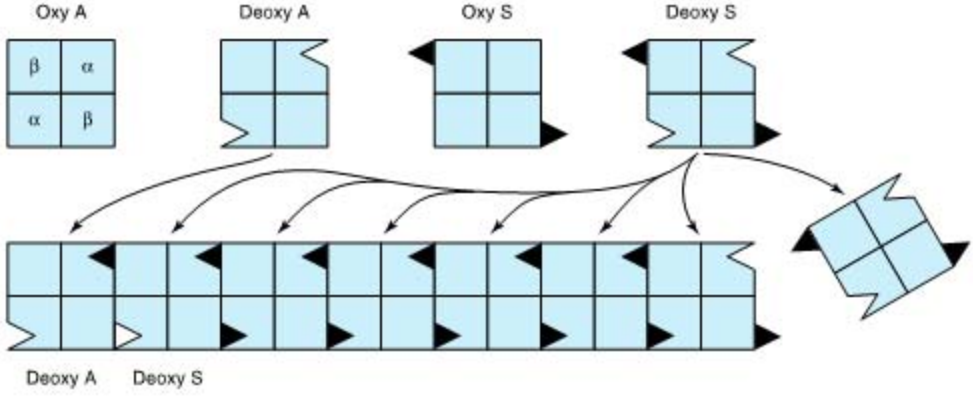


Figure 6–11. Representation of the sticky patch (▲) on hemoglobin S and its "receptor" (△) on deoxyhemoglobin A and deoxyhemoglobin S. The complementary surfaces allow deoxyhemoglobin S to polymerize into a fibrous structure, but the presence of deoxyhemoglobin A will terminate the polymerization by failing to provide sticky patches. (Modified and reproduced, with permission, from Stryer L: *Biochemistry,* 4th ed. Freeman, 1995.)

BIOMEDICAL IMPLICATIONS

Myoglobinuria

Following massive crush injury, myoglobin released from damaged muscle fibers colors the urine dark red. Myoglobin can be detected in plasma following a myocardial infarction, but assay of serum enzymes (see Chapter 7) provides a more sensitive index of myocardial injury.

Anemias

Anemias, reductions in the number of red blood cells or of hemoglobin in the blood, can reflect impaired synthesis of hemoglobin (eg, in iron deficiency; Chapter 51) or impaired production of erythrocytes (eg, in folic acid or vitamin B_{12} deficiency; Chapter 45). Diagnosis of anemias begins with spectroscopic measurement of blood hemoglobin levels.

Thalassemias

The genetic defects known as thalassemias result from the partial or total absence of one or more α or β chains of hemoglobin. Over 750 different mutations have been identified, but only three are common. Either the α chain (alpha thalassemias) or β chain (beta thalassemias) can be affected. A superscript indicates whether a subunit is completely absent (α^0 or β^0) or whether its synthesis is reduced (α^+ or β^+). Apart from marrow transplantation, treatment is symptomatic.

Certain mutant hemoglobins are common in many populations, and a patient may inherit more than one type. Hemoglobin disorders thus present a complex pattern of clinical phenotypes. The use of DNA probes for their diagnosis is considered in Chapter 40.

Glycosylated Hemoglobin (HbA_{1c})

When blood glucose enters the erythrocytes it glycosylates the ε-amino group of lysine residues and the amino terminals of hemoglobin. The fraction of hemoglobin glycosylated, normally about 5%, is proportionate to blood glucose concentration. Since the half-life of an erythrocyte is typically 60 days, the level of glycosylated hemoglobin (HbA_{1c}) reflects the mean blood glucose concentration over the preceding 6–8 weeks. Measurement of HbA_{1c} therefore provides valuable information for management of diabetes mellitus.

SUMMARY

- Myoglobin is monomeric; hemoglobin is a tetramer of two subunit types ($\alpha_2\beta_2$ in HbA). Despite having different primary structures, myoglobin and the subunits of hemoglobin have nearly identical secondary and tertiary structures.
- Heme, an essentially planar, slightly puckered, cyclic tetrapyrrole, has a central Fe_2^+ linked to all four nitrogen atoms of the heme, to histidine F8, and, in oxyMb and oxyHb, also to O_2.
- The O_2-binding curve for myoglobin is hyperbolic, but for hemoglobin it is sigmoidal, a consequence of cooperative interactions in the tetramer. Cooperativity maximizes the ability of hemoglobin both to load O_2 at the P_{O_2} of the lungs and to deliver O_2 at the P_{O_2} of the tissues.
- Relative affinities of different hemoglobins for oxygen are expressed as P_{50}, the P_{O_2} that half-saturates them with O_2. Hemoglobins saturate at the partial pressures of their respective respiratory organ, eg, the lung or placenta.
- On oxygenation of hemoglobin, the iron, histidine F8, and linked residues move toward the heme ring. Conformational changes that accompany oxygenation include rupture of salt bonds and loosening of quaternary structure, facilitating binding of additional O_2.
- 2,3-Bisphosphoglycerate (BPG) in the central cavity of deoxyHb forms salt bonds with the β subunits that stabilize deoxyHb. On oxygenation, the central cavity contracts, BPG is extruded, and the quaternary structure loosens.
- Hemoglobin also functions in CO_2 and proton transport from tissues to lungs. Release of O_2 from oxyHb at the tissues is accompanied by uptake of protons due to lowering of the pK_a of histidine residues.
- In sickle cell hemoglobin (HbS), Val replaces the β6 Glu of HbA, creating a "sticky patch" that has a complement on deoxyHb (but not on oxyHb). DeoxyHbS polymerizes at low O_2 concentrations, forming fibers that distort erythrocytes into sickle shapes.
- Alpha and beta thalassemias are anemias that result from reduced production of α and β subunits of HbA, respectively.

REFERENCES

Bettati S et al: Allosteric mechanism of haemoglobin: Rupture of salt-bridges raises the oxygen affinity of the T-structure. J Mol Biol 1998;281:581.

Bunn HF: Pathogenesis and treatment of sickle cell disease. N Engl J Med 1997;337:762.

Faustino P et al: Dominantly transmitted beta-thalassemia arising from the production of several aberrant mRNA species and one abnormal peptide. Blood 1998;91:685.

Manning JM et al: Normal and abnormal protein subunit interactions in hemoglobins. J Biol Chem 1998;273:19359.

Mario N, Baudin B, Giboudeau J: Qualitative and quantitative analysis of hemoglobin variants by capillary isoelectric focusing. J Chromatogr B Biomed Sci Appl 1998;706:123.

Reed W, Vichinsky EP: New considerations in the treatment of sickle cell disease. Annu Rev Med 1998;49:461.

Unzai S et al: Rate constants for O_2 and CO binding to the alpha and beta subunits within the R and T states of human hemoglobin. J Biol Chem 1998;273:23150.

Weatherall DJ et al: The hemoglobinopathies. Chapter 181 in *The Metabolic and Molecular Bases of Inherited Disease,* 8th ed. Scriver CR et al (editors). McGraw-Hill, 2000.

Enzymes: Mechanism of Action

7

Victor W. Rodwell, PhD, & Peter J. Kennelly, PhD

BIOMEDICAL IMPORTANCE

Enzymes are biologic polymers that catalyze the chemical reactions which make life as we know it possible. The presence and maintenance of a complete and balanced set of enzymes is essential for the breakdown of nutrients to supply energy and chemical building blocks; the assembly of those building blocks into proteins, DNA, membranes, cells, and tissues; and the harnessing of energy to power cell motility and muscle contraction. With the exception of a few catalytic RNA molecules, or ribozymes, the vast majority of enzymes are proteins. Deficiencies in the quantity or catalytic activity of key enzymes can result from genetic defects, nutritional deficits, or toxins. Defective enzymes can result from genetic mutations or infection by viral or bacterial pathogens (eg, *Vibrio cholerae*). Medical scientists address imbalances in enzyme activity by using pharmacologic agents to inhibit specific enzymes and are investigating gene therapy as a means to remedy deficits in enzyme level or function.

ENZYMES ARE EFFECTIVE & HIGHLY SPECIFIC CATALYSTS

The enzymes that catalyze the conversion of one or more compounds **(substrates)** into one or more different compounds **(products)** enhance the rates of the corresponding noncatalyzed reaction by factors of at least 10^6. Like all catalysts, enzymes are neither consumed nor permanently altered as a consequence of their participation in a reaction.

In addition to being highly efficient, enzymes are also extremely selective catalysts. Unlike most catalysts used in synthetic chemistry, enzymes are specific both for the type of reaction catalyzed and for a single substrate or a small set of closely related substrates. Enzymes are also stereospecific catalysts and typically catalyze reactions only of specific stereoisomers of a given compound—for example, D- but not L-sugars, L- but not D-amino acids. Since they bind substrates through at least "three points of attachment," enzymes can even convert nonchiral substrates to chiral products. Figure 7–1 illustrates why the enzyme-catalyzed reduction of the nonchiral substrate pyruvate produces L-lactate rather a racemic mixture of D- and L-lactate. The exquisite specificity of enzyme catalysts imbues living cells with the ability to simultaneously conduct and independently control a broad spectrum of chemical processes.

ENZYMES ARE CLASSIFIED BY REACTION TYPE & MECHANISM

A system of enzyme nomenclature that is comprehensive, consistent, and at the same time easy to use has proved elusive. The common names for most enzymes derive from their most distinctive characteristic: their ability to catalyze a specific chemical reaction. In general, an enzyme's name consists of a term that identifies the type of reaction catalyzed followed by the suffix *-ase*. For example, dehydrogen*ases* remove hydrogen atoms, prote*ases* hydrolyze proteins, and isomer*ases* catalyze rearrangements in configuration. One or more modifiers usually precede this name. Unfortunately, while many modifiers name the specific substrate involved (xanthine oxidase), others identify the source of the enzyme (pancreatic ribonuclease), specify its mode of regulation (hormone-sensitive lipase), or name a distinguishing characteristic of its mechanism (a cysteine protease). When it was discovered that multiple forms of some enzymes existed, alphanumeric designators were added to distinguish between them (eg, RNA polymerase III; protein kinase Cβ). To address the ambiguity and confusion arising from these inconsistencies in nomenclature and the continuing discovery of new enzymes, the International Union of Biochemists (IUB) developed a complex but unambiguous system of enzyme nomenclature. In the IUB system, each enzyme has a unique name and code number that reflect the type of reaction catalyzed and the substrates involved. Enzymes are grouped into six classes, each with several subclasses. For example, the enzyme commonly called "hexokinase" is designated "ATP:D-hexose-6-phosphotransferase E.C. 2.7.1.1." This identifies hexokinase as a member of class 2 (transferases), subclass 7 (transfer of a phosphoryl group), sub-subclass 1 (alcohol is the phosphoryl acceptor). Finally, the term "hexose-6" indicates that the alcohol phosphorylated is that of carbon six of a hexose. Listed below are the six IUB classes of enzymes and the reactions they catalyze.

1. **Oxidoreductases** catalyze oxidations and reductions.

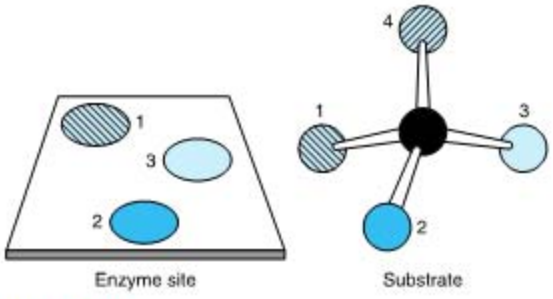


Figure 7–1. Planar representation of the "three-point attachment" of a substrate to the active site of an enzyme. Although atoms 1 and 4 are identical, once atoms 2 and 3 are bound to their complementary sites on the enzyme, only atom 1 can bind. Once bound to an enzyme, apparently identical atoms thus may be distinguishable, permitting a stereospecific chemical change.

2. **Transferases** catalyze transfer of groups such as methyl or glycosyl groups from a donor molecule to an acceptor molecule.
3. **Hydrolases** catalyze the hydrolytic cleavage of C—C, C—O, C—N, P—O, and certain other bonds, including acid anhydride bonds.
4. **Lyases** catalyze cleavage of C—C, C—O, C—N, and other bonds by elimination, leaving double bonds, and also add groups to double bonds.
5. **Isomerases** catalyze geometric or structural changes within a single molecule.
6. **Ligases** catalyze the joining together of two molecules, coupled to the hydrolysis of a pyrophosphoryl group in ATP or a similar nucleoside triphosphate.

Despite the many advantages of the IUB system, texts tend to refer to most enzymes by their older and shorter, albeit sometimes ambiguous names.

PROSTHETIC GROUPS, COFACTORS, & COENZYMES PLAY IMPORTANT ROLES IN CATALYSIS

Many enzymes contain small nonprotein molecules and metal ions that participate directly in substrate binding or catalysis. Termed **prosthetic groups, cofactors,** and **coenzymes,** these extend the repertoire of catalytic capabilities beyond those afforded by the limited number of functional groups present on the aminoacyl side chains of peptides.

Prosthetic Groups Are Tightly Integrated Into an Enzyme's Structure

Prosthetic groups are distinguished by their tight, stable incorporation into a protein's structure by covalent or noncovalent forces. Examples include pyridoxal phosphate, flavin mononucleotide (FMN), flavin dinucleotide (FAD), thiamin pyrophosphate, biotin, and the metal ions of Co, Cu, Mg, Mn, Se, and Zn. Metals are the most common prosthetic groups. The roughly one-third of all enzymes that contain tightly bound metal ions are termed **metalloenzymes.** Metal ions that participate in redox reactions generally are complexed to prosthetic groups such as heme (Chapter 6) or iron-sulfur clusters (Chapter 12). Metals also may facilitate the binding and orientation of substrates, the formation of covalent bonds with reaction intermediates (Co^{2+} in coenzyme B_{12}), or interaction with substrates to render them more **electrophilic** (electron-poor) or **nucleophilic** (electron-rich).

Cofactors Associate Reversibly With Enzymes or Substrates

Cofactors serve functions similar to those of prosthetic groups but bind in a transient, dissociable manner either to the enzyme or to a substrate such as ATP. Unlike the stably associated prosthetic groups, cofactors therefore must be present in the medium surrounding the enzyme for catalysis to occur. The most common cofactors also are metal ions. Enzymes that require a metal ion cofactor are termed **metal-activated enzymes** to distinguish them from the **metalloenzymes** for which metal ions serve as prosthetic groups.

Coenzymes Serve as Substrate Shuttles

Coenzymes serve as recyclable shuttles—or group transfer reagents—that transport many substrates from their point of generation to their point of utilization. Association with the coenzyme also stabilizes substrates such as hydrogen atoms or hydride ions that are unstable in the aqueous environment of the cell. Other chemical moieties transported by coenzymes include methyl groups (folates), acyl groups (coenzyme A), and oligosaccharides (dolichol).

Many Coenzymes, Cofactors, & Prosthetic Groups Are Derivatives of B Vitamins

The water-soluble B vitamins supply important components of numerous coenzymes. Many coenzymes contain, in addition, the adenine, ribose, and phosphoryl moieties of AMP or ADP (Figure 7–2). **Nicotinamide** and **riboflavin** are components of the redox coenzymes

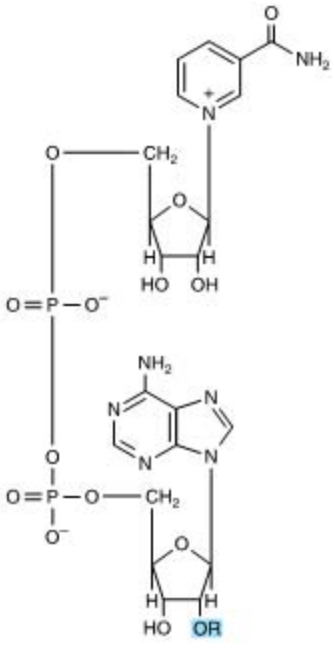


Figure 7–2. Structure of NAD^+ and $NADP^+$. For NAD^+, R = H. For $NADP^+$, $R = PO_3^{2-}$.

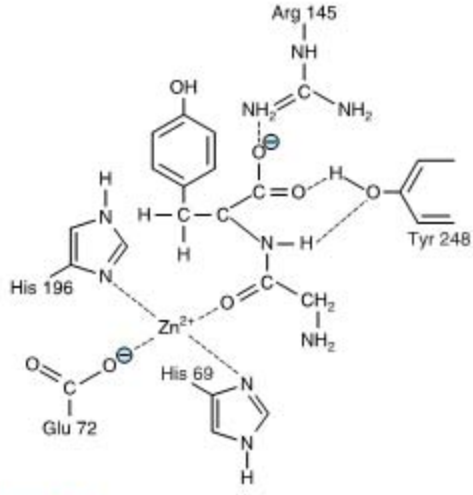


Figure 7–3. Two-dimensional representation of a dipeptide substrate, glycyl-tyrosine, bound within the active site of carboxypeptidase A.

NAD and NADP and FMN and FAD, respectively. **Pantothenic acid** is a component of the acyl group carrier coenzyme A. As its pyrophosphate, **thiamin** participates in decarboxylation of α-keto acids and **folic acid** and **cobamide** coenzymes function in one-carbon metabolism.

CATALYSIS OCCURS AT THE ACTIVE SITE

The extreme substrate specificity and high catalytic efficiency of enzymes reflect the existence of an environment that is exquisitely tailored to a single reaction. Termed the **active site,** this environment generally takes the form of a cleft or pocket. The active sites of multimeric enzymes often are located at the interface between subunits and recruit residues from more than one monomer. The three-dimensional active site both shields substrates from solvent and facilitates catalysis. Substrates bind to the active site at a region complementary to a portion of the substrate that will *not* undergo chemical change during the course of the reaction. This simultaneously aligns portions of the substrate that *will* undergo change with the chemical functional groups of peptidyl aminoacyl residues. The active site also binds and orients cofactors or prosthetic groups. Many amino acyl residues drawn from diverse portions of the polypeptide chain (Figure 7–3) contribute to the extensive size and three-dimensional character of the active site.

ENZYMES EMPLOY MULTIPLE MECHANISMS TO FACILITATE CATALYSIS

Four general mechanisms account for the ability of enzymes to achieve dramatic catalytic enhancement of the rates of chemical reactions.

Catalysis by Proximity

For molecules to react, they must come within bond-forming distance of one another. The higher their concentration, the more frequently they will encounter one another and the greater will be the rate of their reaction. When an enzyme binds substrate molecules in its active site, it creates a region of high local substrate concentration. This environment also orients the substrate molecules spatially in a position ideal for them to interact, resulting in rate enhancements of at least a thousandfold.

Acid-Base Catalysis

The ionizable functional groups of aminoacyl side chains and (where present) of prosthetic groups contribute to catalysis by acting as acids or bases. Acid-base catalysis can be either specific or general. By "specific" we mean only protons (H_3O^+) or OH^- ions. In **specific acid** or **specific base catalysis,** the rate of reaction is sensitive to changes in the concentration of protons but

independent of the concentrations of other acids (proton donors) or bases (proton acceptors) present in solution or at the active site. Reactions whose rates are responsive to *all* the acids or bases present are said to be subject to **general acid** or **general base catalysis.**

Catalysis by Strain

Enzymes that catalyze lytic reactions which involve breaking a covalent bond typically bind their substrates in a conformation slightly unfavorable for the bond that will undergo cleavage. The resulting strain stretches or distorts the targeted bond, weakening it and making it more vulnerable to cleavage.

Covalent Catalysis

The process of **covalent catalysis** involves the formation of a covalent bond between the enzyme and one or more substrates. The modified enzyme then becomes a reactant. Covalent catalysis introduces a new reaction pathway that is energetically more favorable—and therefore faster—than the reaction pathway in homogeneous solution. The chemical modification of the enzyme is, however, transient. On completion of the reaction, the enzyme returns to its original unmodified state. Its role thus remains catalytic. Covalent catalysis is particularly common among enzymes that catalyze group transfer reactions. Residues on the enzyme that participate in covalent catalysis generally are cysteine or serine and occasionally histidine. Covalent catalysis often follows a "ping-pong" mechanism—one in which the first substrate is bound and its product released prior to the binding of the second substrate (Figure 7–4).

SUBSTRATES INDUCE CONFORMATIONAL CHANGES IN ENZYMES

Early in the last century, Emil Fischer compared the highly specific fit between enzymes and their substrates to that of a lock and its key. While the "lock and key model" accounted for the exquisite specificity of enzyme-substrate interactions, the implied rigidity of the enzyme's active site failed to account for the dynamic changes that accompany catalysis. This drawback was addressed by Daniel Koshland's **induced fit** model, which states that when substrates approach and bind to an enzyme they induce a conformational change, a change analogous to placing a hand (substrate) into a glove (enzyme) (Figure 7–5). A corollary is that the enzyme induces reciprocal changes in its substrates, harnessing the energy of binding to facilitate the transformation of substrates into products. The induced fit model has been amply confirmed by biophysical studies of enzyme motion during substrate binding.

HIV PROTEASE ILLUSTRATES ACID-BASE CATALYSIS

Enzymes of the **aspartic protease family,** which includes the digestive enzyme pepsin, the lysosomal cathepsins, and the protease produced by the human immunodeficiency virus (HIV), share a common catalytic mechanism. Catalysis involves two conserved aspartyl residues which act as acid-base catalysts. In the first stage of the reaction, an aspartate functioning as a general base (Asp X, Figure 7–6) extracts a proton from a water molecule, making it more nucleophilic. This resulting nucleophile then attacks the electrophilic carbonyl carbon of the peptide bond targeted for hydrolysis, forming a **tetrahedral transition state intermediate.** A second aspartate (Asp Y, Figure 7–6) then facilitates the decomposition of this tetrahedral intermediate by donating a proton to the amino group produced by rupture of the peptide bond. Two different active site aspartates thus can act simultaneously as a general base or as a general acid. This is possible because their immediate environment favors ionization of one but not the other.

CHYMOTRYPSIN & FRUCTOSE-2,6-BISPHOSPHATASE ILLUSTRATE COVALENT CATALYSIS

Chymotrypsin

While catalysis by aspartic proteases involves the direct hydrolytic attack of water on a peptide bond, catalysis

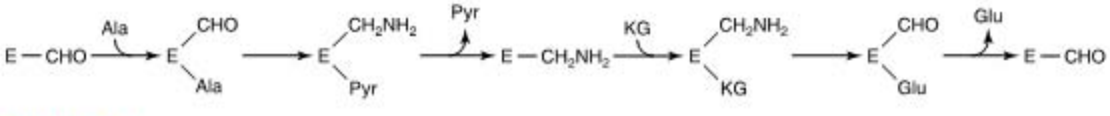


Figure 7–4. Ping-pong mechanism for transamination. E—CHO and E—CH_2NH_2 represent the enzyme-pyridoxal phosphate and enzyme-pyridoxamine complexes, respectively. (Ala, alanine; Pyr, pyruvate; KG, α-ketoglutarate; Glu, glutamate).

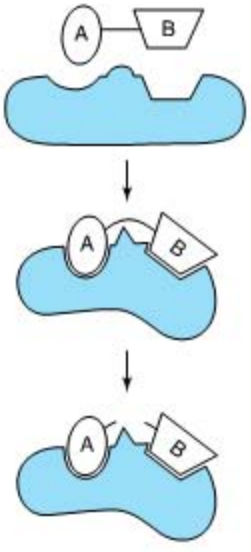


Figure 7–5. Two-dimensional representation of Koshland's induced fit model of the active site of a lyase. Binding of the substrate A—B induces conformational changes in the enzyme that aligns catalytic residues which participate in catalysis and strains the bond between A and B, facilitating its cleavage.

by the **serine protease** chymotrypsin involves prior formation of a covalent acyl enzyme intermediate. A highly reactive seryl residue, serine 195, participates in a charge-relay network with histidine 57 and aspartate 102. Far apart in primary structure, in the active site these residues are within bond-forming distance of one another. Aligned in the order Asp 102-His 57-Ser 195, they constitute a "charge-relay network" that functions as a "proton shuttle."

Binding of substrate initiates proton shifts that in effect transfer the hydroxyl proton of Ser 195 to Asp 102 (Figure 7–7). The enhanced nucleophilicity of the seryl oxygen facilitates its attack on the carbonyl carbon of the peptide bond of the substrate, forming a covalent acyl-enzyme intermediate. The hydrogen on Asp 102 then shuttles through His 57 to the amino group liberated when the peptide bond is cleaved. The portion of the original peptide with a free amino group then leaves the active site and is replaced by a water molecule. The charge-relay network now activates the water molecule by withdrawing a proton through His 57 to Asp 102. The resulting hydroxide ion attacks the acyl-enzyme in-

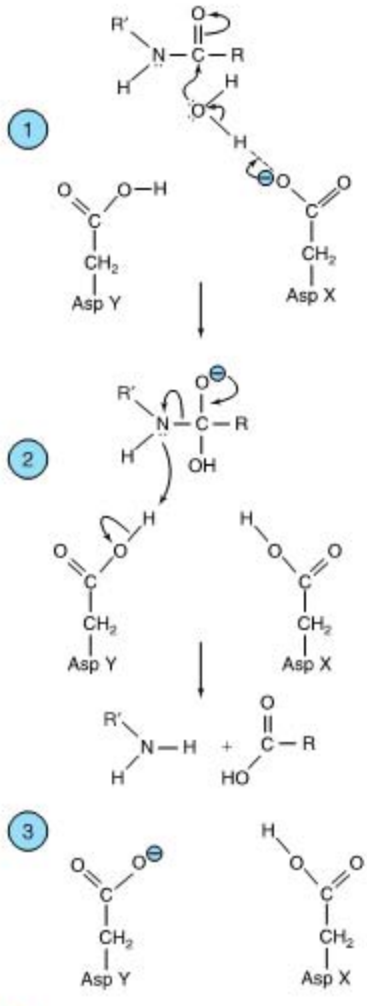


Figure 7–6. Mechanism for catalysis by an aspartic protease such as HIV protease. Curved arrows indicate directions of electron movement. ① Aspartate X acts as a base to activate a water molecule by abstracting a proton. ② The activated water molecule attacks the peptide bond, forming a transient tetrahedral intermediate. ③ Aspartate Y acts as an acid to facilitate breakdown of the tetrahedral intermediate and release of the split products by donating a proton to the newly formed amino group. Subsequent shuttling of the proton on Asp X to Asp Y restores the protease to its initial state.

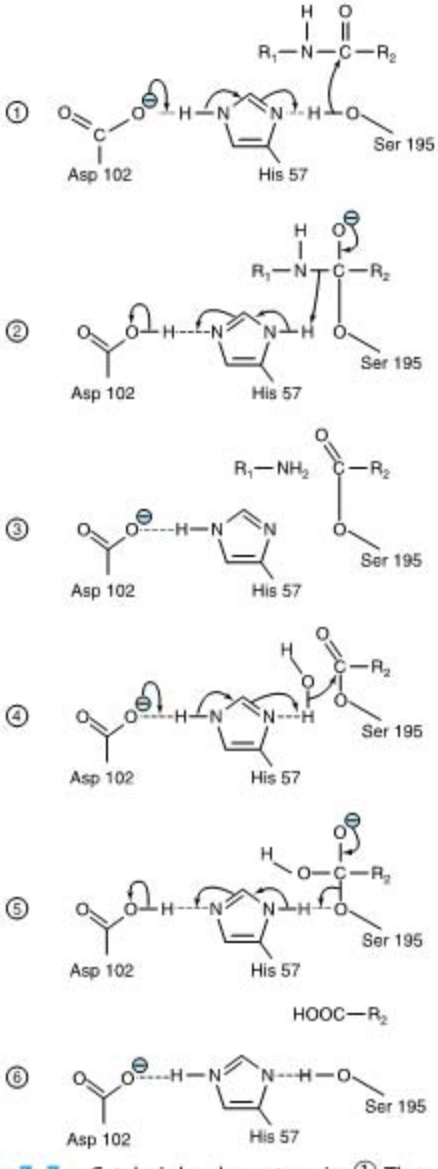


Figure 7–7. Catalysis by chymotrypsin. ① The charge-relay system removes a proton from Ser 195, making it a stronger nucleophile. ② Activated Ser 195 attacks the peptide bond, forming a transient tetrahedral intermediate. ③ Release of the amino terminal peptide is facilitated by donation of a proton to the newly formed amino group by His 57 of the charge-relay system, yielding an acyl-Ser 195 intermediate. ④ His 57 and Asp 102 collaborate to activate a water molecule, which attacks the acyl-Ser 195, forming a second tetrahedral intermediate. ⑤ The charge-relay system donates a proton to Ser 195, facilitating breakdown of tetrahedral intermediate to release the carboxyl terminal peptide ⑥.

termediate and a reverse proton shuttle returns a proton to Ser 195, restoring its original state. While modified during the process of catalysis, chymotrypsin emerges unchanged on completion of the reaction. Trypsin and elastase employ a similar catalytic mechanism, but the numbers of the residues in their Ser-His-Asp proton shuttles differ.

Fructose-2,6-Bisphosphatase

Fructose-2,6-bisphosphatase, a regulatory enzyme of gluconeogenesis (Chapter 19), catalyzes the hydrolytic release of the phosphate on carbon 2 of fructose 2,6-bisphosphate. Figure 7–8 illustrates the roles of seven active site residues. Catalysis involves a "catalytic triad" of one Glu and two His residues and a covalent phosphohistidyl intermediate.

CATALYTIC RESIDUES ARE HIGHLY CONSERVED

Members of an enzyme family such as the aspartic or serine proteases employ a similar mechanism to catalyze a common reaction type but act on different substrates. Enzyme families appear to arise through gene duplication events that create a second copy of the gene which encodes a particular enzyme. The proteins encoded by the two genes can then evolve independently to recognize different substrates—resulting, for example, in chymotrypsin, which cleaves peptide bonds on the carboxyl terminal side of large hydrophobic amino acids; and trypsin, which cleaves peptide bonds on the carboxyl terminal side of basic amino acids. The common ancestry of enzymes can be inferred from the presence of specific amino acids in the same position in each family member. These residues are said to be **conserved residues.** Proteins that share a large number of conserved residues are said to be **homologous** to one another. Table 7–1 illustrates the primary structural conservation of two components of the charge-relay network for several serine proteases. Among the most highly conserved residues are those that participate directly in catalysis.

ISOZYMES ARE DISTINCT ENZYME FORMS THAT CATALYZE THE SAME REACTION

Higher organisms often elaborate several physically distinct versions of a given enzyme, each of which catalyzes the same reaction. Like the members of other protein families, these protein catalysts or **isozymes** arise through gene duplication. Isozymes may exhibit subtle differences in properties such as sensitivity to

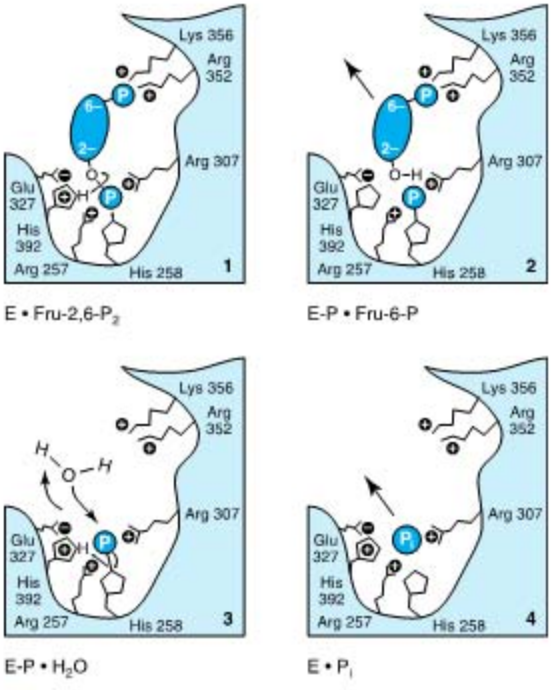


Figure 7–8. Catalysis by fructose-2,6-bisphosphatase. (1) Lys 356 and Arg 257, 307, and 352 stabilize the quadruple negative charge of the substrate by charge-charge interactions. Glu 327 stabilizes the positive charge on His 392. (2) The nucleophile His 392 attacks the C-2 phosphoryl group and transfers it to His 258, forming a phosphoryl-enzyme intermediate. Fructose 6-phosphate leaves the enzyme. (3) Nucleophilic attack by a water molecule, possibly assisted by Glu 327 acting as a base, forms inorganic phosphate. (4) Inorganic orthophosphate is released from Arg 257 and Arg 307. (Reproduced, with permission, from Pilkis SJ et al: 6-Phosphofructo-2-kinase/fructose-2,6-bisphosphatase: A metabolic signaling enzyme. Annu Rev Biochem 1995;64:799.)

particular regulatory factors (Chapter 9) or substrate affinity (eg, hexokinase and glucokinase) that adapt them to specific tissues or circumstances. Some isozymes may also enhance survival by providing a "back-up" copy of an essential enzyme.

THE CATALYTIC ACTIVITY OF ENZYMES FACILITATES THEIR DETECTION

The minute quantities of enzymes present in cells complicate determination of their presence and concentration. However, the ability to rapidly transform thousands of molecules of a specific substrate into products imbues each enzyme with the ability to reveal its presence. Assays of the catalytic activity of enzymes are frequently used in research and clinical laboratories. Under appropriate conditions (see Chapter 8), the rate of the catalytic reaction being monitored is proportionate to the amount of enzyme present, which allows its concentration to be inferred.

Enzyme-Linked Immunoassays

The sensitivity of enzyme assays can also be exploited to detect proteins that lack catalytic activity. **Enzyme-linked immunoassays** (ELISAs) use antibodies covalently linked to a "reporter enzyme" such as alkaline phosphatase or horseradish peroxidase, enzymes whose products are readily detected. When serum or other samples to be tested are placed in a plastic microtiter plate, the proteins adhere to the plastic surface and are immobilized. Any remaining absorbing areas of the well are then "blocked" by adding a nonantigenic protein such as bovine serum albumin. A solution of antibody covalently linked to a reporter enzyme is then added. The antibodies adhere to the immobilized antigen and these are themselves immobilized. Excess free antibody molecules are then removed by washing. The presence and quantity of bound antibody are then determined by adding the substrate for the reporter enzyme.

Table 7–1. Amino acid sequences in the neighborhood of the catalytic sites of several bovine proteases. Regions shown are those on either side of the catalytic site seryl (S) and histidyl (H) residues.

Enzyme	Sequence Around Serine Ⓢ															Sequence Around Histidine Ⓗ										
Trypsin	D	S	C	Q	D	G	Ⓢ	G	G	P	V	V	C	S	G	K	V	V	S	A	A	Ⓗ	C	Y	K	S	G
Chymotrypsin A	S	S	C	M	G	D	Ⓢ	G	G	P	L	V	C	K	K	N	V	V	T	A	A	Ⓗ	G	G	V	T	T
Chymotrypsin B	S	S	C	M	G	D	Ⓢ	G	G	P	L	V	C	Q	K	N	V	V	T	A	A	Ⓗ	C	G	V	T	T
Thrombin	D	A	C	E	G	D	Ⓢ	G	G	P	F	V	M	K	S	P	V	L	T	A	A	Ⓗ	C	L	L	Y	P

NAD(P)$^+$-Dependent Dehydrogenases Are Assayed Spectrophotometrically

The physicochemical properties of the reactants in an enzyme-catalyzed reaction dictate the options for the assay of enzyme activity. Spectrophotometric assays exploit the ability of a substrate or product to absorb light. The reduced coenzymes NADH and NADPH, written as NAD(P)H, absorb light at a wavelength of 340 nm, whereas their oxidized forms NAD(P)$^+$ do not (Figure 7–9). When NAD(P)$^+$ is reduced, the absorbance at 340 nm therefore increases in proportion to—and at a rate determined by—the quantity of NAD(P)H produced. Conversely, for a dehydrogenase that catalyzes the oxidation of NAD(P)H, a decrease in absorbance at 340 nm will be observed. In each case, the rate of change in optical density at 340 nm will be proportionate to the quantity of enzyme present.

Many Enzymes Are Assayed by Coupling to a Dehydrogenase

The assay of enzymes whose reactions are not accompanied by a change in absorbance or fluorescence is generally more difficult. In some instances, the product or remaining substrate can be transformed into a more readily detected compound. In other instances, the reaction product may have to be separated from unreacted substrate prior to measurement—a process facilitated by the use of radioactive substrates. An alternative strategy is to devise a synthetic substrate whose product absorbs light. For example, *p*-nitrophenyl phosphate is an artificial substrate for certain phosphatases and for chymotrypsin that does not absorb visible light. However, following hydrolysis, the resulting *p*-nitrophenylate anion absorbs light at 419 nm.

Another quite general approach is to employ a "coupled" assay (Figure 7–10). Typically, a dehydrogenase whose substrate is the product of the enzyme of interest is added in catalytic excess. The rate of appearance or disappearance of NAD(P)H then depends on the rate of the enzyme reaction to which the dehydrogenase has been coupled.

THE ANALYSIS OF CERTAIN ENZYMES AIDS DIAGNOSIS

Of the thousands of different enzymes present in the human body, those that fulfill functions indispensable to cell vitality are present throughout the body tissues. Other enzymes or isozymes are expressed only in specific cell types, during certain periods of development, or in response to specific physiologic or pathophysiologic changes. Analysis of the presence and distribution of enzymes and isozymes—whose expression is normally tissue-, time-, or circumstance-specific—often aids diagnosis.

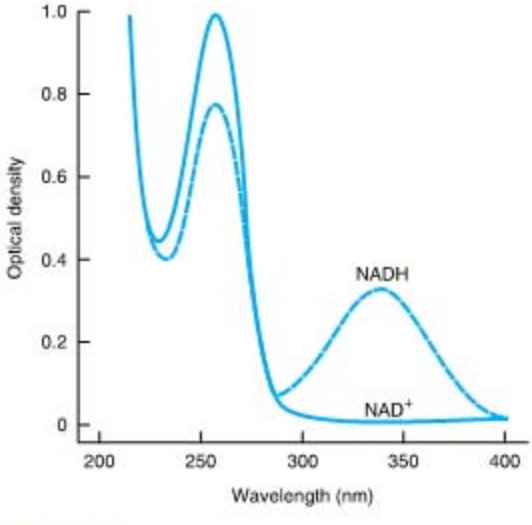


Figure 7–9. Absorption spectra of NAD$^+$ and NADH. Densities are for a 44 mg/L solution in a cell with a 1 cm light path. NADP$^+$ and NADPH have spectrums analogous to NAD$^+$ and NADH, respectively.

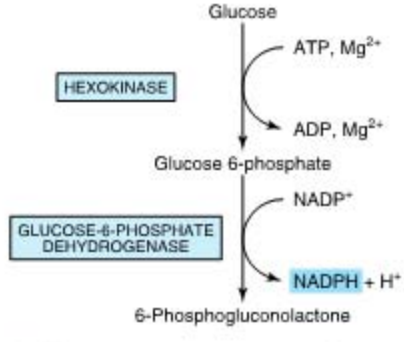


Figure 7–10. Coupled enzyme assay for hexokinase activity. The production of glucose 6-phosphate by hexokinase is coupled to the oxidation of this product by glucose-6-phosphate dehydrogenase in the presence of added enzyme and NADP$^+$. When an excess of glucose-6-phosphate dehydrogenase is present, the rate of formation of NADPH, which can be measured at 340 nm, is governed by the rate of formation of glucose 6-phosphate by hexokinase.

Nonfunctional Plasma Enzymes Aid Diagnosis & Prognosis

Certain enzymes, proenzymes, and their substrates are present at all times in the circulation of normal individuals and perform a physiologic function in the blood. Examples of these **functional plasma enzymes** include lipoprotein lipase, pseudocholinesterase, and the proenzymes of blood coagulation and blood clot dissolution (Chapters 9 and 51). The majority of these enzymes are synthesized in and secreted by the liver.

Plasma also contains numerous other enzymes that perform no known physiologic function in blood. These apparently **nonfunctional plasma enzymes** arise from the routine normal destruction of erythrocytes, leukocytes, and other cells. Tissue damage or necrosis resulting from injury or disease is generally accompanied by increases in the levels of several nonfunctional plasma enzymes. Table 7–2 lists several enzymes used in diagnostic enzymology.

Isozymes of Lactate Dehydrogenase Are Used to Detect Myocardial Infarctions

L-Lactate dehydrogenase is a tetrameric enzyme whose four subunits occur in two isoforms, designated H (for heart) and M (for muscle). The subunits can combine as shown below to yield catalytically active isozymes of L-lactate dehydrogenase:

Lactate Dehydrogenase Isozyme	Subunits
I_1	HHHH
I_2	HHHM
I_3	HHMM
I_4	HMMM
I_5	MMMM

Distinct genes whose expression is differentially regulated in various tissues encode the H and M subunits. Since heart expresses the H subunit almost exclusively, isozyme I_1 predominates in this tissue. By contrast, isozyme I_5 predominates in liver. Small quantities of lactate dehydrogenase are normally present in plasma. Following a myocardial infarction or in liver disease, the damaged tissues release characteristic lactate dehydrogenase isoforms into the blood. The resulting elevation in the levels of the I_1 or I_5 isozymes is detected by separating the different oligomers of lactate dehydrogenase by electrophoresis and assaying their catalytic activity (Figure 7–11).

Table 7–2. Principal serum enzymes used in clinical diagnosis. Many of the enzymes are not specific for the disease listed.

Serum Enzyme	Major Diagnostic Use
Aminotransferases	
Aspartate aminotransferase (AST, or SGOT)	Myocardial infarction
Alanine aminotransferase (ALT, or SGPT)	Viral hepatitis
Amylase	Acute pancreatitis
Ceruloplasmin	Hepatolenticular degeneration (Wilson's disease)
Creatine kinase	Muscle disorders and myocardial infarction
γ-Glutamyl transpeptidase	Various liver diseases
Lactate dehydrogenase (isozymes)	Myocardial infarction
Lipase	Acute pancreatitis
Phosphatase, acid	Metastatic carcinoma of the prostate
Phosphatase, alkaline (isozymes)	Various bone disorders, obstructive liver diseases

ENZYMES FACILITATE DIAGNOSIS OF GENETIC DISEASES

While many diseases have long been known to result from alterations in an individual's DNA, tools for the detection of genetic mutations have only recently become widely available. These techniques rely upon the catalytic efficiency and specificity of enzyme catalysts. For example, the **polymerase chain reaction** (PCR) relies upon the ability of enzymes to serve as catalytic amplifiers to analyze the DNA present in biologic and forensic samples. In the PCR technique, a thermostable DNA polymerase, directed by appropriate oligonucleotide primers, produces thousands of copies of a sample of DNA that was present initially at levels too low for direct detection.

The detection of **restriction fragment length polymorphisms** (RFLPs) facilitates prenatal detection of hereditary disorders such as sickle cell trait, beta-thalassemia, infant phenylketonuria, and Huntington's disease. Detection of RFLPs involves cleavage of double-stranded DNA by restriction endonucleases, which can detect subtle alterations in DNA that affect their recognized sites. Chapter 40 provides further details concerning the use of PCR and restriction enzymes for diagnosis.

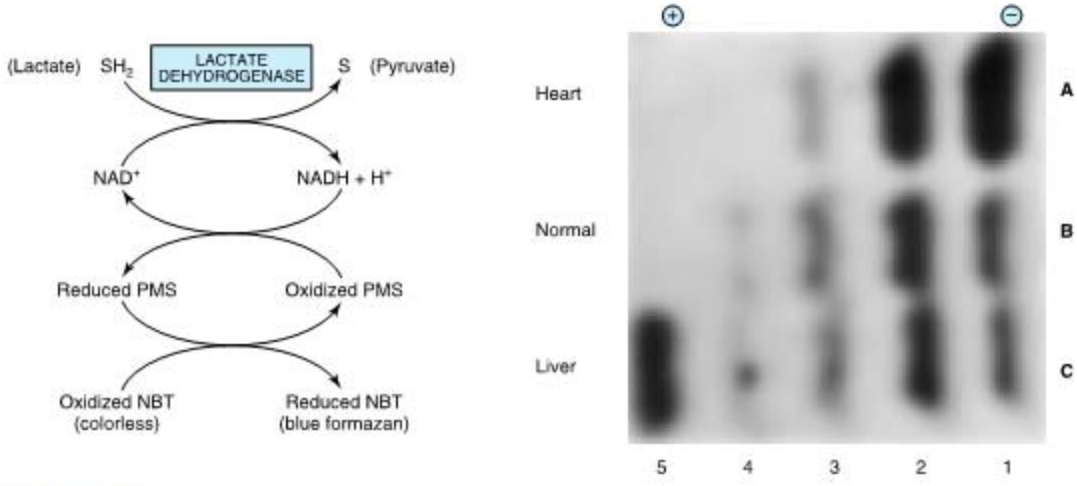


Figure 7–11. Normal and pathologic patterns of lactate dehydrogenase (LDH) isozymes in human serum. LDH isozymes of serum were separated by electrophoresis and visualized using the coupled reaction scheme shown on the left. (NBT, nitroblue tetrazolium; PMS, phenazine methylsulfate). At right is shown the stained electropherogram. Pattern A is serum from a patient with a myocardial infarct; B is normal serum; and C is serum from a patient with liver disease. Arabic numerals denote specific LDH isozymes.

RECOMBINANT DNA PROVIDES AN IMPORTANT TOOL FOR STUDYING ENZYMES

Recombinant DNA technology has emerged as an important asset in the study of enzymes. Highly purified samples of enzymes are necessary for the study of their structure and function. The isolation of an individual enzyme, particularly one present in low concentration, from among the thousands of proteins present in a cell can be extremely difficult. If the gene for the enzyme of interest has been cloned, it generally is possible to produce large quantities of its encoded protein in *Escherichia coli* or yeast. However, not all animal proteins can be expressed in active form in microbial cells, nor do microbes perform certain posttranslational processing tasks. For these reasons, a gene may be expressed in cultured animal cell systems employing the baculovirus expression vector to transform cultured insect cells. For more details concerning recombinant DNA techniques, see Chapter 40.

Recombinant Fusion Proteins Are Purified by Affinity Chromatography

Recombinant DNA technology can also be used to create modified proteins that are readily purified by affinity chromatography. The gene of interest is linked to an oligonucleotide sequence that encodes a carboxyl or amino terminal extension to the encoded protein. The resulting modified protein, termed a **fusion protein,** contains a domain tailored to interact with a specific affinity support. One popular approach is to attach an oligonucleotide that encodes six consecutive histidine residues. The expressed "His tag" protein binds to chromatographic supports that contain an immobilized divalent metal ion such as Ni^{2+}. Alternatively, the substrate-binding domain of glutathione S-transferase (GST) can serve as a "GST tag." Figure 7–12 illustrates the purification of a GST-fusion protein using an affinity support containing bound glutathione. Fusion proteins also often encode a cleavage site for a highly specific protease such as thrombin in the region that links the two portions of the protein. This permits removal of the added fusion domain following affinity purification.

Site-Directed Mutagenesis Provides Mechanistic Insights

Once the ability to express a protein from its cloned gene has been established, it is possible to employ **site-directed mutagenesis** to change specific aminoacyl residues by altering their codons. Used in combination with kinetic analyses and x-ray crystallography, this approach facilitates identification of the specific roles of given aminoacyl residues in substrate binding and catalysis. For example, the inference that a particular aminoacyl residue functions as a general acid can be tested by replacing it with an aminoacyl residue incapable of donating a proton.

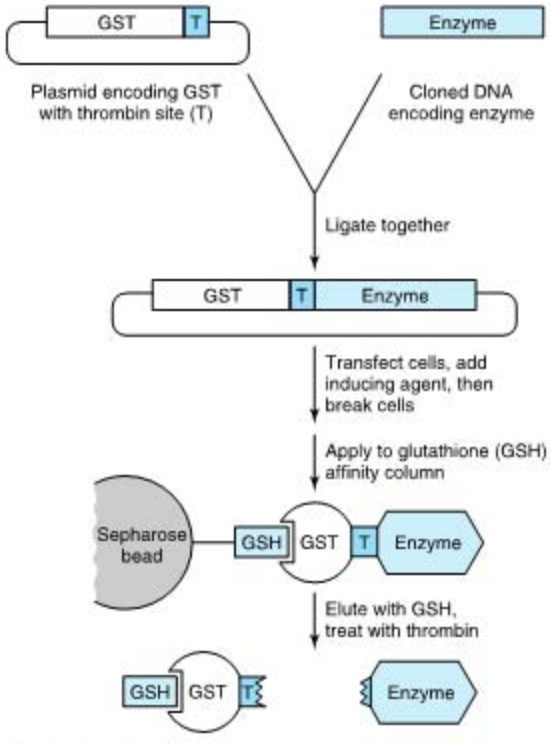


Figure 7–12. Use of glutathione S-transferase (GST) fusion proteins to purify recombinant proteins. (GSH, glutathione.)

SUMMARY

- Enzymes are highly effective and extremely specific catalysts.
- Organic and inorganic prosthetic groups, cofactors, and coenzymes play important roles in catalysis. Coenzymes, many of which are derivatives of B vitamins, serve as "shuttles."
- Catalytic mechanisms employed by enzymes include the introduction of strain, approximation of reactants, acid-base catalysis, and covalent catalysis.
- Aminoacyl residues that participate in catalysis are highly conserved among all classes of a given enzyme activity.
- Substrates and enzymes induce mutual conformational changes in one another that facilitate substrate recognition and catalysis.
- The catalytic activity of enzymes reveals their presence, facilitates their detection, and provides the basis for enzyme-linked immunoassays.
- Many enzymes can be assayed spectrophotometrically by coupling them to an $NAD(P)^+$-dependent dehydrogenase.
- Assay of plasma enzymes aids diagnosis and prognosis. For example, a myocardial infarction elevates serum levels of lactate dehydrogenase isozyme I_1.
- Restriction endonucleases facilitate diagnosis of genetic diseases by revealing restriction fragment length polymorphisms.
- Site-directed mutagenesis, used to change residues suspected of being important in catalysis or substrate binding, provides insights into the mechanisms of enzyme action.
- Recombinant fusion proteins such as His-tagged or GST fusion enzymes are readily purified by affinity chromatography.

REFERENCES

Conyers GB et al: Metal requirements of a diadenosine pyrophosphatase from *Bartonella bacilliformis*. Magnetic resonance and kinetic studies of the role of Mn^{2+}. Biochemistry 2000; 39:2347.

Fersht A: *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*. Freeman, 1999.

Suckling CJ: *Enzyme Chemistry*. Chapman & Hall, 1990.

Walsh CT: *Enzymatic Reaction Mechanisms*. Freeman, 1979.

Enzymes: Kinetics

8

Victor W. Rodwell, PhD, & Peter J. Kennelly, PhD

BIOMEDICAL IMPORTANCE

Enzyme kinetics is the field of biochemistry concerned with the quantitative measurement of the rates of enzyme-catalyzed reactions and the systematic study of factors that affect these rates. Kinetic analyses permit scientists to reconstruct the **number** and **order** of the individual steps by which enzymes transform substrates into products. The study of enzyme kinetics also represents the principal way to identify potential therapeutic agents that selectively enhance or inhibit the rates of specific enzyme-catalyzed processes. Together with site-directed mutagenesis and other techniques that probe protein structure, kinetic analysis can also reveal details of the catalytic mechanism. A complete, balanced set of enzyme activities is of fundamental importance for maintaining homeostasis. An understanding of enzyme kinetics thus is important for understanding how physiologic stresses such as anoxia, metabolic acidosis or alkalosis, toxins, and pharmacologic agents affect that balance.

CHEMICAL REACTIONS ARE DESCRIBED USING BALANCED EQUATIONS

A **balanced chemical equation** lists the initial chemical species (substrates) present and the new chemical species (products) formed for a particular chemical reaction, all in their correct proportions or **stoichiometry.** For example, balanced equation (1) below describes the reaction of one molecule each of substrates A and B to form one molecule each of products P and Q.

$$A + B \rightleftarrows P + Q \qquad (1)$$

The double arrows indicate reversibility, an intrinsic property of all chemical reactions. Thus, for reaction (1), if A and B can form P and Q, then P and Q can also form A and B. Designation of a particular reactant as a "substrate" or "product" is therefore somewhat arbitrary since the products for a reaction written in one direction are the substrates for the reverse reaction. The term "products" is, however, often used to designate the reactants whose formation is thermodynamically favored. Reactions for which thermodynamic factors strongly favor formation of the products to which the arrow points often are represented with a single arrow as if they were "irreversible":

$$A + B \rightarrow P + Q \qquad (2)$$

Unidirectional arrows are also used to describe reactions in living cells where the products of reaction (2) are immediately consumed by a subsequent enzyme-catalyzed reaction. The rapid removal of product P or Q therefore precludes occurrence of the reverse reaction, rendering equation (2) **functionally irreversible under physiologic conditions.**

CHANGES IN FREE ENERGY DETERMINE THE DIRECTION & EQUILIBRIUM STATE OF CHEMICAL REACTIONS

The Gibbs free energy change ΔG (also called either the free energy or Gibbs energy) describes both the *direction* in which a chemical reaction will tend to proceed and the concentrations of reactants and products that will be present at equilibrium. ΔG for a chemical reaction equals the sum of the free energies of formation of the reaction products ΔG_P minus the sum of the free energies of formation of the substrates ΔG_S. ΔG^0 denotes the change in free energy that accompanies transition from the standard state, one-molar concentrations of substrates and products, to equilibrium. A more useful biochemical term is $\Delta G^{0'}$, which defines ΔG^0 at a standard state of 10^{-7} M protons, pH 7.0 (Chapter 10). If the free energy of the products is lower than that of the substrates, the signs of ΔG^0 and $\Delta G^{0'}$ will be negative, indicating that the reaction as written is favored in the direction left to right. Such reactions are referred to as **spontaneous.** The **sign** and the **magnitude** of the free energy change determine how far the reaction will proceed. Equation (3)—

$$\Delta G^0 = -RT \ln K_{eq} \qquad (3)$$

—illustrates the relationship between the equilibrium constant K_{eq} and ΔG^0, where R is the gas constant (1.98 cal/mol/°K or 8.31 J/mol/°K) and T is the absolute temperature in degrees Kelvin. K_{eq} is equal to the product of the concentrations of the reaction products, each raised to the power of their stoichiometry, divided by the product of the substrates, each raised to the power of their stoichiometry.

For the reaction A + B → P + Q—

$$K_{eq} = \frac{[P][Q]}{[A][B]} \quad (4)$$

and for reaction (5)

$$A + A \rightleftarrows P \quad (5)$$

$$K_{eq} = \frac{[P]}{[A]^2} \quad (6)$$

—ΔG^0 may be calculated from equation (3) if the concentrations of substrates and products present at equilibrium are known. If ΔG^0 is a negative number, K_{eq} will be greater than unity and the concentration of products at equilibrium will exceed that of substrates. If ΔG^0 is positive, K_{eq} will be less than unity and the formation of substrates will be favored.

Notice that, since ΔG^0 is a function exclusively of the initial and final states of the reacting species, it can provide information only about the *direction* and *equilibrium state* of the reaction. $\Delta \mathbf{G^0}$ is independent of the mechanism of the reaction and therefore provides no information concerning ***rates*** of reactions. Consequently—and as explained below—although a reaction may have a large negative ΔG^0 or $\Delta G^{0'}$, it may nevertheless take place at a negligible rate.

THE RATES OF REACTIONS ARE DETERMINED BY THEIR ACTIVATION ENERGY

Reactions Proceed via Transition States

The concept of the **transition state** is fundamental to understanding the chemical and thermodynamic basis of catalysis. Equation (7) depicts a displacement reaction in which an entering group E displaces a leaving group L, attached initially to R.

$$E + R{-}L \rightleftarrows E{-}R + L \quad (7)$$

Midway through the displacement, the bond between R and L has weakened but has not yet been completely severed, and the new bond between E and R is as yet incompletely formed. This transient intermediate—in which neither free substrate nor product exists—is termed the **transition state, E···R···L.** Dotted lines represent the "partial" bonds that are undergoing formation and rupture.

Reaction (7) can be thought of as consisting of two "partial reactions," the first corresponding to the formation (F) and the second to the subsequent decay (D) of the transition state intermediate. As for all reactions, characteristic changes in free energy, ΔG_F, and ΔG_D are associated with each partial reaction.

$$E + R{-}L \rightleftarrows E{\cdots}R{\cdots}L \qquad \Delta G_F \quad (8)$$

$$E{\cdots}R{\cdots}L \rightleftarrows E{-}R + L \qquad \Delta G_D \quad (9)$$

$$E + R{-}L \rightleftarrows E{-}R + L \quad \Delta G = \Delta G_F + \Delta G_D \quad (8\text{-}10)$$

For the overall reaction (10), ΔG is the sum of ΔG_F and ΔG_D. As for any equation of two terms, it is not possible to infer from ΔG either the sign or the magnitude of ΔG_F or ΔG_D.

Many reactions involve multiple transition states, each with an associated change in free energy. For these reactions, the overall ΔG represents the sum of *all* of the free energy changes associated with the formation and decay of *all* of the transition states. **Therefore, it is not possible to infer from the overall ΔG the number or type of transition states through which the reaction proceeds.** Stated another way: overall thermodynamics tells us nothing about kinetics.

ΔG_F Defines the Activation Energy

Regardless of the sign or magnitude of ΔG, ΔG_F for the overwhelming majority of chemical reactions has a positive sign. The formation of transition state intermediates therefore requires surmounting of energy barriers. For this reason, ΔG_F is often termed the **activation energy,** E_{act}, the energy required to surmount a given energy barrier. The ease—and hence the frequency—with which this barrier is overcome is inversely related to E_{act}. The thermodynamic parameters that determine how *fast* a reaction proceeds thus are the ΔG_F values for formation of the transition states through which the reaction proceeds. For a simple reaction, where $\propto$ means "proportionate to,"

$$\text{Rate} \propto e^{\frac{-E_{act}}{RT}} \quad (11)$$

The activation energy for the reaction proceeding in the opposite direction to that drawn is equal to $-\Delta G_D$.

NUMEROUS FACTORS AFFECT THE REACTION RATE

The **kinetic theory**—also called the **collision theory**—of chemical kinetics states that for two molecules to react they must (1) approach within bond-forming distance of one another, or "collide"; and (2) must possess sufficient kinetic energy to overcome the energy barrier for reaching the transition state. It therefore follows

that anything which increases the *frequency* or *energy* of collision between substrates will increase the rate of the reaction in which they participate.

Temperature

Raising the temperature increases the kinetic energy of molecules. As illustrated in Figure 8–1, the total number of molecules whose kinetic energy exceeds the energy barrier E_{act} (vertical bar) for formation of products increases from low (A), through intermediate (B), to high (C) temperatures. Increasing the kinetic energy of molecules also increases their motion and therefore the frequency with which they collide. This combination of more frequent and more highly energetic and productive collisions increases the reaction rate.

Reactant Concentration

The frequency with which molecules collide is directly proportionate to their concentrations. For two different molecules A and B, the frequency with which they collide will double if the concentration of either A or B is doubled. If the concentrations of both A and B are doubled, the probability of collision will increase fourfold.

For a chemical reaction proceeding at constant temperature that involves one molecule each of A and B,

$$A + B \rightarrow P \tag{12}$$

the number of molecules that possess kinetic energy sufficient to overcome the activation energy barrier will be a constant. The number of collisions with sufficient energy to produce product P therefore will be directly proportionate to the number of collisions between A and B and thus to their molar concentrations, denoted by square brackets.

$$\text{Rate} \propto [A][B] \tag{13}$$

Similarly, for the reaction represented by

$$A + 2B \rightarrow P \tag{14}$$

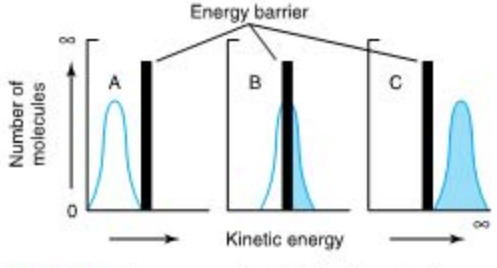


Figure 8–1. The energy barrier for chemical reactions.

which can also be written as

$$A + B + B \rightarrow P \tag{15}$$

the corresponding rate expression is

$$\text{Rate} \propto [A][B][B] \tag{16}$$

or

$$\text{Rate} \propto [A][B]^2 \tag{17}$$

For the general case when n molecules of A react with m molecules of B,

$$nA + mB \rightarrow P \tag{18}$$

the rate expression is

$$\text{Rate} \propto [A]^n[B]^m \tag{19}$$

Replacing the proportionality constant with an equal sign by introducing a proportionality or **rate constant k** characteristic of the reaction under study gives equations (20) and (21), in which the subscripts 1 and −1 refer to the rate constants for the forward and reverse reactions, respectively.

$$\text{Rate}_1 = k_1[A]^n[B]^m \tag{20}$$

$$\text{Rate}_{-1} = k_{-1}[P] \tag{21}$$

K_{eq} Is a Ratio of Rate Constants

While all chemical reactions are to some extent reversible, at equilibrium the *overall* concentrations of reactants and products remain constant. At equilibrium, the rate of conversion of substrates to products therefore equals the rate at which products are converted to substrates.

$$\text{Rate}_1 = \text{Rate}_{-1} \tag{22}$$

Therefore,

$$k_1[A]^n[B]^m = k_{-1}[P] \tag{23}$$

and

$$\frac{k_1}{k_{-1}} = \frac{[P]}{[A]^n[B]^m} \tag{24}$$

The ratio of k_1 to k_{-1} is termed the equilibrium constant, K_{eq}. The following important properties of a system at equilibrium must be kept in mind:

(1) The equilibrium constant is a ratio of the reaction *rate constants* (not the reaction *rates*).

(2) At equilibrium, the reaction *rates* (not the *rate constants*) of the forward and back reactions are equal.

(3) Equilibrium is a *dynamic* state. Although there is no *net* change in the concentration of substrates or products, individual substrate and product molecules are continually being interconverted.

(4) The numeric value of the equilibrium constant K_{eq} can be calculated either from the concentrations of substrates and products at equilibrium or from the ratio k_1/k_{-1}.

THE KINETICS OF ENZYMATIC CATALYSIS

Enzymes Lower the Activation Energy Barrier for a Reaction

All enzymes accelerate reaction rates by providing transition states with a lowered ΔG_F for formation of the transition states. However, they may differ in the way this is achieved. Where the mechanism or the sequence of chemical steps at the active site is essentially the same as those for the same reaction proceeding in the absence of a catalyst, **the environment of the active site lowers ΔG_F** by stabilizing the transition state intermediates. As discussed in Chapter 7, stabilization can involve (1) acid-base groups suitably positioned to transfer protons to or from the developing transition state intermediate, (2) suitably positioned charged groups or metal ions that stabilize developing charges, or (3) the imposition of steric strain on substrates so that their geometry approaches that of the transition state. HIV protease (Figure 7–6) illustrates catalysis by an enzyme that lowers the activation barrier by stabilizing a transition state intermediate.

Catalysis by enzymes that proceeds via a *unique* reaction mechanism typically occurs when the transition state intermediate forms a covalent bond with the enzyme **(covalent catalysis).** The catalytic mechanism of the serine protease chymotrypsin (Figure 7–7) illustrates how an enzyme utilizes covalent catalysis to provide a unique reaction pathway.

ENZYMES DO NOT AFFECT K_{eq}

Enzymes accelerate reaction rates by lowering the activation barrier ΔG_F. While they may undergo transient modification during the process of catalysis, enzymes emerge unchanged at the completion of the reaction. The presence of an enzyme therefore has no effect on ΔG^0 for the *overall* reaction, which is a function solely of the *initial and final states* of the reactants. Equation (25) shows the relationship between the equilibrium constant for a reaction and the standard free energy change for that reaction:

$$\Delta G^0 = -RT \ln K_{eq} \tag{25}$$

If we include the presence of the enzyme (E) in the calculation of the equilibrium constant for a reaction,

$$A + B + Enz \rightleftarrows P + Q + Enz \tag{26}$$

the expression for the equilibrium constant,

$$K_{eq} = \frac{[P][Q][Enz]}{[A][B][Enz]} \tag{27}$$

reduces to one identical to that for the reaction in the absence of the enzyme:

$$K_{eq} = \frac{[P][Q]}{[A][B]} \tag{28}$$

Enzymes therefore have no effect on K_{eq}.

MULTIPLE FACTORS AFFECT THE RATES OF ENZYME-CATALYZED REACTIONS

Temperature

Raising the temperature increases the rate of both uncatalyzed and enzyme-catalyzed reactions by increasing the kinetic energy and the collision frequency of the reacting molecules. However, heat energy can also increase the kinetic energy of the enzyme to a point that exceeds the energy barrier for disrupting the noncovalent interactions that maintain the enzyme's three-dimensional structure. The polypeptide chain then begins to unfold, or **denature,** with an accompanying rapid loss of catalytic activity. The temperature range over which an enzyme maintains a stable, catalytically competent conformation depends upon—and typically moderately exceeds—the normal temperature of the cells in which it resides. Enzymes from humans generally exhibit stability at temperatures up to 45–55 °C. By contrast, enzymes from the thermophilic microorganisms that reside in volcanic hot springs or undersea hydrothermal vents may be stable up to or above 100 °C.

The **Q_{10}, or temperature coefficient,** is the factor by which the rate of a biologic process increases for a 10 °C increase in temperature. For the temperatures over which enzymes are stable, the rates of most biologic processes typically double for a 10 °C rise in temperature (Q_{10} = 2). Changes in the rates of enzyme-catalyzed reactions that accompany a rise or fall in body temperature constitute a prominent survival feature for "cold-blooded" life forms such as lizards or fish, whose body temperatures are dictated by the external environment. However, for mammals and other homeothermic organisms, changes in enzyme reaction rates with temperature assume physiologic importance only in circumstances such as fever or hypothermia.

Hydrogen Ion Concentration

The rate of almost all enzyme-catalyzed reactions exhibits a significant dependence on hydrogen ion concentration. Most intracellular enzymes exhibit optimal activity at pH values between 5 and 9. The relationship of activity to hydrogen ion concentration (Figure 8–2) reflects the balance between enzyme denaturation at high or low pH and effects on the charged state of the enzyme, the substrates, or both. For enzymes whose mechanism involves acid-base catalysis, the residues involved must be in the appropriate state of protonation for the reaction to proceed. The binding and recognition of substrate molecules with dissociable groups also typically involves the formation of salt bridges with the enzyme. The most common charged groups are the negative carboxylate groups and the positively charged groups of protonated amines. Gain or loss of critical charged groups thus will adversely affect substrate binding and thus will retard or abolish catalysis.

ASSAYS OF ENZYME-CATALYZED REACTIONS TYPICALLY MEASURE THE INITIAL VELOCITY

Most measurements of the rates of enzyme-catalyzed reactions employ relatively short time periods, conditions that approximate **initial rate conditions.** Under these conditions, only traces of product accumulate, hence the rate of the reverse reaction is negligible. The **initial velocity (v_i)** of the reaction thus is essentially that of the rate of the forward reaction. Assays of enzyme activity almost always employ a large (10^5–10^7) molar excess of substrate over enzyme. Under these conditions, v_i is proportionate to the concentration of enzyme. Measuring the initial velocity therefore permits one to estimate the quantity of enzyme present in a biologic sample.

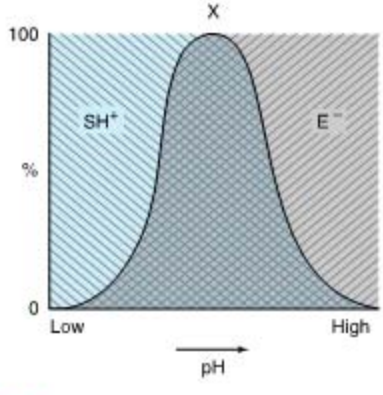


Figure 8–2. Effect of pH on enzyme activity. Consider, for example, a negatively charged enzyme (EH^-) that binds a positively charged substrate (SH^+). Shown is the proportion (%) of SH^+ [\\\] and of EH^- [///] as a function of pH. Only in the cross-hatched area do both the enzyme and the substrate bear an appropriate charge.

SUBSTRATE CONCENTRATION AFFECTS REACTION RATE

In what follows, enzyme reactions are treated as if they had only a single substrate and a single product. While most enzymes have more than one substrate, the principles discussed below apply with equal validity to enzymes with multiple substrates.

For a typical enzyme, as substrate concentration is increased, v_i increases until it reaches a maximum value V_{max} (Figure 8–3). When further increases in substrate concentration do not further increase v_i, the enzyme is said to be "saturated" with substrate. Note that the shape of the curve that relates activity to substrate concentration (Figure 8–3) is hyperbolic. At any given instant, only substrate molecules that are combined with the enzyme as an ES complex can be transformed into product. Second, the equilibrium constant for the formation of the enzyme-substrate complex is not infinitely large. Therefore, even when the substrate is present in excess (points A and B of Figure 8–4), only a fraction of the enzyme may be present as an ES complex. At points A or B, increasing or decreasing [S] therefore will increase or decrease the number of ES complexes with a corresponding change in v_i. At point C (Figure 8–4), essentially all the enzyme is present as the ES complex. Since no free enzyme remains available for forming ES, further increases in [S] cannot increase the rate of the reaction. Under these saturating conditions, v_i depends solely on—and thus is limited by—the rapidity with which free enzyme is released to combine with more substrate.

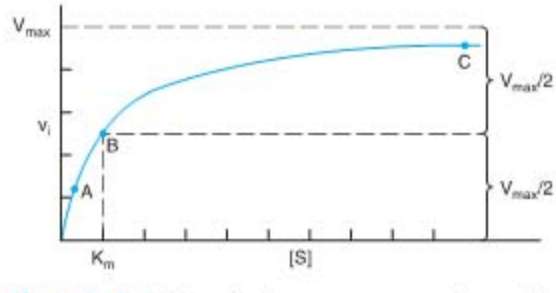


Figure 8–3. Effect of substrate concentration on the initial velocity of an enzyme-catalyzed reaction.

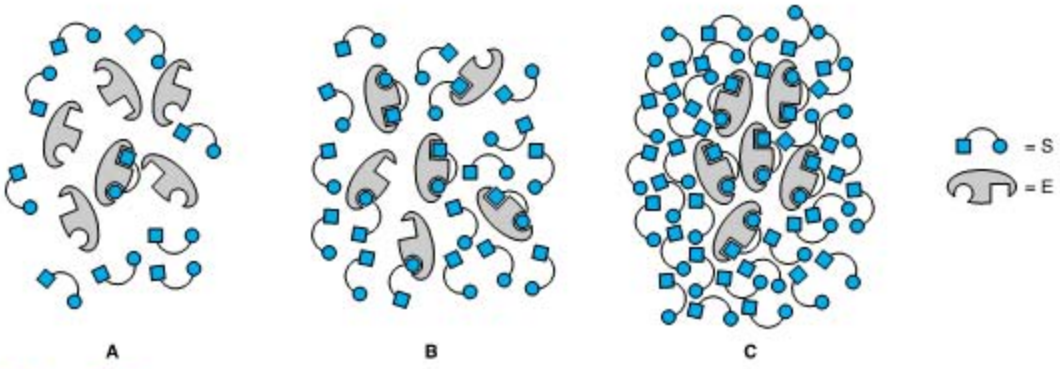


Figure 8–4. Representation of an enzyme at low **(A),** at high **(C),** and at a substrate concentration equal to K_m **(B).** Points A, B, and C correspond to those points in Figure 8–3.

THE MICHAELIS-MENTEN & HILL EQUATIONS MODEL THE EFFECTS OF SUBSTRATE CONCENTRATION

The Michaelis-Menten Equation

The Michaelis-Menten equation (29) illustrates in mathematical terms the relationship between initial reaction velocity v_i and substrate concentration [S], shown graphically in Figure 8–3.

$$v_i = \frac{V_{max}[S]}{K_m + [S]} \qquad (29)$$

The Michaelis constant K_m is the substrate concentration at which v_i is half the maximal velocity ($V_{max}/2$) attainable at a particular concentration of enzyme. K_m thus has the dimensions of substrate concentration. The dependence of initial reaction velocity on [S] and K_m may be illustrated by evaluating the Michaelis-Menten equation under three conditions.

(1) When [S] is much less than K_m (point A in Figures 8–3 and 8–4), the term K_m + [S] is essentially equal to K_m. Replacing K_m + [S] with K_m reduces equation (29) to

$$v_i = \frac{V_{max}[S]}{K_m + [S]} \qquad v_i = \frac{V_{max}[S]}{K_m} = \left(\frac{V_{max}}{K_m}\right)[S] \qquad (30)$$

where ≃ means "approximately equal to." Since V_{max} and K_m are both constants, their ratio is a constant. In other words, when [S] is considerably below K_m, $v_i \propto k[S]$. The initial reaction velocity therefore is directly proportionate to [S].

(2) When [S] is much greater than K_m (point C in Figures 8–3 and 8–4), the term K_m + [S] is essentially equal to [S]. Replacing K_m + [S] with [S] reduces equation (29) to

$$v_i = \frac{V_{max}[S]}{K_m + [S]} \qquad v_i \simeq \frac{V_{max}[S]}{[S]} \simeq V_{max} \qquad (31)$$

Thus, when [S] greatly exceeds K_m, the reaction velocity is maximal (V_{max}) and unaffected by further increases in substrate concentration.

(3) When [S] = K_m (point B in Figures 8–3 and 8–4).

$$v_i = \frac{V_{max}[S]}{K_m + [S]} = \frac{V_{max}[S]}{2[S]} = \frac{V_{max}}{2} \qquad (32)$$

Equation (32) states that when [S] equals K_m, the initial velocity is half-maximal. Equation (32) also reveals that K_m is—and may be determined experimentally from—the substrate concentration at which the initial velocity is half-maximal.

A Linear Form of the Michaelis-Menten Equation Is Used to Determine K_m & V_{max}

The direct measurement of the numeric value of V_{max} and therefore the calculation of K_m often requires impractically high concentrations of substrate to achieve saturating conditions. A linear form of the Michaelis-Menten equation circumvents this difficulty and permits V_{max} and K_m to be extrapolated from initial velocity data obtained at less than saturating concentrations of substrate. Starting with equation (29),

$$v_i = \frac{V_{max}[S]}{K_m + [S]} \qquad (29)$$

invert

$$\frac{1}{v_i} = \frac{K_m + [S]}{V_{max}[S]} \quad (33)$$

factor

$$\frac{1}{v_i} = \frac{K_m}{V_{max}[S]} + \frac{[S]}{V_{max}[S]} \quad (34)$$

and simplify

$$\frac{1}{v_i} = \left(\frac{K_m}{V_{max}}\right)\frac{1}{[S]} + \frac{1}{V_{max}} \quad (35)$$

Equation (35) is the equation for a straight line, $y = ax + b$, where $y = 1/v_i$ and $x = 1/[S]$. A plot of $1/v_i$ as y as a function of $1/[S]$ as x therefore gives a straight line whose y intercept is $1/V_{max}$ and whose slope is K_m/V_{max}. Such a plot is called a **double reciprocal** or **Lineweaver-Burk plot** (Figure 8–5). Setting the y term of equation (36) equal to zero and solving for x reveals that the x intercept is $-1/K_m$.

$$0 = ax + b;\ \text{therefore, } x = \frac{-b}{a} = \frac{-1}{K_m} \quad (36)$$

K_m is thus most easily calculated from the x intercept.

K_m May Approximate a Binding Constant

The affinity of an enzyme for its substrate is the inverse of the dissociation constant K_d for dissociation of the enzyme substrate complex ES.

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \quad (37)$$

$$K_d = \frac{k_{-1}}{k_1} \quad (38)$$

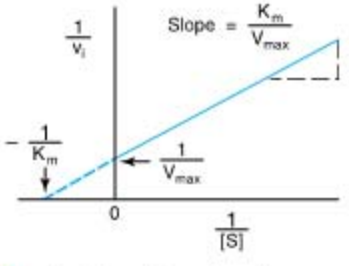


Figure 8–5. Double reciprocal or Lineweaver-Burk plot of $1/v_i$ versus $1/[S]$ used to evaluate K_m and V_{max}.

Stated another way, the smaller the tendency of the enzyme and its substrate to *dissociate,* the *greater* the affinity of the enzyme for its substrate. While the Michaelis constant K_m often approximates the dissociation constant K_d, this is by no means always the case. For a typical enzyme-catalyzed reaction,

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \xrightarrow{k_2} E + P \quad (39)$$

the value of [S] that gives $v_i = V_{max}/2$ is

$$[S] = \frac{k_{-1} + k_2}{k_1} = K_m \quad (40)$$

When $k_{-1} \gg k_2$, then

$$k_{-1} + k_2 = k_{-1} \quad (41)$$

and

$$[S] = \frac{k_1}{k_{-1}} = K_d \quad (42)$$

Hence, $1/K_m$ only approximates $1/K_d$ under conditions where the association and dissociation of the ES complex is rapid relative to the rate-limiting step in catalysis. For the many enzyme-catalyzed reactions for which $k_{-1} + k_2$ is *not* approximately equal to k_{-1}, $1/K_m$ will underestimate $1/K_d$.

The Hill Equation Describes the Behavior of Enzymes That Exhibit Cooperative Binding of Substrate

While most enzymes display the simple **saturation kinetics** depicted in Figure 8–3 and are adequately described by the Michaelis-Menten expression, some enzymes bind their substrates in a cooperative fashion analogous to the binding of oxygen by hemoglobin (Chapter 6). Cooperative behavior may be encountered for multimeric enzymes that bind substrate at multiple sites. For enzymes that display positive cooperativity in binding substrate, the shape of the curve that relates changes in v_i to changes in [S] is sigmoidal (Figure 8–6). Neither the Michaelis-Menten expression nor its derived double-reciprocal plots can be used to evaluate cooperative saturation kinetics. Enzymologists therefore employ a graphic representation of the **Hill equation** originally derived to describe the cooperative binding of O_2 by hemoglobin. Equation (43) represents the Hill equation arranged in a form that predicts a straight line, where k′ is a complex constant.

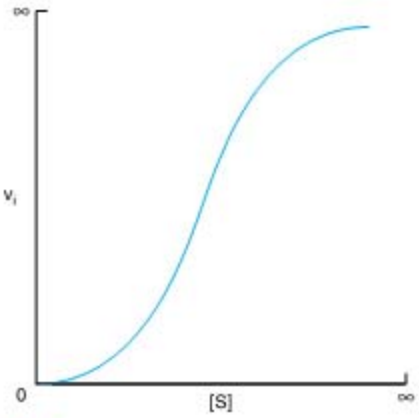


Figure 8–6. Representation of sigmoid substrate saturation kinetics.

$$\frac{\log v_1}{V_{max} - v_1} = n \log[S] - \log k' \qquad (43)$$

Equation (43) states that when [S] is low relative to k′, the initial reaction velocity increases as the nth power of [S].

A graph of log $v_i/(V_{max} - v_i)$ versus log[S] gives a straight line (Figure 8–7), where the slope of the line **n** is the **Hill coefficient,** an empirical parameter whose value is a function of the number, kind, and strength of the interactions of the multiple substrate-binding sites on the enzyme. When n = 1, all binding sites behave independently, and simple Michaelis-Menten kinetic behavior is observed. If n is greater than 1, the enzyme is said to exhibit positive cooperativity. Binding of the first substrate molecule then enhances the affinity of the enzyme for binding additional substrate. The greater the value for n, the higher the degree of cooperativity and the more sigmoidal will be the plot of v_i versus [S]. A perpendicular dropped from the point where the *y* term log $v_i/(V_{max} - v_i)$ is zero intersects the *x* axis at a substrate concentration termed $\mathbf{S_{50}}$, the substrate concentration that results in half-maximal velocity. S_{50} thus is analogous to the P_{50} for oxygen binding to hemoglobin (Chapter 6).

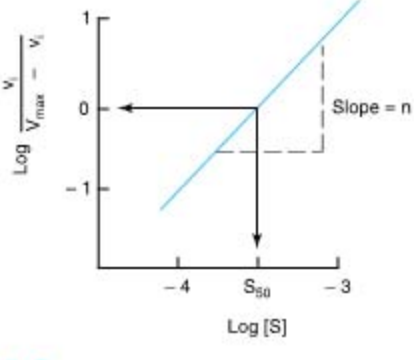


Figure 8–7. A graphic representation of a linear form of the Hill equation is used to evaluate S_{50}, the substrate concentration that produces half-maximal velocity, and the degree of cooperativity **n.**

KINETIC ANALYSIS DISTINGUISHES COMPETITIVE FROM NONCOMPETITIVE INHIBITION

Inhibitors of the catalytic activities of enzymes provide both pharmacologic agents and research tools for study of the mechanism of enzyme action. Inhibitors can be classified based upon their site of action on the enzyme, on whether or not they chemically modify the enzyme, or on the kinetic parameters they influence. Kinetically, we distinguish two classes of inhibitors based upon whether raising the substrate concentration does or does not overcome the inhibition.

Competitive Inhibitors Typically Resemble Substrates

The effects of competitive inhibitors can be overcome by raising the concentration of the substrate. Most frequently, in competitive inhibition the inhibitor, **I,** binds to the substrate-binding portion of the active site and blocks access by the substrate. The structures of most classic competitive inhibitors therefore tend to resemble the structures of a substrate and thus are termed **substrate analogs.** Inhibition of the enzyme succinate dehydrogenase by malonate illustrates competitive inhibition by a substrate analog. Succinate dehydrogenase catalyzes the removal of one hydrogen atom from each of the two methylene carbons of succinate (Figure 8–8). Both succinate and its structural analog malonate ($^-OOC—CH_2—COO^-$) can bind to the active site of succinate dehydrogenase, forming an ES or an EI complex, respectively. However, since malonate contains

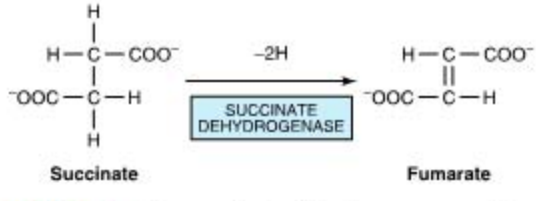


Figure 8–8. The succinate dehydrogenase reaction.

only one methylene carbon, it cannot undergo dehydrogenation. The formation and dissociation of the EI complex is a dynamic process described by

$$\text{EnzI} \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} \text{Enz} + \text{I} \qquad (44)$$

for which the equilibrium constant K_i is

$$K_1 = \frac{[\text{Enz}][\text{I}]}{[\text{EnzI}]} = \frac{k_1}{k_{-1}} \qquad (45)$$

In effect, **a competitive inhibitor acts by decreasing the number of free enzyme molecules available to bind substrate, ie, to form ES, and thus eventually to form product,** as described below:

(46)

A competitive inhibitor and substrate exert reciprocal effects on the concentration of the EI and ES complexes. Since binding substrate removes free enzyme available to combine with inhibitor, increasing the [S] decreases the concentration of the EI complex and raises the reaction velocity. The extent to which [S] must be increased to completely overcome the inhibition depends upon the concentration of inhibitor present, its affinity for the enzyme K_i, and the K_m of the enzyme for its substrate.

Double Reciprocal Plots Facilitate the Evaluation of Inhibitors

Double reciprocal plots distinguish between competitive and noncompetitive inhibitors and simplify evaluation of inhibition constants K_i. v_i is determined at several substrate concentrations both in the presence and in the absence of inhibitor. For classic competitive inhibition, the lines that connect the experimental data points meet at the *y* axis (Figure 8–9). Since the *y* intercept is equal to $1/V_{max}$, this pattern indicates that **when 1/[S] approaches 0, v_i is independent of the presence of inhibitor.** Note, however, that the intercept on the *x* axis does vary with *inhibitor* concentration—and that since $-1/K_m'$ is smaller than $1/K_m$, K_m' (the "apparent K_m") becomes larger in the presence of increasing concentrations of inhibitor. Thus, **a competitive inhibitor has no effect on V_{max} but raises K'_m, the apparent K_m for the substrate.**

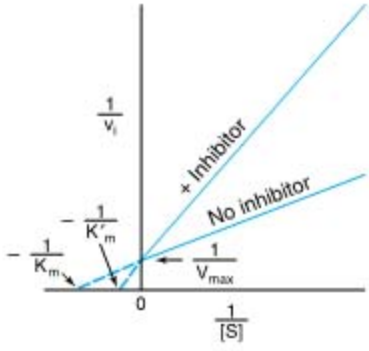


Figure 8–9. Lineweaver-Burk plot of competitive inhibition. Note the complete relief of inhibition at high [S] (ie, low 1/[S]).

For simple competitive inhibition, the intercept on the *x* axis is

$$x = \frac{-1}{K_m}\left(1 + \frac{[\text{I}]}{K_i}\right) \qquad (47)$$

Once K_m has been determined in the absence of inhibitor, K_i can be calculated from equation (47). K_i values are used to compare different inhibitors of the same enzyme. The lower the value for K_i, the more effective the inhibitor. For example, the statin drugs that act as competitive inhibitors of HMG-CoA reductase (Chapter 26) have K_i values several orders of magnitude lower than the K_m for the substrate HMG-CoA.

Simple Noncompetitive Inhibitors Lower V_{max} but Do Not Affect K_m

In noncompetitive inhibition, binding of the inhibitor does not affect binding of substrate. Formation of both EI and EIS complexes is therefore possible. However, while the enzyme-inhibitor complex can still bind substrate, its efficiency at transforming substrate to product, reflected by V_{max}, is decreased. Noncompetitive inhibitors bind enzymes at sites distinct from the substrate-binding site and generally bear little or no structural resemblance to the substrate.

For simple noncompetitive inhibition, E and EI possess identical affinity for substrate, and the EIS complex generates product at a negligible rate (Figure 8–10). More complex noncompetitive inhibition occurs when binding of the inhibitor *does* affect the apparent affinity of the enzyme for substrate, causing the lines to intercept in either the third or fourth quadrants of a double reciprocal plot (not shown).

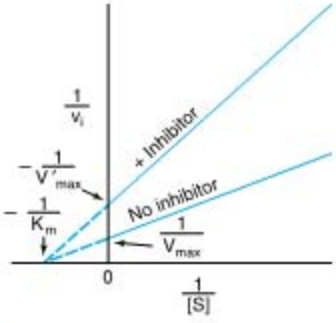


Figure 8–10. Lineweaver-Burk plot for simple noncompetitive inhibition.

Irreversible Inhibitors "Poison" Enzymes

In the above examples, the inhibitors form a dissociable, dynamic complex with the enzyme. Fully active enzyme can therefore be recovered simply by removing the inhibitor from the surrounding medium. However, a variety of other inhibitors act irreversibly by chemically modifying the enzyme. These modifications generally involve making or breaking covalent bonds with aminoacyl residues essential for substrate binding, catalysis, or maintenance of the enzyme's functional conformation. Since these covalent changes are relatively stable, an enzyme that has been "poisoned" by an irreversible inhibitor remains inhibited even after removal of the remaining inhibitor from the surrounding medium.

MOST ENZYME-CATALYZED REACTIONS INVOLVE TWO OR MORE SUBSTRATES

While many enzymes have a single substrate, many others have two—and sometimes more than two—substrates and products. The fundamental principles discussed above, while illustrated for single-substrate enzymes, apply also to multisubstrate enzymes. The mathematical expressions used to evaluate multisubstrate reactions are, however, complex. While detailed kinetic analysis of multisubstrate reactions exceeds the scope of this chapter, two-substrate, two-product reactions (termed "Bi-Bi" reactions) are considered below.

Sequential or Single Displacement Reactions

In **sequential reactions,** both substrates must combine with the enzyme to form a ternary complex before catalysis can proceed (Figure 8–11, top). Sequential reactions are sometimes referred to as single displacement

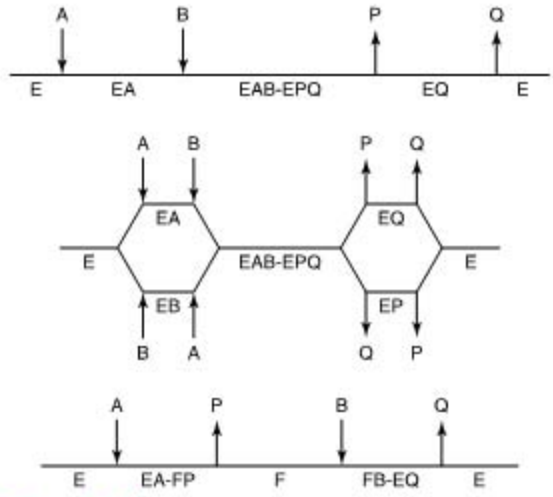


Figure 8–11. Representations of three classes of Bi-Bi reaction mechanisms. Horizontal lines represent the enzyme. Arrows indicate the addition of substrates and departure of products. **Top:** An ordered Bi-Bi reaction, characteristic of many NAD(P)H-dependent oxidoreductases. **Center:** A random Bi-Bi reaction, characteristic of many kinases and some dehydrogenases. **Bottom:** A ping-pong reaction, characteristic of aminotransferases and **serine proteases.**

reactions because the group undergoing transfer is usually passed directly, in a single step, from one substrate to the other. Sequential Bi-Bi reactions can be further distinguished based on whether the two substrates add in a **random** or in a **compulsory** order. For random-order reactions, either substrate A or substrate B may combine first with the enzyme to form an EA or an EB complex (Figure 8–11, center). For compulsory-order reactions, A must first combine with E before B can combine with the EA complex. One explanation for a compulsory-order mechanism is that the addition of A induces a conformational change in the enzyme that aligns residues which recognize and bind B.

Ping-Pong Reactions

The term **"ping-pong"** applies to mechanisms in which one or more products are released from the enzyme before all the substrates have been added. Ping-pong reactions involve covalent catalysis and a transient, modified form of the enzyme (Figure 7–4). Ping-pong Bi-Bi reactions are **double displacement reactions.** The group undergoing transfer is first displaced from substrate A by the enzyme to form product

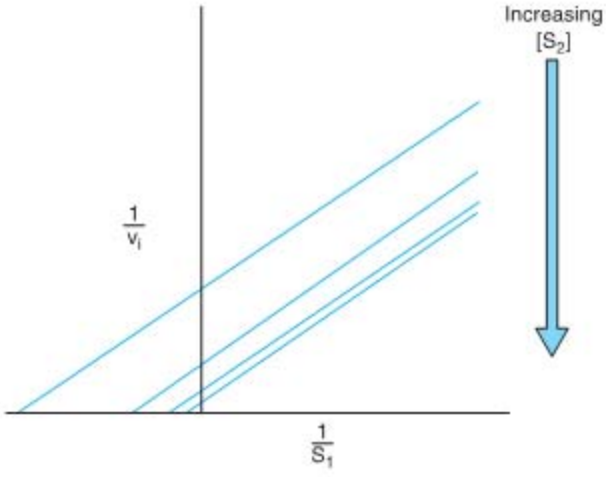


Figure 8–12. Lineweaver-Burk plot for a two-substrate ping-pong reaction. An increase in concentration of one substrate (S_1) while that of the other substrate (S_2) is maintained constant changes both the *x* and *y* intercepts, but not the slope.

P and a modified form of the enzyme (F). The subsequent group transfer from F to the second substrate B, forming product Q and regenerating E, constitutes the second displacement (Figure 8–11, bottom).

Most Bi-Bi Reactions Conform to Michaelis-Menten Kinetics

Most Bi-Bi reactions conform to a somewhat more complex form of Michaelis-Menten kinetics in which V_{max} refers to the reaction rate attained when both substrates are present at saturating levels. Each substrate has its own characteristic K_m value which corresponds to the concentration that yields half-maximal velocity when the second substrate is present at saturating levels. As for single-substrate reactions, double-reciprocal plots can be used to determine V_{max} and K_m. v_i is measured as a function of the concentration of one substrate (the variable substrate) while the concentration of the other substrate (the fixed substrate) is maintained constant. If the lines obtained for several fixed-substrate concentrations are plotted on the same graph, it is possible to distinguish between a ping-pong enzyme, which yields parallel lines, and a sequential mechanism, which yields a pattern of intersecting lines (Figure 8–12).

Product inhibition studies are used to complement kinetic analyses and to distinguish between ordered and random Bi-Bi reactions. For example, in a random-order Bi-Bi reaction, each product will be a competitive inhibitor regardless of which substrate is designated the variable substrate. However, for a sequential mechanism (Figure 8–11, bottom), only product Q will give the pattern indicative of competitive inhibition when A is the variable substrate, while only product P will produce this pattern with B as the variable substrate. The other combinations of product inhibitor and variable substrate will produce forms of complex noncompetitive inhibition.

SUMMARY

- The study of enzyme kinetics—the factors that affect the rates of enzyme-catalyzed reactions—reveals the individual steps by which enzymes transform substrates into products.
- ΔG, the overall change in free energy for a reaction, is independent of reaction mechanism and provides no information concerning *rates* of reactions.
- Enzymes do not affect K_{eq}. K_{eq}, a ratio of reaction *rate constants,* may be calculated from the concentrations of substrates and products at equilibrium or from the ratio k_1/k_{-1}.
- Reactions proceed via transition states in which $\Delta G_{\ddagger}$ is the activation energy. Temperature, hydrogen ion concentration, enzyme concentration, substrate concentration, and inhibitors all affect the rates of enzyme-catalyzed reactions.
- A measurement of the rate of an enzyme-catalyzed reaction generally employs initial rate conditions, for which the essential absence of product precludes the reverse reaction.
- A linear form of the Michaelis-Menten equation simplifies determination of K_m and V_{max}.
- A linear form of the Hill equation is used to evaluate the cooperative substrate-binding kinetics exhibited by some multimeric enzymes. The slope **n,** the Hill coefficient, reflects the number, nature, and strength of the interactions of the substrate-binding sites. A

value of **n** greater than 1 indicates positive cooperativity.

- The effects of competitive inhibitors, which typically resemble substrates, are overcome by raising the concentration of the substrate. Noncompetitive inhibitors lower V_{max} but do not affect K_m.
- Substrates may add in a random order (either substrate may combine first with the enzyme) or in a compulsory order (substrate A must bind before substrate B).
- In ping-pong reactions, one or more products are released from the enzyme before all the substrates have added.

REFERENCES

Fersht A: *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding.* Freeman, 1999.

Schultz AR: *Enzyme Kinetics: From Diastase to Multi-enzyme Systems.* Cambridge Univ Press, 1994.

Segel IH: *Enzyme Kinetics.* Wiley Interscience, 1975.

Enzymes: Regulation of Activities

9

Victor W. Rodwell, PhD, & Peter J. Kennelly, PhD

BIOMEDICAL IMPORTANCE

The 19th-century physiologist Claude Bernard enunciated the conceptual basis for metabolic regulation. He observed that living organisms respond in ways that are both quantitatively and temporally appropriate to permit them to survive the multiple challenges posed by changes in their external and internal environments. Walter Cannon subsequently coined the term "homeostasis" to describe the ability of animals to maintain a constant intracellular environment despite changes in their external environment. We now know that organisms respond to changes in their external and internal environment by balanced, coordinated changes in the rates of specific metabolic reactions. Many human diseases, including cancer, diabetes, cystic fibrosis, and Alzheimer's disease, are characterized by regulatory dysfunctions triggered by pathogenic agents or genetic mutations. For example, many oncogenic viruses elaborate protein-tyrosine kinases that modify the regulatory events which control patterns of gene expression, contributing to the initiation and progression of cancer. The toxin from *Vibrio cholerae,* the causative agent of cholera, disables sensor-response pathways in intestinal epithelial cells by ADP-ribosylating the GTP-binding proteins (G-proteins) that link cell surface receptors to adenylyl cyclase. The consequent activation of the cyclase triggers the flow of water into the intestines, resulting in massive diarrhea and dehydration. *Yersinia pestis,* the causative agent of plague, elaborates a protein-tyrosine phosphatase that hydrolyzes phosphoryl groups on key cytoskeletal proteins. Knowledge of factors that control the rates of enzyme-catalyzed reactions thus is essential to an understanding of the molecular basis of disease. This chapter outlines the patterns by which metabolic processes are controlled and provides illustrative examples. Subsequent chapters provide additional examples.

REGULATION OF METABOLITE FLOW CAN BE ACTIVE OR PASSIVE

Enzymes that operate at their maximal rate cannot respond to an increase in substrate concentration, and can respond only to a precipitous decrease in substrate concentration. For most enzymes, therefore, the average intracellular concentration of their substrate tends to be close to the K_m value, so that changes in substrate concentration generate corresponding changes in metabolite flux (Figure 9–1). Responses to changes in substrate level represent an important but *passive* means for coordinating metabolite flow and maintaining homeostasis in quiescent cells. However, they offer limited scope for responding to changes in environmental variables. The mechanisms that regulate enzyme activity in an *active* manner in response to internal and external signals are discussed below.

Metabolite Flow Tends to Be Unidirectional

Despite the existence of short-term oscillations in metabolite concentrations and enzyme levels, living cells exist in a dynamic steady state in which the mean concentrations of metabolic intermediates remain relatively constant over time (Figure 9–2). While all chemical reactions are to some extent reversible, in living cells the reaction products serve as substrates for—and are removed by—other enzyme-catalyzed reactions. Many nominally reversible reactions thus occur unidirectionally. This succession of coupled metabolic reactions is accompanied by an overall change in free energy that favors unidirectional metabolite flow (Chapter 10). The unidirectional flow of metabolites through a pathway with a large overall negative change in free energy is analogous to the flow of water through a pipe in which one end is lower than the other. Bends or kinks in the pipe simulate individual enzyme-catalyzed steps with a small negative or positive change in free energy. Flow of water through the pipe nevertheless remains unidirectional due to the overall change in height, which corresponds to the overall change in free energy in a pathway (Figure 9–3).

COMPARTMENTATION ENSURES METABOLIC EFFICIENCY & SIMPLIFIES REGULATION

In eukaryotes, anabolic and catabolic pathways that interconvert common products may take place in specific subcellular compartments. For example, many of the enzymes that degrade proteins and polysaccharides reside inside organelles called lysosomes. Similarly, fatty acid biosynthesis occurs in the cytosol, whereas fatty

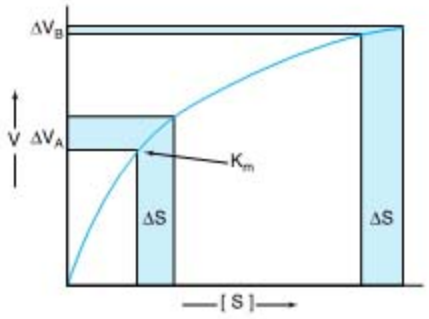


Figure 9–1. Differential response of the rate of an enzyme-catalyzed reaction, ΔV, to the same incremental change in substrate concentration at a substrate concentration of K_m (ΔV_A) or far above K_m (ΔV_B).

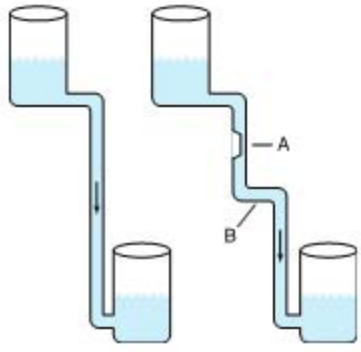


Figure 9–3. Hydrostatic analogy for a pathway with a rate-limiting step **(A)** and a step with a ΔG value near zero **(B).**

acid oxidation takes place within mitochondria (Chapters 21 and 22). Segregation of certain metabolic pathways within specialized cell types can provide further physical compartmentation. Alternatively, possession of one or more *unique intermediates* can permit apparently opposing pathways to coexist even in the absence of physical barriers. For example, despite many shared intermediates and enzymes, both glycolysis and gluconeogenesis are favored energetically. This cannot be true if *all* the reactions were the same. If one pathway was favored energetically, the other would be accompanied by a change in free energy G equal in magnitude but opposite in sign. Simultaneous spontaneity of both pathways results from substitution of one or more reactions by different reactions favored thermodynamically in the opposite direction. The glycolytic enzyme phosphofructokinase (Chapter 17) is replaced by the gluconeogenic enzyme fructose-1,6-bisphosphatase (Chapter 19). The ability of enzymes to discriminate between the structurally similar coenzymes NAD^+ and $NADP^+$ also results in a form of compartmentation, since it segregates the electrons of NADH that are destined for ATP generation from those of NADPH that participate in the reductive steps in many biosynthetic pathways.

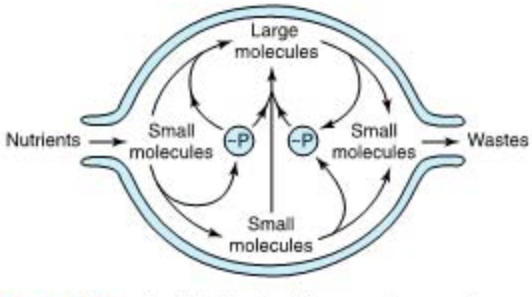


Figure 9–2. An idealized cell in steady state. Note that metabolite flow is unidirectional.

Controlling an Enzyme That Catalyzes a Rate-Limiting Reaction Regulates an Entire Metabolic Pathway

While the flux of metabolites through metabolic pathways involves catalysis by numerous enzymes, active control of homeostasis is achieved by regulation of only a small number of enzymes. The ideal enzyme for regulatory intervention is one whose quantity or catalytic efficiency dictates that the reaction it catalyzes is slow relative to all others in the pathway. Decreasing the catalytic efficiency or the quantity of the catalyst for the "bottleneck" or **rate-limiting reaction** immediately reduces metabolite flux through the entire pathway. Conversely, an increase in either its quantity or catalytic efficiency enhances flux through the pathway as a whole. For example, acetyl-CoA carboxylase catalyzes the synthesis of malonyl-CoA, the first committed reaction of fatty acid biosynthesis (Chapter 21). When synthesis of malonyl-CoA is inhibited, subsequent reactions of fatty acid synthesis cease due to lack of substrates. Enzymes that catalyze rate-limiting steps serve as natural "governors" of metabolic flux. Thus, they constitute efficient targets for regulatory intervention by drugs. For example, inhibition by "statin" drugs of HMG-CoA reductase, which catalyzes the rate-limiting reaction of cholesterogenesis, curtails synthesis of cholesterol.

REGULATION OF ENZYME QUANTITY

The catalytic capacity of the rate-limiting reaction in a metabolic pathway is the product of the concentration of enzyme molecules and their intrinsic catalytic efficiency. It therefore follows that catalytic capacity can be

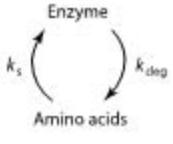
influenced both by changing the quantity of enzyme present and by altering its intrinsic catalytic efficiency.

Control of Enzyme Synthesis

Enzymes whose concentrations remain essentially constant over time are termed **constitutive enzymes.** By contrast, the concentrations of many other enzymes depend upon the presence of **inducers,** typically substrates or structurally related compounds, that initiate their synthesis. *Escherichia coli* grown on glucose will, for example, only catabolize lactose after addition of a β-galactoside, an inducer that initiates synthesis of a β-galactosidase and a galactoside permease (Figure 39–3). Inducible enzymes of humans include tryptophan pyrrolase, threonine dehydrase, tyrosine-α-ketoglutarate aminotransferase, enzymes of the urea cycle, HMG-CoA reductase, and cytochrome P450. Conversely, an excess of a metabolite may curtail synthesis of its cognate enzyme via **repression.** Both induction and repression involve cis elements, specific DNA sequences located upstream of regulated genes, and trans-acting regulatory proteins. The molecular mechanisms of induction and repression are discussed in Chapter 39.

Control of Enzyme Degradation

The absolute quantity of an enzyme reflects the net balance between enzyme synthesis and enzyme degradation, where k_s and k_{deg} represent the rate constants for the overall processes of synthesis and degradation, respectively. Changes in both the k_s and k_{deg} of specific enzymes occur in human subjects.



Protein turnover represents the net result of enzyme synthesis and degradation. By measuring the rates of incorporation of ^{15}N-labeled amino acids into protein and the rates of loss of ^{15}N from protein, Schoenheimer deduced that body proteins are in a state of "dynamic equilibrium" in which they are continuously synthesized and degraded. Mammalian proteins are degraded both by ATP and ubiquitin-dependent pathways and by ATP-independent pathways (Chapter 29). Susceptibility to proteolytic degradation can be influenced by the presence of ligands such as substrates, coenzymes, or metal ions that alter protein conformation. Intracellular ligands thus can influence the rates at which specific enzymes are degraded.

Enzyme levels in mammalian tissues respond to a wide range of physiologic, hormonal, or dietary factors. For example, glucocorticoids increase the concentration of tyrosine aminotransferase by stimulating k_s, and glucagon—despite its antagonistic physiologic effects—increases k_s fourfold to fivefold. Regulation of liver arginase can involve changes either in k_s or in k_{deg}. After a protein-rich meal, liver arginase levels rise and arginine synthesis decreases (Chapter 29). Arginase levels also rise in starvation, but here arginase degradation decreases, whereas k_s remains unchanged. Similarly, injection of glucocorticoids and ingestion of tryptophan both elevate levels of tryptophan oxygenase. While the hormone raises k_s for oxygenase synthesis, tryptophan specifically lowers k_{deg} by stabilizing the oxygenase against proteolytic digestion.

MULTIPLE OPTIONS ARE AVAILABLE FOR REGULATING CATALYTIC ACTIVITY

In humans, the induction of protein synthesis is a complex multistep process that typically requires hours to produce significant changes in overall enzyme level. By contrast, changes in intrinsic catalytic efficiency effected by binding of dissociable ligands (**allosteric regulation**) or by **covalent modification** achieve regulation of enzymic activity within seconds. Changes in protein level serve long-term adaptive requirements, whereas changes in catalytic efficiency are best suited for rapid and transient alterations in metabolite flux.

ALLOSTERIC EFFECTORS REGULATE CERTAIN ENZYMES

Feedback inhibition refers to inhibition of an enzyme in a biosynthetic pathway by an end product of that pathway. For example, for the biosynthesis of D from A catalyzed by enzymes Enz_1 through Enz_3,

$$A \xrightarrow{Enz_1} B \xrightarrow{Enz_2} C \xrightarrow{Enz_3} D$$

high concentrations of D inhibit conversion of A to B. Inhibition results not from the "backing up" of intermediates but from the ability of D to bind to and inhibit Enz_1. Typically, D binds at an **allosteric site** spatially distinct from the catalytic site of the target enzyme. Feedback inhibitors thus are allosteric effectors and typically bear little or no structural similarity to the substrates of the enzymes they inhibit. In this example, the feedback inhibitor D acts as a **negative allosteric effector** of Enz_1.

In a branched biosynthetic pathway, the initial reactions participate in the synthesis of several products. Figure 9–4 shows a hypothetical branched biosynthetic pathway in which curved arrows lead from feedback inhibitors to the enzymes whose activity they inhibit. The sequences $S_3 \rightarrow A$, $S_4 \rightarrow B$, $S_4 \rightarrow C$, and $S_3 \rightarrow \rightarrow D$ each represent linear reaction sequences that are feedback-inhibited by their end products. The pathways of nucleotide biosynthesis (Chapter 34) provide specific examples.

The kinetics of feedback inhibition may be competitive, noncompetitive, partially competitive, or mixed. Feedback inhibitors, which frequently are the small molecule building blocks of macromolecules (eg, amino acids for proteins, nucleotides for nucleic acids), typically inhibit the first committed step in a particular biosynthetic sequence. A much-studied example is inhibition of bacterial aspartate transcarbamoylase by CTP (see below and Chapter 34).

Multiple feedback loops can provide additional fine control. For example, as shown in Figure 9–5, the presence of excess product B decreases the requirement for substrate S_2. However, S_2 is also required for synthesis of A, C, and D. Excess B should therefore also curtail synthesis of all four end products. To circumvent this potential difficulty, each end product typically only partially inhibits catalytic activity. The effect of an excess of two or more end products may be strictly additive or, alternatively, may be greater than their individual effect (cooperative feedback inhibition).

Aspartate Transcarbamoylase Is a Model Allosteric Enzyme

Aspartate transcarbamoylase (ATCase), the catalyst for the first reaction unique to pyrimidine biosynthesis (Figure 34–7), is feedback-inhibited by cytidine triphosphate (CTP). Following treatment with mercurials, ATCase loses its sensitivity to inhibition by CTP but retains its full activity for synthesis of carbamoyl aspartate. This suggests that CTP is bound at a different (allosteric) site from either substrate. ATCase consists of multiple catalytic and regulatory subunits. Each catalytic subunit contains four aspartate (substrate) sites and each regulatory subunit at least two CTP (regulatory) sites (Chapter 34).

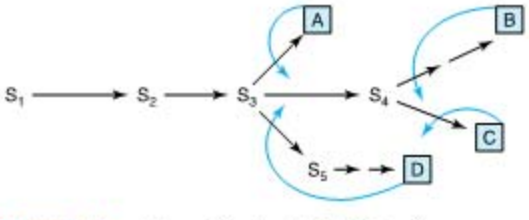


Figure 9–4. Sites of feedback inhibition in a branched biosynthetic pathway. S_1–S_5 are intermediates in the biosynthesis of end products A–D. Straight arrows represent enzymes catalyzing the indicated conversions. Curved arrows represent feedback loops and indicate sites of feedback inhibition by specific end products.

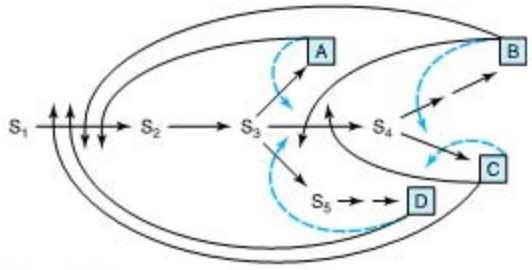


Figure 9–5. Multiple feedback inhibition in a branched biosynthetic pathway. Superimposed on simple feedback loops (dashed, curved arrows) are multiple feedback loops (solid, curved arrows) that regulate enzymes common to biosynthesis of several end products.

Allosteric & Catalytic Sites Are Spatially Distinct

The lack of structural similarity between a feedback inhibitor and the substrate for the enzyme whose activity it regulates suggests that these effectors are not **isosteric** with a substrate but **allosteric** ("occupy another space"). Jacques Monod therefore proposed the existence of allosteric sites that are physically distinct from the catalytic site. **Allosteric enzymes** thus are those whose activity at the active site may be modulated by the presence of effectors at an allosteric site. This hypothesis has been confirmed by many lines of evidence, including x-ray crystallography and site-directed mutagenesis, demonstrating the existence of spatially distinct active and allosteric sites on a variety of enzymes.

Allosteric Effects May Be on K_m or on V_{max}

To refer to the kinetics of allosteric inhibition as "competitive" or "noncompetitive" with substrate carries misleading mechanistic implications. We refer instead to two classes of regulated enzymes: K-series and V-series enzymes. For K-series allosteric enzymes, the substrate saturation kinetics are competitive in the sense that K_m is raised without an effect on V_{max}. For V-series allosteric enzymes, the allosteric inhibitor lowers V_{max}

without affecting the K_m. Alterations in K_m or V_{max} probably result from conformational changes at the catalytic site induced by binding of the allosteric effector at the allosteric site. For a K-series allosteric enzyme, this conformational change may weaken the bonds between substrate and substrate-binding residues. For a V-series allosteric enzyme, the primary effect may be to alter the orientation or charge of catalytic residues, lowering V_{max}. Intermediate effects on K_m and V_{max}, however, may be observed consequent to these conformational changes.

FEEDBACK REGULATION IS NOT SYNONYMOUS WITH FEEDBACK INHIBITION

In both mammalian and bacterial cells, end products "feed back" and control their own synthesis, in many instances by feedback inhibition of an early biosynthetic enzyme. We must, however, distinguish between **feedback regulation,** a phenomenologic term devoid of mechanistic implications, and **feedback inhibition,** a mechanism for regulation of enzyme activity. For example, while dietary cholesterol decreases hepatic synthesis of cholesterol, this feedback **regulation** does not involve feedback **inhibition.** HMG-CoA reductase, the rate-limiting enzyme of cholesterologenesis, is affected, but cholesterol does not feedback-inhibit its activity. Regulation in response to dietary cholesterol involves curtailment by cholesterol or a cholesterol metabolite of the expression of the gene that encodes HMG-CoA reductase (enzyme repression) (Chapter 26).

MANY HORMONES ACT THROUGH ALLOSTERIC SECOND MESSENGERS

Nerve impulses—and binding of hormones to cell surface receptors—elicit changes in the rate of enzyme-catalyzed reactions within target cells by inducing the release or synthesis of specialized allosteric effectors called **second messengers.** The primary or "first" messenger is the hormone molecule or nerve impulse. Second messengers include 3′,5′-cAMP, synthesized from ATP by the enzyme adenylyl cyclase in response to the hormone epinephrine, and Ca^{2+}, which is stored inside the endoplasmic reticulum of most cells. Membrane depolarization resulting from a nerve impulse opens a membrane channel that releases calcium ion into the cytoplasm, where it binds to and activates enzymes involved in the regulation of contraction and the mobilization of stored glucose from glycogen. Glucose then supplies the increased energy demands of muscle contraction. Other second messengers include 3′,5′-cGMP and polyphosphoinositols, produced by the hydrolysis of inositol phospholipids by hormone-regulated phospholipases.

REGULATORY COVALENT MODIFICATIONS CAN BE REVERSIBLE OR IRREVERSIBLE

In mammalian cells, the two most common forms of covalent modification are **partial proteolysis** and **phosphorylation.** Because cells lack the ability to reunite the two portions of a protein produced by hydrolysis of a peptide bond, proteolysis constitutes an irreversible modification. By contrast, phosphorylation is a reversible modification process. The phosphorylation of proteins on seryl, threonyl, or tyrosyl residues, catalyzed by protein kinases, is thermodynamically spontaneous. Equally spontaneous is the hydrolytic removal of these phosphoryl groups by enzymes called protein phosphatases.

PROTEASES MAY BE SECRETED AS CATALYTICALLY INACTIVE PROENZYMES

Certain proteins are synthesized and secreted as inactive precursor proteins known as **proproteins.** The proproteins of enzymes are termed **proenzymes** or **zymogens.** Selective proteolysis converts a proprotein by one or more successive proteolytic "clips" to a form that exhibits the characteristic activity of the mature protein, eg, its enzymatic activity. Proteins synthesized as proproteins include the hormone insulin (proprotein = proinsulin), the digestive enzymes pepsin, trypsin, and chymotrypsin (proproteins = pepsinogen, trypsinogen, and chymotrypsinogen, respectively), several factors of the blood clotting and blood clot dissolution cascades (see Chapter 51), and the connective tissue protein collagen (proprotein = procollagen).

Proenzymes Facilitate Rapid Mobilization of an Activity in Response to Physiologic Demand

The synthesis and secretion of proteases as catalytically inactive proenzymes protects the tissue of origin (eg, the pancreas) from autodigestion, such as can occur in pancreatitis. Certain physiologic processes such as digestion are intermittent but fairly regular and predictable. Others such as blood clot formation, clot dissolution, and tissue repair are brought "on line" only in response to pressing physiologic or pathophysiologic need. The processes of blood clot formation and dissolution clearly must be temporally coordinated to achieve homeostasis. Enzymes needed intermittently but rapidly often are secreted in an initially inactive form since the secretion process or new synthesis of the required proteins might be insufficiently rapid for response to a pressing pathophysiologic demand such as the loss of blood.

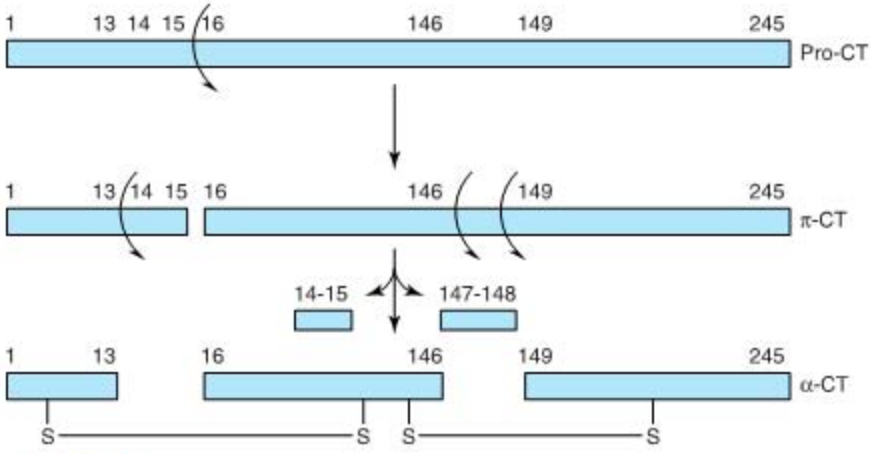


Figure 9–6. Selective proteolysis and associated conformational changes form the active site of chymotrypsin, which includes the Asp102-His57-Ser195 catalytic triad. Successive proteolysis forms prochymotrypsin (pro-CT), π-chymotrypsin (π-CT), and ultimately α-chymotrypsin (α-CT), an active protease whose three peptides remain associated by covalent inter-chain disulfide bonds.

Activation of Prochymotrypsin Requires Selective Proteolysis

Selective proteolysis involves one or more highly specific proteolytic clips that may or may not be accompanied by separation of the resulting peptides. Most importantly, selective proteolysis often results in conformational changes that "create" the catalytic site of an enzyme. Note that while His 57 and Asp 102 reside on the B peptide of α-chymotrypsin, Ser 195 resides on the C peptide (Figure 9–6). The conformational changes that accompany selective proteolysis of prochymotrypsin (chymotrypsinogen) align the three residues of the charge-relay network, creating the catalytic site. Note also that contact and catalytic residues can be located on different peptide chains but still be within bond-forming distance of bound substrate.

REVERSIBLE COVALENT MODIFICATION REGULATES KEY MAMMALIAN ENZYMES

Mammalian proteins are the targets of a wide range of covalent modification processes. Modifications such as glycosylation, hydroxylation, and fatty acid acylation introduce new structural features into newly synthesized proteins that tend to persist for the lifetime of the protein. Among the covalent modifications that regulate protein function (eg, methylation, adenylylation), the most common by far is phosphorylation-dephosphorylation. **Protein kinases** phosphorylate proteins by catalyzing transfer of the terminal phosphoryl group of ATP to the hydroxyl groups of seryl, threonyl, or tyrosyl residues, forming *O*-phosphoseryl, *O*-phosphothreonyl, or *O*-phosphotyrosyl residues, respectively (Figure 9–7). Some protein kinases target the side chains of histidyl, lysyl, arginyl, and aspartyl residues. The unmodified form of the protein can be regenerated by hydrolytic removal of phosphoryl groups, catalyzed by **protein phosphatases.**

A typical mammalian cell possesses over 1000 phosphorylated proteins and several hundred protein kinases and protein phosphatases that catalyze their interconversion. The ease of interconversion of enzymes between their phospho- and dephospho- forms in part

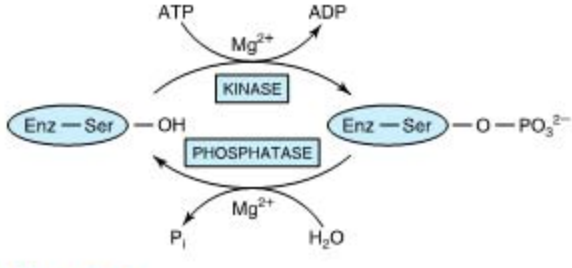


Figure 9–7. Covalent modification of a regulated enzyme by phosphorylation-dephosphorylation of a seryl residue.

accounts for the frequency of phosphorylation-dephosphorylation as a mechanism for regulatory control. Phosphorylation-dephosphorylation permits the functional properties of the affected enzyme to be altered only for as long as it serves a specific need. Once the need has passed, the enzyme can be converted back to its original form, poised to respond to the next stimulatory event. A second factor underlying the widespread use of protein phosphorylation-dephosphorylation lies in the chemical properties of the phosphoryl group itself. In order to alter an enzyme's functional properties, any modification of its chemical structure must influence the protein's three-dimensional configuration. The high charge density of protein-bound phosphoryl groups—generally –2 at physiologic pH—and their propensity to form salt bridges with arginyl residues make them potent agents for modifying protein structure and function. Phosphorylation generally targets amino acids distant from the catalytic site itself. Consequent conformational changes then influence an enzyme's intrinsic catalytic efficiency or other properties. In this sense, the sites of phosphorylation and other covalent modifications can be considered another form of allosteric site. However, in this case the "allosteric ligand" binds covalently to the protein.

PROTEIN PHOSPHORYLATION IS EXTREMELY VERSATILE

Protein phosphorylation-dephosphorylation is a highly versatile and selective process. Not all proteins are subject to phosphorylation, and of the many hydroxyl groups on a protein's surface, only one or a small subset are targeted. While the most common enzyme function affected is the protein's catalytic efficiency, phosphorylation can also alter the affinity for substrates, location within the cell, or responsiveness to regulation by allosteric ligands. Phosphorylation can increase an enzyme's catalytic efficiency, converting it to its active form in one protein, while phosphorylation of another converts it into an intrinsically inefficient, or inactive, form (Table 9–1).

Many proteins can be phosphorylated at multiple sites or are subject to regulation both by phosphorylation-dephosphorylation and by the binding of allosteric ligands. Phosphorylation-dephosphorylation at any one site can be catalyzed by multiple protein kinases or protein phosphatases. Many protein kinases and most protein phosphatases act on more than one protein and are themselves interconverted between active and inactive forms by the binding of second messengers or by covalent modification by phosphorylation-dephosphorylation.

The interplay between protein kinases and protein phosphatases, between the functional consequences of phosphorylation at different sites, or between phosphorylation sites and allosteric sites provides the basis for regulatory networks that integrate multiple environmental input signals to evoke an appropriate coordinated cellular response. In these sophisticated regulatory networks, individual enzymes respond to different environmental signals. For example, if an enzyme can be phosphorylated at a single site by more than one protein kinase, it can be converted from a catalytically efficient to an inefficient (inactive) form, or vice versa, in response to any one of several signals. If the protein kinase is activated in response to a signal different from the signal that activates the protein phosphatase, the phosphoprotein becomes a decision node. The functional output, generally catalytic activity, reflects the phosphorylation state. This state or degree of phosphorylation is determined by the relative activities of the protein kinase and protein phosphatase, a reflection of the presence and relative strength of the environmental signals that act through each. The ability of many protein kinases and protein phosphatases to target more than one protein provides a means for an environmental signal to coordinately regulate multiple metabolic processes. For example, the enzymes 3-hydroxy-3-methylglutaryl-CoA reductase and acetyl-CoA carboxylase—the rate-controlling enzymes for cholesterol and fatty acid biosynthesis, respectively—are phosphorylated and inactivated by the AMP-activated protein kinase. When this protein kinase is activated either through phosphorylation by yet another protein kinase or in response to the binding of its allosteric activator 5′-AMP, the two major pathways responsible for the synthesis of lipids from acetyl-CoA both are inhibited. Interconvertible enzymes and the enzymes responsible for their interconversion do not act as mere on and off switches working independently of one another.

Table 9–1. Examples of mammalian enzymes whose catalytic activity is altered by covalent phosphorylation-dephosphorylation.

Enzyme	Activity State[1]	
	Low	High
Acetyl-CoA carboxylase	EP	E
Glycogen synthase	EP	E
Pyruvate dehydrogenase	EP	E
HMG-CoA reductase	EP	E
Glycogen phosphorylase	E	EP
Citrate lyase	E	EP
Phosphorylase b kinase	E	EP
HMG-CoA reductase kinase	E	EP

[1]E, dephosphoenzyme; EP, phosphoenzyme.

They form the building blocks of biomolecular computers that maintain homeostasis in cells that carry out a complex array of metabolic processes that must be regulated in response to a broad spectrum of environmental factors.

Covalent Modification Regulates Metabolite Flow

Regulation of enzyme activity by phosphorylation-dephosphorylation has analogies to regulation by feedback inhibition. Both provide for short-term, readily reversible regulation of metabolite flow in response to specific physiologic signals. Both act without altering gene expression. Both act on early enzymes of a protracted, often biosynthetic metabolic sequence, and both act at allosteric rather than catalytic sites. Feedback inhibition, however, involves a single protein and lacks hormonal and neural features. By contrast, regulation of mammalian enzymes by phosphorylation-dephosphorylation involves several proteins and ATP and is under direct neural and hormonal control.

SUMMARY

- Homeostasis involves maintaining a relatively constant intracellular and intra-organ environment despite wide fluctuations in the external environment via appropriate changes in the rates of biochemical reactions in response to physiologic need.
- The substrates for most enzymes are usually present at a concentration close to K_m. This facilitates passive control of the rates of product formation response to changes in levels of metabolic intermediates.
- Active control of metabolite flux involves changes in the concentration, catalytic activity, or both of an enzyme that catalyzes a committed, rate-limiting reaction.
- Selective proteolysis of catalytically inactive proenzymes initiates conformational changes that form the active site. Secretion as an inactive proenzyme facilitates rapid mobilization of activity in response to injury or physiologic need and may protect the tissue of origin (eg, autodigestion by proteases).
- Binding of metabolites and second messengers to sites distinct from the catalytic site of enzymes triggers conformational changes that alter V_{max} or the K_m.
- Phosphorylation by protein kinases of specific seryl, threonyl, or tyrosyl residues—and subsequent dephosphorylation by protein phosphatases—regulates the activity of many human enzymes. The protein kinases and phosphatases that participate in regulatory cascades which respond to hormonal or second messenger signals constitute a "bio-organic computer" that can process and integrate complex environmental information to produce an appropriate and comprehensive cellular response.

REFERENCES

Bray D: Protein molecules as computational elements in living cells. Nature 1995;376:307.

Graves DJ, Martin BL, Wang JH: *Co- and Post-translational Modification of Proteins: Chemical Principles and Biological Effects.* Oxford Univ Press, 1994.

Johnson LN, Barford D: The effect of phosphorylation on the structure and function of proteins. Annu Rev Biophys Biomol Struct 1993;22:199.

Marks F (editor): *Protein Phosphorylation.* VCH Publishers, 1996.

Pilkis SJ et al: 6-Phosphofructo-2-kinase/fructose-2,6-bisphosphatase: A metabolic signaling enzyme. Annu Rev Biochem 1995;64:799.

Scriver CR et al (editors): *The Metabolic and Molecular Bases of Inherited Disease,* 8th ed. McGraw-Hill, 2000.

Sitaramayya A (editor): *Introduction to Cellular Signal Transduction.* Birkhauser, 1999.

Stadtman ER, Chock PB (editors): *Current Topics in Cellular Regulation.* Academic Press, 1969 to the present.

Weber G (editor): *Advances in Enzyme Regulation.* Pergamon Press, 1963 to the present.