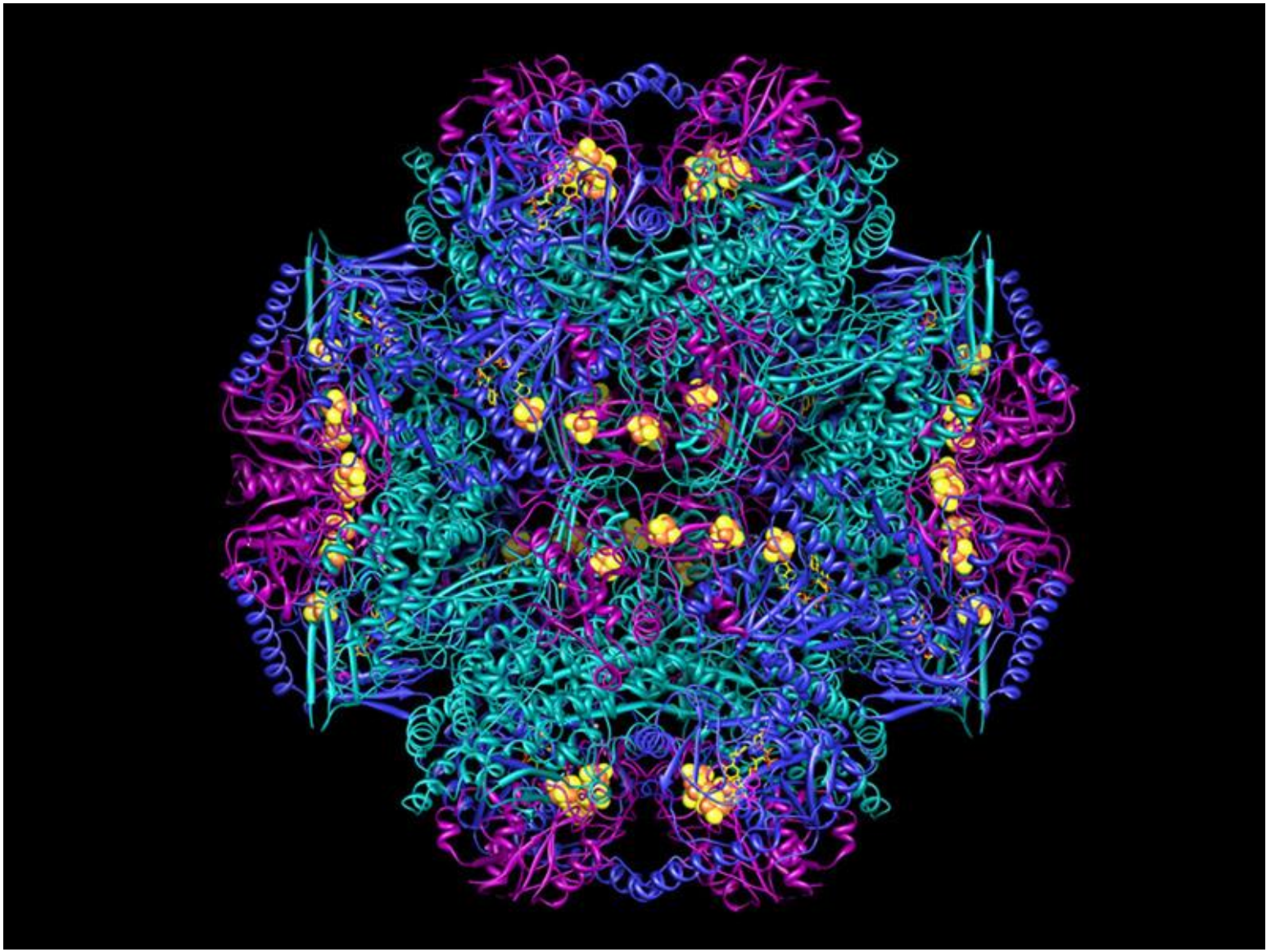
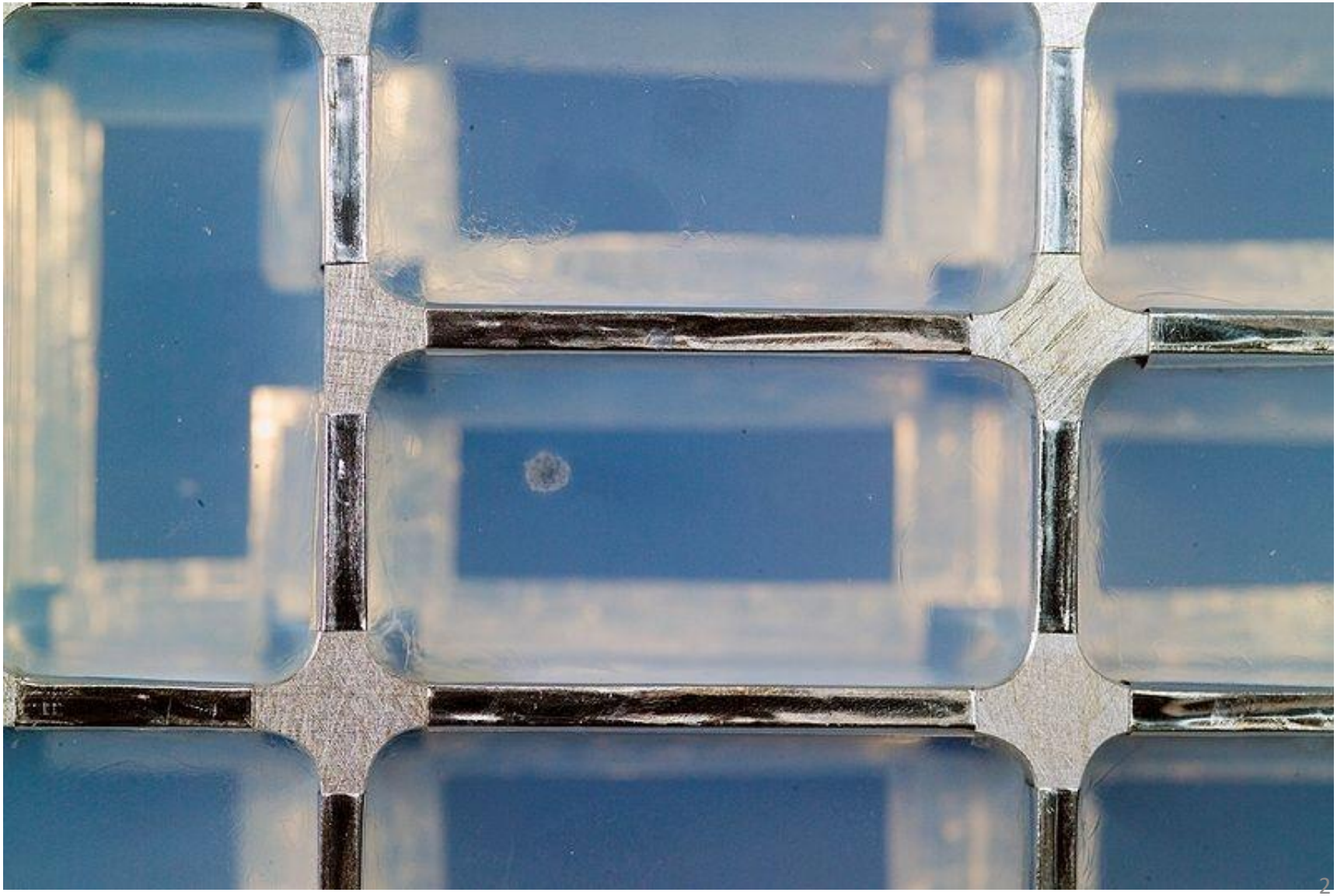
AMINO ACIDS, PEPTIDES AND PROTEINS



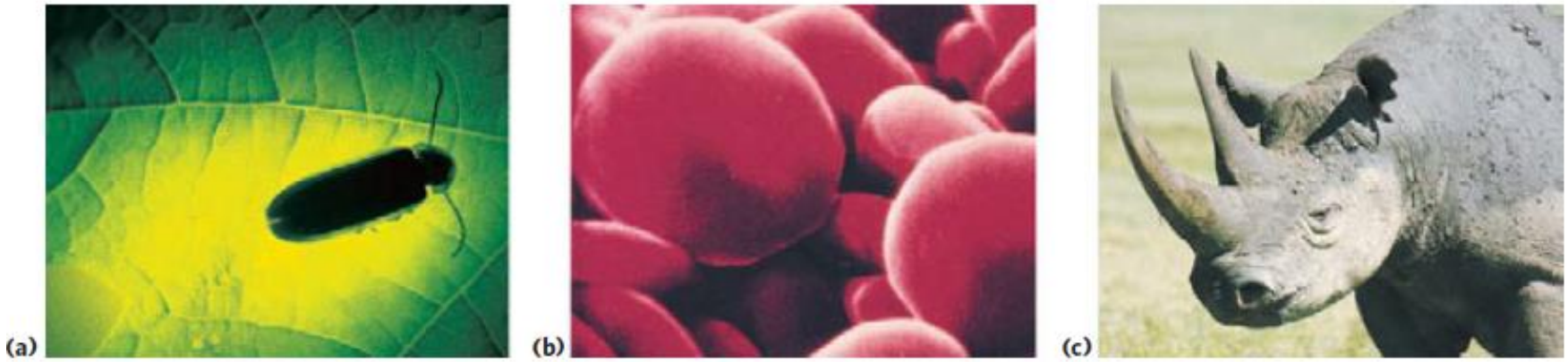
AMINO ACIDS, PEPTIDES AND PROTEINS I: Amino acids



❖Proteins mediate virtually every process that takes place in a cell, exhibiting an almost endless diversity of functions.

❖To explore the molecular mechanism of a biological process, a biochemist almost inevitably studies one or more proteins.

❖Proteins are the most abundant biological macromolecules, occurring in all cells and all parts of cells.



❖Proteins also occur in great variety; thousands of different kinds, ranging in size from relatively small peptides to huge polymers with molecular weights in the millions, may be found in a single cell.

❖Moreover, proteins exhibit enormous diversity of biological function.

❖All proteins, whether from the most ancient lines of bacteria or from the most complex forms of life, are constructed from the same ubiquitous set of 20 amino acids, covalently linked in characteristic linear sequences.

❖Because each of these amino acids has a side chain with distinctive chemical properties, this group of 20 precursor molecules may be regarded as the alphabet in which the language of protein structure is written.

❖What is most remarkable is that cells can produce proteins with strikingly different properties and activities by joining the same 20 amino acids in many different combinations and sequences.

❖From these building blocks different organisms can make such widely diverse products as enzymes, hormones, antibodies, transporters, muscle fibers, the lens protein of the eye, feathers, spider webs, rhinoceros horn, milk proteins, antibiotics, mushroom poisons, and myriad other substances having distinct biological activities.

❖Among these protein products, the enzymes are the most varied and specialized.

❖Virtually all cellular reactions are catalyzed by enzymes.

Amino Acids

❖Proteins are polymers of amino acids, with each **amino acid residue** joined to its neighbor by a specific type of covalent bond.

❖Proteins can be broken down (hydrolyzed) to their constituent amino acids by a variety of methods, and the earliest studies of proteins naturally focused on the free amino acids derived from them.

❖Twenty different amino acids are commonly found in proteins.

❖All the amino acids have trivial or common names, in some cases derived from the source from which they were first isolated.

❖Asparagine was first found in asparagus, and glutamate in wheat gluten; tyrosine was first isolated from cheese (its name is derived from the Greek tyros, “cheese”); and glycine (Greek glykos, “sweet”) was so named because of its sweet taste.

TABLE 3–1 Properties and Conventions Associated with the Common Amino Acids Found in Proteins

Amino acid	Abbreviation/ symbol	M_r	pK_a values: pK_1 (—COOH)	pK_2 ($—NH_3^+$)	pK_R (R group)	pI	Hydropathy index*	Occurrence in proteins (%)†
Nonpolar, aliphatic R groups								
Glycine	Gly G	75	2.34	9.60		5.97	−0.4	7.2
Alanine	Ala A	89	2.34	9.69		6.01	1.8	7.8
Proline	Pro P	115	1.99	10.96		6.48	1.6	5.2
Valine	Val V	117	2.32	9.62		5.97	4.2	6.6
Leucine	Leu L	131	2.36	9.60		5.98	3.8	9.1
Isoleucine	Ile I	131	2.36	9.68		6.02	4.5	5.3
Methionine	Met M	149	2.28	9.21		5.74	1.9	2.3
Aromatic R groups								
Phenylalanine	Phe F	165	1.83	9.13		5.48	2.8	3.9
Tyrosine	Tyr Y	181	2.20	9.11	10.07	5.66	−1.3	3.2
Tryptophan	Trp W	204	2.38	9.39		5.89	−0.9	1.4
Polar, uncharged R groups								
Serine	Ser S	105	2.21	9.15		5.68	−0.8	6.8
Threonine	Thr T	119	2.11	9.62		5.87	−0.7	5.9
Cysteine	Cys C	121	1.96	10.28	8.18	5.07	2.5	1.9
Asparagine	Asn N	132	2.02	8.80		5.41	−3.5	4.3
Glutamine	Gln Q	146	2.17	9.13		5.65	−3.5	4.2
Positively charged R groups								
Lysine	Lys K	146	2.18	8.95	10.53	9.74	−3.9	5.9
Histidine	His H	155	1.82	9.17	6.00	7.59	−3.2	2.3
Arginine	Arg R	174	2.17	9.04	12.48	10.76	−4.5	5.1
Negatively charged R groups								
Aspartate	Asp D	133	1.88	9.60	3.65	2.77	−3.5	5.3
Glutamate	Glu E	147	2.19	9.67	4.25	3.22	−3.5	6.3

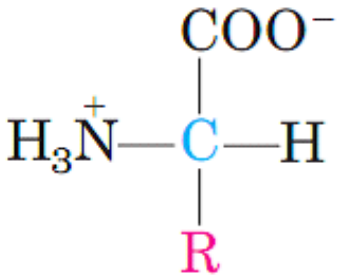
*A scale combining hydrophobicity and hydrophilicity of R groups; it can be used to measure the tendency of an amino acid to seek an aqueous environment (− values) or a hydrophobic environment (+ values). See Chapter 11. From Kyte, J. & Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157,** 105–132.

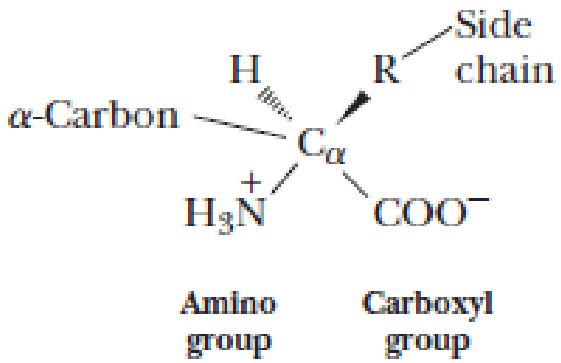
†Average occurrence in more than 1,150 proteins. From Doolittle, R.F. (1989) Redundancies in protein sequences. In *Prediction of Protein Structure and the Principles of Protein Conformation* (Fasman, G.D., ed.), pp. 599–623, Plenum Press, New York.

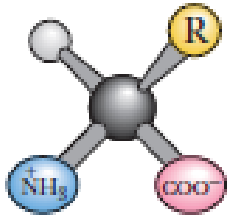
❖All 20 of the common amino acids are α-amino acids.

❖They have a carboxyl group and an amino group bonded to the same carbon atom.

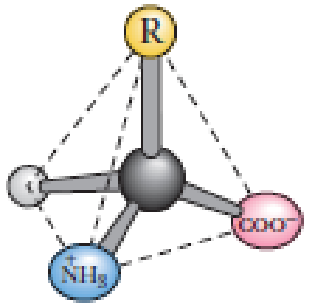
❖They differ from each other in their side chains, or R groups, which vary in structure, size, and electric charge, and which influence the solubility of the amino acids in water.







Ball-and-stick model



Amino acids are tetrahedral structures

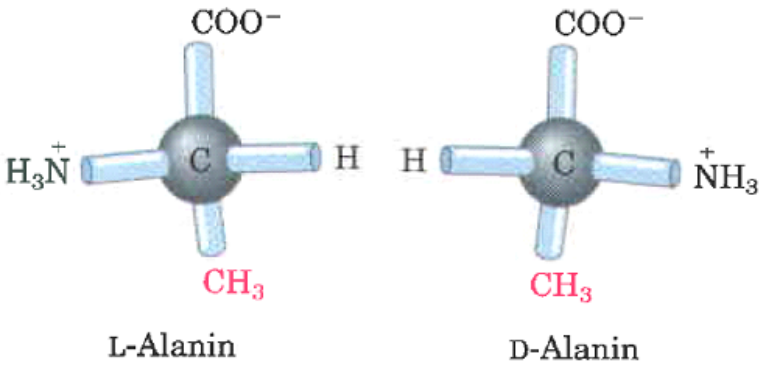
❖In addition to these 20 amino acids there are many less common ones.

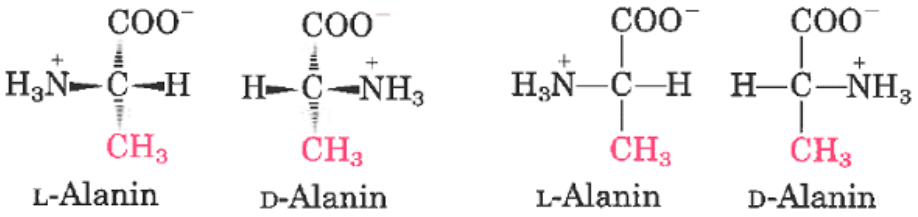
❖Some are residues modified after a protein has been synthesized; others are amino acids present in living organisms but not as constituents of proteins.

❖For all the common amino acids except glycine, the α-carbon is bonded to four different groups: a carboxyl group, an amino group, an R group, and a hydrogen atom.

❖The α-carbon atom is thus a chiral center. Because of the tetrahedral arrangement of the bonding orbitals around the α-carbon atom, the four different groups can occupy two unique spatial arrangements, and thus amino acids have two possible stereoisomers.

❖Since they are non-superimposable mirror images of each other, the two forms represent a class of stereoisomers called enantiomers

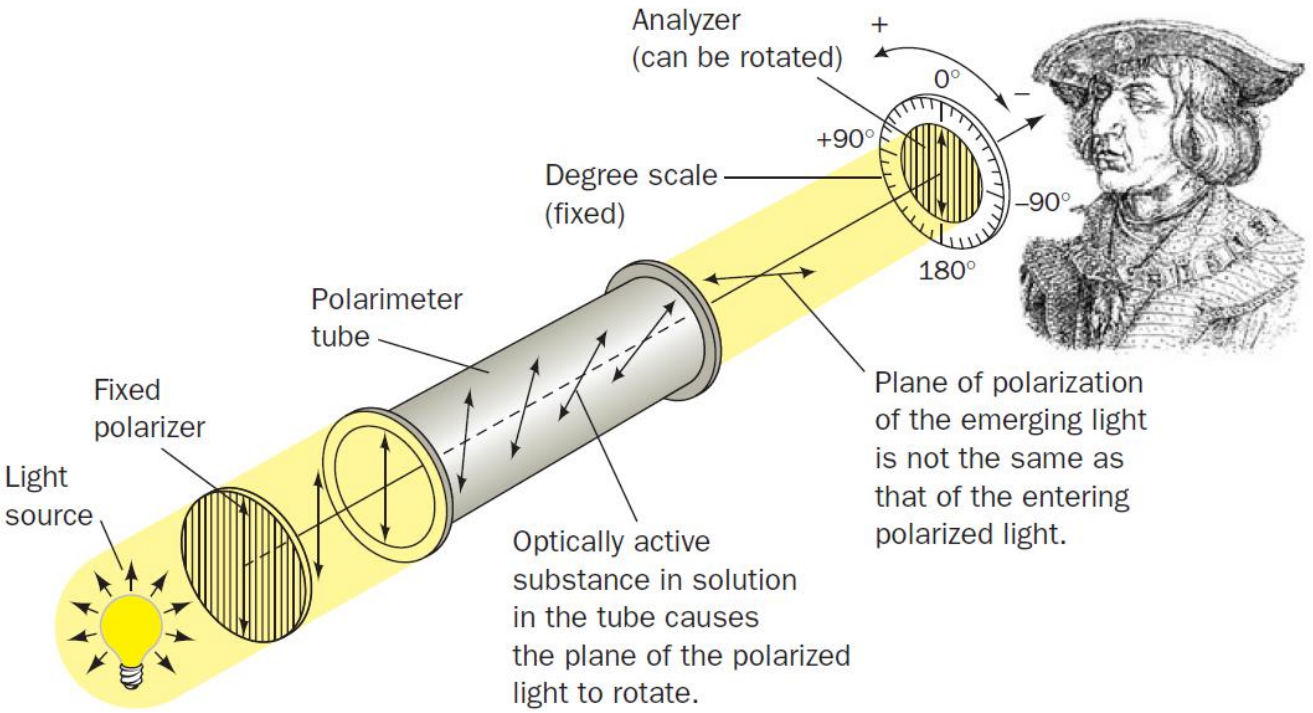




❖With the exception of glycine, all the amino acids recovered from polypeptides are optically active; that is, they rotate the plane of polarized light.

❖The direction and angle of rotation can be measured using an instrument known as a **polarimeter.**

❖Optically active molecules are asymmetric; that is, they are not superimposable on their mirror image.

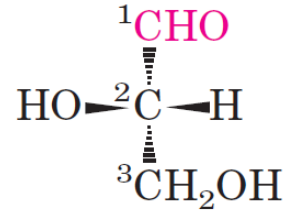


❖The absolute configurations of simple sugars and amino acids are specified by the **D, L system**, based on the absolute configuration of the three-carbon sugar glyceraldehyde, a convention proposed by Emil Fischer in 1891.

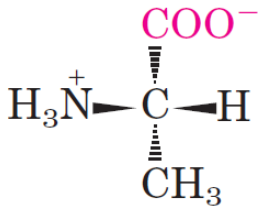
❖For all chiral compounds, stereoisomers having a configuration related to that of L-glyceraldehyde are designated L, and stereoisomers related to D-glyceraldehyde are designated D.

❖Historically, the similar L and D designations were used for levorotatory (rotating plane-polarized light to the left) and dextrorotatory (rotating light to the right).

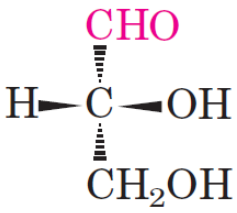
❖However, not all L-amino acids are levorotatory and vice versa.



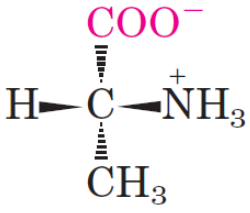
L-Glyceraldehyde



L-Alanine



D-Glyceraldehyde



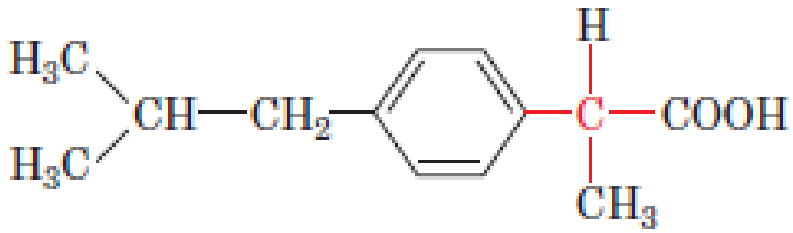
D-Alanine

❖The importance of stereochemistry in living systems is also a concern of the pharmaceutical industry. Many drugs are chemically synthesized as racemic mixtures, although only one enantiomer has biological activity.

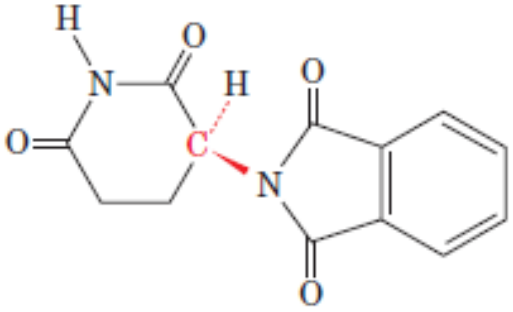
❖In most cases, the opposite enantiomer is biologically inert and is therefore packaged along with its active counterpart. This is true, for example, of the anti-inflammatory agent ibuprofen, only one enantiomer of which is physiologically active.

❖Occasionally, the inactive enantiomer of a useful drug produces harmful effects and must therefore be eliminated from the racemic mixture.

❖The most striking example of this is the drug thalidomide, a mild sedative whose inactive enantiomer causes severe birth defects.



Ibuprofen

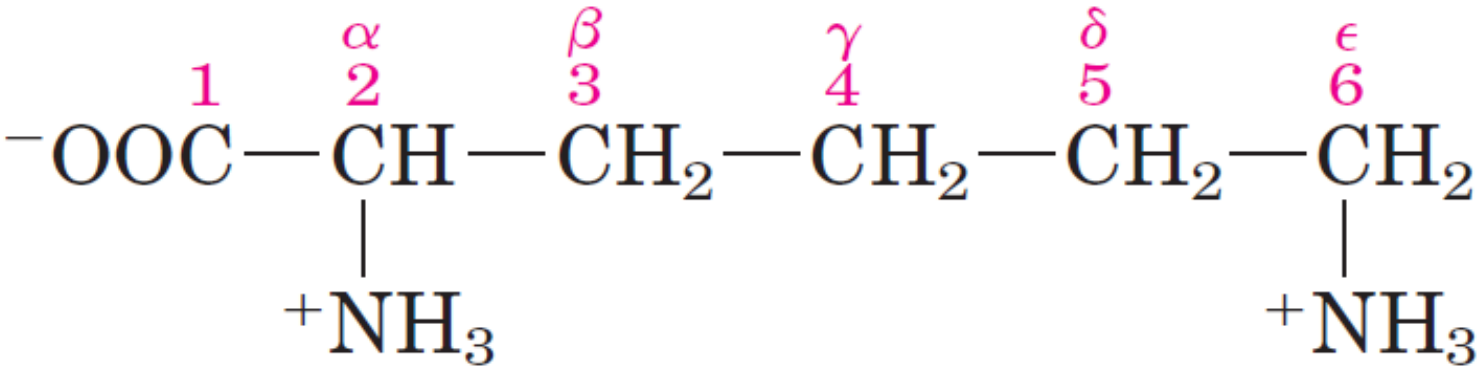


Thalidomide

❖Two conventions are used to identify the carbons in an amino acid—a practice that can be confusing.

❖The additional carbons in an R group are commonly designated β, γ, δ, ε and so forth, proceeding out from the α carbon.

❖For most other organic molecules, carbon atoms are simply numbered from one end, giving highest priority (C-1) to the carbon with the substituent containing the atom of highest atomic number.



Lysine

❖Nearly all biological compounds with a chiral center occur naturally in only one stereoisomeric form, either D or L.

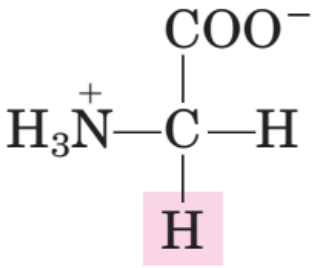
❖The amino acid residues in protein molecules are exclusively L stereoisomers.

❖D-Amino acid residues have been found only in a few, generally small peptides, including some peptides of bacterial cell walls and certain peptide antibiotics.

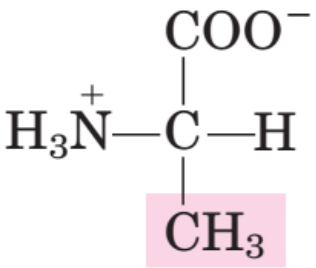
❖Amino acids can be classified based on the properties of their R groups, particularly their **polarity**, or tendency to interact with water at biological pH (near pH 7.0).

❖Amino acids which contain nonpolar, aliphatic R groups are glycine, alanine, proline, valine, leucine, isoleucine and methionine.

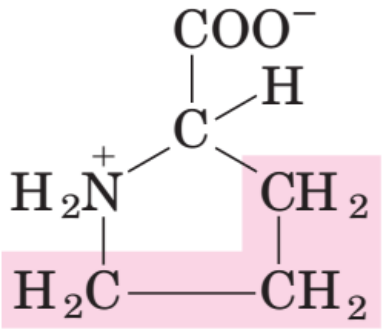
Nonpolar, aliphatic R groups



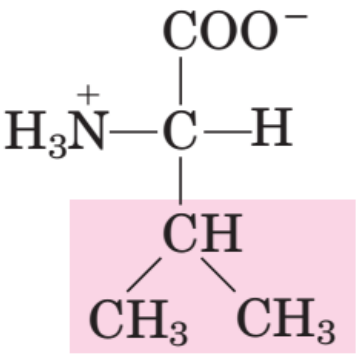
Glycine



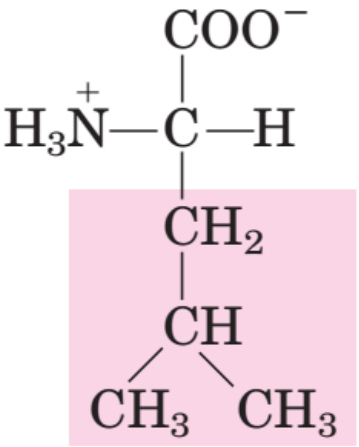
Alanine



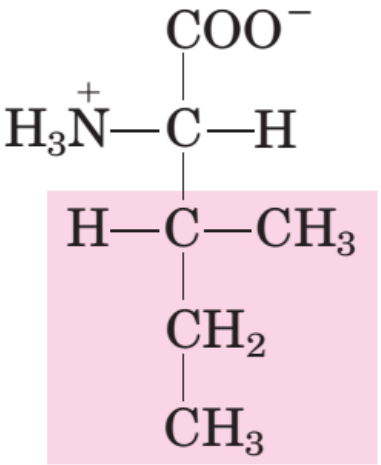
Proline



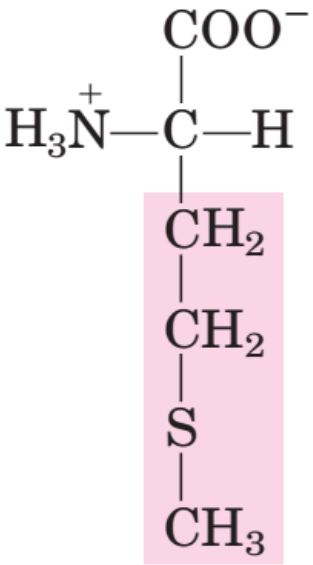
Valine



Leucine



Isoleucine



Methionine

Amino acids with nonpolar, aliphatic R Groups

❖ The side chains of **alanine**, **valine**, **leucine**, and **isoleucine** tend to cluster together within proteins, stabilizing protein structure by means of hydrophobic interactions.

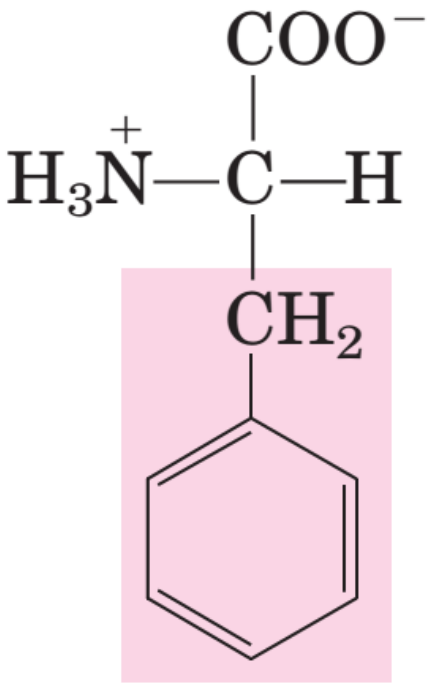
❖**Glycine** has the simplest structure. Although it is formally nonpolar, its very small side chain makes no real contribution to hydrophobic interactions.

❖**Methionine**, one of the two sulfur-containing amino acids, has a nonpolar thioether group in its side chain.

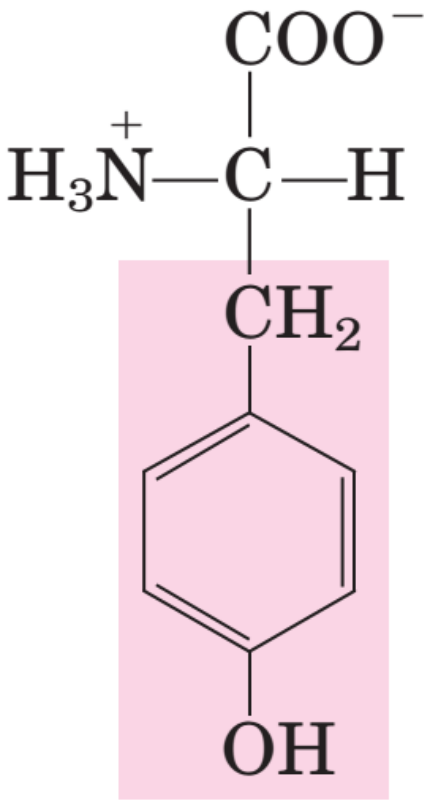
❖**Proline** has an aliphatic side chain with a distinctive cyclic structure.

❖The secondary amino (imino) group of proline residues is held in a rigid conformation that reduces the structural flexibility of polypeptide regions containing proline.

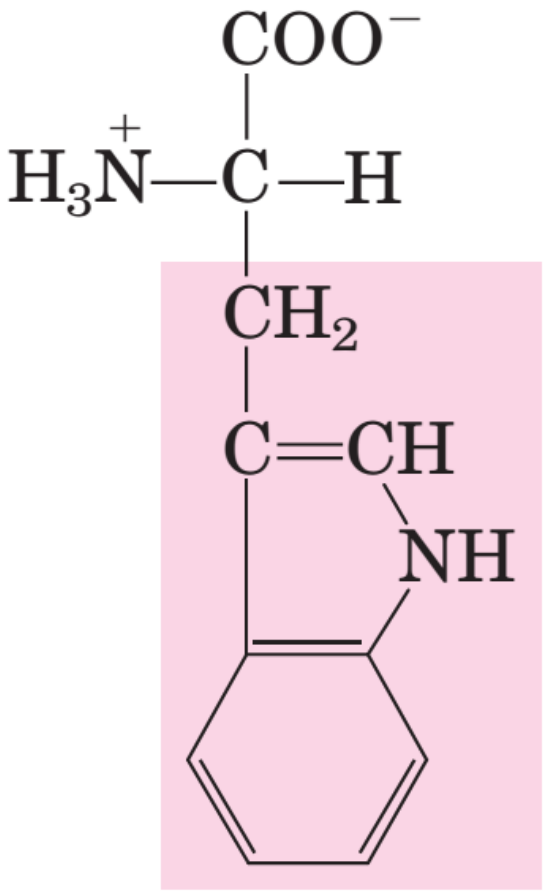
Aromatic R groups



Phenylalanine



Tyrosine



Tryptophan

Amino acids with aromatic R Groups

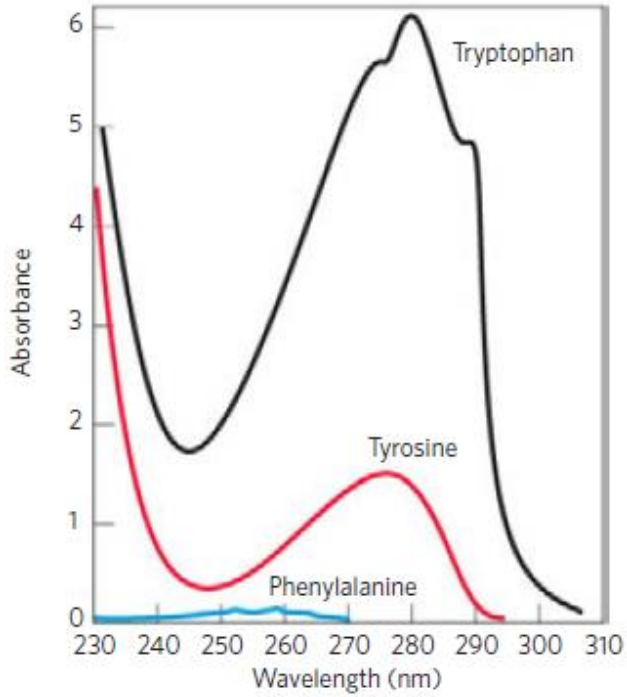
❖**Phenylalanine**, **tyrosine**, and **tryptophan**, with their aromatic side chains, are relatively nonpolar (hydrophobic). All can participate in hydrophobic interactions.

❖The hydroxyl group of tyrosine can form hydrogen bonds, and it is an important functional group in some enzymes.

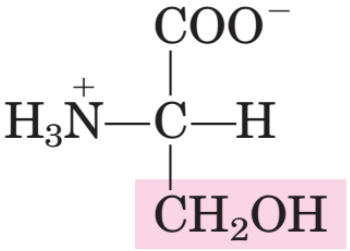
❖Tyrosine and tryptophan are significantly more polar than phenylalanine, because of the tyrosine hydroxyl group and the nitrogen of the tryptophan indole ring.

❖Tryptophan and tyrosine, and to a much lesser extent phenylalanine, absorb ultraviolet light.

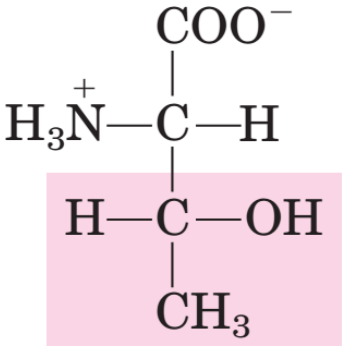
❖This accounts for the characteristic strong absorbance of light by most proteins at a wavelength of 280 nm, a property exploited by researchers in the characterization of proteins.



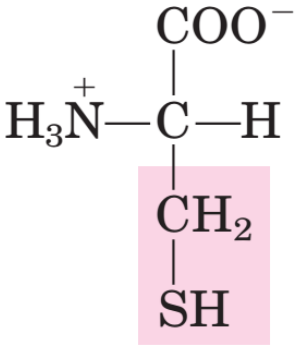
Polar, uncharged R groups



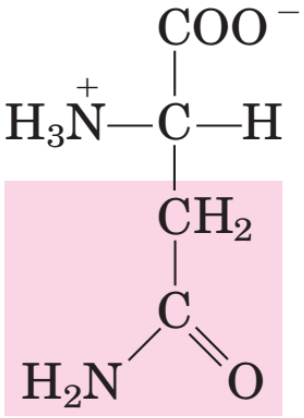
Serine



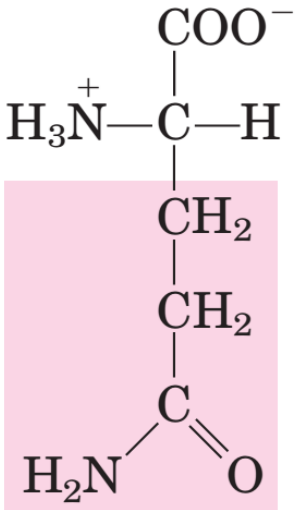
Threonine



Cysteine



Asparagine



Glutamine

Amino acids with polar, uncharged R Groups

❖The R groups of these amino acids are more soluble in water, or more hydrophilic, than those of the nonpolar amino acids, because they contain functional groups that form hydrogen bonds with water.

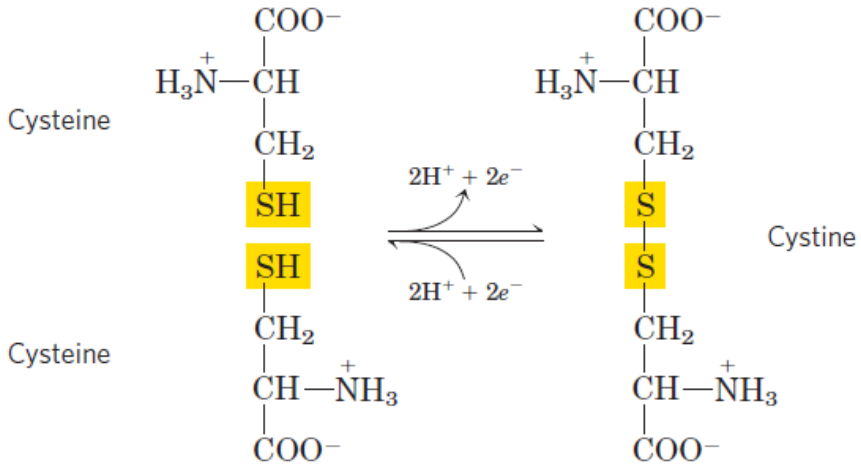
❖This class of amino acids includes **serine**, **threonine**, **cysteine**, **asparagine**, and **glutamine**.

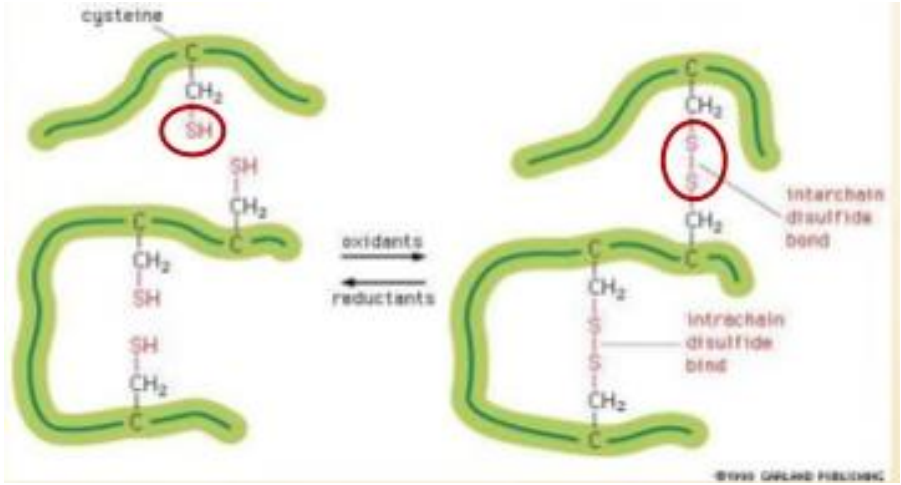
❖The polarity of serine and threonine is contributed by their hydroxyl groups; that of cysteine by its sulfhydryl group; and that of asparagine and glutamine by their amide groups.

❖Asparagine and glutamine are the amides of two other amino acids also found in proteins, aspartate and glutamate, respectively, to which asparagine and glutamine are easily hydrolyzed by acid or base.

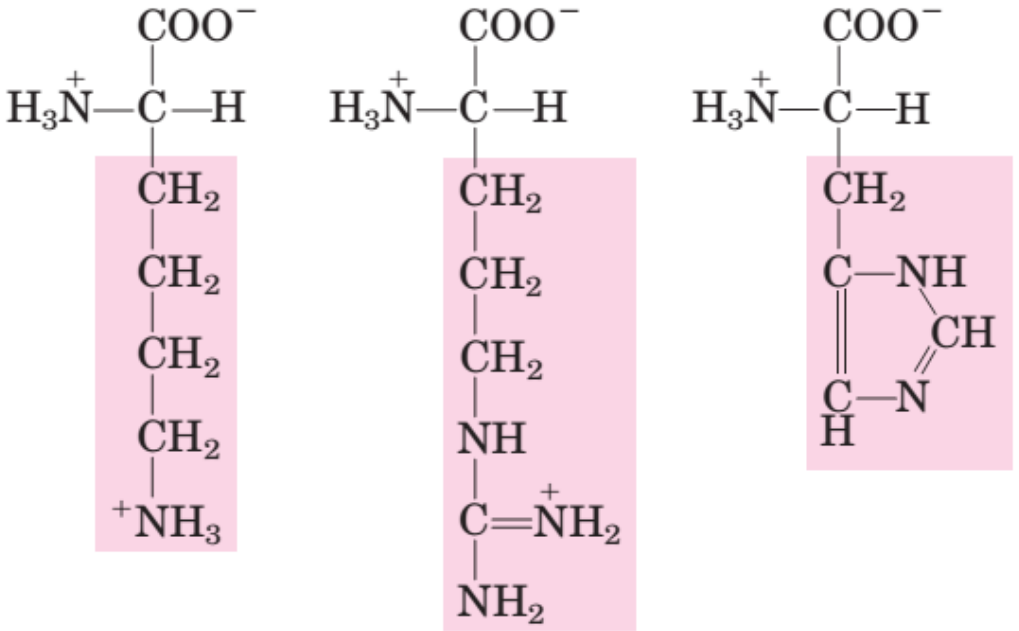
❖Cysteine is readily oxidized to form a covalently linked dimeric amino acid called cystine, in which two cysteine molecules or residues are joined by a disulfide bond.

❖The disulfide-linked residues are strongly hydrophobic (nonpolar). Disulfide bonds play a special role in the structures of many proteins by forming covalent links between parts of a protein molecule or between two different polypeptide chains



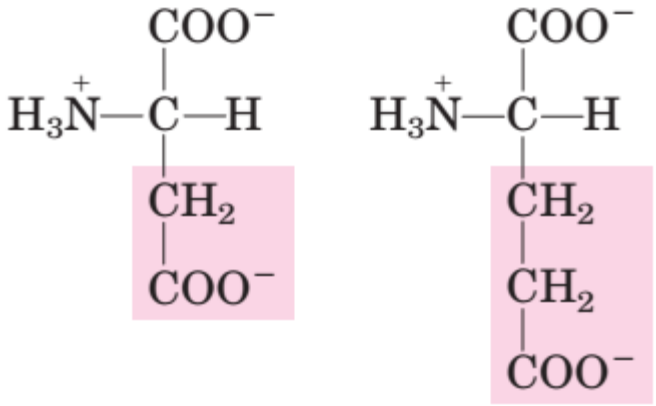


Positively charged R groups



Lysine Arginine Histidine

Negatively charged R groups



Aspartate Glutamate

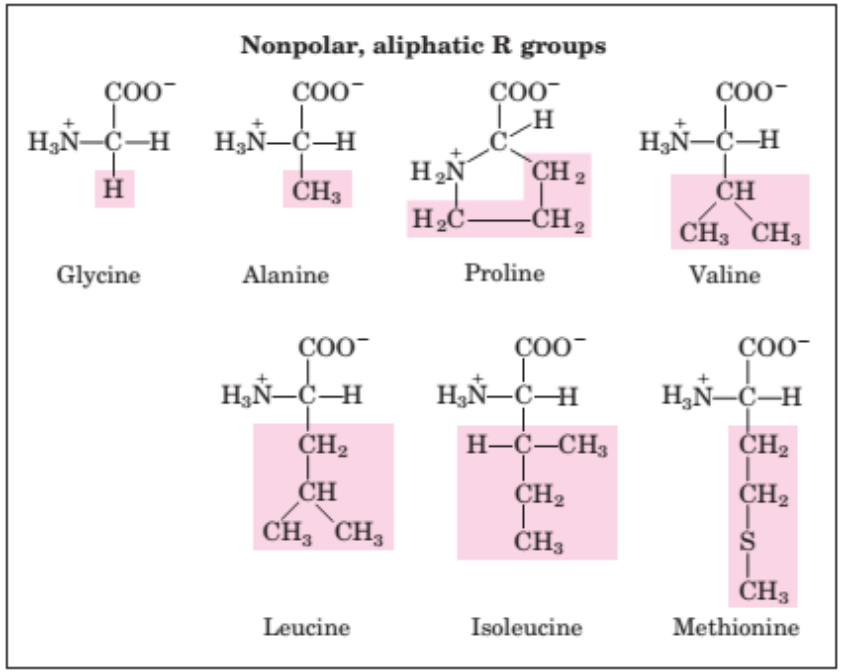
Amino acids with polar, charged R Groups

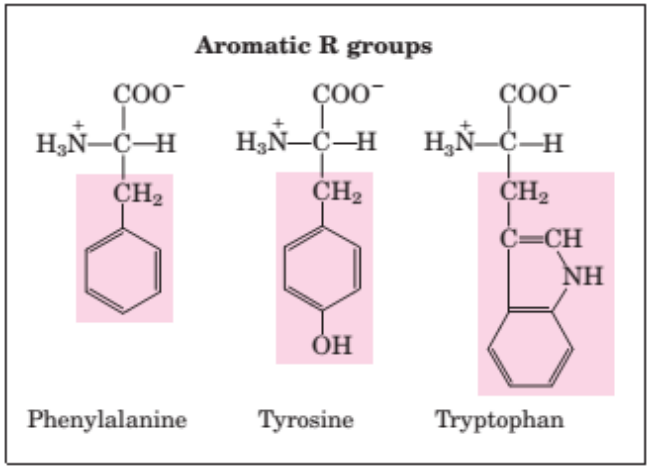
❖The most hydrophilic R groups are those that are either positively or negatively charged.

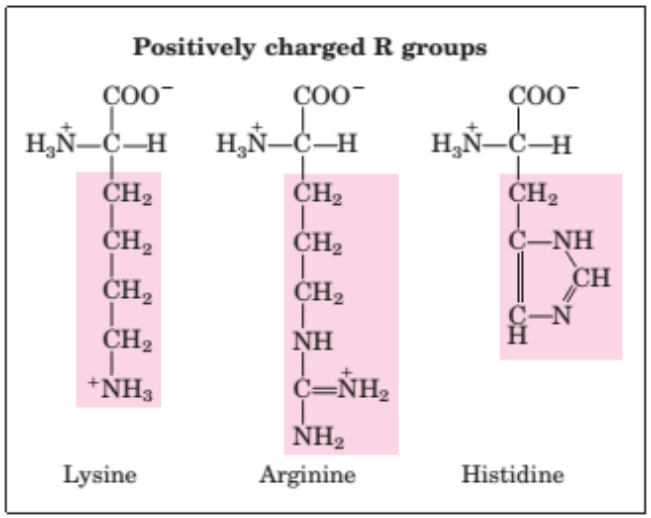
❖The amino acids in which the R groups have significant positive charge at pH 7.0 are **lysine**, which has a second primary amino group at the position on its aliphatic chain; **arginine**, which has a positively charged guanidinium group; and histidine, which has an imidazole group.

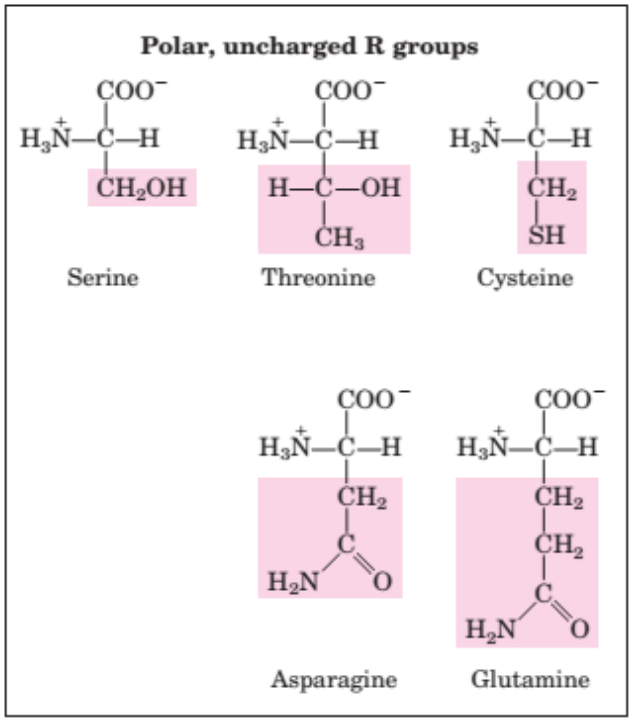
❖**Histidine** is the only common amino acid having an ionizable side chain with a pKa near neutrality. In many enzyme-catalyzed reactions, a His residue facilitates the reaction by serving as a proton donor/acceptor.

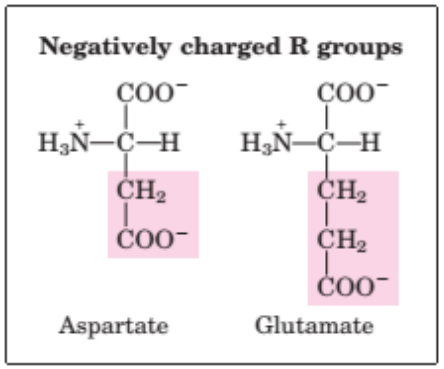
❖Negatively Charged (Acidic) R Groups: The two amino acids having R groups with a net negative charge at pH 7.0 are **aspartate** and **glutamate**, each of which has a second carboxyl group.


Nonpolar, aliphatic R groups
Glycine
Alanine
Proline
Valine
Leucine
Isoleucine
Methionine


Aromatic R groups
Phenylalanine
Tyrosine
Tryptophan


Positively charged R groups
Lysine
Arginine
Histidine


Polar, uncharged R groups
Serine
Threonine
Cysteine
Asparagine
Glutamine


Negatively charged R groups
Aspartate
Glutamate

❖There are alternative ways to classify the 20 common amino acids.

❖For example, it would be reasonable to imagine that the amino acids could be described as hydrophobic, hydrophilic, or amphipathic:,

Hydrophobic:		**Hydrophilic:**		**Amphipathic:**
Alanine	Proline	Arginine	Glutamine	Lysine
Glycine	Valine	Asparagine	Histidine	Methionine
Isoleucine		Aspartic acid	Serine	Tryptophan
Leucine		Cysteine	Threonine	Tyrosine
Phenylalanine		Glutamic acid		

❖Methionine is the least polar of the amphipathic amino acids, but its thioether sulfur can be an effective metal ligand in proteins.

❖Cysteine can deprotonate at pH values greater than 7, and the thiolate anion is the most potent nucleophile that can be generated among the 20 common acids.

❖In addition to the 20 common amino acids, proteins may contain residues created by modification of common residues already incorporated into a polypeptide.

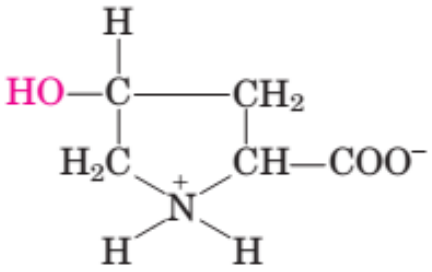
❖Among these uncommon amino acids are **4-hydroxyproline**, a derivative of proline, and **5-hydroxylysine**, derived from lysine.

❖The former is found in plant cell wall proteins, and both are found in collagen, a fibrous protein of connective tissues.

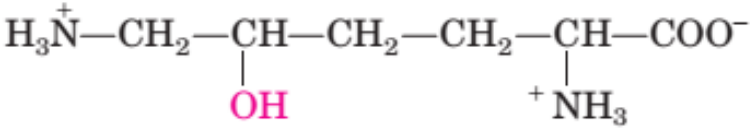
❖**6-N-Methyllysine** is a constituent of myosin, a contractile protein of muscle.

❖Some 300 additional amino acids have been found in cells. They have a variety of functions but are not constituents of proteins.

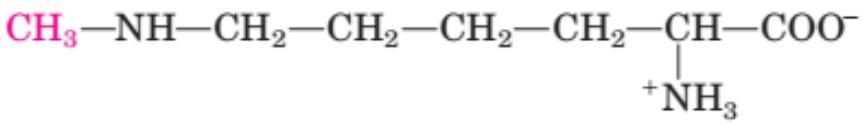
❖ **Ornithine** and **citrulline** deserve special note because they are key intermediates (metabolites) in the biosynthesis of arginine and in the urea cycle.



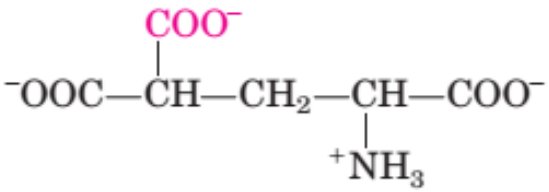
4-Hydroxyproline



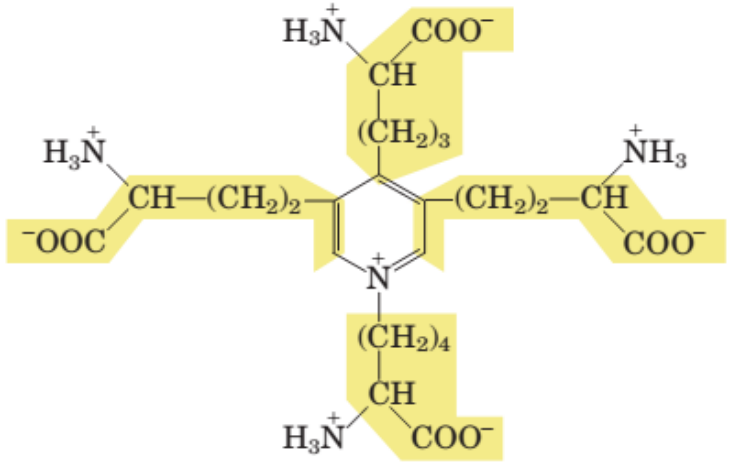
5-Hydroxylysine



6-*N*-Methyllysine

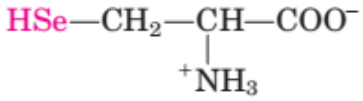


γ-Carboxyglutamate

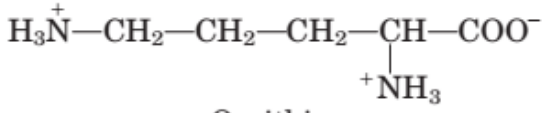


Desmosine

(a)

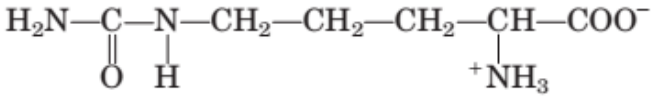


Selenocysteine



Ornithine

(b)



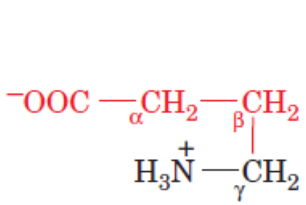
Citrulline

❖Amino acids and their derivatives often function as chemical messengers for communication between cells.

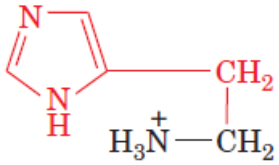
❖For example, glycine, γ-aminobutyric acid (GABA), and dopamine (a tyrosine derivative) are neurotransmitters, substances released by nerve cells to alter the behavior of their neighbors.

❖Histamine (the decarboxylation product of histidine) is a potent local mediator of allergic reactions.

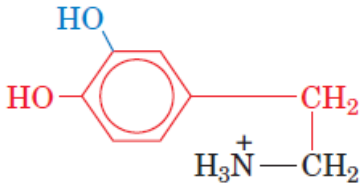
❖Thyroxine (another tyrosine derivative) is an iodine-containing thyroid hormone that generally stimulates vertebrate metabolism.



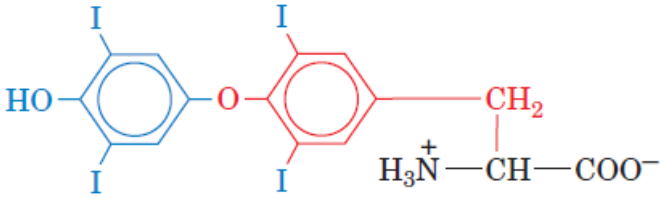
γ-Aminobutyric acid (GABA)



Histamine

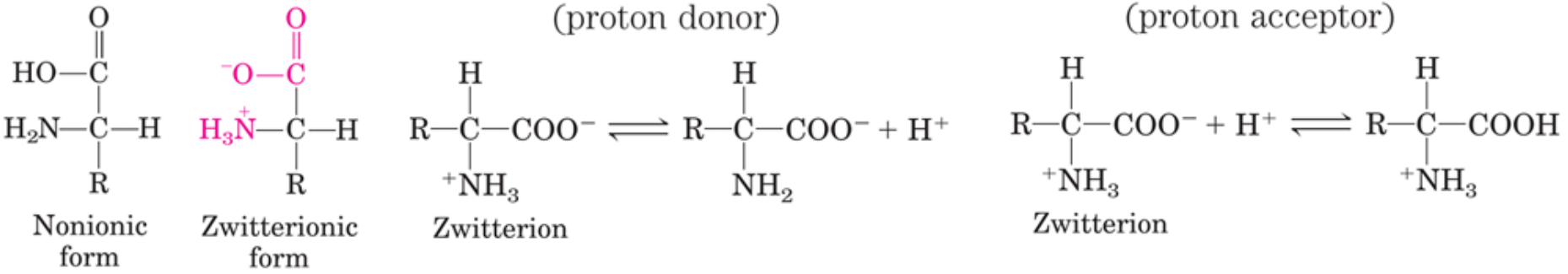


Dopamine



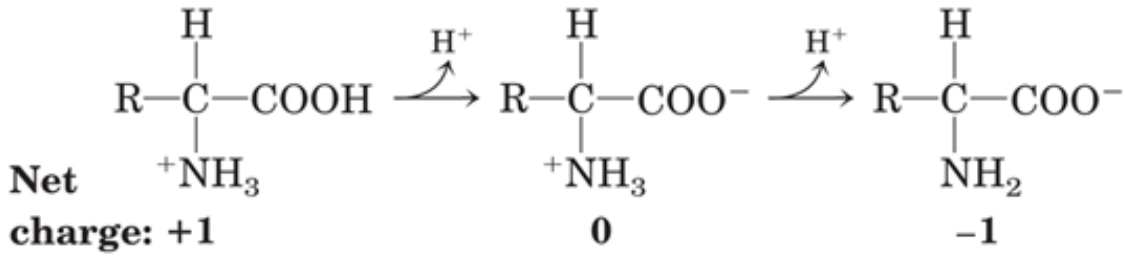
Thyroxine

❖When an amino acid is dissolved in water, it exists in solution as the dipolar ion, or **zwitterion** (German for "hybrid ion"). A zwitterion can act as either an acid (proton donor) or a base (proton acceptor).



❖Substances having this dual nature are **amphoteric** and are often called **ampholytes** (from "amphoteric electrolytes").

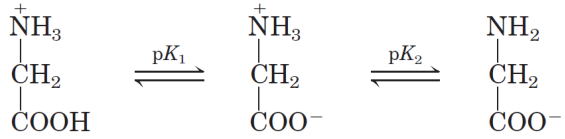
❖A simple monoamino monocarboxylic α-amino acid, such as alanine, is a diprotic acid when fully protonated; it has two groups, the —COOH group and the —NH_3^+ group, that can yield protons:



❖Acid-base titration involves the gradual addition or removal of protons.

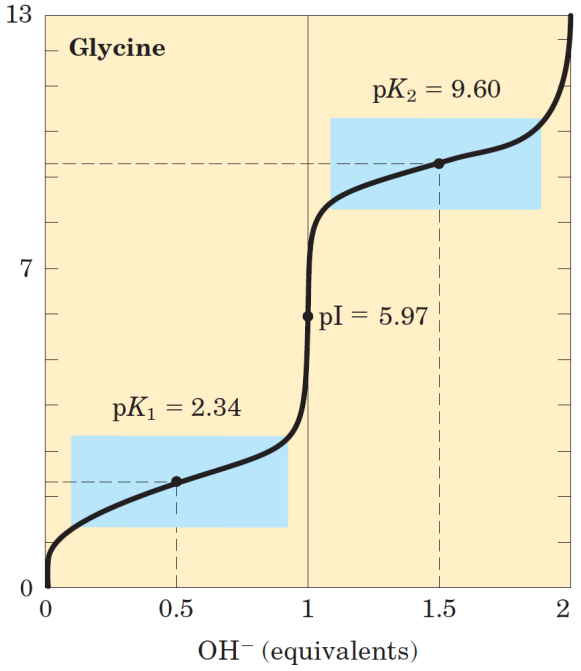
❖The plot of titration of glycine has two distinct stages, corresponding to deprotonation of two different groups on glycine. Each of the two stages resembles in shape the titration curve of a monoprotic acid, such as acetic acid.

❖At very low pH, the predominant ionic species of glycine is the fully protonated form.



❖In the first stage of the titration, the —COOH group of glycine loses its proton.

❖At the midpoint of this stage, equimolar concentrations of the proton-donor and proton-acceptor species are present.



❖As in the titration of any weak acid, a point of inflection is reached at this midpoint where the pH is equal to the pKa of the protonated group being titrated.

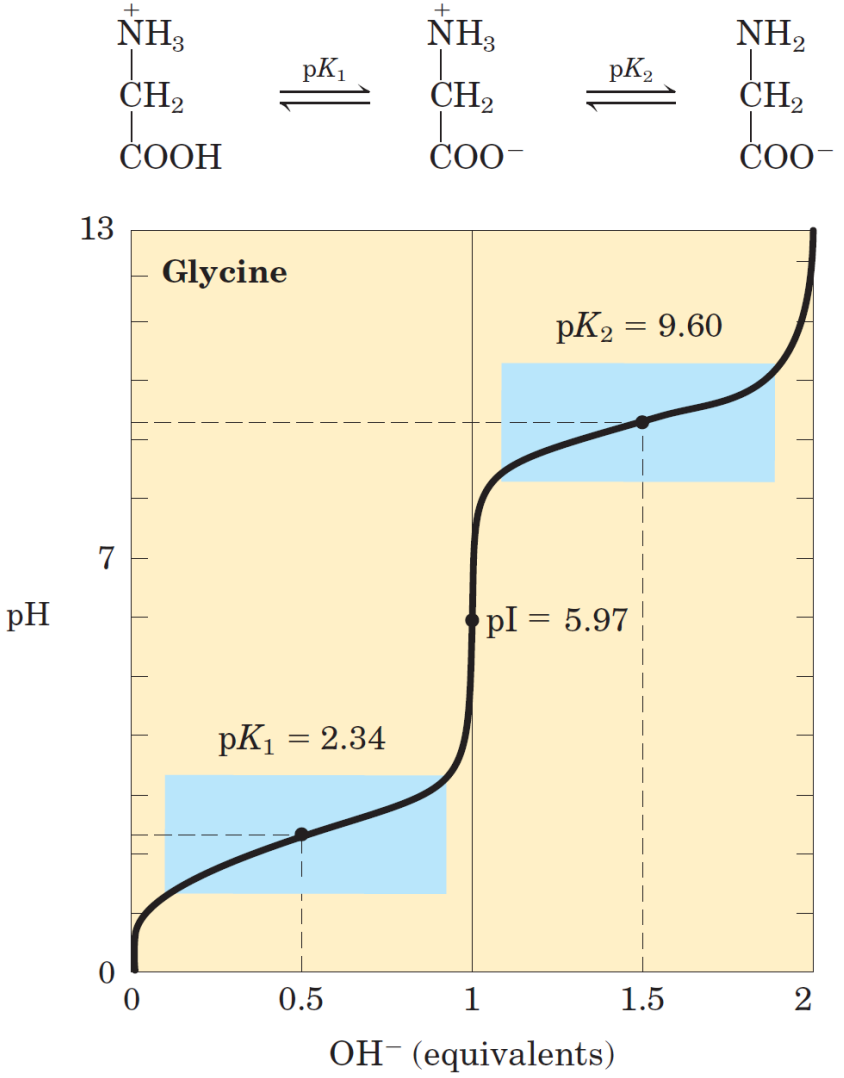
❖For glycine, the pH at the midpoint is 2.34, thus its —COOH group has a pKa of 2.34.

❖The pKa is a measure of the tendency of a group to give up a proton.

❖As the titration of glycine proceeds, another important point is reached at pH 5.97.

❖Here there is another point of inflection, at which removal of the first proton is essentially complete and removal of the second has just begun.

❖At this pH glycine is present largely as the dipolar ion (zwitterion).



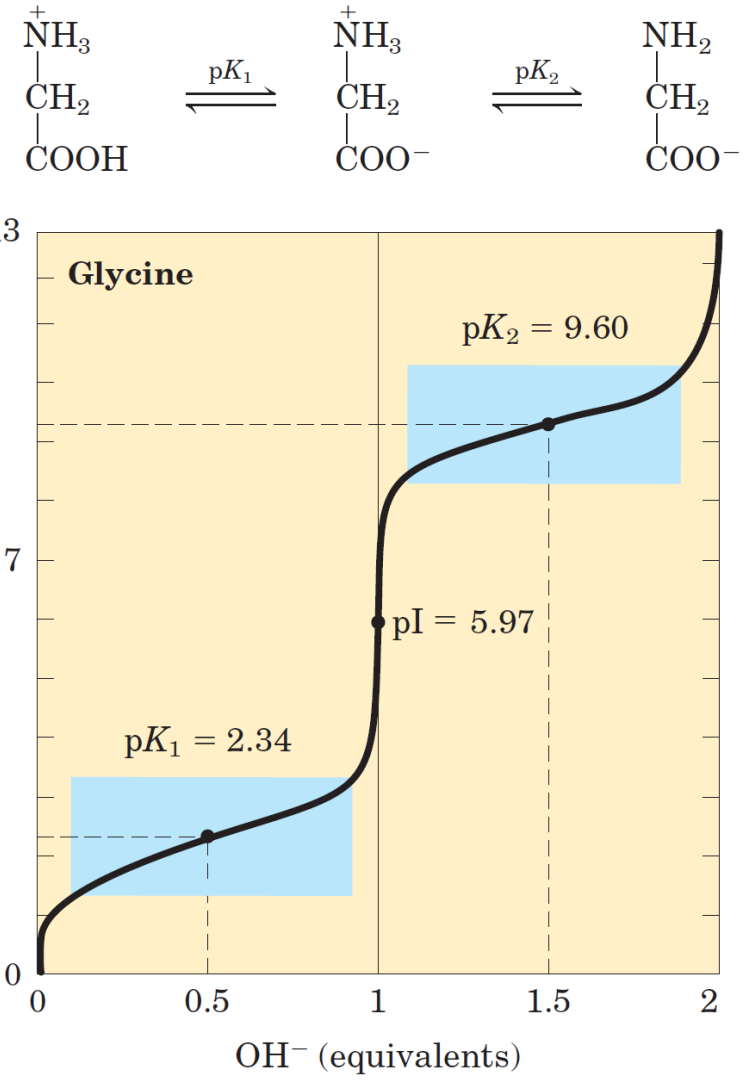
❖The second stage of the titration corresponds to the removal of a proton from the $—NH_3^+$ group of glycine.

❖The pH at the midpoint of this stage is 9.60, equal to the pKa for the $—NH_3^+$ group.

❖The titration is essentially complete at a pH of about 12, at which point the predominant form of glycine is negatively charged form.

❖From the titration curve of glycine we can derive several important pieces of information.

❖First, it gives a quantitative measure of the pKa of each of the two ionizing groups

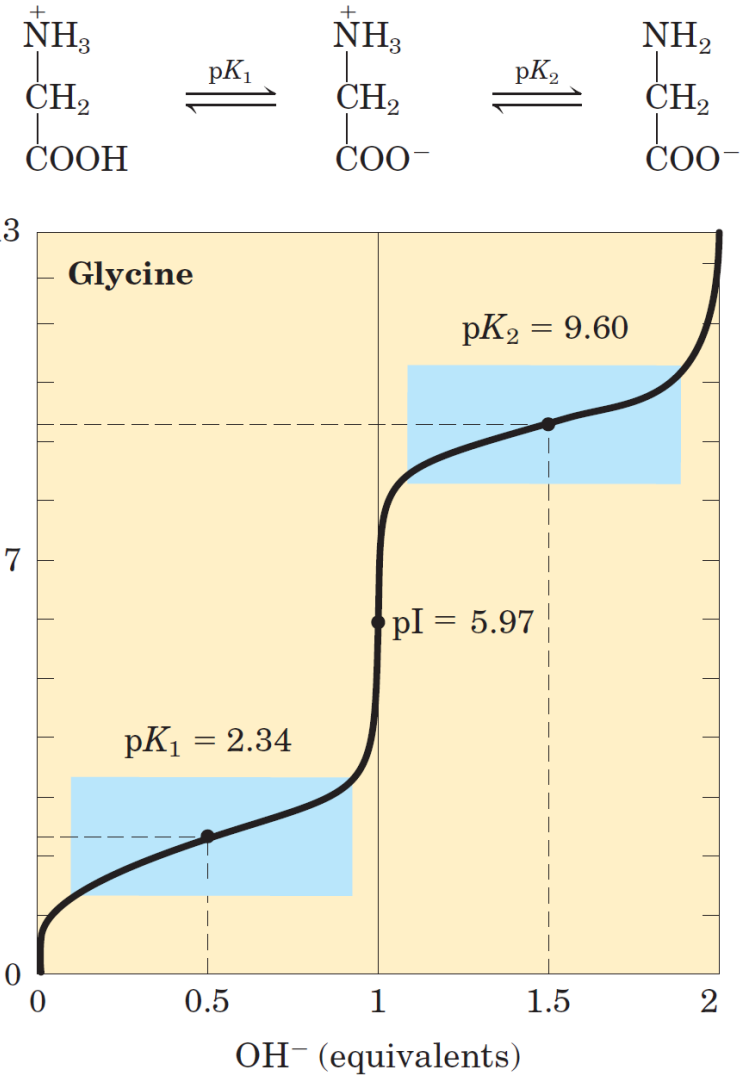


❖The second piece of information provided by the titration curve of glycine is that this amino acid has two regions of buffering power.

❖One of these is the relatively flat portion of the curve, extending for approximately 1 pH unit on either side of the first pKa of 2.34, indicating that glycine is a good buffer near this pH.

❖The other buffering zone is centered around pH 9.60.

❖Within the buffering ranges of glycine, the Henderson-Hasselbalch equation can be used to calculate the proportions of proton-donor and proton-acceptor species of glycine required to make a buffer at a given pH.

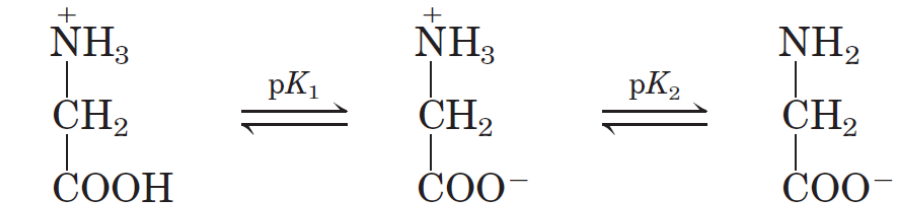


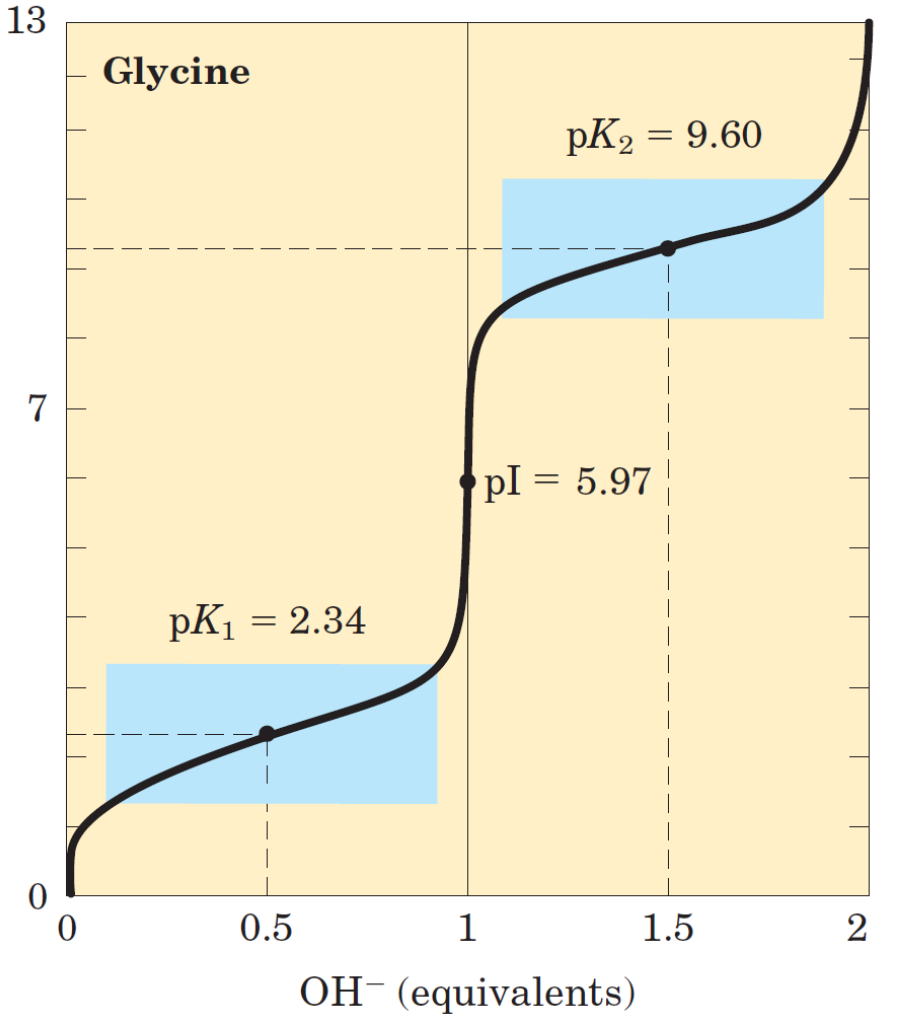
❖Another important piece of information derived from the titration curve of an amino acid is the relationship between its net charge and the pH of the solution.

❖At pH 5.97, the point of inflection between the two stages in its titration curve, glycine is present predominantly as its dipolar form, fully ionized but with no net electric charge.

❖The characteristic pH at which the net electric charge is zero is called the **isoelectric point** or **isoelectric pH**, designated **pI**.

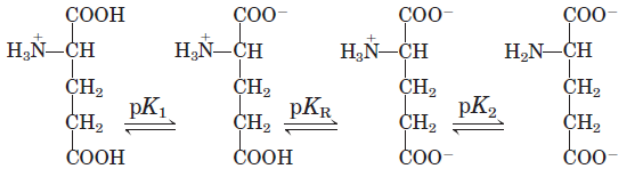
$$\text{pI} = \frac{1}{2}(\text{p}K_1 + \text{p}K_2) = \frac{1}{2}(2.34 + 9.60) = 5.97$$

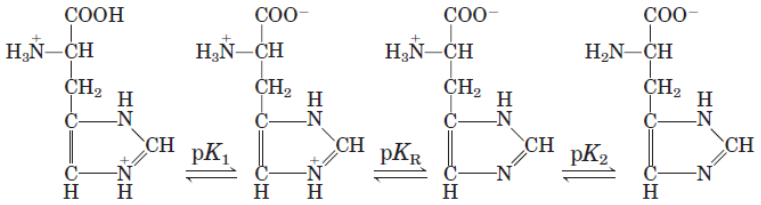


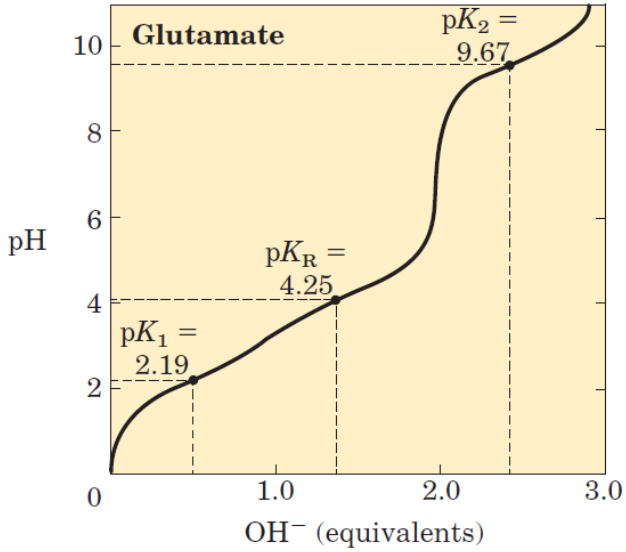


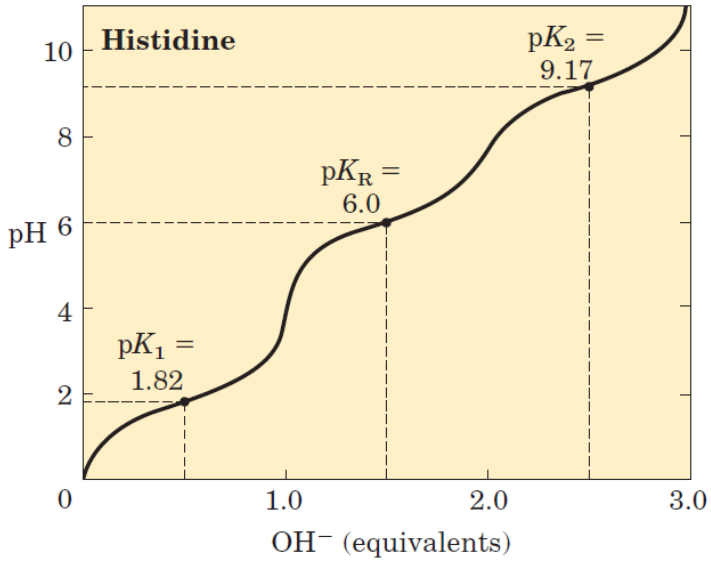
❖ All amino acids with a single α-amino group, a single α-carboxyl group, and an R group that does not ionize have titration curves resembling that of glycine. These amino acids have very similar, although not identical, pKa values.

❖Amino acids with an ionizable R group have more complex titration curves, with three stages corresponding to the three possible ionization steps

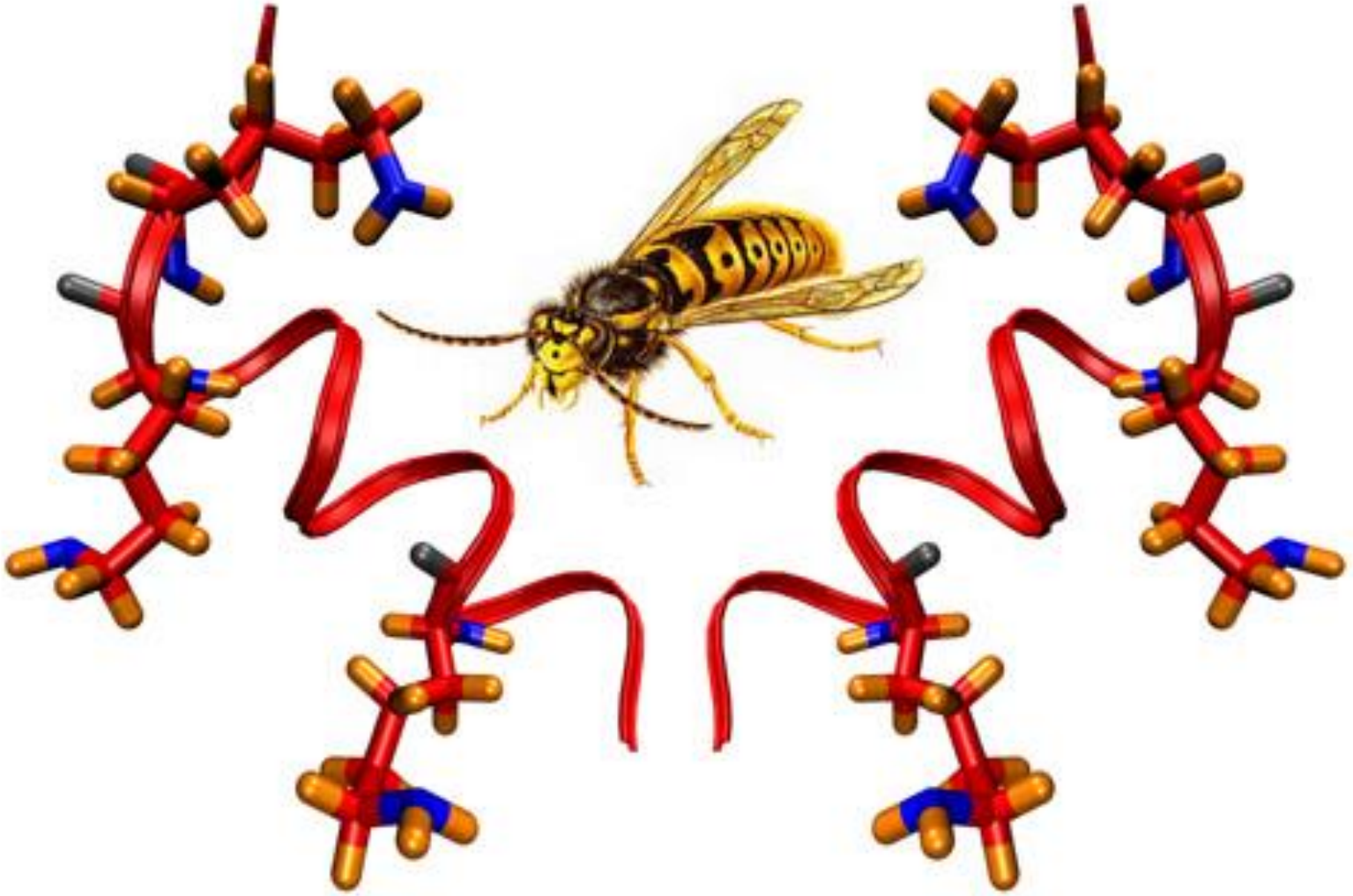








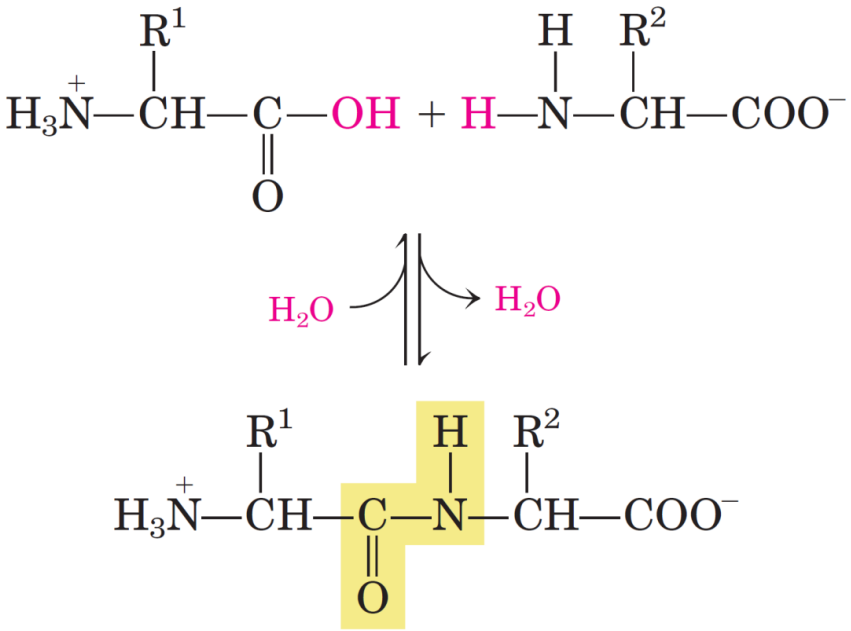
AMINO ACIDS, PEPTIDES AND PROTEINS II:
Peptides and Proteins

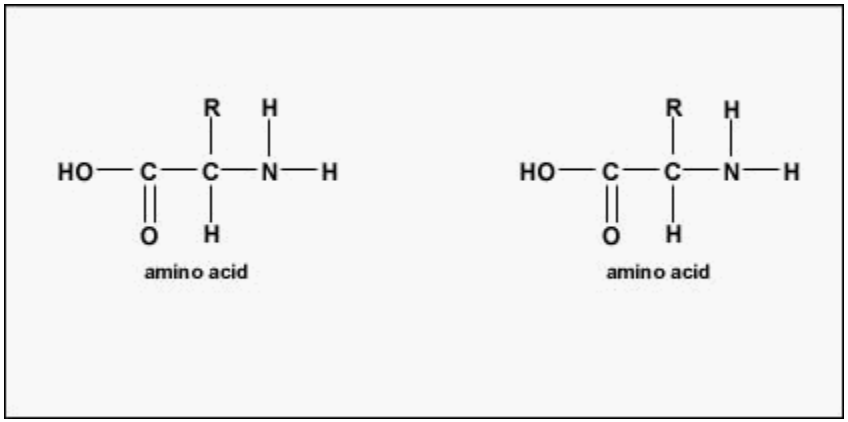


❖Peptides are chains of amino acids.

❖Two amino acid molecules can be covalently joined through a substituted amide linkage, termed a **peptide bond**, to yield a dipeptide.

❖Such a linkage is formed by removal of the elements of water (dehydration) from the α-carboxyl group of one amino acid and the α -amino group of another.





❖Peptide bond formation is an example of a condensation reaction.

❖Under standard biochemical conditions, the equilibrium for the reaction favors the amino acids over the dipeptide.

❖To make the reaction thermodynamically more favorable, the carboxyl group must be chemically modified or activated so that the hydroxyl group can be more readily eliminated.

❖Three amino acids can be joined by two peptide bonds to form a **tripeptide**; similarly, amino acids can be linked to form **tetrapeptides**, **pentapeptides**, and so forth.

❖When a few amino acids are joined in this fashion, the structure is called an **oligopeptide**.

❖When many amino acids are joined, the product is called a **polypeptide**.

❖Molecules referred to as **proteins** generally consist of 100 or more amino acids.

❖However, polypeptide chains which have lower than 100 amino acids can also be termed as proteins as long as they have definite three dimensional conformations and functions.

❖Although the terms “protein” and “polypeptide” are sometimes used interchangeably, molecules referred to as polypeptides generally have molecular weights below 10,000, and those called proteins have higher molecular weights.

❖An amino acid unit in a peptide is often called a residue (the part left over after losing a hydrogen atom from its amino group and the hydroxyl moiety from its carboxyl group).

❖In a peptide, the amino acid residue at the end with a free -amino group is the amino-terminal (or N-terminal) residue; the residue at the other end, which has a free carboxyl group, is the carboxyl-terminal (C-terminal) residue.

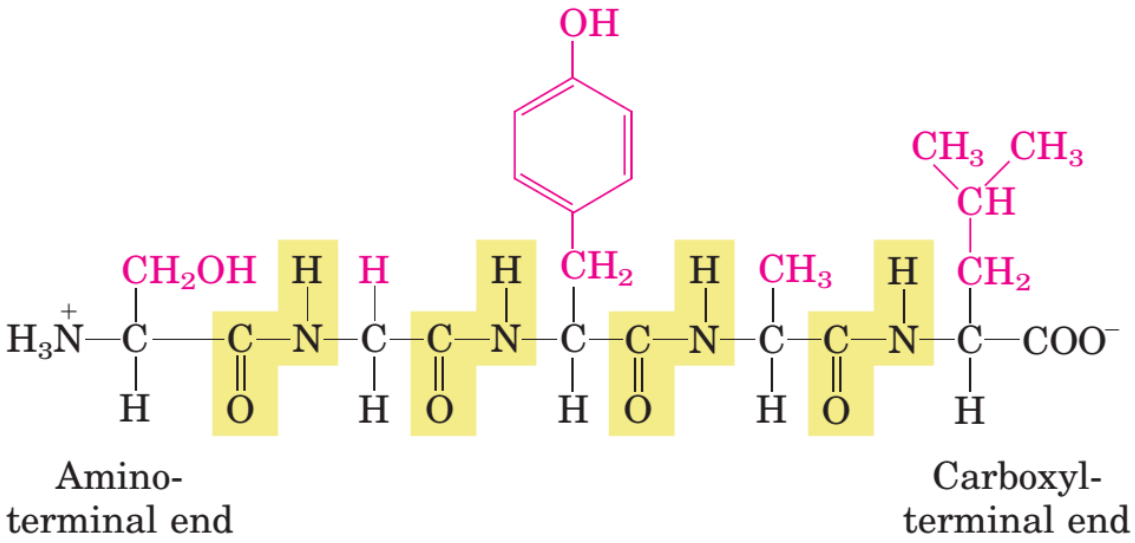
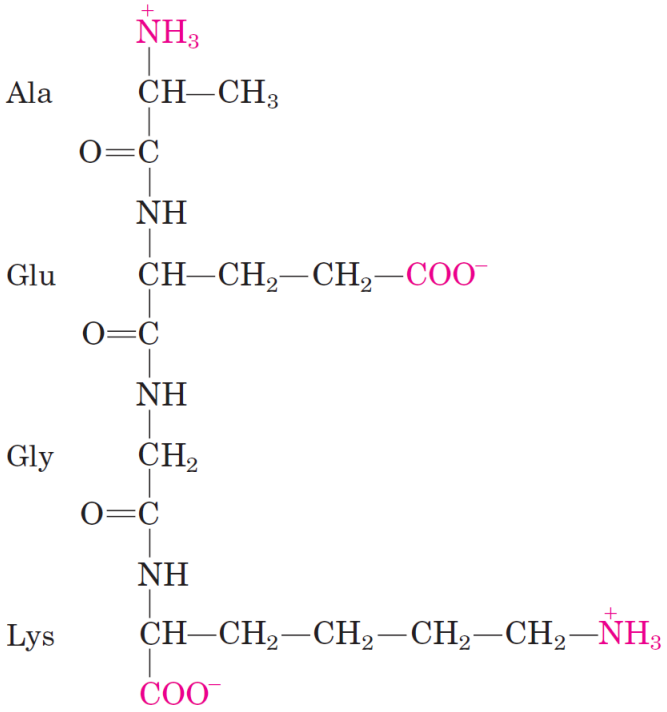


FIGURE 3–14 The pentapeptide serylglycyltyrosylalanylleucine, or Ser–Gly–Tyr–Ala–Leu. Peptides are named beginning with the amino-terminal residue, which by convention is placed at the left. The peptide bonds are shaded in yellow; the R groups are in red.



❖Peptides contain only one free α-amino group and one free α-carboxyl group, at opposite ends of the chain.

❖ These groups ionize as they do in free amino acids.

❖The R groups of some amino acids can also ionize.

❖The α-amino and α-carboxyl groups of all nonterminal amino acids are covalently joined in the peptide bonds, which do not ionize and thus do not contribute to the total acid-base behavior of peptides.

❖Like free amino acids, peptides have characteristic titration curves and a characteristic isoelectric pH (pI) at which they do not move in an electric field.

❖No generalizations can be made about the molecular weights of biologically active peptides and proteins in relation to their functions.

❖Many small peptides exert their effects at very low concentrations. For example, a number of vertebrate hormones are small peptides.

❖These include oxytocin (nine amino acid residues), which is secreted by the posterior pituitary and stimulates uterine contractions; bradykinin (nine residues), which inhibits inflammation of tissues.

❖Some extremely toxic mushroom poisons, such as amanitin, are also small peptides, as are many antibiotics.

❖Slightly larger are small polypeptides and oligopeptides such as the pancreatic hormone insulin and glucagon.

❖Molecular weights of proteins also very diverse.

❖Some proteins consist of a single polypeptide chain, but others, called **multisubunit proteins**, have two or more polypeptides associated noncovalently.

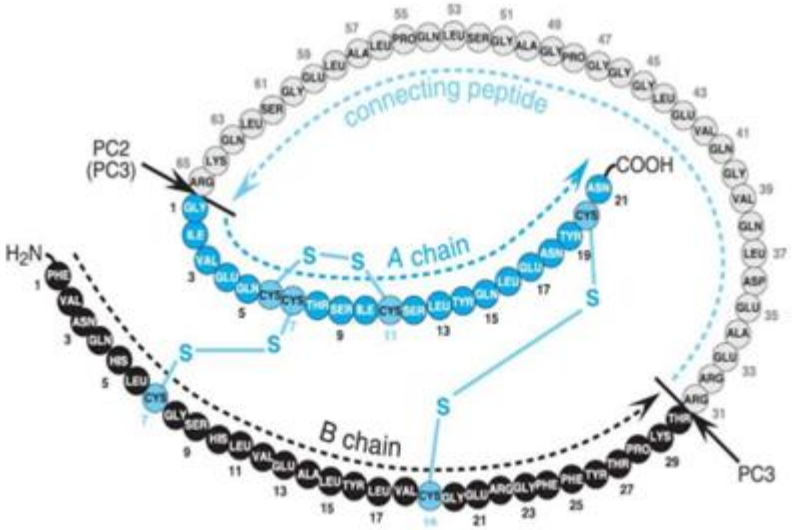
❖The individual polypeptide chains in a multisubunit protein may be identical or different.

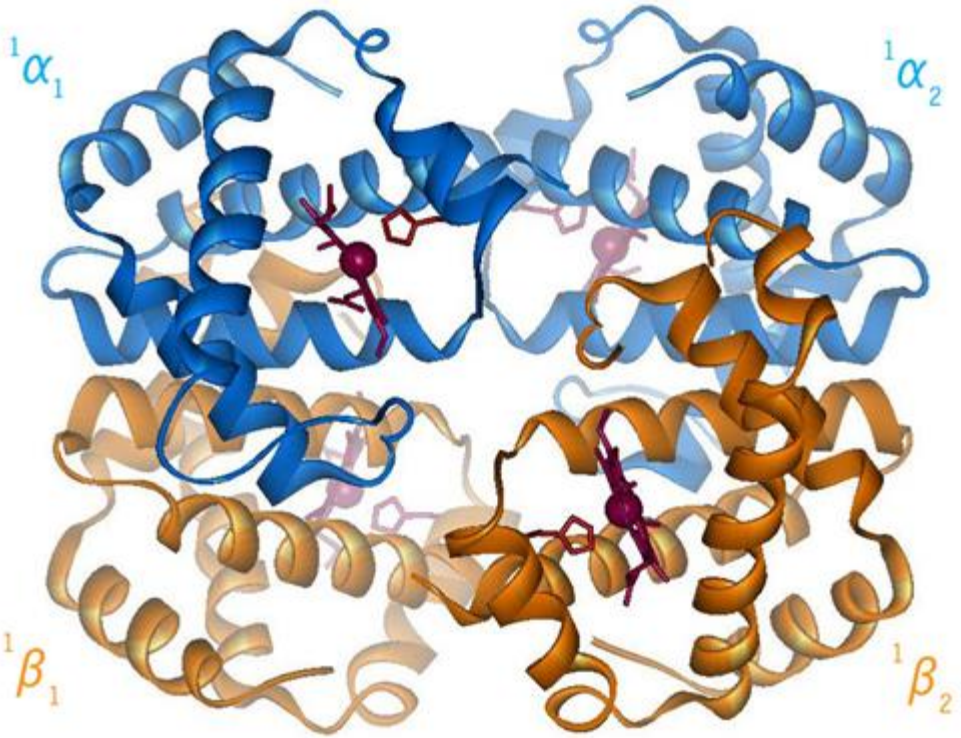
❖If at least two are identical the protein is said to be **oligomeric**, and the identical units (consisting of one or more polypeptide chains) are referred to as **protomers**.

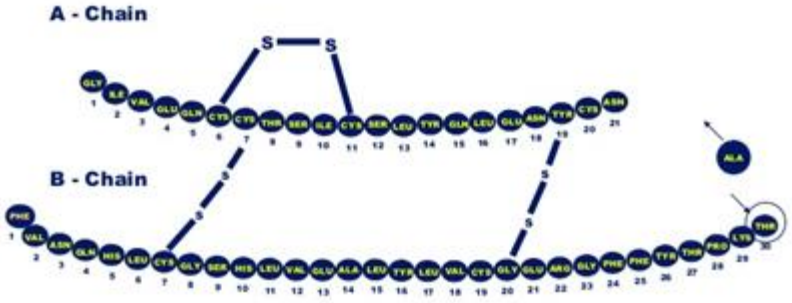
❖Hemoglobin, for example, has four polypeptide subunits: two identical α chains and two identical β chains, all four held together by noncovalent interactions. So, hemoglobin can be considered either a tetramer of four polypeptide subunits or a dimer of αβ protomers

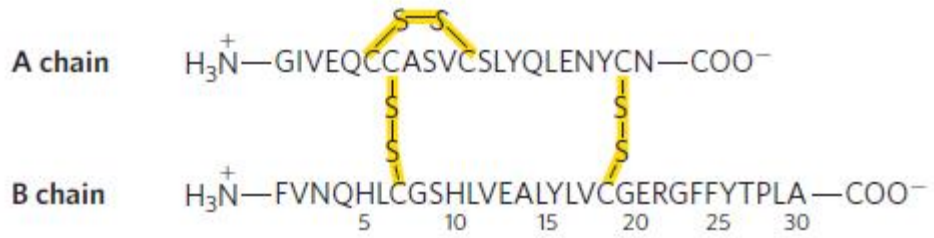
❖A few proteins contain two or more polypeptide chains linked covalently. For example, the two polypeptide chains of insulin are linked by disulfide bonds. In such cases, the individual polypeptides are not considered subunits but are commonly referred to simply as **chains**.

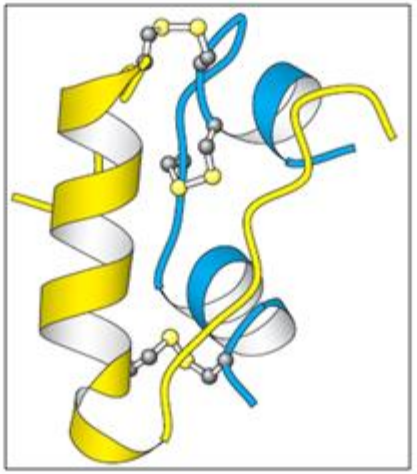
Insulin structure











❖The amino acid composition of proteins is also highly variable.

❖The 20 common amino acids almost never occur in equal amounts in a protein. Some amino acids may occur only once or not at all in a given type of protein

❖Hydrolysis of peptides or proteins with acid yields a mixture of free α-amino acids. When completely hydrolyzed, each type of protein yields a characteristic proportion or mixture of the different amino acids.

❖Complete hydrolysis alone is not sufficient for an exact analysis of amino acid composition, however, because some side reactions occur during the procedure.

❖For example, the amide bonds in the side chains of asparagine and glutamine are cleaved by acid treatment, yielding aspartate and glutamate, respectively. The side chain of tryptophan is almost completely degraded by acid hydrolysis

❖Some proteins contain only amino acids, these are considered simple proteins.

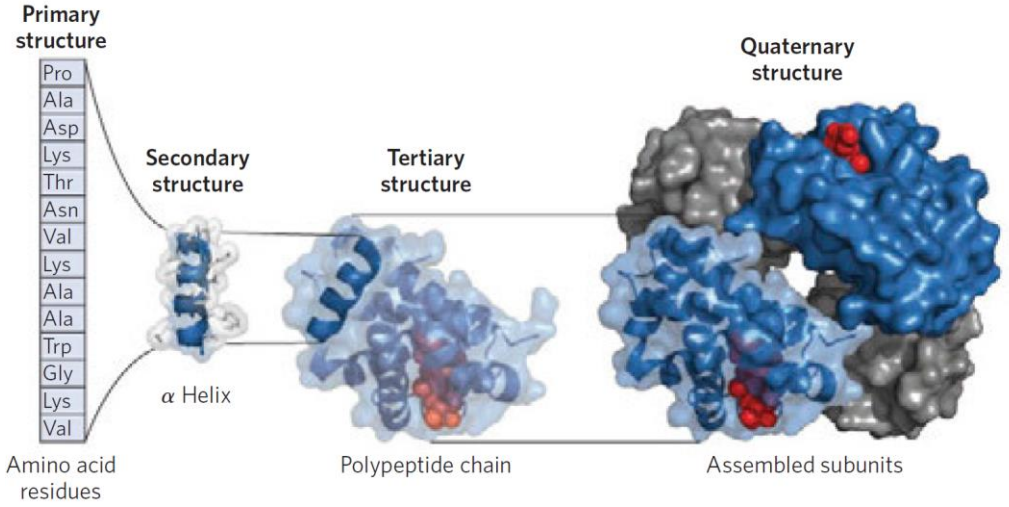
❖However, some proteins contain permanently associated chemical components in addition to amino acids; these are called **conjugated proteins**.

❖The non–amino acid part of a conjugated protein is usually called its **prosthetic group**.

❖Conjugated proteins are classified on the basis of the chemical nature of their prosthetic groups; for example, **lipoproteins** contain lipids and **glycoproteins** contain sugar groups, and **metalloproteins** contain a specific metal.

❖Some proteins contain more than one prosthetic group and usually the prosthetic group plays an important role in the protein's biological function.

❖Proteins have to fold three dimensional structures to make its function. Four levels of protein structure are commonly defined



❖A description of all covalent bonds (mainly peptide bonds and disulfide bonds) linking amino acid residues in a polypeptide chain is its **primary structure**.

❖The most important element of primary structure is the sequence of amino acid residues.

❖**Secondary structure** refers to particularly stable arrangements of amino acid residues giving rise to recurring structural patterns.

❖**Tertiary structure** describes all aspects of the three-dimensional folding of a polypeptide.

❖When a protein has two or more polypeptide subunits, their arrangement in space is referred to as **quaternary structure**.

❖Differences in primary structure can be especially informative.

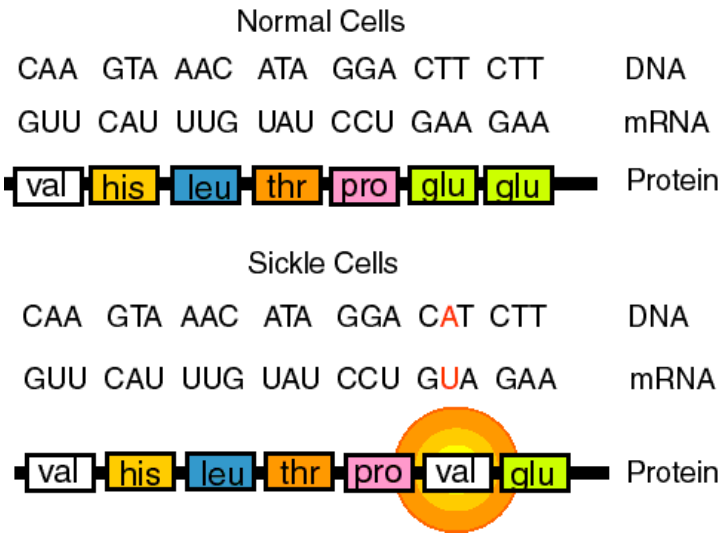
❖Each protein has a distinctive number and sequence of amino acid residues.

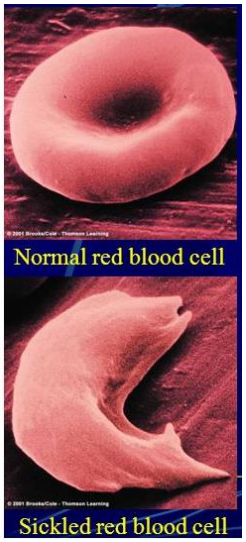
❖The primary structure of a protein determines how it folds up into its unique three-dimensional structure, and this in turn determines the function of the protein.

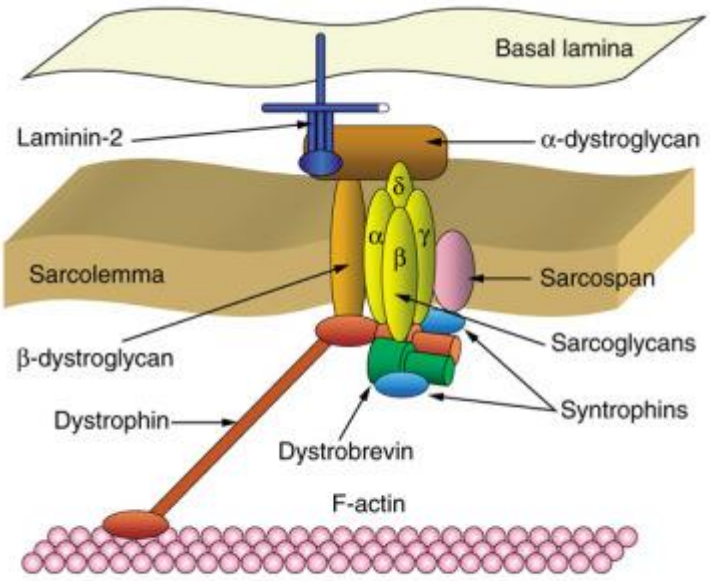
❖Proteins with different functions always have different amino acid sequences.

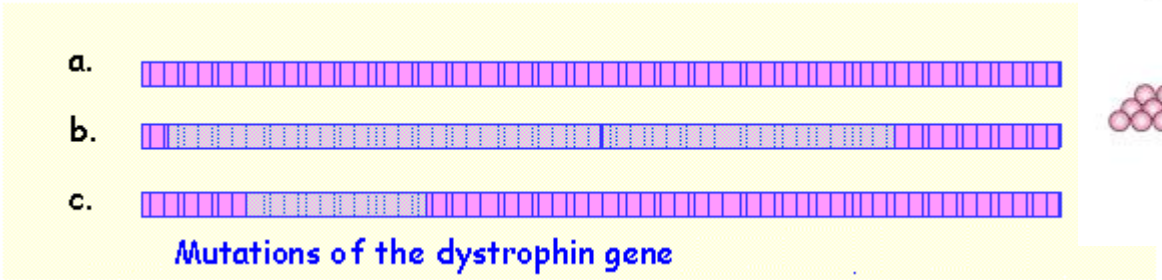
❖Thousands of human genetic diseases have been traced to the production of defective proteins. Perhaps one-third of these proteins are defective because of a single change in their amino acid sequence.

❖The defect can range from a single change in the amino acid sequence (as in sickle cell anemia) to deletion of a larger portion of the polypeptide chain (as in most cases of Duchenne muscular dystrophy).









❖On comparing functionally similar proteins from different species, we find that these proteins often have similar amino acid sequences.

❖The amino acid sequence is not absolutely fixed for a particular protein.

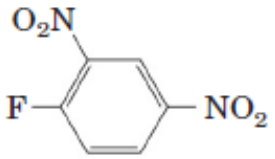
❖An estimated 20% to 30% of the proteins in humans are **polymorphic**, having amino acid sequence variants in the human population.

❖Many of these variations in sequence have little or no effect on the function of the protein.

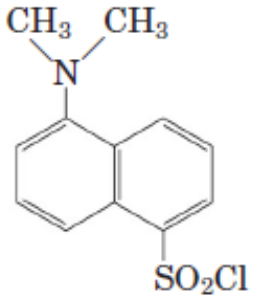
❖Furthermore, proteins that carry out a broadly similar function in distantly related species can differ greatly in overall size and amino acid sequence.

❖Although the amino acid sequence in some regions of the primary structure might vary considerably without affecting biological function, most proteins contain crucial regions that are essential to their function and whose sequence is therefore conserved.

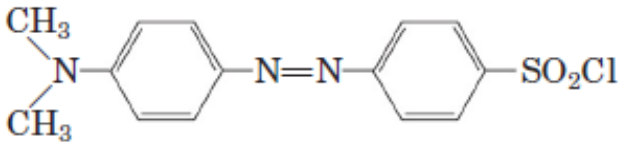
❖The amino acid sequences of thousands of different proteins from many species have been determined using principles first developed by Sanger.



FDNB



Dansyl chloride



Dabsyl chloride

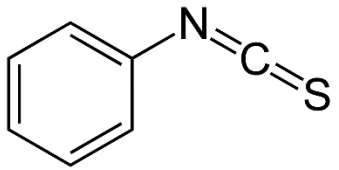
❖Sanger developed the reagent 1-fluoro-2,4-dinitrobenzene (FDNB) for this purpose; other reagents used to label the amino-terminal residue, dansyl chloride and dabsyl chloride, yield derivatives that are more easily detectable than the dinitrophenyl derivatives.

❖After the amino-terminal residue is labeled with one of these reagents, the polypeptide is hydrolyzed to its constituent amino acids and the labeled amino acid is identified.

❖Because the hydrolysis stage destroys the polypeptide, this procedure cannot be used to sequence a polypeptide beyond its amino-terminal residue.

❖The chemical sequencing process itself is based on a two-step process developed by Pehr Edman.

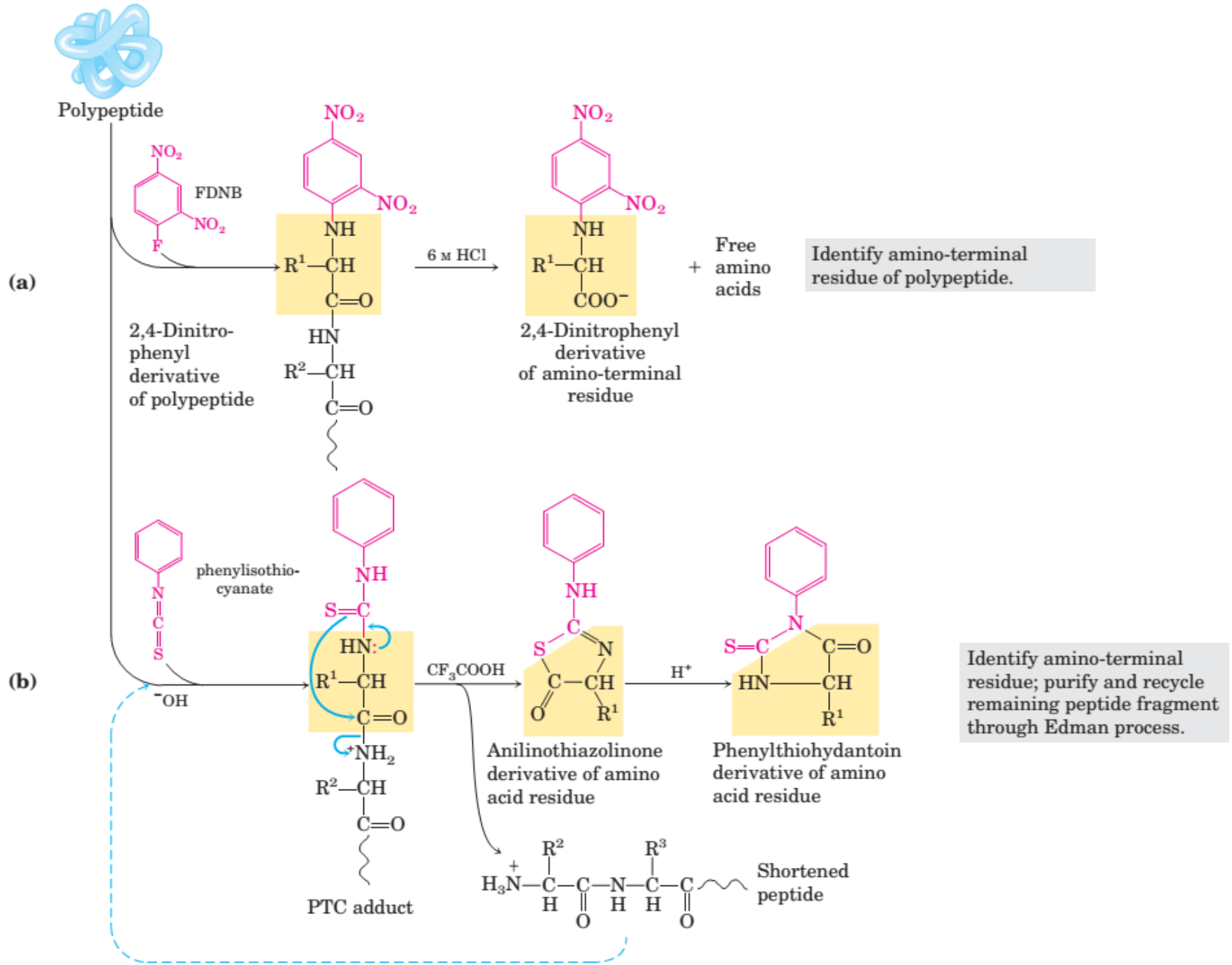
❖The Edman degradation procedure labels and removes only the amino-terminal residue from a peptide, leaving all other peptide bonds intact. The peptide is reacted with phenylisothiocyanate under mildly alkaline conditions, which converts the amino-terminal amino acid to a phenylthiocarbamoyl (PTC) adduct.



❖The peptide bond next to the PTC adduct is then cleaved in a step carried out in anhydrous trifluoroacetic acid, with removal of the aminoterminal amino acid as an anilinothiazolinone derivative.

❖The derivatized amino acid is extracted with organic solvents, converted to the more stable phenylthiohydantoin derivative by treatment with aqueous acid, and then identified.

❖The use of sequential reactions carried out under first basic and then acidic conditions provides a means of controlling the entire process.


Polypeptide
(a)
FDNB
2,4-Dinitrophenyl derivative of polypeptide
6 M HCl
2,4-Dinitrophenyl derivative of amino-terminal residue
+ Free amino acids
Identify amino-terminal residue of polypeptide.
(b)
phenylisothiocyanate
$^-$OH
PTC adduct
CF_3COOH
Anilinothiazolinone derivative of amino acid residue
H^+
Phenylthiohydantoin derivative of amino acid residue
Shortened peptide
Identify amino-terminal residue; purify and recycle remaining peptide fragment through Edman process.

❖To determine the sequence of large proteins, early developers of sequencing protocols had to devise methods to eliminate disulfide bonds and to cleave proteins precisely into smaller polypeptides.

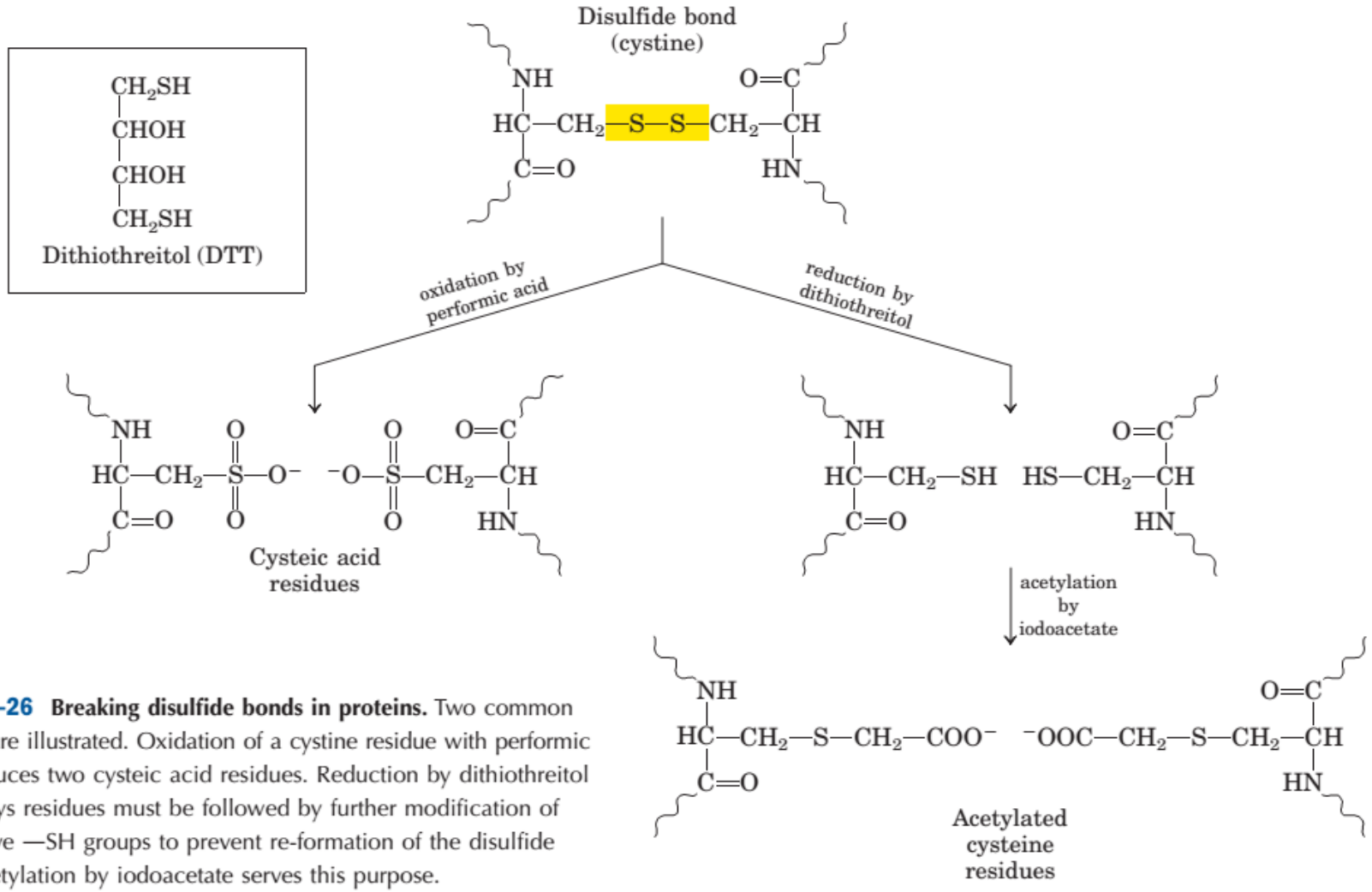


FIGURE 3–26 Breaking disulfide bonds in proteins. Two common methods are illustrated. Oxidation of a cystine residue with performic acid produces two cysteic acid residues. Reduction by dithiothreitol to form Cys residues must be followed by further modification of the reactive —SH groups to prevent re-formation of the disulfide bond. Acetylation by iodoacetate serves this purpose.

❖Enzymes called **proteases** catalyze the hydrolytic cleavage of peptide bonds.

❖Some proteases cleave only the peptide bond adjacent to particular amino acid residues and thus fragment a polypeptide chain in a predictable and reproducible way.

❖A few chemical reagents also cleave the peptide bond adjacent to specific residues.

TABLE 3–7 The Specificity of Some Common Methods for Fragmenting Polypeptide Chains

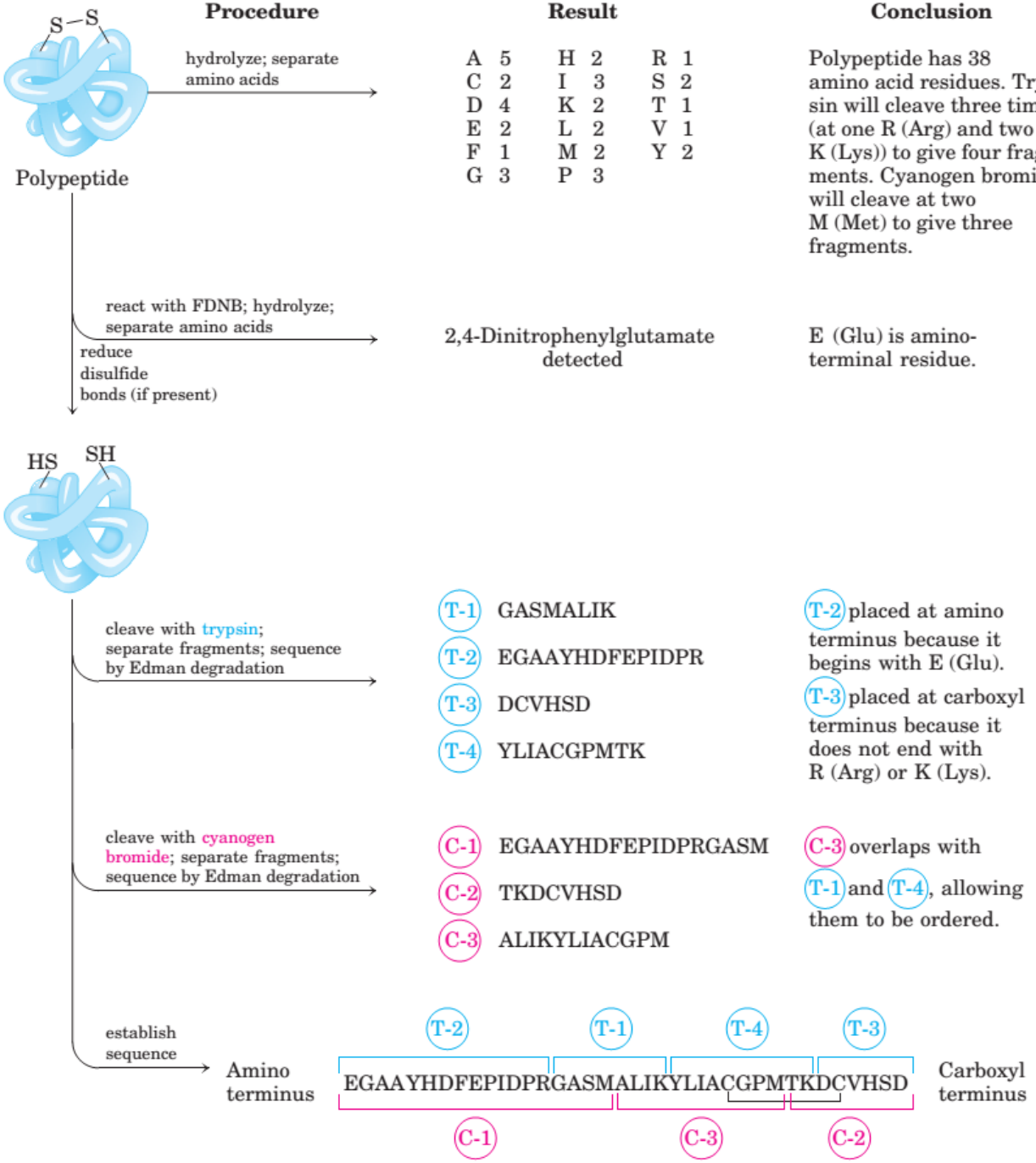
*Reagent (biological source)**	*Cleavage points*†
Trypsin (bovine pancreas)	Lys, Arg (C)
Submaxillarus protease (mouse submaxillary gland)	Arg (C)
Chymotrypsin (bovine pancreas)	Phe, Trp, Tyr (C)
Staphylococcus aureus V8 protease (bacterium *S. aureus*)	Asp, Glu (C)
Asp-*N*-protease (bacterium *Pseudomonas fragi*)	Asp, Glu (N)
Pepsin (porcine stomach)	Phe, Trp, Tyr (N)
Endoproteinase Lys C (bacterium *Lysobacter enzymogenes*)	Lys (C)
Cyanogen bromide	Met (C)

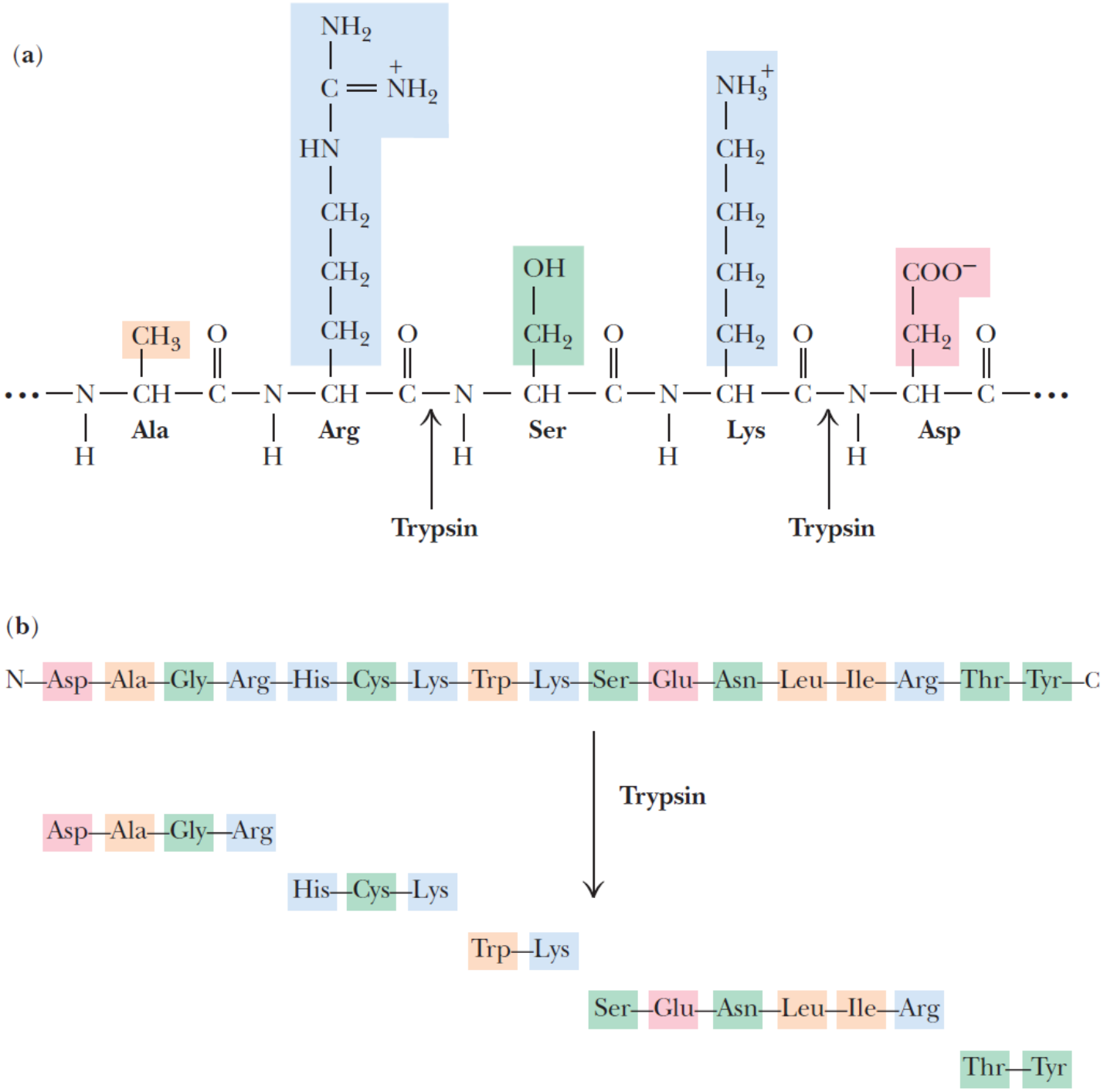
❖Among proteases, the digestive enzyme trypsin catalyzes the hydrolysis of only those peptide bonds in which the carbonyl group is contributed by either a Lys or an Arg residue, regardless of the length or amino acid sequence of the chain.

❖The number of smaller peptides produced by trypsin cleavage can thus be predicted from the total number of Lys or Arg residues in the original polypeptide, as determined by hydrolysis of an intact sample.

❖The fragments produced by trypsin (or other enzyme or chemical) action are then separated by chromatographic or electrophoretic methods.

❖In classical sequencing, a large protein would be cleaved into fragments twice, using a different protease or cleavage reagent each time so that the fragment endpoints would be different.


Procedure
Result
Conclusion
S—S
hydrolyze; separate amino acids
A 5 H 2 R 1
C 2 I 3 S 2
D 4 K 2 T 1
E 2 L 2 V 1
F 1 M 2 Y 2
G 3 P 3
Polypeptide has 38 amino acid residues. Trypsin will cleave three times (at one R (Arg) and two K (Lys)) to give four fragments. Cyanogen bromide will cleave at two M (Met) to give three fragments.
Polypeptide
react with FDNB; hydrolyze; separate amino acids
2,4-Dinitrophenylglutamate detected
E (Glu) is amino-terminal residue.
reduce disulfide bonds (if present)
HS SH
cleave with trypsin; separate fragments; sequence by Edman degradation
T-1 GASMALIK
T-2 EGAAYHDFEPIDPR
T-3 DCVHSD
T-4 YLIACGPMTK
T-2 placed at amino terminus because it begins with E (Glu).
T-3 placed at carboxyl terminus because it does not end with R (Arg) or K (Lys).
cleave with cyanogen bromide; separate fragments; sequence by Edman degradation
C-1 EGAAYHDFEPIDPRGASM
C-2 TKDCVHSD
C-3 ALIKYLIACGPM
C-3 overlaps with T-1 and T-4, allowing them to be ordered.
establish sequence
Amino terminus
T-2 T-1 T-4 T-3
EGAAYHDFEPIDPRGASMALIKYLIACGPMTKDCVHSD
C-1 C-3 C-2
Carboxyl terminus


(a)
Ala
Arg
Ser
Lys
Asp
Trypsin
Trypsin
(b)
N—Asp—Ala—Gly—Arg—His—Cys—Lys—Trp—Lys—Ser—Glu—Asn—Leu—Ile—Arg—Thr—Tyr—C
Trypsin
Asp—Ala—Gly—Arg
His—Cys—Lys
Trp—Lys
Ser—Glu—Asn—Leu—Ile—Arg
Thr—Tyr

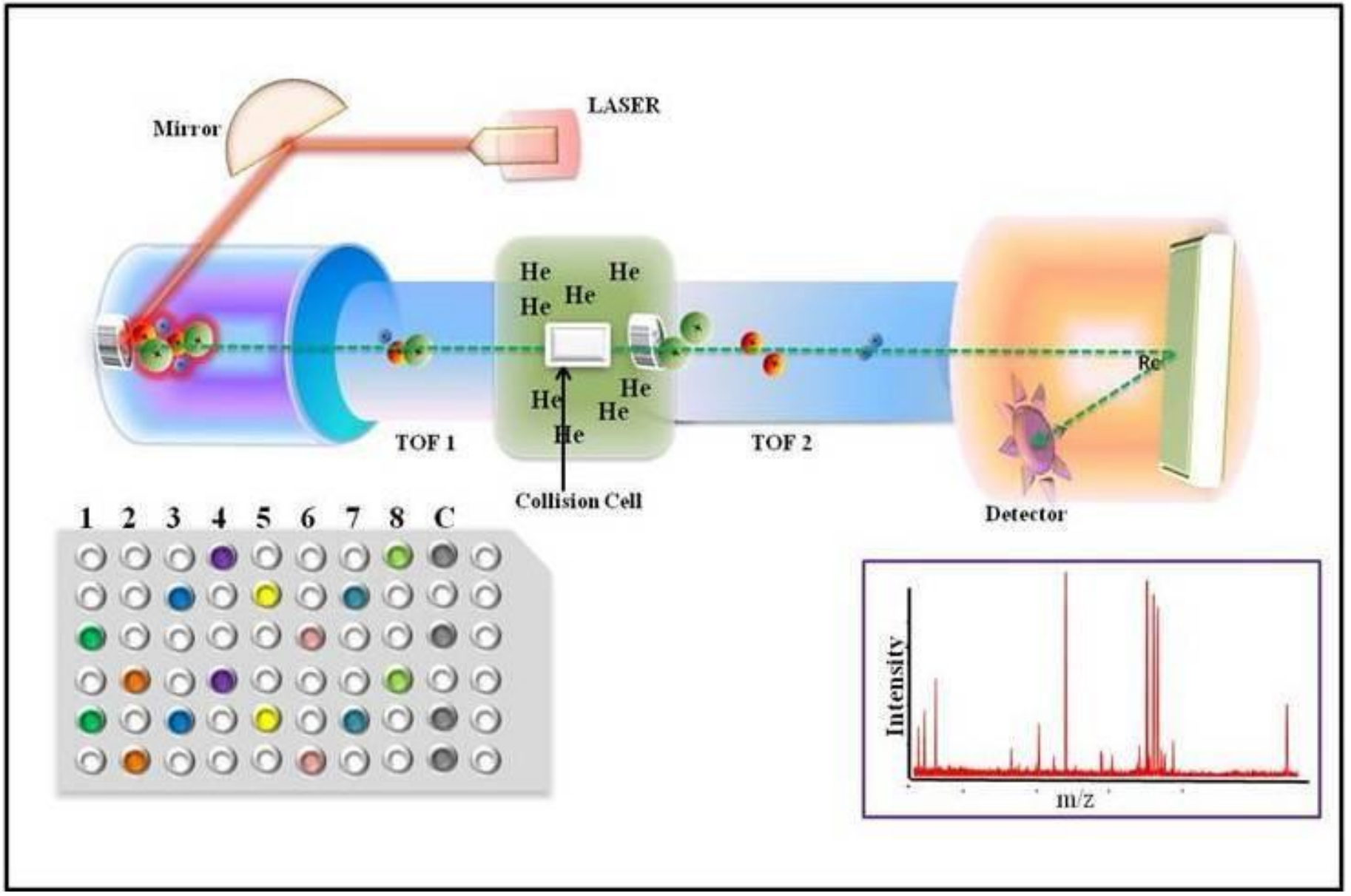
❖Amino acid sequences can also be deduced by other methods.

❖New methods based on mass spectrometry permit the sequencing of short polypeptides (20 to 30 amino acid residues) in just a few minutes.

❖In matrix-assisted laser desorption/ ionization mass spectrometry, or MALDI MS, proteins are placed in a light-absorbing matrix.

❖With a short pulse of laser light, the proteins are ionized and then desorbed from the matrix into the vacuum system

❖MALDI MS has been successfully used to measure the mass of a wide range of macromolecules.


Mirror
LASER
He
He
He
He
He
He
He
He
TOF 1
TOF 2
Collision Cell
Detector
Re
1 2 3 4 5 6 7 8 C
Intensity
m/z

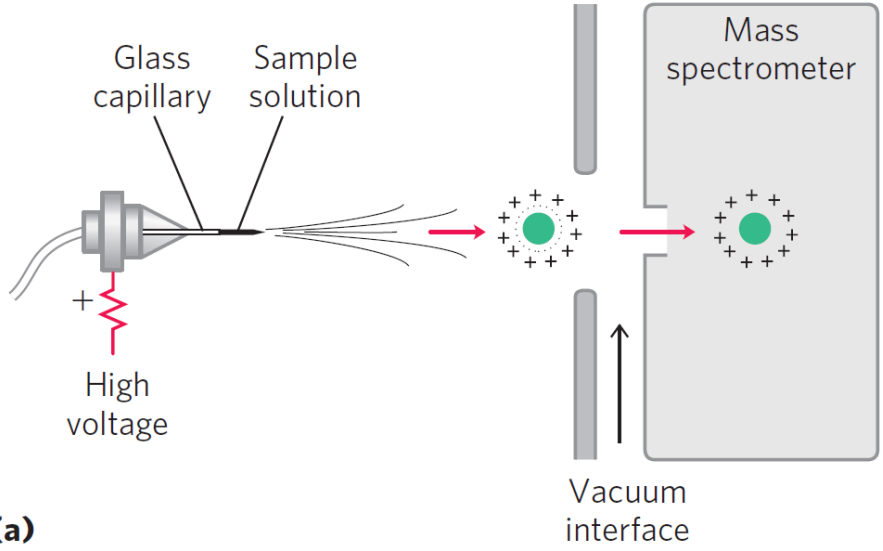
❖In a second and equally successful method, macromolecules in solution are forced directly from the liquid to gas phase.

❖A solution of analytes is passed through a charged needle that is kept at a high electrical potential, dispersing the solution into a fine mist of charged microdroplets.

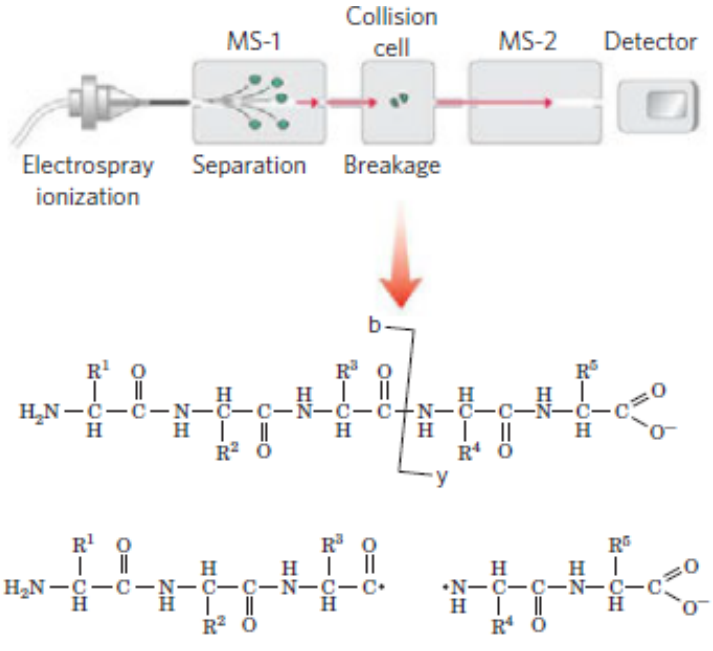
❖The solvent surrounding the macromolecules rapidly evaporates, leaving multiply charged macromolecular ions in the gas phase.

❖This technique is called **electrospray ionization mass spectrometry, or ESI MS**. Protons added during passage through the needle give additional charge to the macromolecule. The m/z of the molecule can be analyzed in the vacuum chamber.

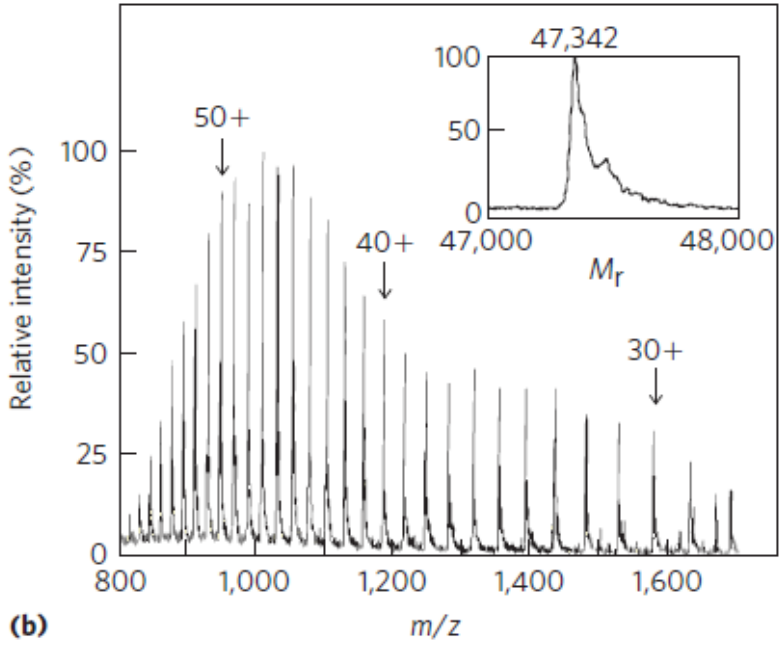
❖**Tandem MS**, or **MS/MS** can also be used to sequence short stretches of polypeptide.


Glass capillary
Sample solution
Mass spectrometer
High voltage
Vacuum interface

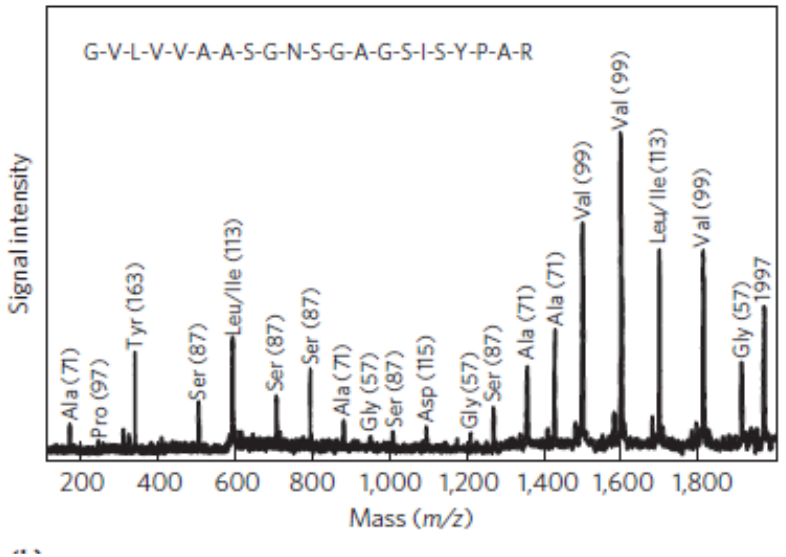
(a)


MS-1
Collision cell
MS-2
Detector
Electrospray ionization
Separation
Breakage
b
y

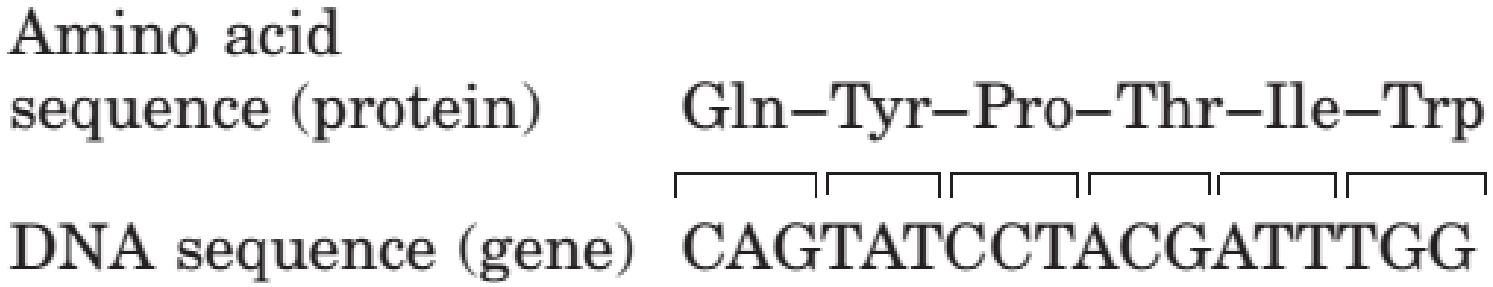
(a)


47,342
50+
40+
30+
Relative intensity (%)
m/z
M_r

(b)


G-V-L-V-V-A-A-S-G-N-S-G-A-G-S-I-S-Y-P-A-R
Signal intensity
Mass (m/z)

❖In addition, with the development of rapid DNA sequencing methods, researchers can deduce the sequence of a polypeptide by determining the sequence of nucleotides in the gene that codes for it.



❖Most proteins are now sequenced in this indirect way, however if the gene has not been isolated, direct sequencing of peptides is necessary

❖The complete sequence of an organism's DNA, its genome, is now available for organisms ranging from viruses to bacteria to multicellular eukaryotes.

❖Genes are being discovered by the millions, including many that encode proteins with no known function.

❖Many peptides are potentially useful as pharmacologic agents, and their production is of considerable commercial importance.

❖There are three ways to obtain a peptide:

❖(1) purification from tissue, a task often made difficult by the vanishingly low concentrations of some peptides;

❖(2) genetic engineering; or

❖(3) direct chemical synthesis.

❖Powerful techniques now make direct chemical synthesis an attractive option in many cases.

❖In addition to commercial applications, the synthesis of specific peptide portions of larger proteins is an increasingly important tool for the study of protein structure and function

R. Bruce Merrifield

❖Knowledge of the sequence of amino acids in a protein can offer insights into its three-dimensional structure and its function, cellular location, and evolution.

❖Most of these insights are derived by searching for similarities between a protein of interest and previously studied proteins.

❖Thousands of sequences are known and available in databases accessible through the Internet.

❖A comparison of a newly obtained sequence with this large bank of stored sequences often reveals relationships both surprising and enlightening.

❖Exactly how the amino acid sequence determines three-dimensional structure is not understood in detail, nor can we always predict function from sequence.

❖However, protein families that have some shared structural or functional features can be readily identified on the basis of amino acid sequence similarities.

❖Individual proteins are assigned to **families** based on the degree of similarity in amino acid sequence.

❖Members of a family are usually identical across 25% or more of their sequences, and proteins in these families generally share at least some structural and functional characteristics.

❖A number of similar substructures, or “domains” occur in many functionally unrelated proteins.

❖Evolutionary relationships can also be inferred from the structural and functional similarities within protein families.

❖Each protein's function relies on its three-dimensional structure, which in turn is determined largely by its primary structure.

❖Protein sequences are beginning to tell us how the proteins evolved and, ultimately, how life evolved on this planet.

❖If two organisms are closely related, the sequences of their genes and proteins should be similar.

❖The sequences increasingly diverge as the evolutionary distance between two organisms increases.

❖The members of protein families are called **homologous proteins**, or **homologs**.

❖The concept of a homolog can be further refined. If two proteins in a family (that is, two homologs) are present in the same species, they are referred to as **paralogs**. Homologs from different species are called
❖**orthologs**.

❖Each protein's function relies on its three-dimensional structure, which in turn is determined largely by its primary structure.

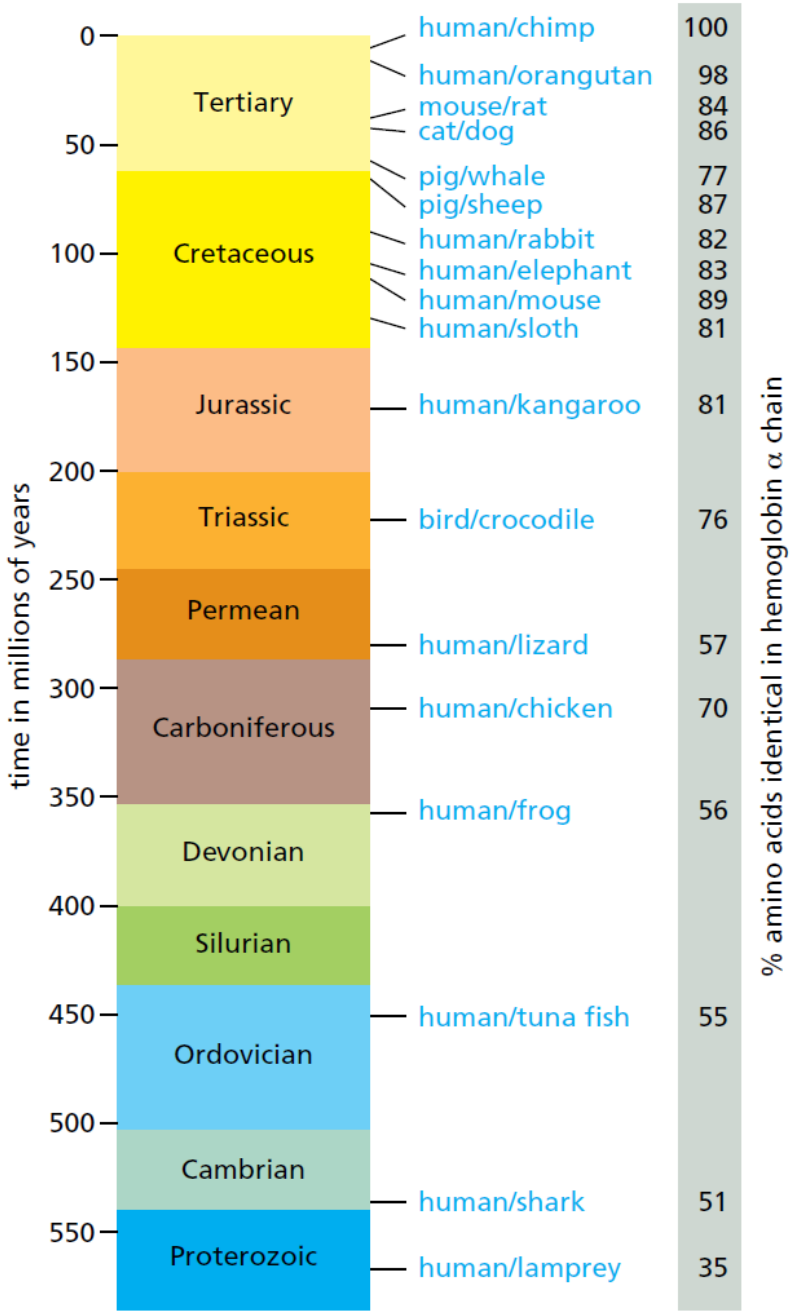
❖Protein sequences are beginning to tell us how the proteins evolved and, ultimately, how life evolved on this planet.

❖If two organisms are closely related, the sequences of their genes and proteins should be similar.

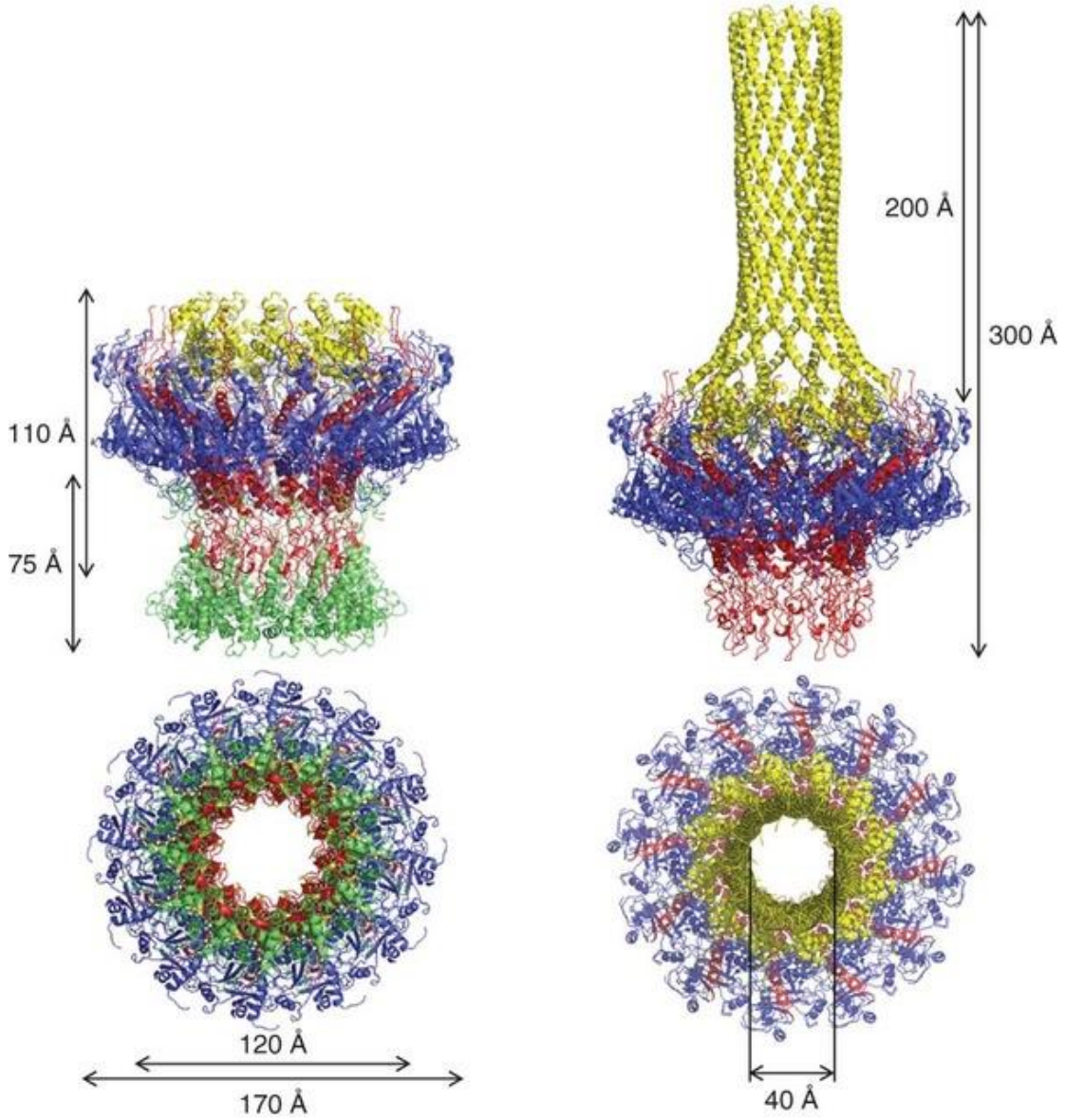
❖The sequences increasingly diverge as the evolutionary distance between two organisms increases.

❖The members of protein families are called **homologous proteins**, or **homologs**.

❖The concept of a homolog can be further refined. If two proteins in a family (that is, two homologs) are present in the same species, they are referred to as **paralogs**. Homologs from different species are called
❖**orthologs**.


time in millions of years
0
50
100
150
200
250
300
350
400
450
500
550
Tertiary
Cretaceous
Jurassic
Triassic
Permean
Carboniferous
Devonian
Silurian
Ordovician
Cambrian
Proterozoic
human/chimp
human/orangutan
mouse/rat
cat/dog
pig/whale
pig/sheep
human/rabbit
human/elephant
human/mouse
human/sloth
human/kangaroo
bird/crocodile
human/lizard
human/chicken
human/frog
human/tuna fish
human/shark
human/lamprey
100
98
84
86
77
87
82
83
89
81
81
76
57
70
56
55
51
35
% amino acids identical in hemoglobin α chain

AMINO ACIDS PEPTIDES AND PROTEINS III: Three Dimensional Structures of Proteins



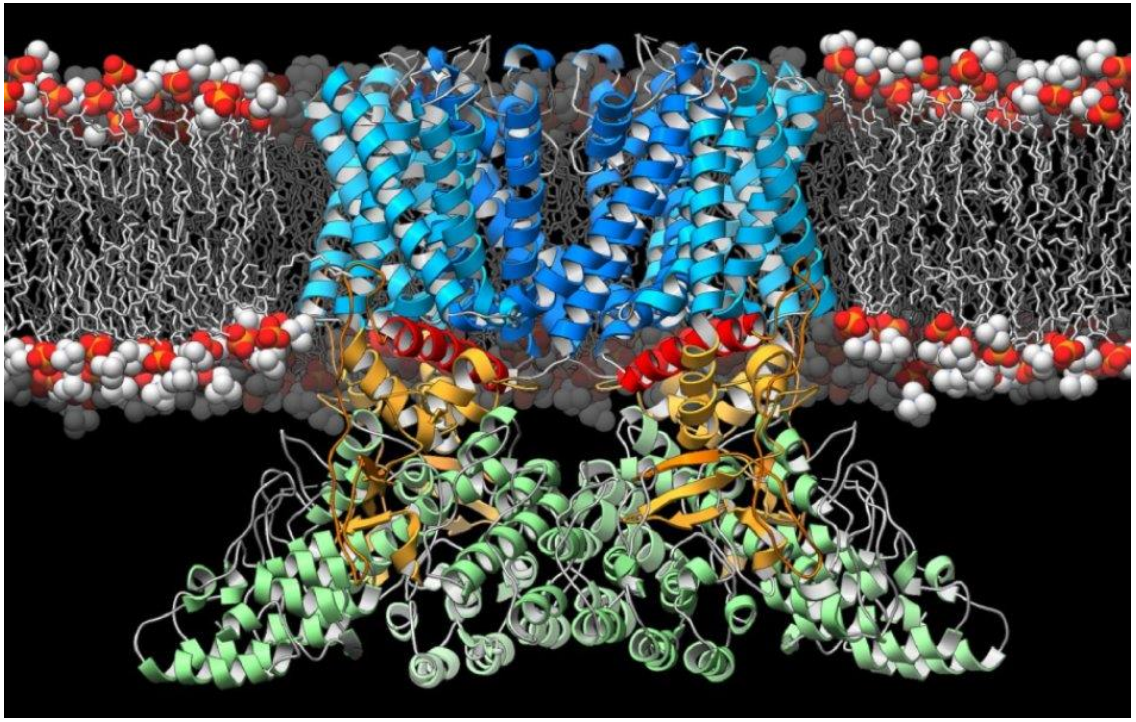
❖The covalent backbone of a typical protein contains hundreds of individual bonds.

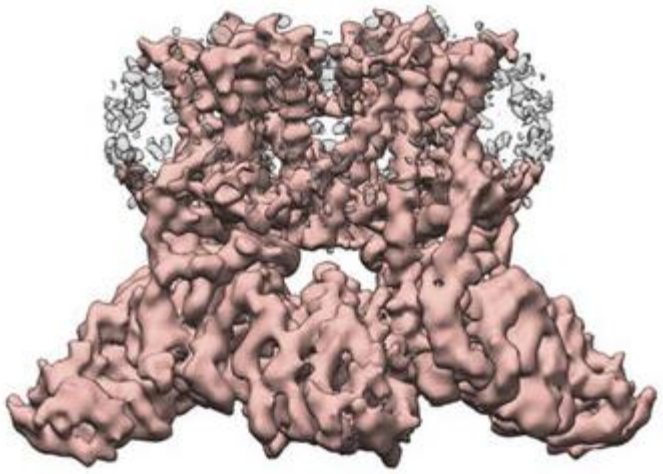
❖Because free rotation is possible around many of these bonds, the protein can assume an unlimited number of conformations.

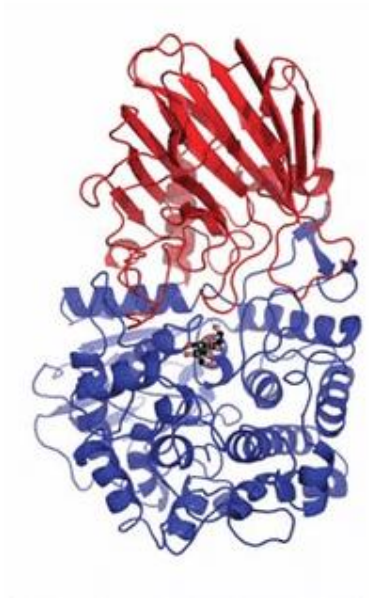
❖However, each protein has a specific chemical or structural function, strongly suggesting that each has a unique three-dimensional structure.

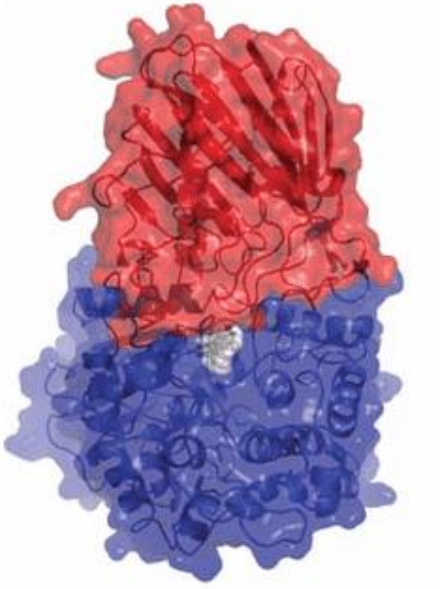
❖These themes should not be taken to imply that proteins have static, unchanging three-dimensional structures.

❖Protein function often entails an interconversion between two or more structural forms.









❖The spatial arrangement of atoms in a protein is called its **conformation**.

❖The possible conformations of a protein or protein segment include any structural state it can achieve without breaking covalent bonds.

❖A change in conformation could occur, for example, by rotation about single bonds.

❖Of the many conformations that are theoretically possible in a protein, one or (more commonly) a few generally predominate under biological conditions.

❖The conformations existing under a given set of conditions are usually the ones that are thermodynamically the most stable, having the lowest Gibbs free energy (G).

❖Proteins in any of their functional, folded conformations are called **native** proteins.

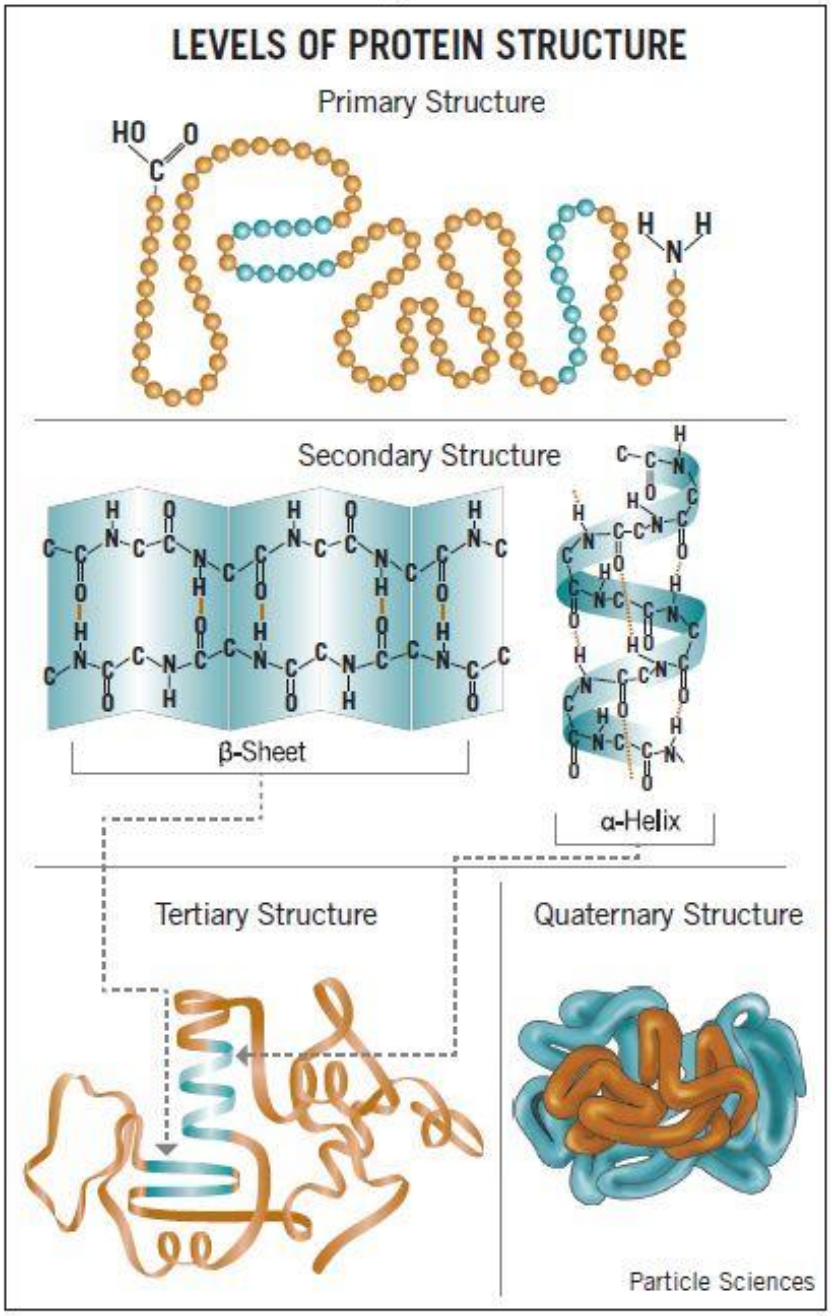
❖A description of all covalent bonds linking amino acid residues in a polypeptide chain is its **primary structure**. Primary structure of a protein defines its three dimensional structure.

❖**Secondary structure** refers to particularly stable arrangements of amino acid residues giving rise to recurring structural patterns.

❖**Tertiary structure** describes all aspects of the three-dimensional folding of a polypeptide.

❖When a protein has two or more polypeptide subunits, their arrangement in space is referred to as **quaternary structure**.

Figure 2



❖In the context of protein structure, the term **stability** can be defined as the tendency to maintain a native conformation.

❖A given polypeptide chain can theoretically assume countless conformations, and as a result the unfolded state of a protein is characterized by a high degree of conformational entropy.

❖This entropy, and the hydrogen-bonding interactions of many groups in the polypeptide chain with the solvent (water), tend to maintain the unfolded state.

❖The chemical interactions that counteract these effects and stabilize the native conformation include disulfide (covalent) bonds and the weak (noncovalent) interactions: hydrogen bonds, hydrophobic interactions and Van der Waals interactions .

❖Many proteins do not have disulfide bonds.

❖In eukaryotes, disulfide bonds are found primarily in secreted, extracellular proteins.

❖Disulfide bonds are also uncommon in bacterial proteins.

❖However, thermophilic bacteria, as well as the archaea, typically have many proteins with disulfide bonds, which stabilize proteins; this is presumably an adaptation to life at high temperatures.

❖For all proteins of all organisms, weak interactions are especially important in the folding of polypeptide chains into their secondary and tertiary structures.

❖The association of multiple polypeptides to form quaternary structures also relies on these weak interactions.

❖On carefully examining the contribution of weak interactions to protein stability, we find that hydrophobic interactions generally predominate.

❖When water surrounds a hydrophobic molecule, the optimal arrangement of hydrogen bonds results in a highly structured shell, or solvation layer.

❖The increased order of the water molecules in the solvation layer correlates with an unfavorable decrease in the entropy of the water.

❖However, when nonpolar groups cluster together, the extent of the solvation layer decreases, because each group no longer presents its entire surface to the solution. The result is a favorable increase in entropy.

❖Hydrophobic amino acid side chains therefore tend to cluster in a protein's interior, away from water

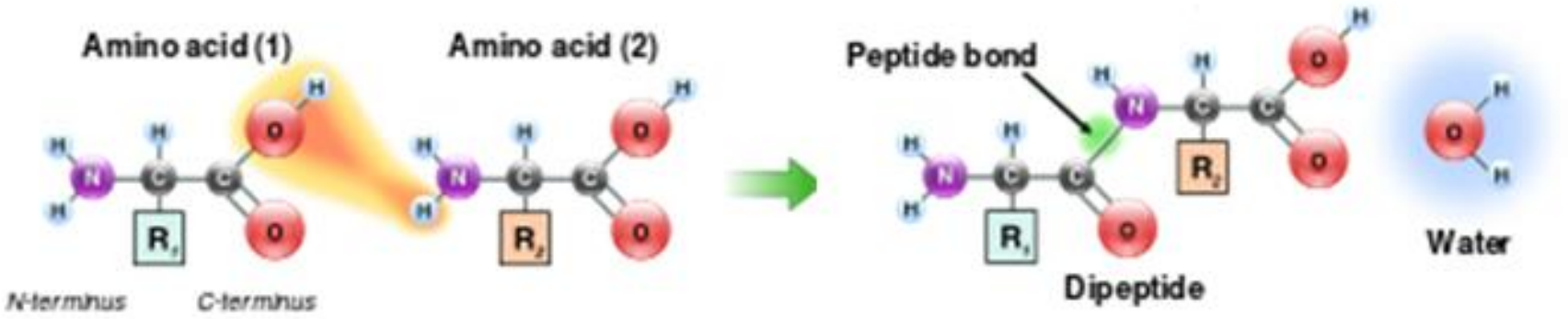
❖Most of the structural patterns outlined here reflect two simple rules:

- (1) hydrophobic residues are largely buried in the protein interior, away from water, and

- (2) the number of hydrogen bonds and ionic interactions within the protein is maximized, thus reducing the number of hydrogen-bonding and ionic groups that are not paired with a suitable partner.

❖Proteins within membranes and proteins that are intrinsically disordered follow different rules.

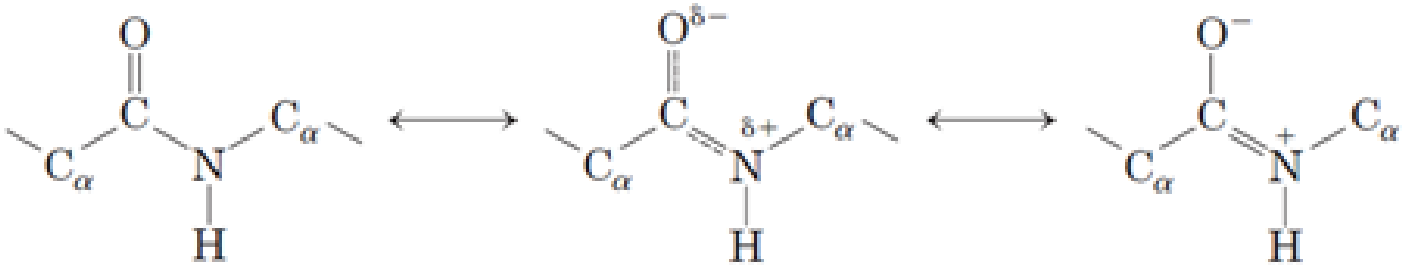
Peptide Bond



❖The α carbons of adjacent amino acid residues are separated by three covalent bonds, arranged as C_{α}—C—N—$C_{\alpha'}$.

❖X-ray diffraction studies showed that the peptide C—N bond is somewhat shorter than the C—N bond in a simple amine.

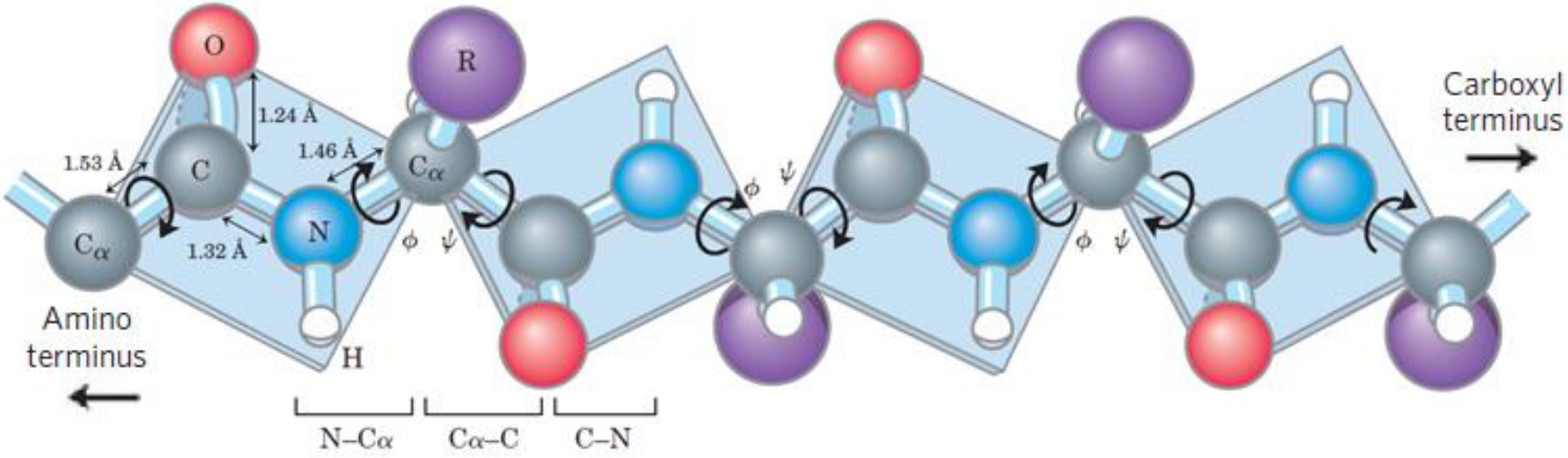
❖This indicated a resonance or partial sharing of two pairs of electrons between the carbonyl oxygen and the amide nitrogen

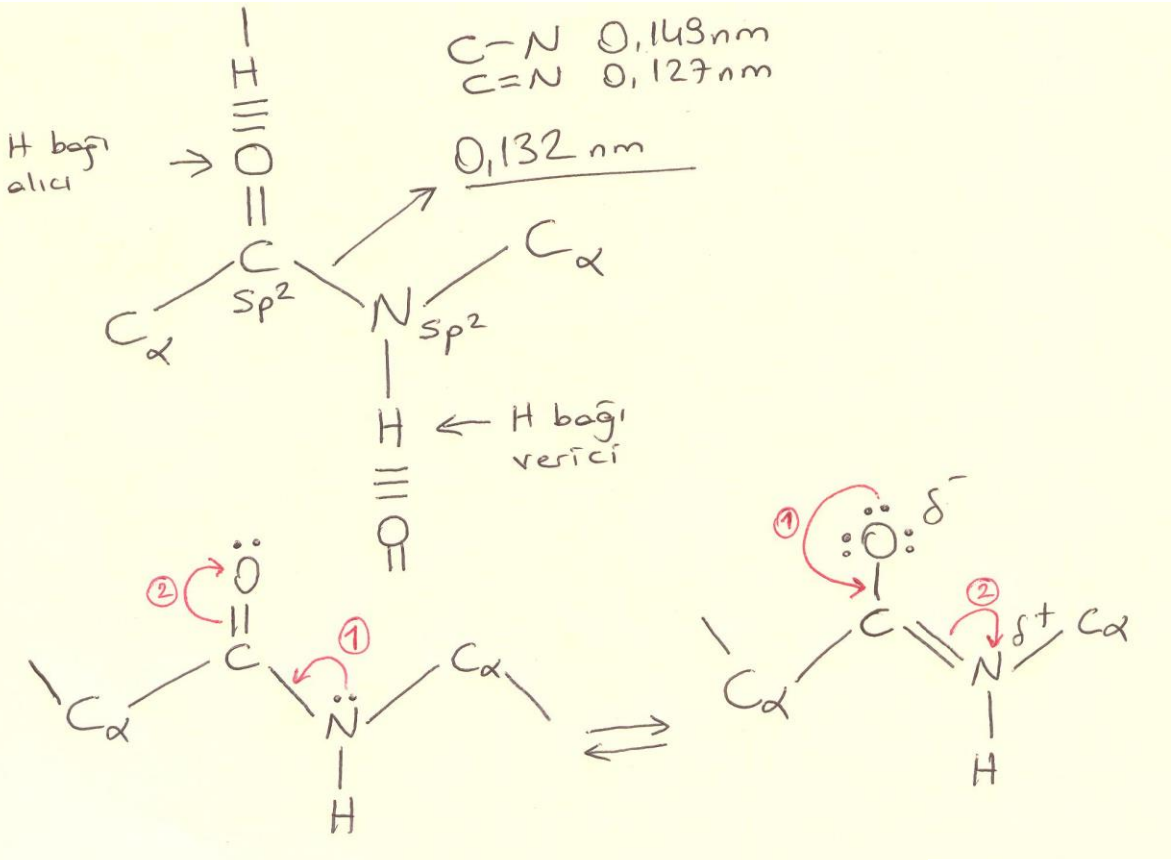


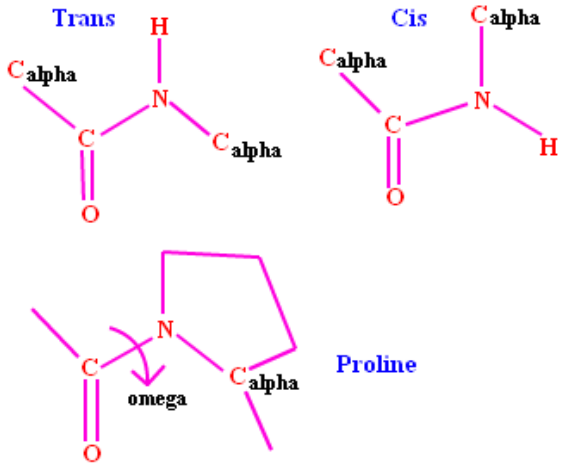
❖The six atoms of the peptide group lie in a single plane, with the oxygen atom of the carbonyl group trans to the hydrogen atom of the amide nitrogen.

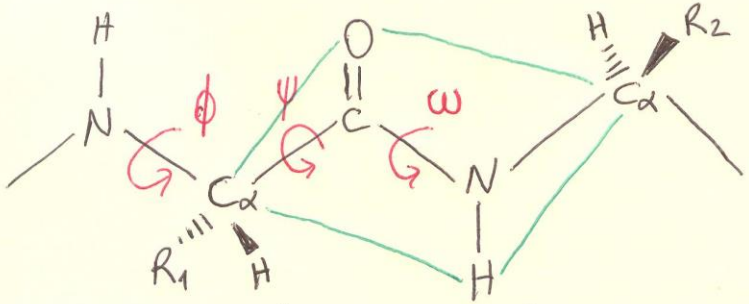
❖The peptide C—N bonds, because of their partial double-bond character, cannot rotate freely.

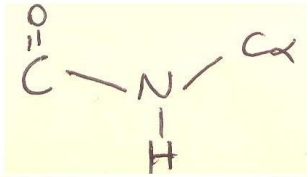
❖Rotation is permitted about the N—C_α and the C_α—C bonds. The rigid peptide bonds limit the range of conformations possible for a polypeptide chain.



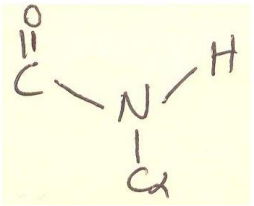






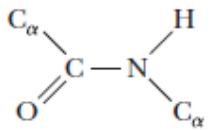


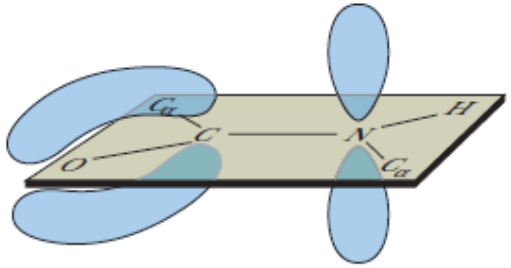
ω 180° trans
%99.6



ω 0° cis

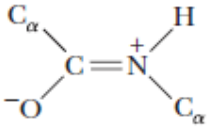
(a)

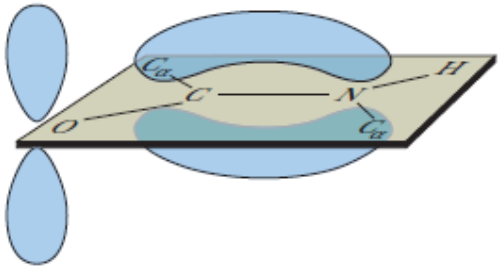




A pure double bond between C and O would permit free rotation around the C—N bond.

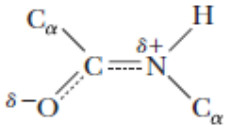
(b)

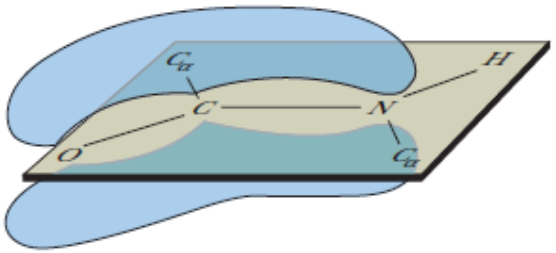




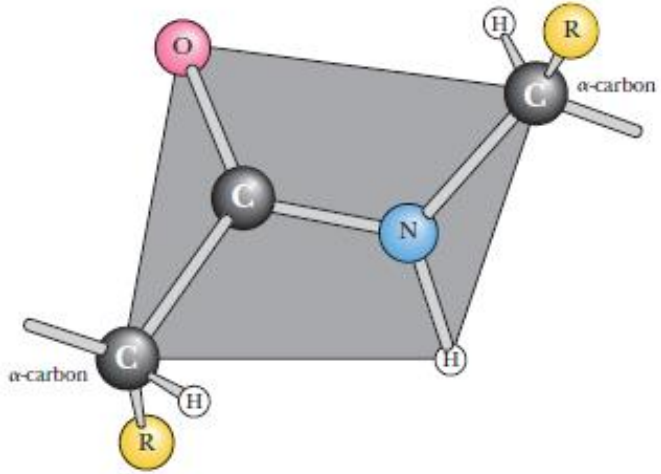
The other extreme would prohibit C—N bond rotation but would place too great a charge on O and N.

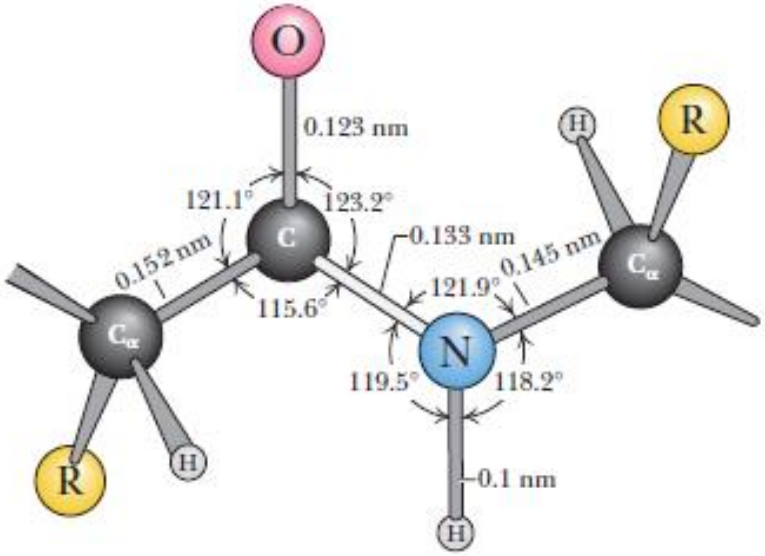
(c)





The true electron density is intermediate. The barrier to C—N bond rotation of about 88 kJ/mol is enough to keep the amide group planar.

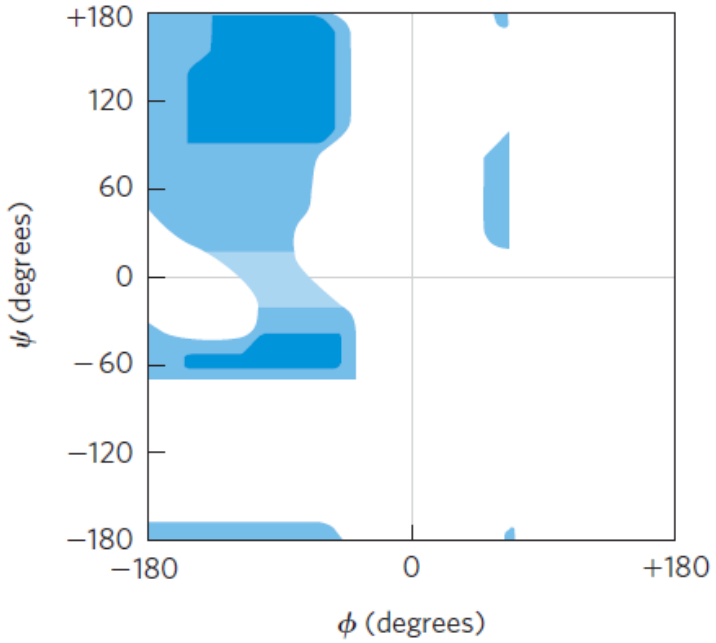




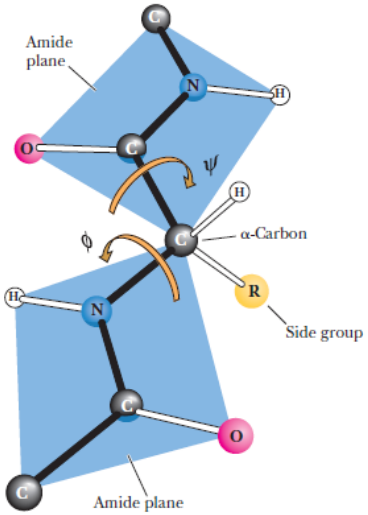
❖Peptide conformation is defined by three dihedral angles (also known as torsion angles) called ϕ (phi), ψ (psi), and ω (omega), reflecting rotation about each of the three repeating bonds in the peptide backbone. A dihedral angle is the angle at the intersection of two planes.

❖In principle, ϕ and ψ can have any value between –180° and +180°, but many values are prohibited by steric interference between atoms in the polypeptide backbone and amino acid side chains.

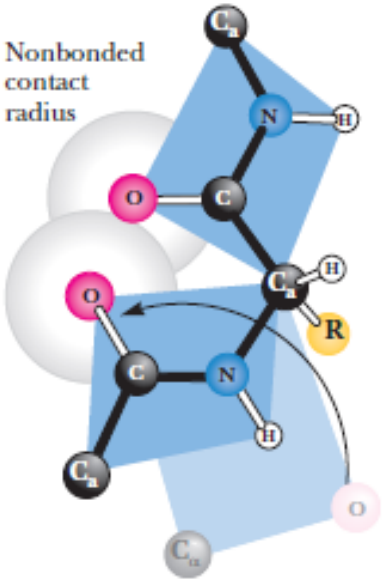
❖Allowed values for ϕ and ψ become evident when ϕ is plotted versus ψ in a Ramachandran plot.



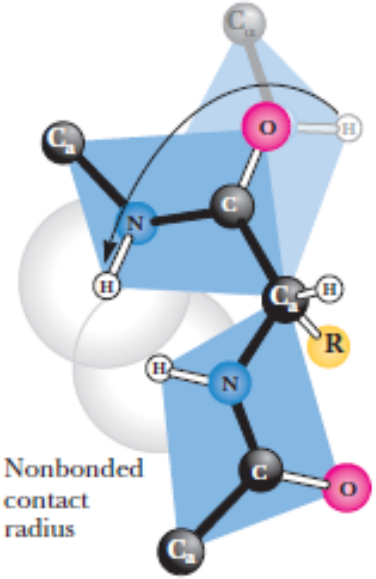
Structure	ϕ	ψ
α Helix	−57°	−47°
β Conformation		
Antiparallel	−139°	+135°
Parallel	−119°	+113°
Collagen triple helix	−51°	+153°
β Turn type I		
i + 1*	−60°	−30°
i + 2*	−90°	0°
β Turn type II		
i + 1	−60°	+120°
i + 2	+80°	0°



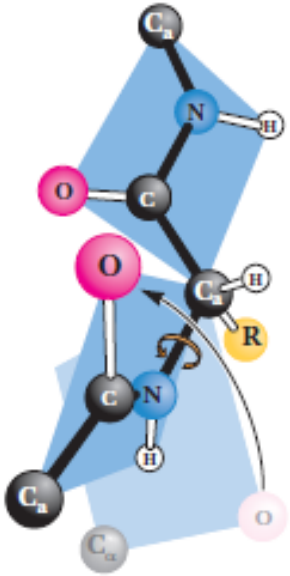
$\phi = 180°, \psi = 180°$



$\phi = 0°, \psi = 180°$

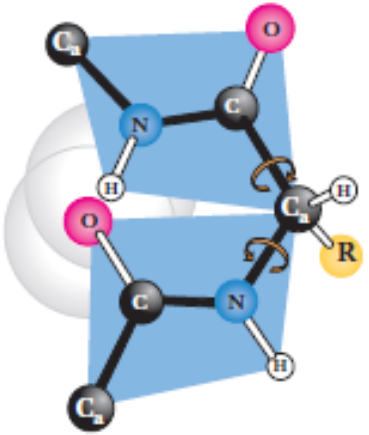


$\phi = 180°, \psi = 0°$



$\phi = -60°, \psi = 180°$

A further ϕ rotation of 120° removes the bulky carbonyl group as far as possible from the side chain



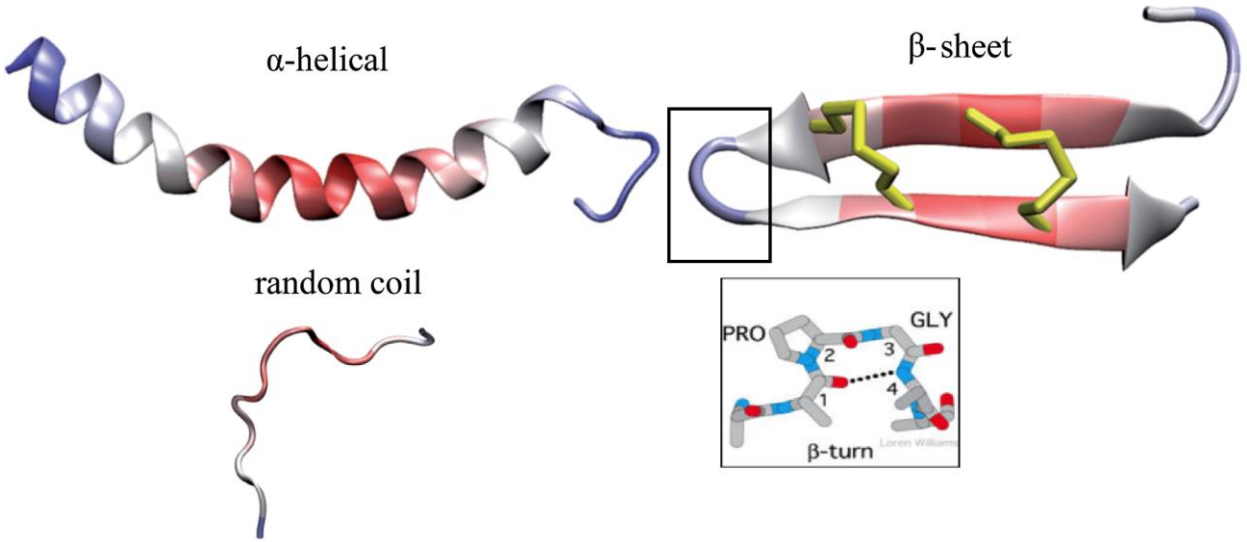
$\phi = 0°, \psi = 0°$

Protein Secondary Structure

❖The term **secondary structure** refers to any chosen segment of a polypeptide chain and describes the local spatial arrangement of its main-chain atoms.

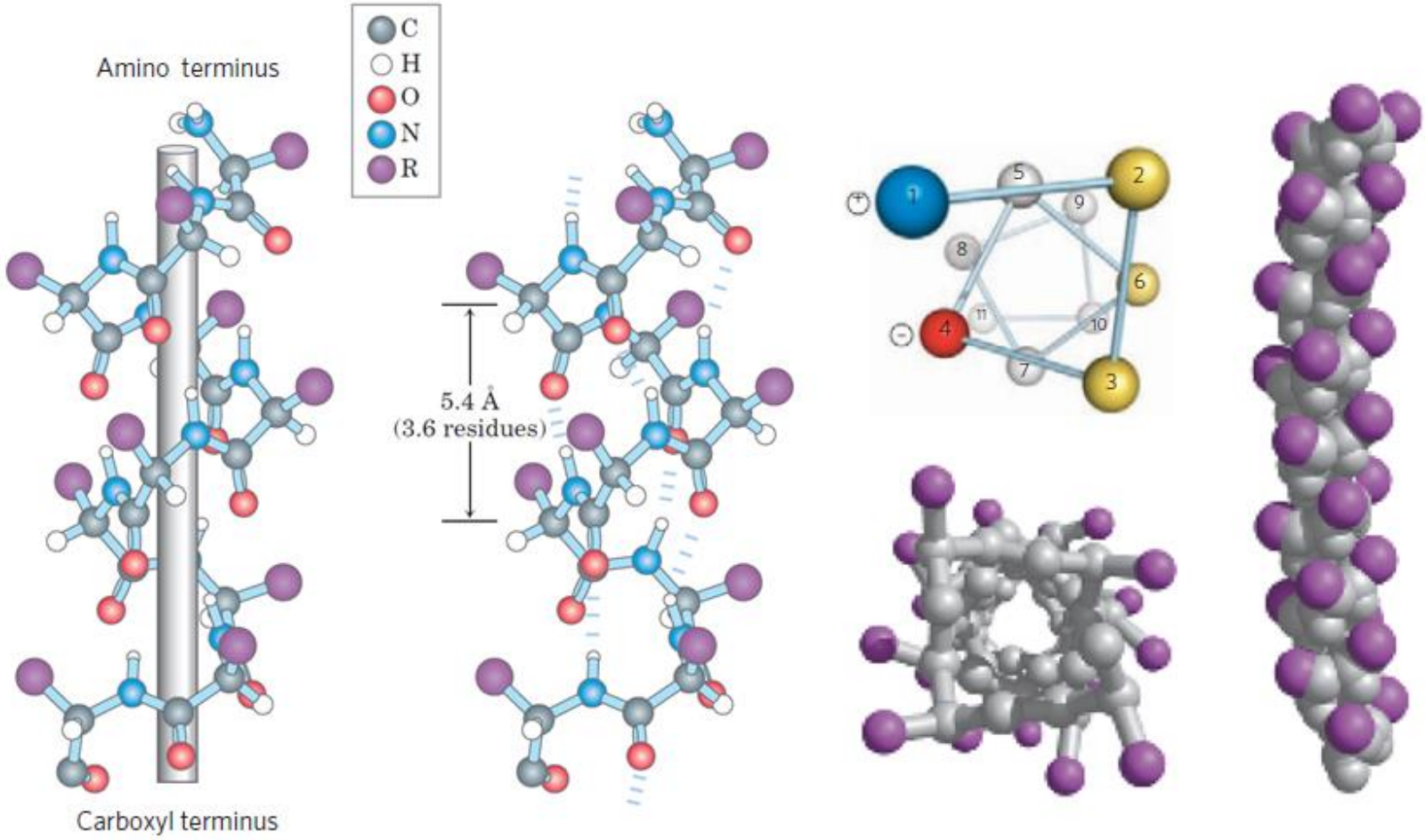
❖There are a few types of secondary structure that are particularly stable and occur widely in proteins.

❖The most prominent are the α helix and β conformations; another common type is the β turn. Where a regular pattern is not found, the secondary structure is sometimes referred to as undefined or as a random coil.



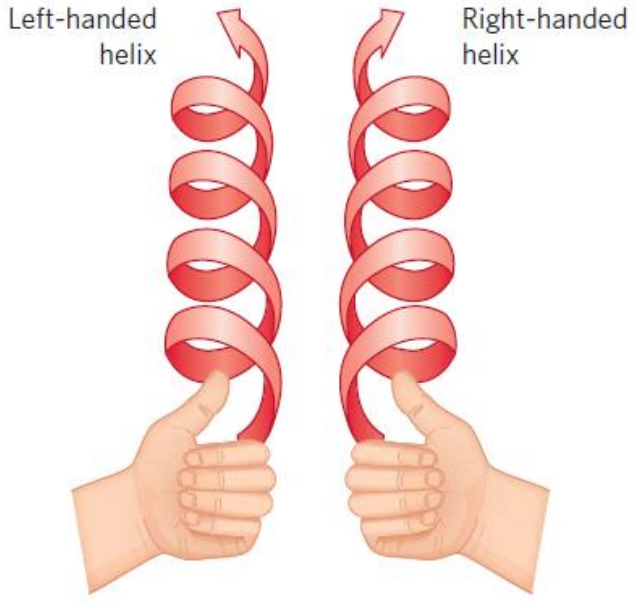
❖The α helix is a common protein secondary structure.

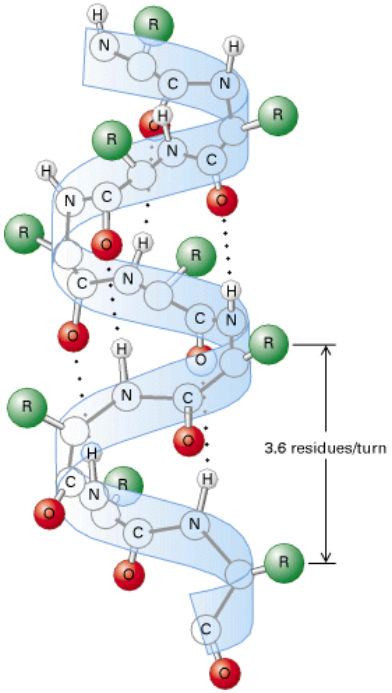
❖In this structure, the polypeptide backbone is tightly wound around an imaginary axis, and the R groups of the amino acid residues protrude outward from the helical backbone.

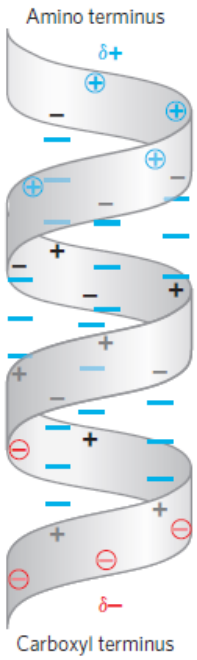


❖The backbone atoms of the amino acid residues in the prototypical α helix have a characteristic set of dihedral angles that define the α -helix conformation, and each helical turn includes 3.6 amino acid residues.

❖Right-handed α helix is the common form. Extended left-handed α helices are theoretically less stable and have not been observed in proteins.







❖The structure is stabilized by a hydrogen bond between the hydrogen atom attached to the electronegative nitrogen atom of a peptide linkage and the electronegative carbonyl oxygen atom of the fourth amino acid on the amino-terminal side of that peptide bond.

❖ Further experiments have shown that an α helix can form in polypeptides consisting of either L- or D-amino acids. However, all residues must be of one stereoisomeric series; a D-amino acid will disrupt a regular structure consisting of L-amino acids, and vice versa.

❖The position of an amino acid residue relative to its neighbors is also important.

❖For example, if a polypeptide chain has a long block of Glu residues, this segment of the chain will not form an α helix at pH 7.0.

❖Glu residues repel each other so strongly that they prevent formation of the α helix.

❖The bulk and shape of Asn, Ser, Thr, and Cys residues can also destabilize an α helix if they are close together in the chain.

❖The twist of an α helix ensures that critical interactions occur between an amino acid side chain and the side chain three residues away on either side of it.

❖Positively charged amino acids are often found three residues away from negatively charged amino acids, permitting the formation of an ion pair.

❖Two aromatic amino acid residues are often similarly spaced, resulting in a hydrophobic interaction.

❖A constraint on the formation of the α helix is the presence of Pro or Gly residues.

❖In proline, the nitrogen atom is part of a rigid ring, and rotation about the N—C_α bond is not possible. Thus, a Pro residue introduces a destabilizing kink in an α helix.

❖Glycine occurs infrequently in α helices for a different reason: it has more conformational flexibility than the other amino acid residues.

❖Polymers of glycine tend to take up coiled structures quite different from an α helix.

❖A final factor affecting the stability of an α helix is the identity of the amino acid residues near the ends of the α -helical segment of the polypeptide.

❖A small electric dipole exists in each peptide bond. These dipoles are aligned through the hydrogen bonds of the helix, resulting in a net dipole along the helical axis that increases with helix length.

❖For this reason, negatively charged amino acids are often found near the amino terminus of the helical segment.

❖In summary, five types of constraints affect the stability of an α helix:

❖(1) the intrinsic propensity of an amino acid residue to form an α helix;

❖(2) the interactions between R groups, particularly those spaced three (or four) residues apart;

❖(3) the bulkiness of adjacent R groups;

❖(4) the occurrence of Pro and Gly residues; and

❖(5) interactions between amino acid residues at the ends of the helical segment and the electric dipole inherent to the α helix.

❖The **β conformation** organizes polypeptide chains into sheets.

❖In the β conformation, the backbone of the polypeptide chain is extended into a zigzag rather than helical structure.

❖The arrangement of several segments side by side, all of which are in the conformation, is called a sheet.

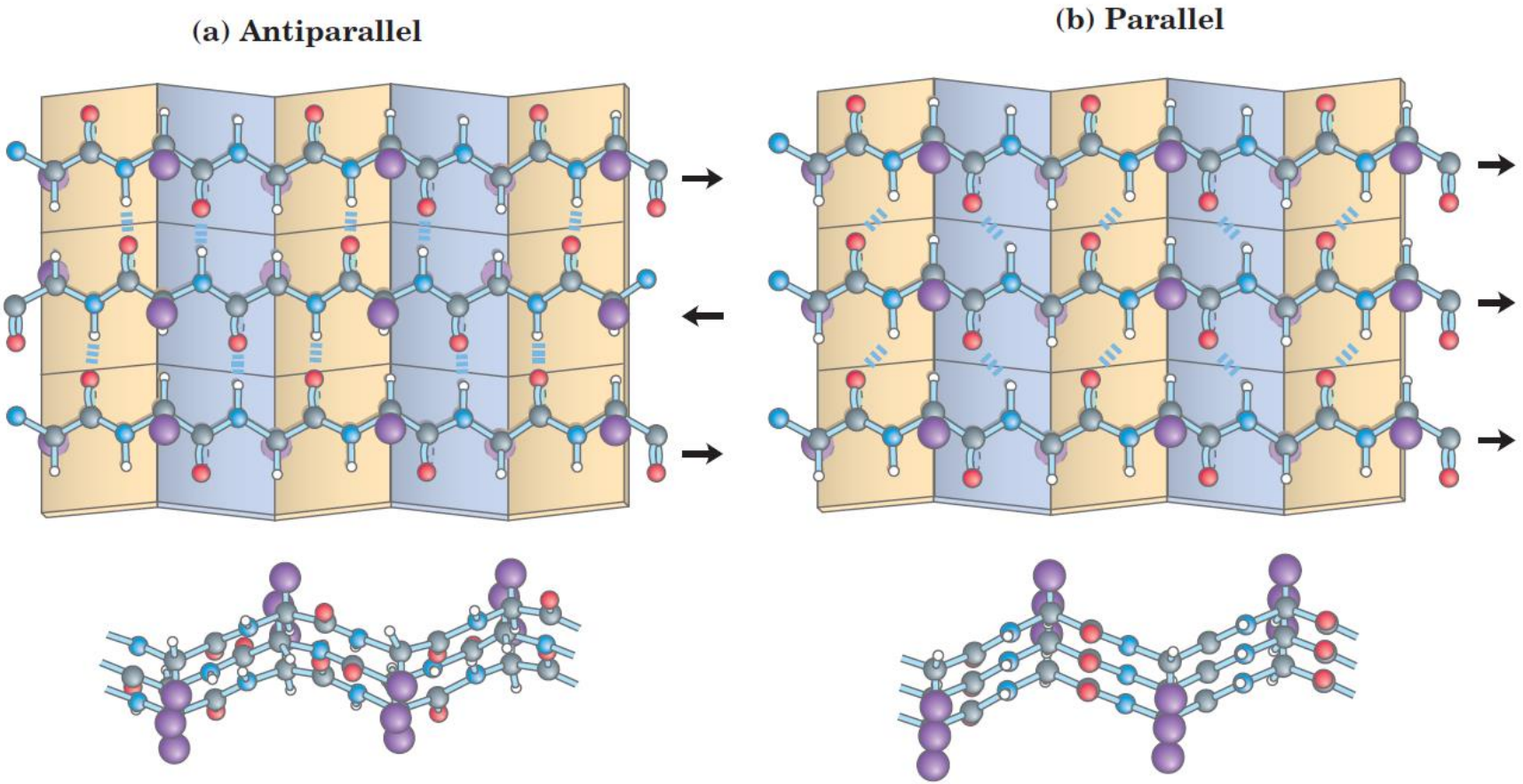
❖The zigzag structure of the individual polypeptide segments gives rise to a pleated appearance of the overall sheet.

❖Hydrogen bonds form between adjacent segments of polypeptide chain within the sheet.

❖The R groups of adjacent amino acids protrude from the zigzag structure in opposite directions, creating the alternating pattern

❖The adjacent polypeptide chains in a **β sheet** can be either parallel or antiparallel (having the same or opposite amino-to-carboxyl orientations, respectively).

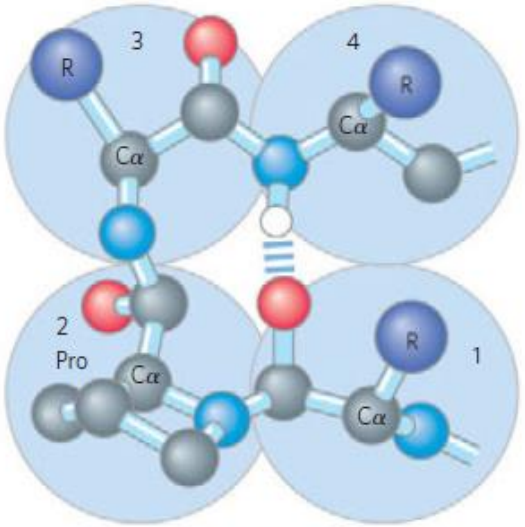
❖The structures are somewhat similar, although the repeat period is shorter for the parallel conformation and the hydrogen-bonding patterns are different.



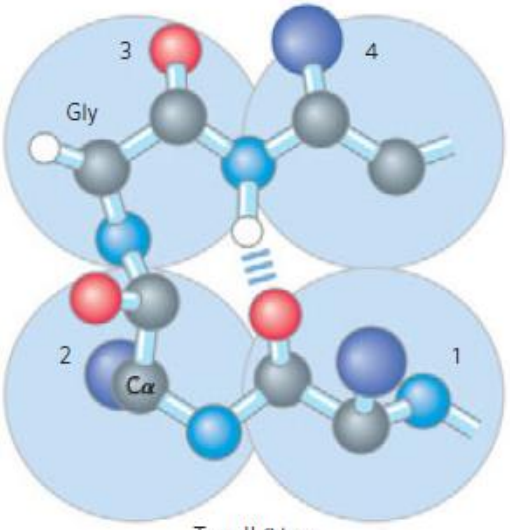
❖In globular proteins, which have a compact folded structure, some amino acid residues are in turns or loops where the polypeptide chain reverses direction.

❖These are the connecting elements that link successive runs of α helix or β conformation.

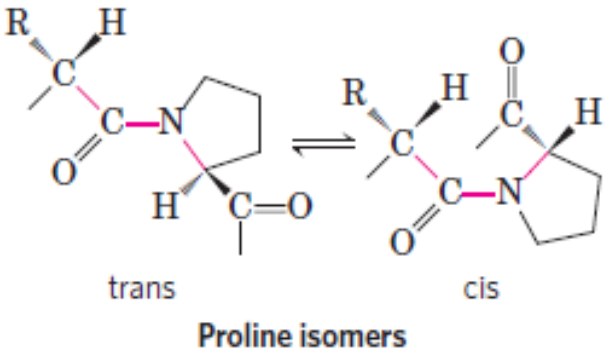
❖Particularly common are β turns that connect the ends of two adjacent segments of an antiparallel β sheet.



Type I β turn

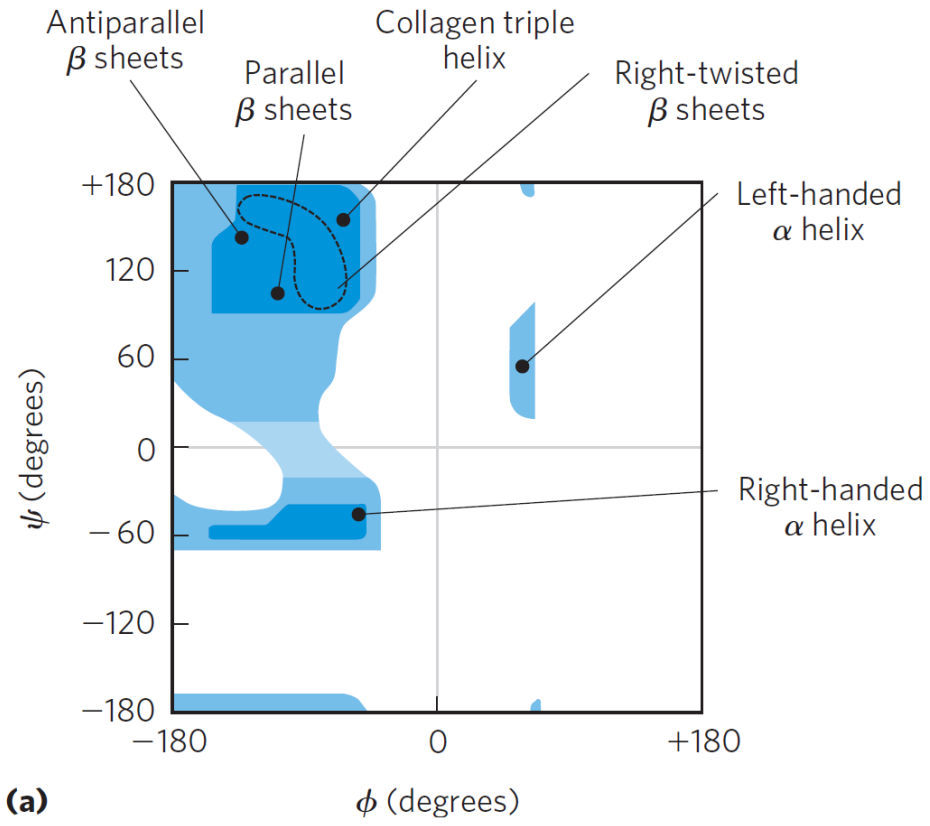


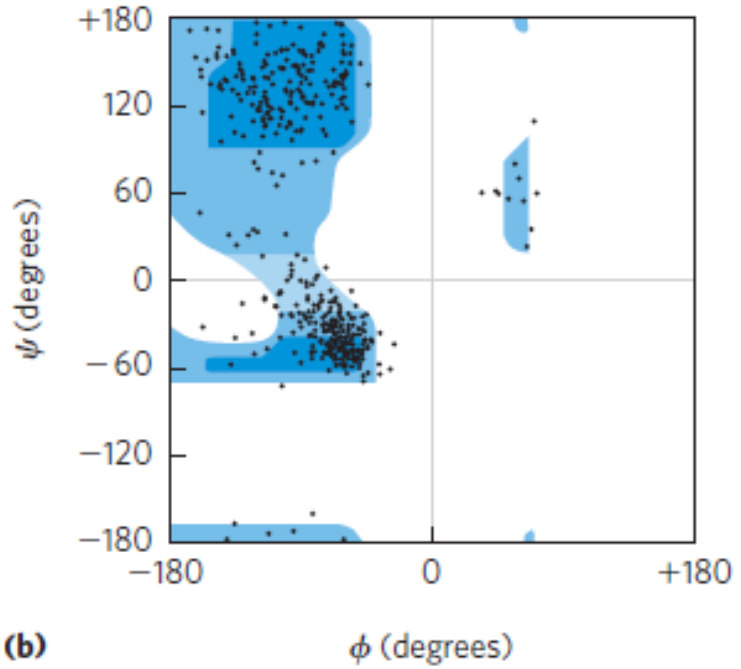
Type II β turn



Proline isomers

❖Every type of secondary structure can be completely described by the dihedral angles ϕ and ψ associated with each residue.





Protein Tertiary and Quaternary Structure

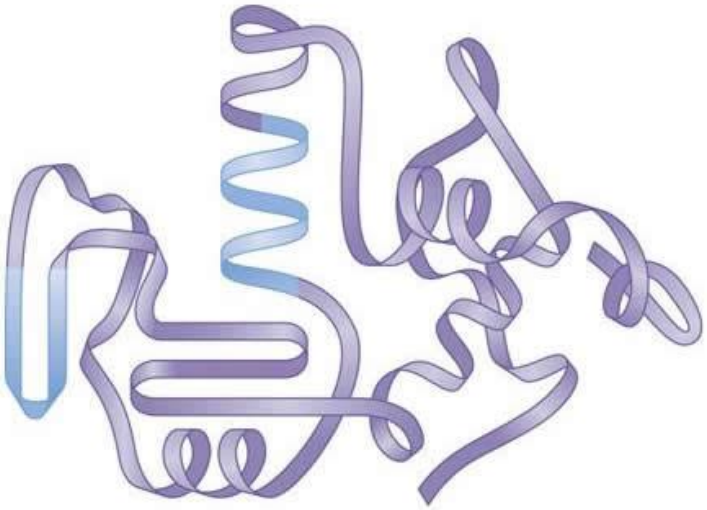
❖The overall three-dimensional arrangement of all atoms in a protein is referred to as the protein's tertiary structure.

❖Whereas the term "secondary structure" refers to the spatial arrangement of amino acid residues that are adjacent in a segment of a polypeptide, tertiary structure includes longer-range aspects of amino acid sequence.

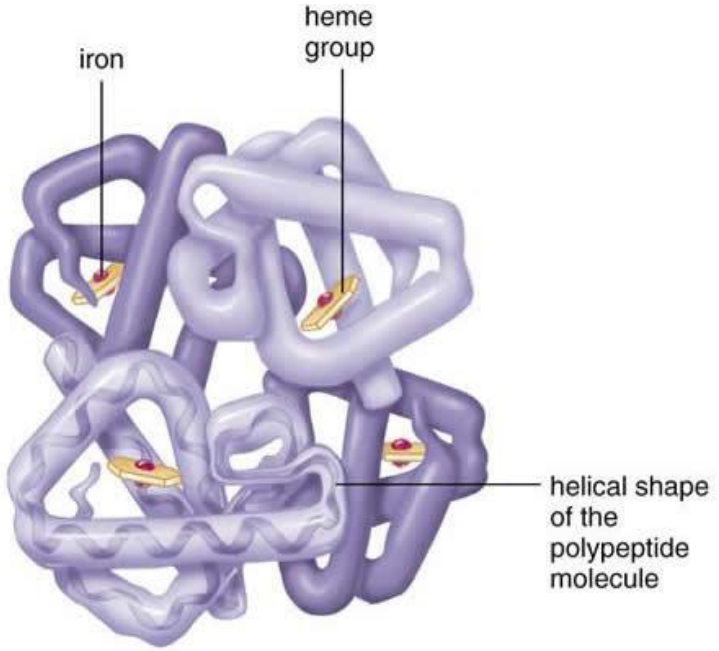
❖Amino acids that are far apart in the polypeptide sequence and are in different types of secondary structure may interact within the completely folded structure of a protein.

❖Interacting segments of polypeptide chains are held in their characteristic tertiary positions by several kinds of weak interactions (and sometimes by covalent bonds such as disulfide cross-links) between the segments.

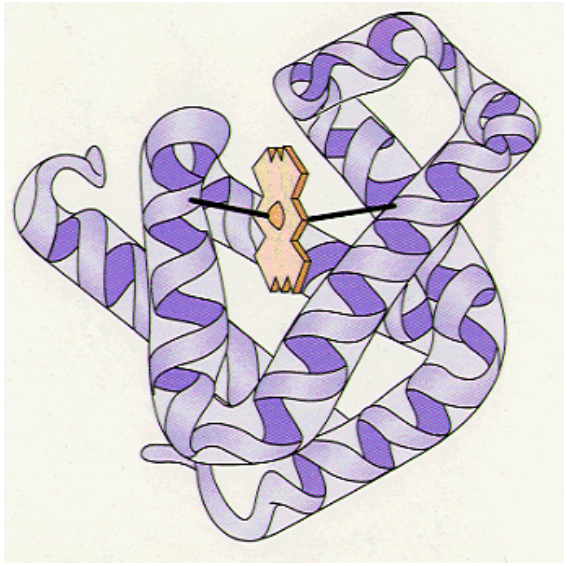
❖Some proteins contain two or more separate polypeptide chains, or subunits, which may be identical or different. The arrangement of these protein subunits in three-dimensional complexes constitutes quaternary structure.

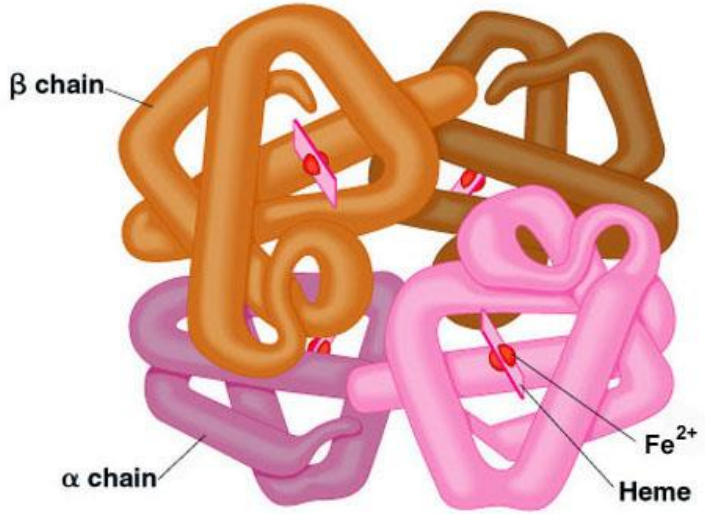


d. Tertiary structure



e. Quaternary structure



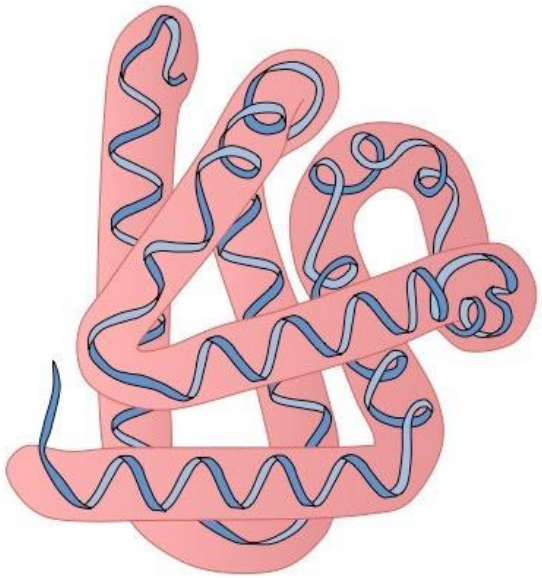


❖In considering these higher levels of structure, it is useful to designate two major groups into which many proteins can be classified:

❖**fibrous proteins**, with polypeptide chains arranged in long strands or sheets, and

❖**globular proteins**, with polypeptide chains folded into a spherical or globular shape.

Fibrous Protein



Globular Protein

- The two groups are structurally distinct.
  - Fibrous proteins usually consist largely of a single type of secondary structure, and their tertiary structure is relatively simple.
  - Globular proteins often contain several types of secondary structure.
- The two groups also differ functionally:
  - the structures that provide support, shape, and external protection to vertebrates are made of fibrous proteins,
  - whereas most enzymes and regulatory proteins are globular proteins.

❖Fibrous proteins are adapted for a structural function

❖α-Keratin, collagen, and silk fibroin nicely illustrate the relationship between protein structure and biological function.

❖Fibrous proteins share properties that give strength and/or flexibility to the structures in which they occur.

❖In each case, the fundamental structural unit is a simple repeating element of secondary structure.

❖All fibrous proteins are insoluble in water, a property conferred by a high concentration of hydrophobic amino acid residues both in the interior of the protein and on its surface.

❖These hydrophobic surfaces are largely buried as many similar polypeptide chains are packed together to form elaborate supramolecular complexes.

❖Fibrous proteins are adapted for a structural function

❖α-Keratin, collagen, and silk fibroin nicely illustrate the relationship between protein structure and biological function.

❖Fibrous proteins share properties that give strength and/or flexibility to the structures in which they occur.

❖In each case, the fundamental structural unit is a simple repeating element of secondary structure.

❖All fibrous proteins are insoluble in water, a property conferred by a high concentration of hydrophobic amino acid residues both in the interior of the protein and on its surface.

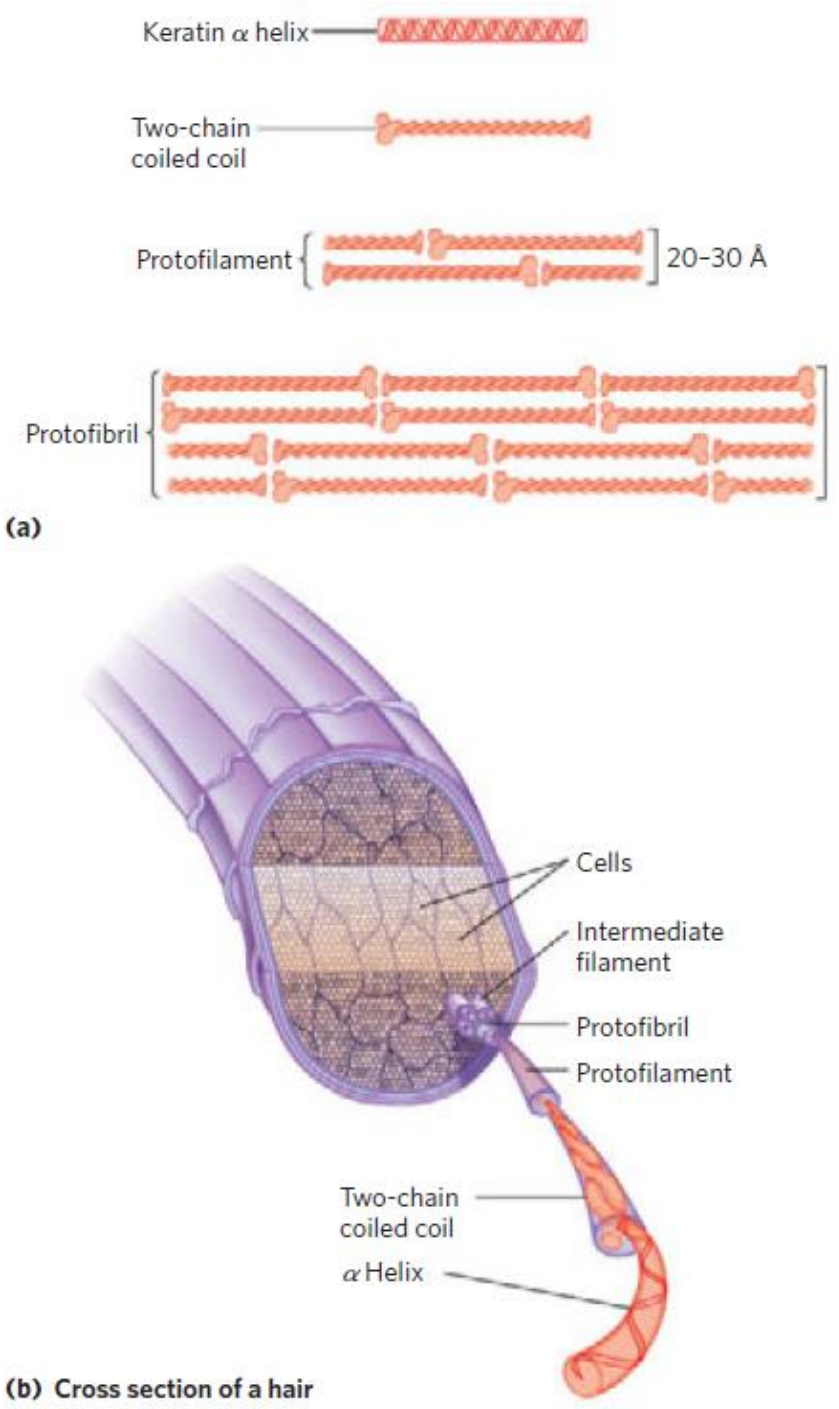
❖These hydrophobic surfaces are largely buried as many similar polypeptide chains are packed together to form elaborate supramolecular complexes.

❖The **α-keratins** have evolved for strength.

❖Found only in mammals, these proteins constitute almost the entire dry weight of hair, wool, nails, claws, quills, horns, hooves, and much of the outer layer of skin.

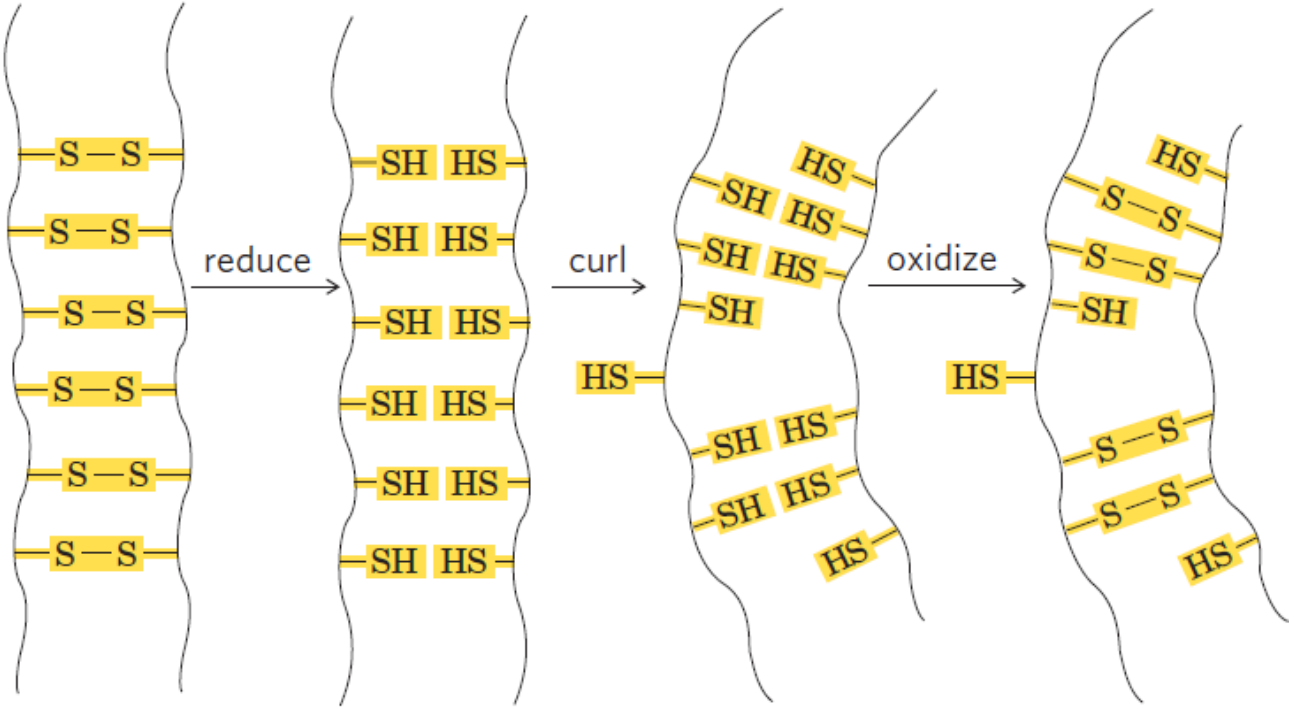
❖The α-keratin helix is a right-handed α helix and were arranged as a coiled coil.

The helical path of the supertwists is lefthanded, and not surprisingly, α-keratin is rich in the hydrophobic residues Ala, Val, Leu, Ile, Met, and Phe.



❖When hair is exposed to moist heat, it can be stretched by stretching of the α helices in the α -keratin of hair to fully extended β conformation. On cooling they spontaneously revert to original form.

❖When a solution of a reducing agent, usually a compound containing a thiol or sulfhydryl group (—SH), is applied with heat a permanent wave (or hair straightening) is achieved.

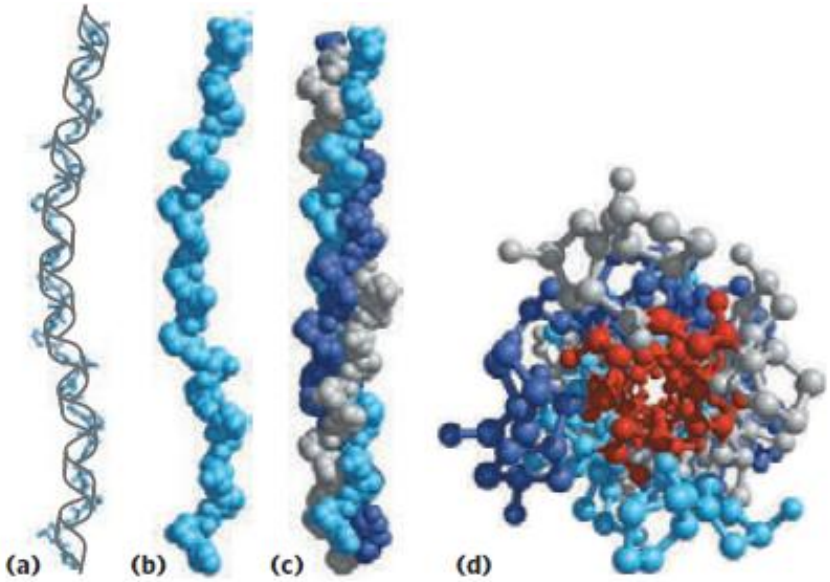


❖Like the α-keratins, collagen has evolved to provide strength.

❖It is found in connective tissue such as tendons, cartilage, the organic matrix of bone, and the cornea of the eye.

❖The collagen helix is a unique secondary structure, quite distinct from the α helix. It is left-handed and has three amino acid residues per turn.

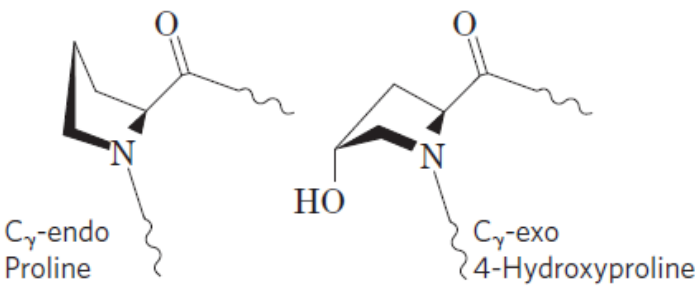
❖There are many types of vertebrate collagen. Typically they contain about 35% Gly, 11% Ala, and 21% Pro and 4-Hyp (4-hydroxyproline).

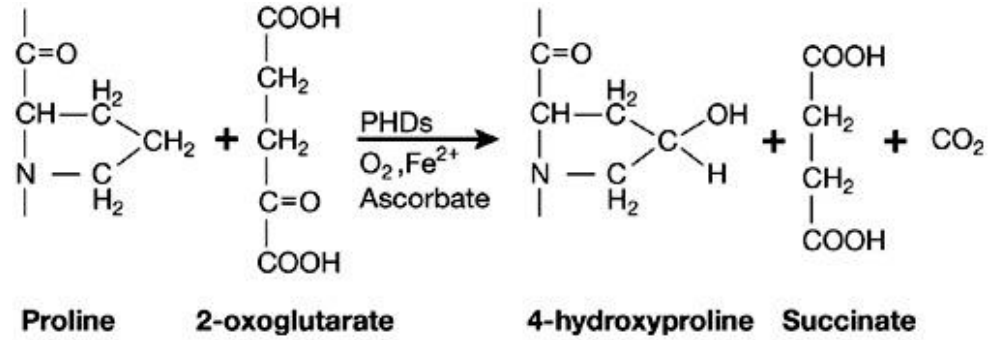


❖Scurvy is caused by lack of vitamin C, or ascorbic acid (ascorbate). Vitamin C is required for, among other things, the hydroxylation of proline and lysine in collagen; scurvy is a deficiency disease characterized by general degeneration of connective tissue.

❖Manifestations of advanced scurvy include numerous small hemorrhages caused by fragile blood vessels, tooth loss, poor wound healing and the reopening of old wounds, bone pain and degeneration, and eventually heart failure.

❖Collagen is constructed of the repeating tripeptide unit Gly–X–Y, where X and Y are generally Pro or 4-Hyp, which plays an essential role in the folding of collagen and in maintaining its structure.





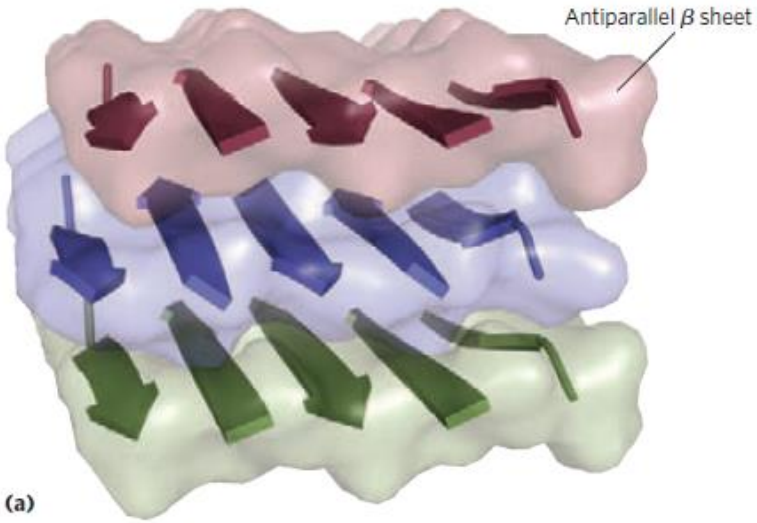
❖Fibroin, the protein of silk, is produced by insects and spiders. Its polypeptide chains are predominantly in the β conformation.

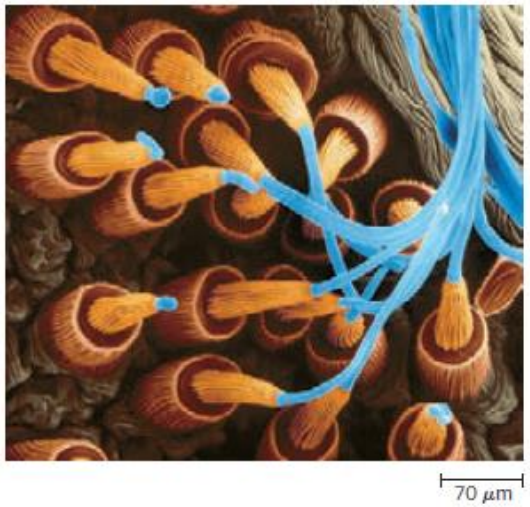
❖Fibroin is rich in Ala and Gly residues, permitting a close packing of β sheets and an interlocking arrangement of R groups.

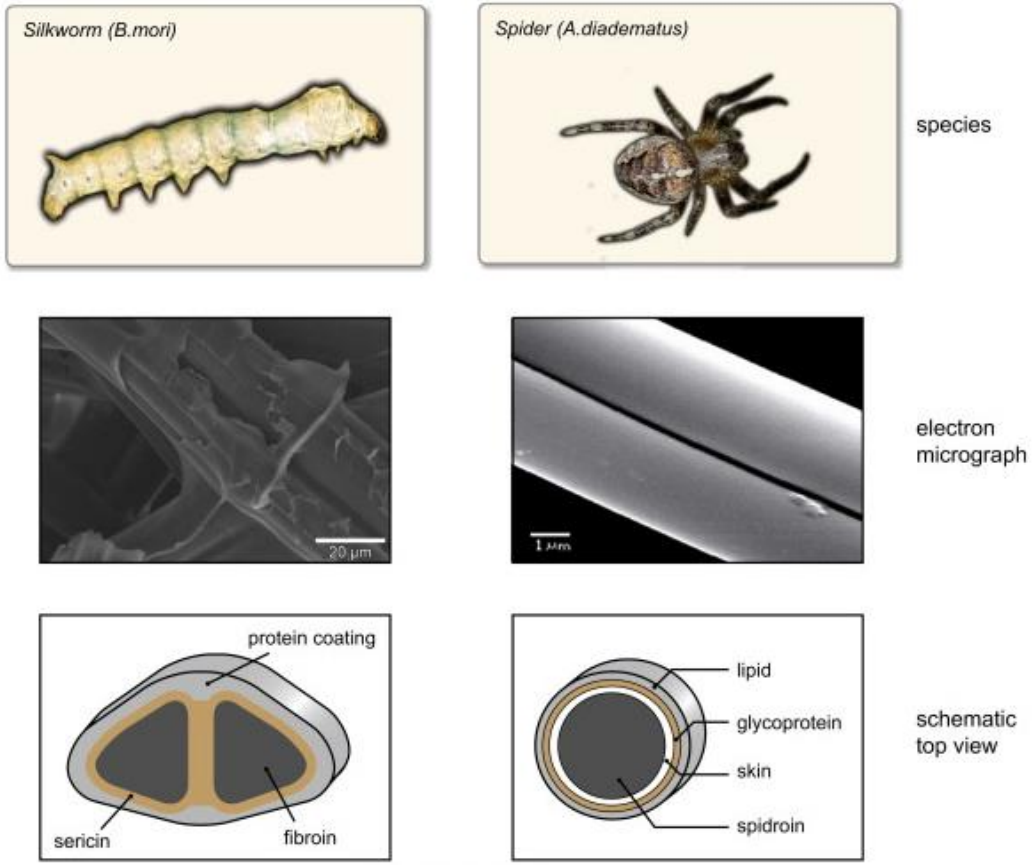
❖The overall structure is stabilized by extensive hydrogen bonding between all peptide linkages in the polypeptides of each β sheet and by the optimization of van der Waals interactions between sheets.

❖Silk does not stretch, because the β conformation is already highly extended.

❖However, the structure is flexible, because the sheets are held together by numerous weak interactions rather than by covalent bonds such as the disulfide bonds in α-keratins.


Antiparallel β sheet
(a)


(b)
70 μm


Silkworm (B.mori)
Spider (A.diadematus)
species
20 μm
1 μm
electron micrograph
protein coating
sericin
fibroin
lipid
glycoprotein
skin
spidroin
schematic top view
sketches are not to scale

❖In a **globular protein**, different segments of the polypeptide chain (or multiple polypeptide chains) fold back on each other, generating a more compact shape than is seen in the fibrous proteins.

❖The folding also provides the structural diversity necessary for proteins to carry out a wide array of biological functions.

❖Globular proteins include enzymes, transport proteins, motor proteins, regulatory proteins, immunoglobulins, and proteins with many other functions.

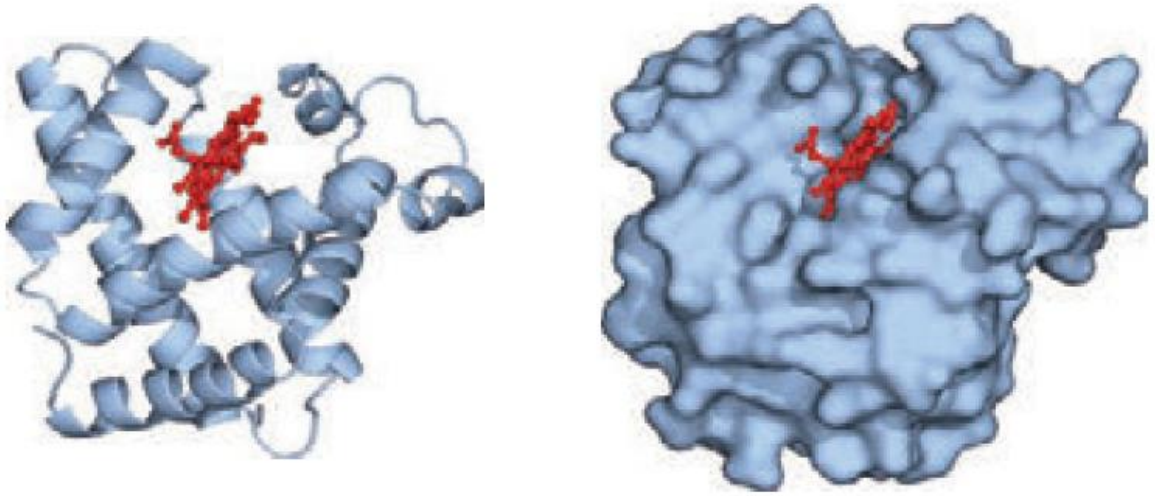
❖The first breakthrough in understanding the three-dimensional structure of a globular protein came from x-ray diffraction studies of myoglobin.

❖Myoglobin contains a single polypeptide chain of 153 amino acid residues of known sequence and a single iron protoporphyrin, or heme, group.

❖The same heme group that is found in myoglobin is found in hemoglobin, the oxygen-binding protein of erythrocytes, and is responsible for the deep red-brown color of both myoglobin and hemoglobin.

❖Myoglobin is particularly abundant in the muscles of diving mammals such as the whale, seal, and porpoise—so abundant that the muscles of these animals are brown.

❖Storage and distribution of oxygen by muscle myoglobin permits diving mammals to remain submerged for long periods.

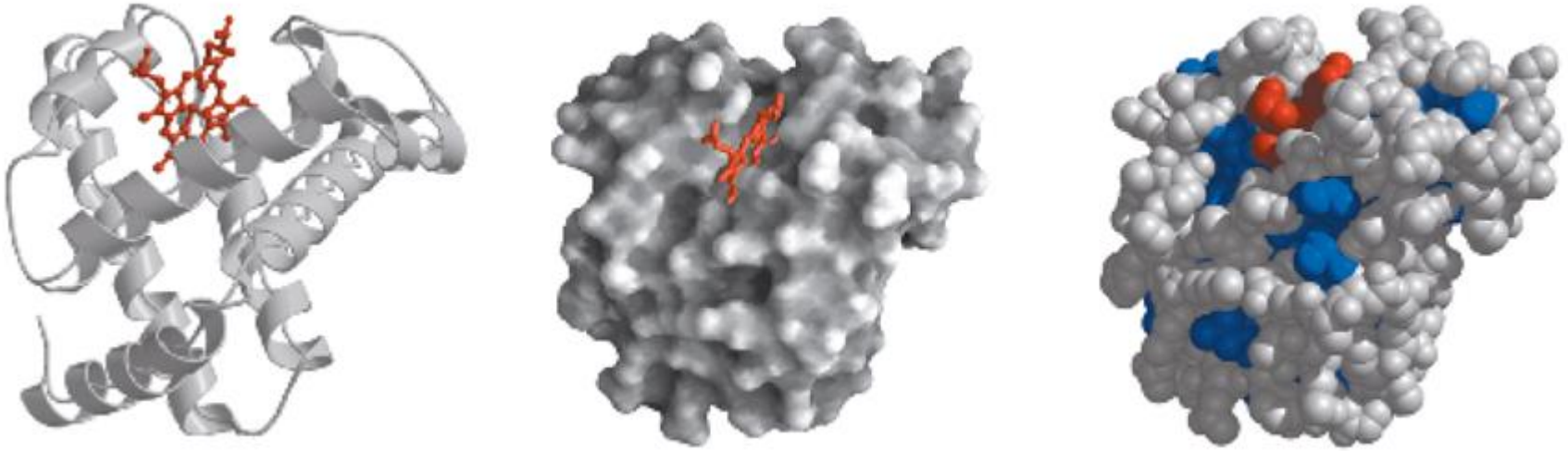


(a) (b)

❖The red group surrounded by protein is heme.

❖The backbone of the myoglobin molecule consists of eight relatively straight segments of α helix interrupted by bends, some of which are β turns.

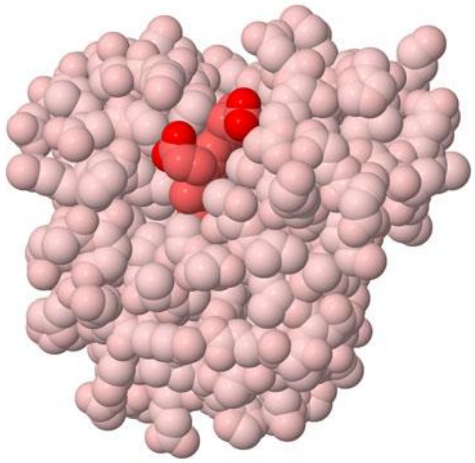
❖Most of the hydrophobic R groups (blue) are in the interior of the molecule, hidden from exposure to water.

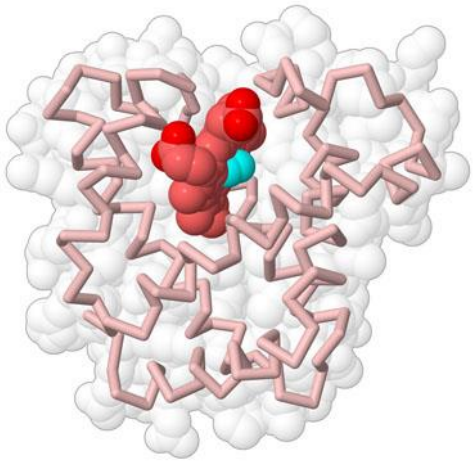


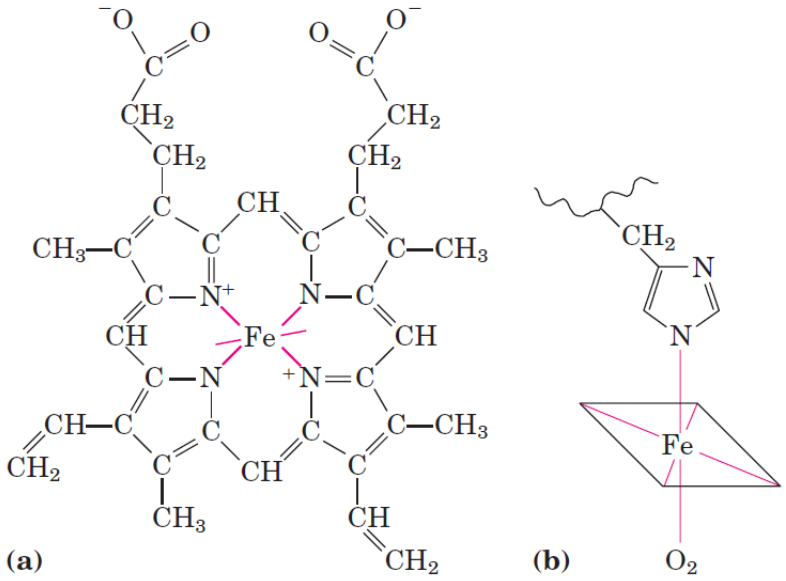
❖The flat heme group rests in a crevice, or pocket, in the myoglobin molecule.

❖The iron atom in the center of the heme group has two bonding (coordination) positions perpendicular to the plane of the heme.

❖One of these is bound to the R group of the His residue at position 93; the other is the site at which an O_2 molecule binds.







❖From what we now know about the tertiary structures of hundreds of globular proteins, it is clear that myoglobin illustrates just one of many ways in which a polypeptide chain can fold.

Protein (total residues)	Residues (%)*	
	α Helix	β Conformation
Chymotrypsin (247)	14	45
Ribonuclease (124)	26	35
Carboxypeptidase (307)	38	17
Cytochrome *c* (104)	39	0
Lysozyme (129)	40	12
Myoglobin (153)	78	0

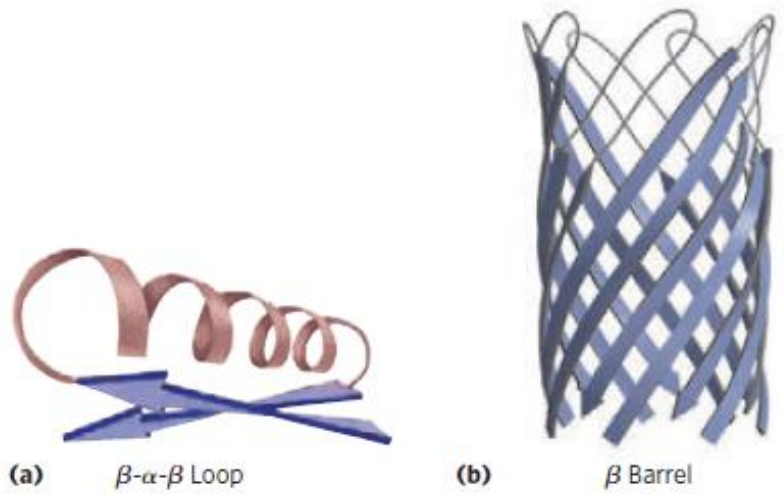
❖For the beginning student, the very complex tertiary structures of globular proteins—some much larger than myoglobin—are best approached by focusing on common structural patterns, recurring in different and often unrelated proteins.

❖To understand a complete three-dimensional structure, we need to analyze its folding patterns.

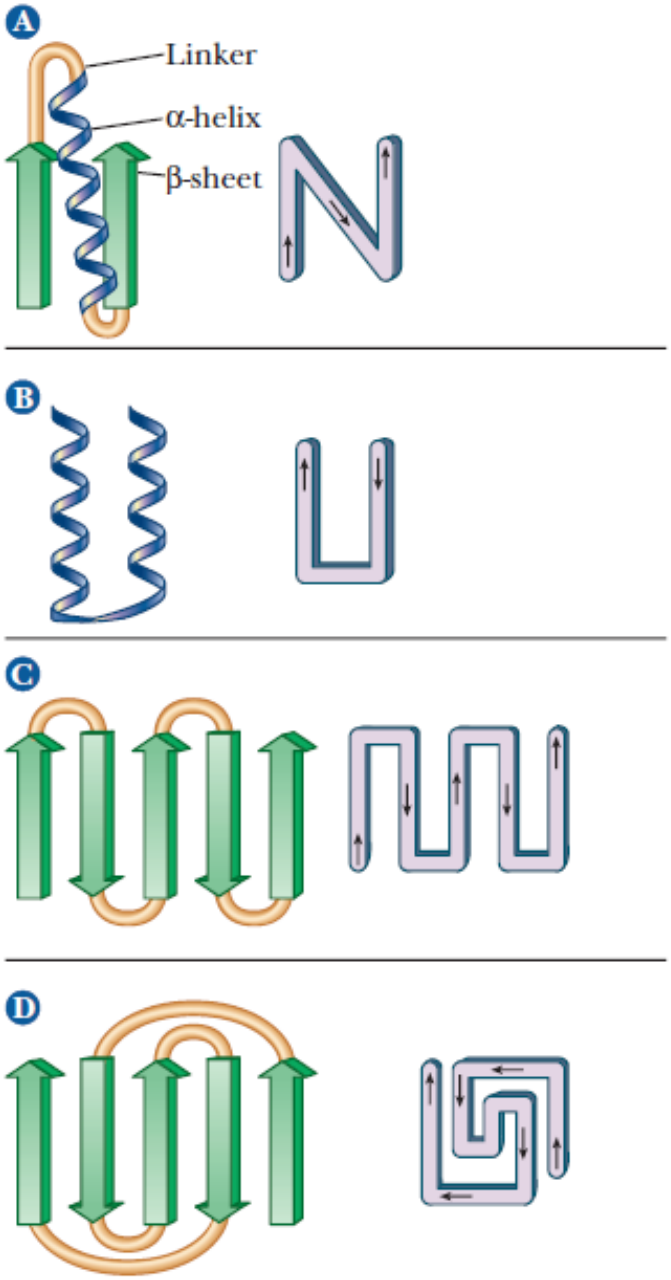
❖The first term is **motif**, also called a **fold**.

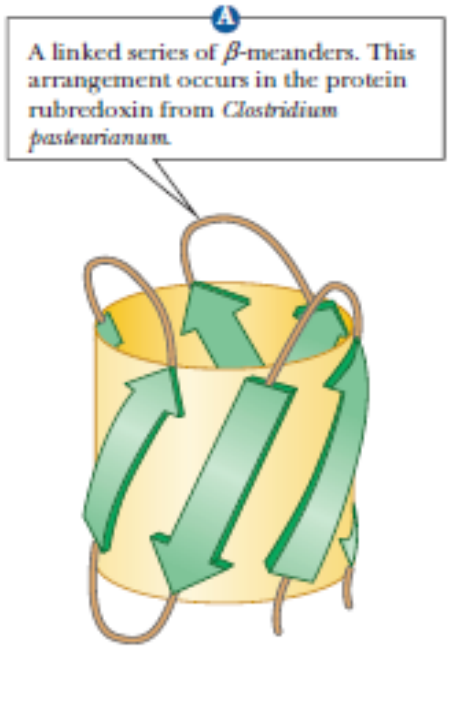
❖A motif or fold is a recognizable folding pattern involving two or more elements of secondary structure and the connection(s) between them.

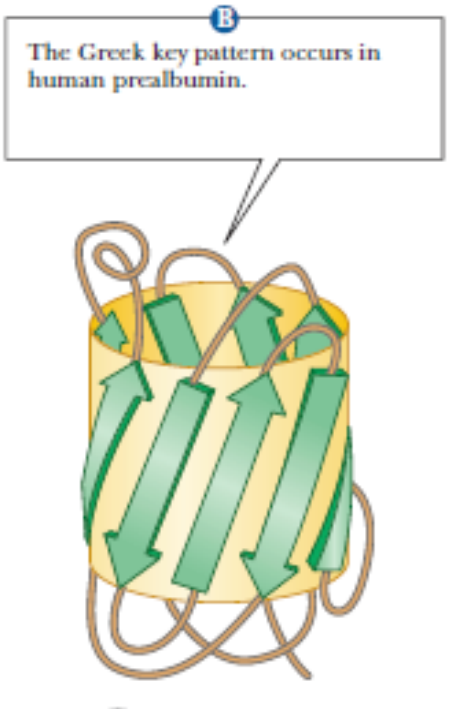
❖A motif can be very simple, such as two elements of secondary structure folded against each other, and represent only a small part of a protein. An example is a β-α-β loop.



❖A motif can also be a very elaborate structure involving scores of protein segments folded together, such as the β barrel


A
Linker
α-helix
β-sheet
B
C
D


A
A linked series of β-meanders. This arrangement occurs in the protein rubredoxin from Clostridium pasteurianum.


B
The Greek key pattern occurs in human prealbumin.

❖The second term for describing structural patterns is **domain**.

❖A domain, as defined as a part of a polypeptide chain that is independently stable or could undergo movements as a single entity with respect to the entire protein.

❖Polypeptides with more than a few hundred amino acid residues often fold into two or more domains, sometimes with different functions.

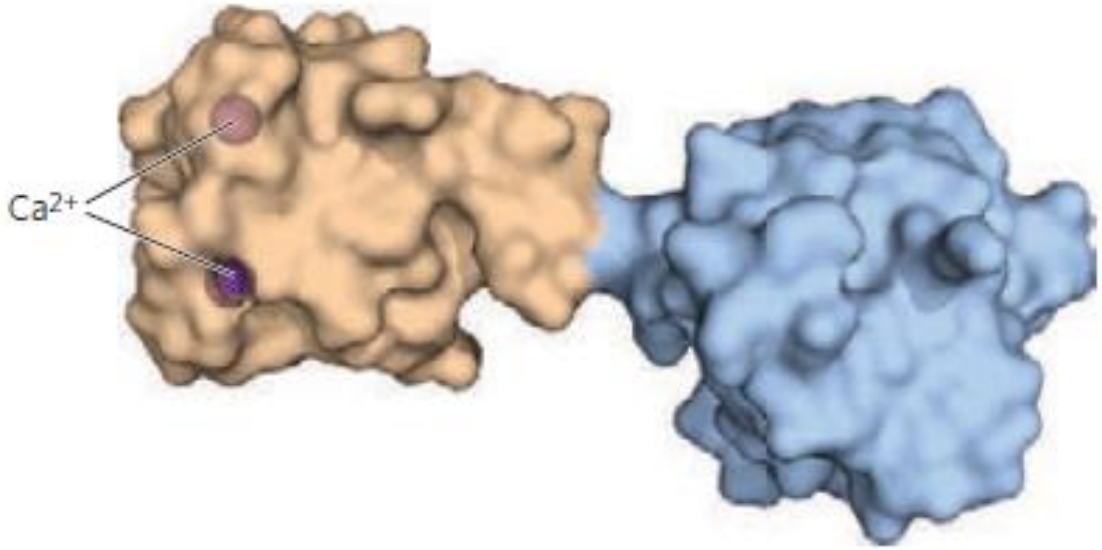


FIGURE 4–19 Structural domains in the polypeptide troponin C. (PDB

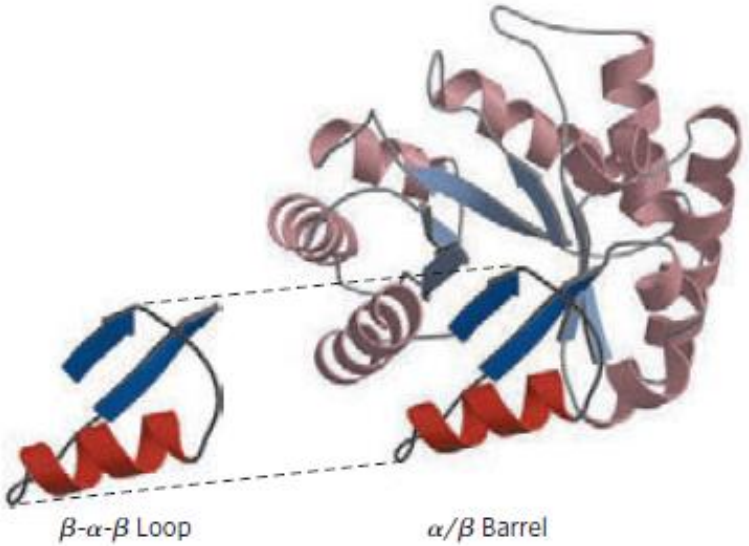
❖Folding of polypeptides is subject to an array of physical and chemical constraints, and several rules have emerged from studies of common protein folding patterns.

❖1. Hydrophobic interactions make a large contribution to the stability of protein structures. Burial of hydrophobic amino acid R groups so as to exclude water requires at least two layers of secondary structure.

❖2. Where they occur together in a protein, α helices and β sheets generally are found in different structural layers.

❖3. Segments adjacent to each other in the amino acid sequence are usually stacked adjacent to each other in the folded structure.

❖4. The β conformation is most stable when the individual segments are twisted slightly in a righthanded sense.



❖Proteins with significant similarity in primary structure and/or with similar tertiary structure and function are said to be in the same **protein family**.

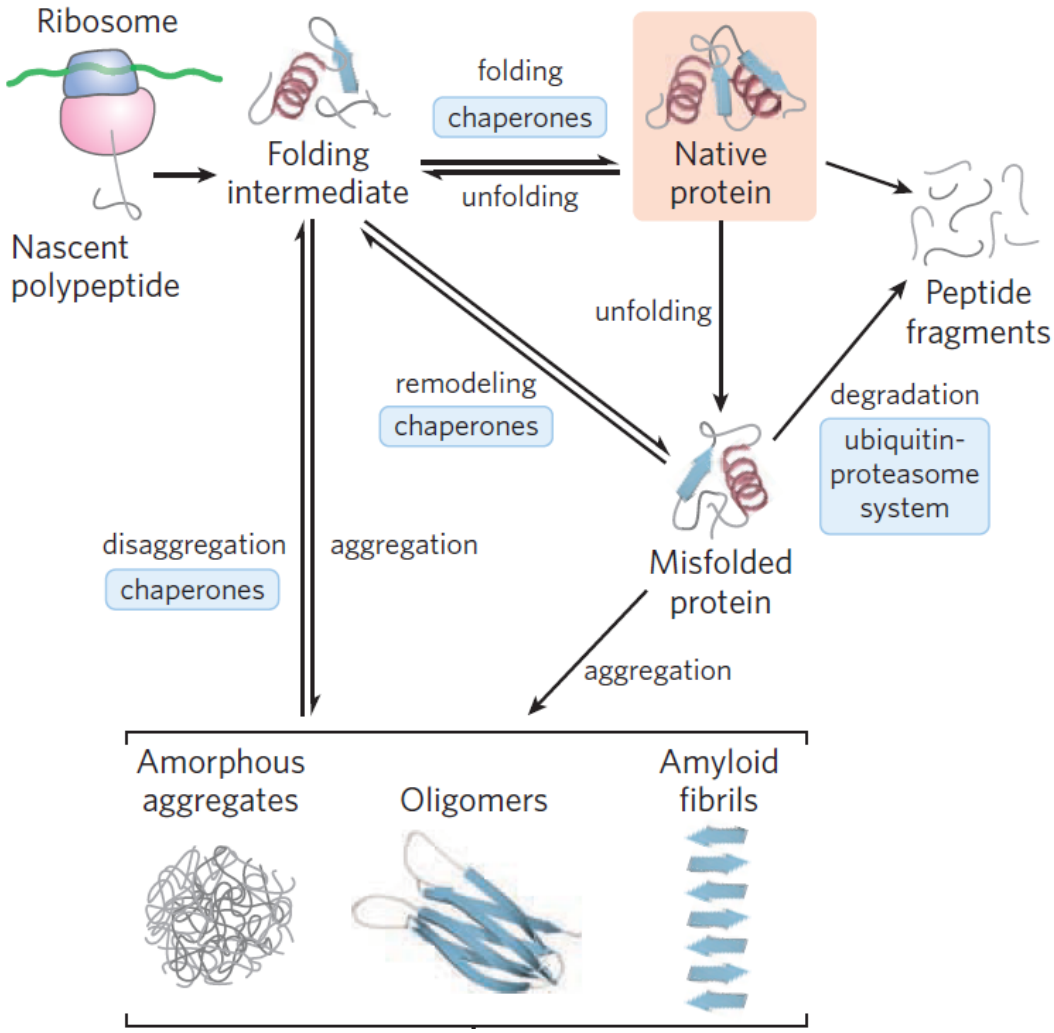
❖A strong evolutionary relationship is usually evident within a protein family.

❖Two or more families with little similarity in amino acid sequence sometimes make use of the same major structural motif and have functional similarities; these families are grouped as **superfamilies**.

Protein denaturation and folding

❖The continual maintenance of the active set of cellular proteins required under a given set of conditions is called **proteostasis**.

❖Cellular proteostasis requires the coordinated function of pathways for protein synthesis and folding, the refolding of proteins that are partially unfolded, and the sequestration and degradation of proteins that have been irreversibly unfolded. In all cells, these networks involve hundreds of enzymes and specialized proteins.



❖A loss of three-dimensional structure sufficient to cause loss of function is called **denaturation**.

❖The denatured state does not necessarily equate with complete unfolding of the protein and randomization of conformation.

❖Most proteins can be denatured by heat, which has complex effects on many weak interactions in a protein (primarily on the hydrogen bonds).

❖If the temperature is increased slowly, a protein's conformation generally remains intact until an abrupt loss of structure (and function) occurs over a narrow temperature range.

❖The effects of heat on proteins are not readily predictable.

❖Proteins can also be denatured by extremes of pH, by certain miscible organic solvents such as alcohol or acetone, by certain solutes such as urea and guanidine hydrochloride, or by detergents.

❖Each of these denaturing agents represents a relatively mild treatment in the sense that no covalent bonds in the polypeptide chain are broken.

❖Organic solvents, urea, and detergents act primarily by disrupting the hydrophobic interactions that make up the stable core of globular proteins; urea also disrupts hydrogen bonds; extremes of pH alter the net charge on the protein, causing electrostatic repulsion and the disruption of some hydrogen bonding.

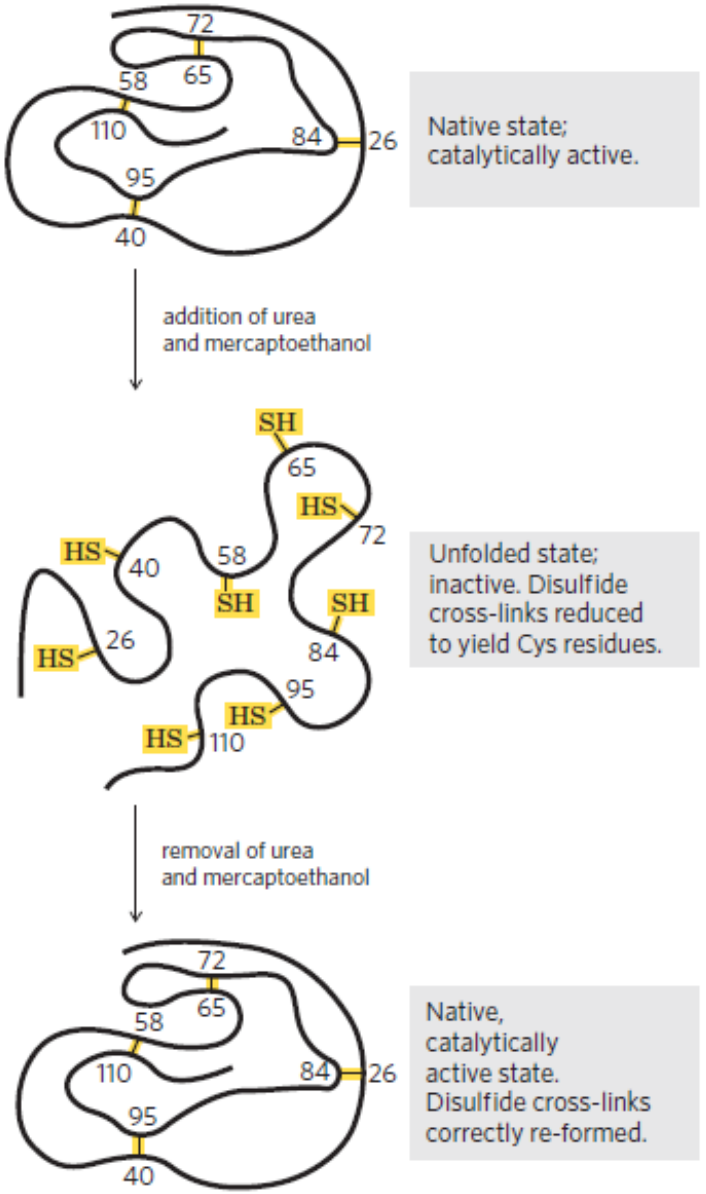
❖Denaturation often leads to protein precipitation, a consequence of protein aggregate formation as exposed hydrophobic surfaces associate.

❖Certain globular proteins denatured by heat, extremes of pH, or denaturing reagents will regain their native structure and their biological activity if returned to conditions in which the native conformation is stable. This process is called **renaturation**.

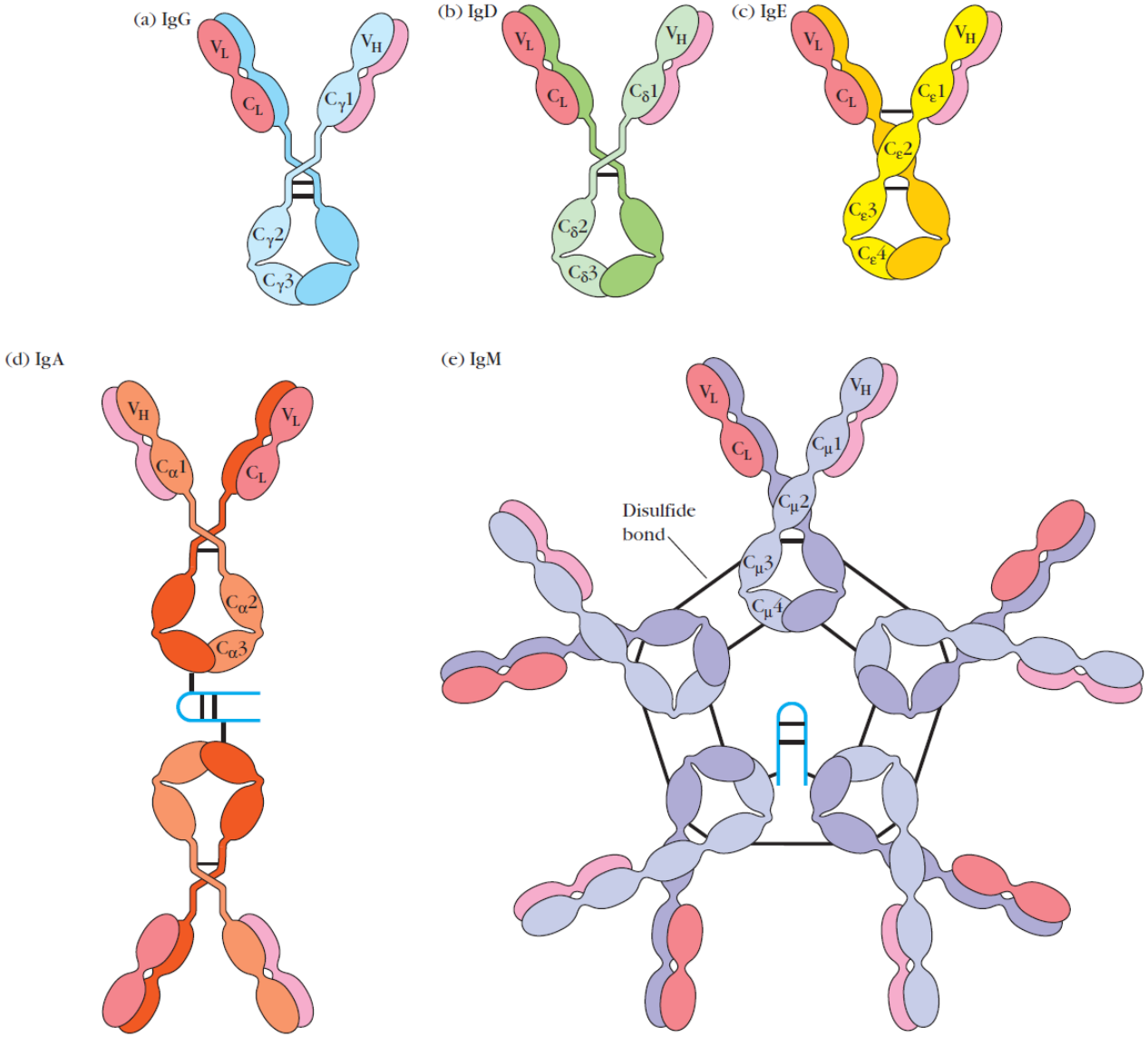
❖The folding pathway of a large polypeptide chain is unquestionably complicated.

❖Not all proteins fold spontaneously as they are synthesized in the cell.

❖Folding for many proteins requires **chaperones**, proteins that interact with partially folded or improperly folded polypeptides, facilitating correct folding pathways or providing microenvironments in which folding can occur.



AMINO ACIDS, PEPTIDES AND PROTEINS IV: Functions of Proteins



❖Knowing the three-dimensional structure of a protein is an important part of understanding how the protein functions.

❖However, the structure shown in two dimensions on a page is deceptively static.

❖Proteins are dynamic molecules whose functions almost invariably depend on interactions with other molecules.

❖These interactions are affected in physiologically important ways by sometimes subtle, sometimes striking changes in protein conformation.

❖They are the basis of complex physiological processes such as oxygen transport, immune function, and muscle contraction.

❖The functions of many proteins involve the reversible binding of other molecules.

❖A molecule bound reversibly by a protein is called a **ligand**.

❖The transient nature of protein-ligand interactions is critical to life, allowing an organism to respond rapidly and reversibly to changing environmental and metabolic circumstances.

❖A ligand binds at a site on the protein called the **binding site**, which is complementary to the ligand in size, shape, charge, and hydrophobic or hydrophilic character.

❖Furthermore, the interaction is specific: the protein can discriminate among the thousands of different molecules in its environment and selectively bind only one or a few.

❖Proteins are flexible.

❖Changes in conformation may be subtle, reflecting molecular vibrations and small movements of amino acid residues throughout the protein.

❖They may also be quite dramatic, with major segments of the protein structure moving as much as several nanometers.

❖Specific conformational changes are frequently essential to a protein's function.

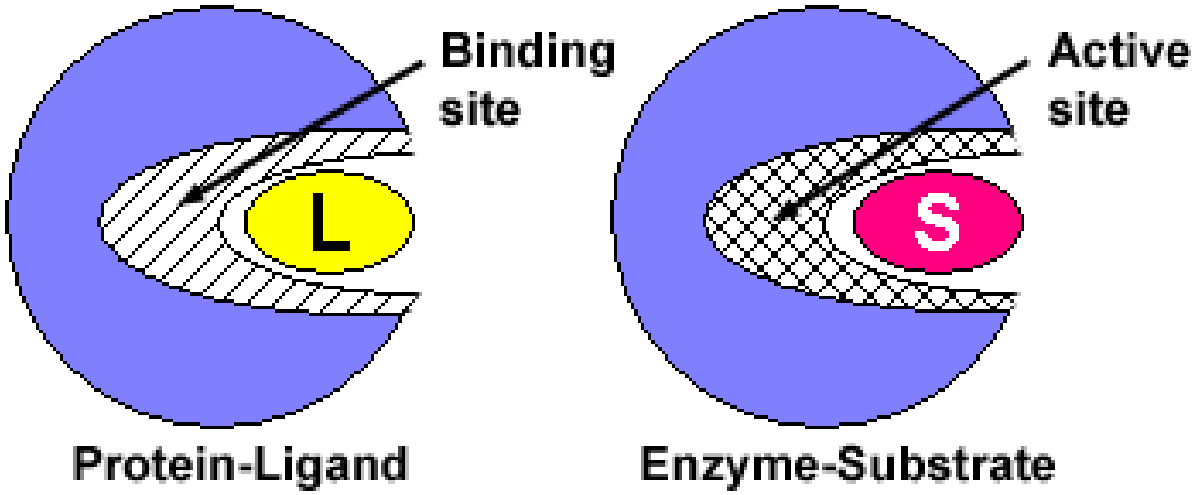
❖The binding of a protein and ligand is often coupled to a conformational change in the protein that makes the binding site more complementary to the ligand, permitting tighter binding.

❖The structural adaptation that occurs between protein and ligand is called **induced fit**.

❖Enzymes represent a special case of protein function.

❖Enzymes bind and chemically transform other molecules—they catalyze reactions.

❖The molecules acted upon by enzymes are called reaction **substrates** rather than ligands, and the ligand-binding site is called the **catalytic site** or **active site**.



Oxygen Binding Proteins

❖Oxygen is poorly soluble in aqueous solutions and cannot be carried to tissues in sufficient quantity if it is simply dissolved in blood serum.

❖The evolution of larger, multicellular animals depended on the evolution of proteins that could transport and store oxygen.

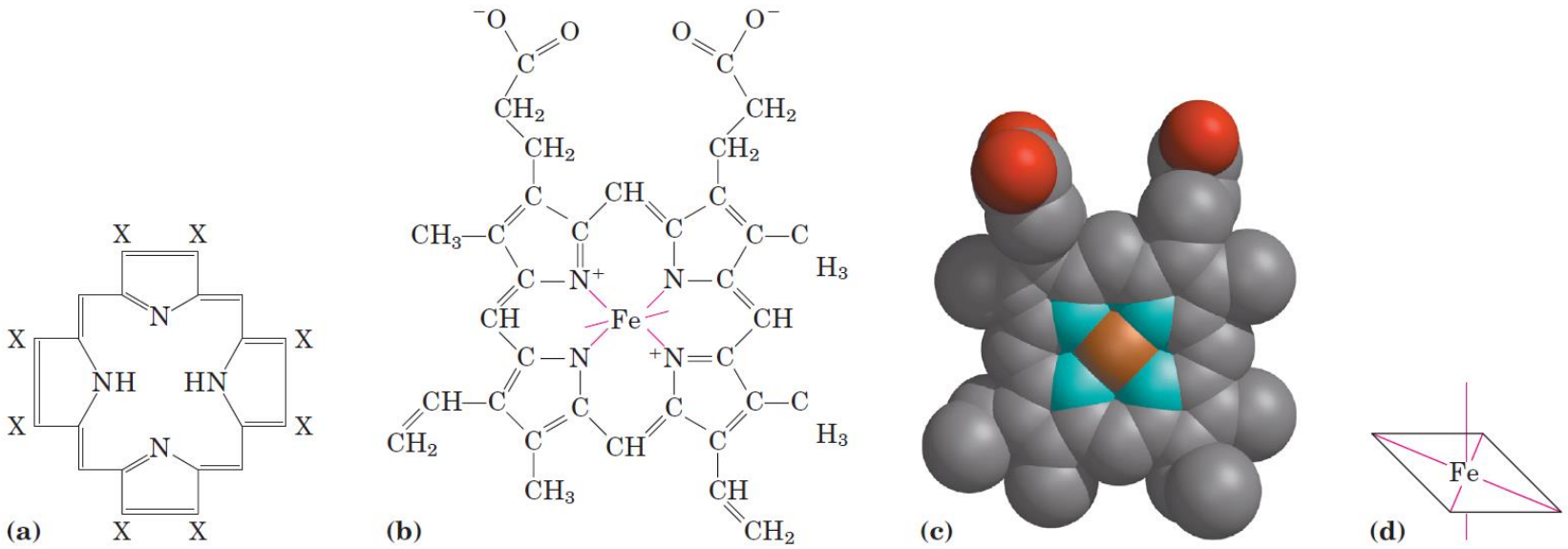
❖However, none of the amino acid side chains in proteins are suited for the reversible binding of oxygen molecules.

❖This role is filled by certain transition metals, among them iron and copper, that have a strong tendency to bind oxygen.

❖However, free iron promotes the formation of highly reactive oxygen species such as hydroxyl radicals that can damage DNA and other macromolecules.

❖Iron used in cells is therefore bound in forms that sequester it and/or make it less reactive.

❖In multicellular organisms iron is often incorporated into a protein bound prosthetic group called **heme**.

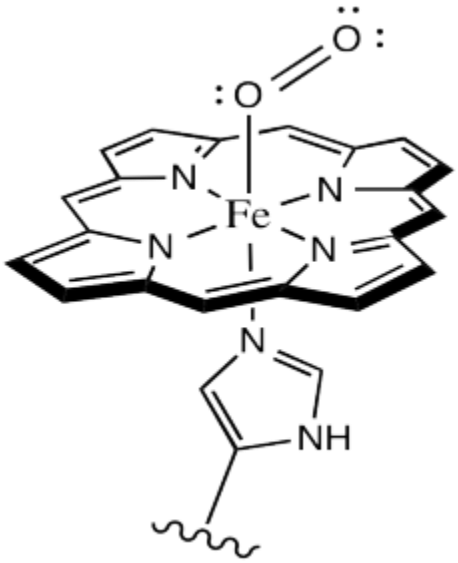


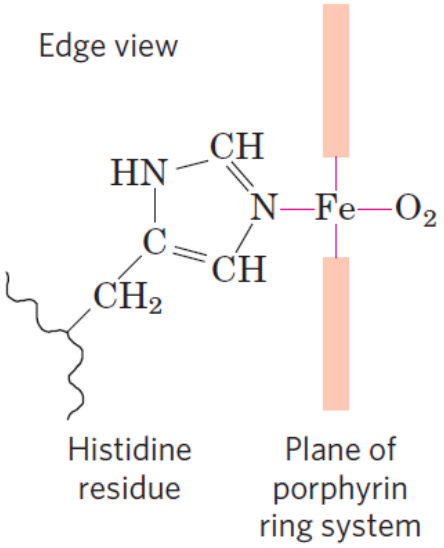
❖Heme consists of a complex organic ring structure, **protoporphyrin**, to which is bound a single iron atom in its ferrous (Fe^{2+}) state .

❖The iron atom has six coordination bonds, four to nitrogen atoms that are part of the flat porphyrin ring system and two perpendicular to the porphyrin.

❖The coordinated nitrogen atoms (which have an electron-donating character) help prevent conversion of the heme iron to the ferric (Fe^{3+}) state.

❖Iron in the Fe^{2+} state binds oxygen reversibly; in the Fe^{3+} state it does not bind oxygen.





❖When oxygen binds, the electronic properties of heme iron change; this accounts for the change in color from the dark purple of oxygen-depleted venous blood to the bright red of oxygen-rich arterial blood.

❖Some small molecules, such as carbon monoxide (CO) and nitric oxide (NO), coordinate to heme iron with greater affinity than does O_2.

❖When a molecule of CO is bound to heme, O_2 is excluded, which is why CO is highly toxic to aerobic organisms.

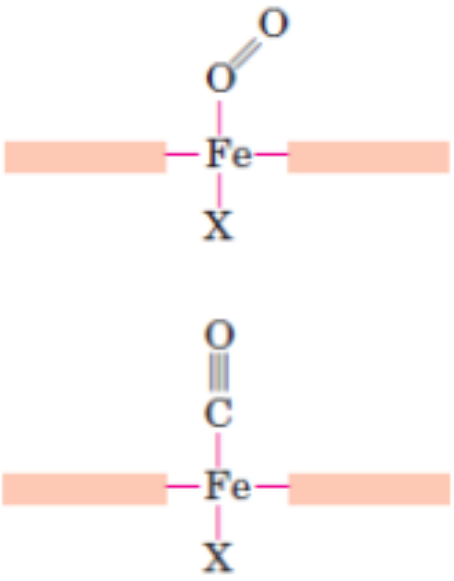
❖By surrounding and sequestering heme, oxygen-binding proteins regulate the access of CO and other small molecules to the heme iron.

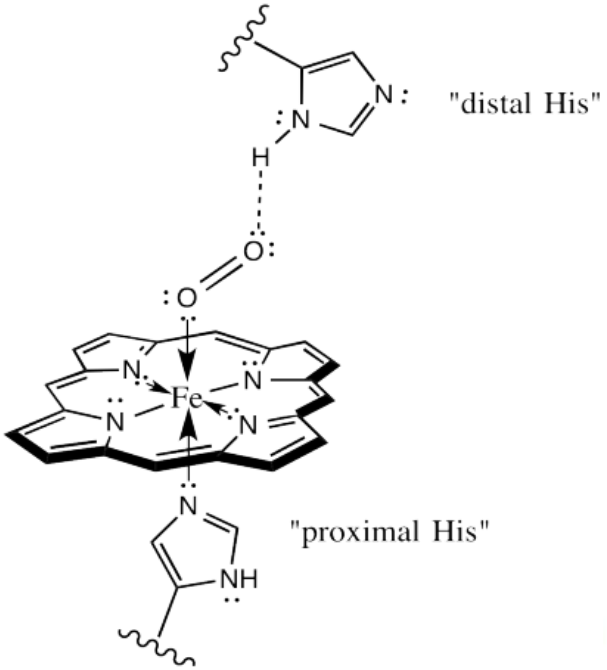
❖Carbon monoxide binds to free heme molecules more than 20,000 times better than does O_2, but it binds only about 200 times better than O_2 when the heme is bound in protein.

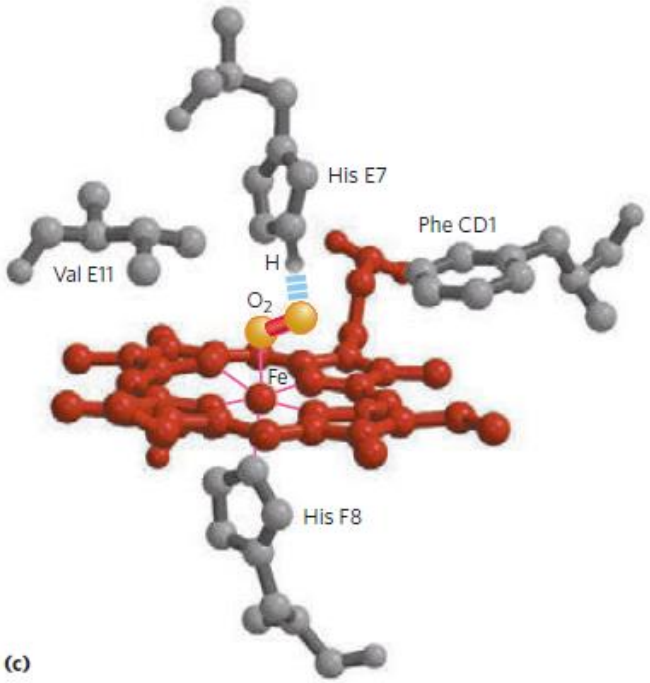
❖The difference may be partly explained by steric hindrance. When O_2 binds to free heme, the axis of the oxygen molecule is positioned at an angle to the Fe-O bond.

❖In contrast, when CO binds to free heme, the Fe, C, and O atoms lie in a straight line.

❖To stabilize this conformation, His E7 residue, called the distal His (as distinct from the proximal His, His F8), forms a hydrogen bond with O_2.







❖The **globins** are a widespread family of proteins, all having similar primary and tertiary structures.

❖In humans and other mammals, there are at least four kinds of globins.

❖The monomeric **myoglobin** facilitates oxygen diffusion in muscle tissue. Myoglobin is particularly abundant in the muscles of diving marine mammals such as seals and whales.

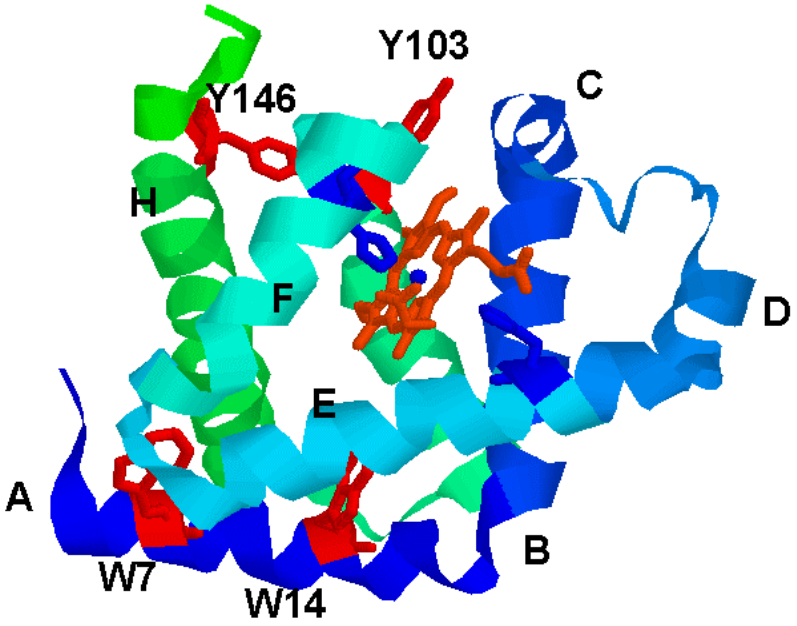
❖The tetrameric **hemoglobin** is responsible for oxygen transport in the blood stream.

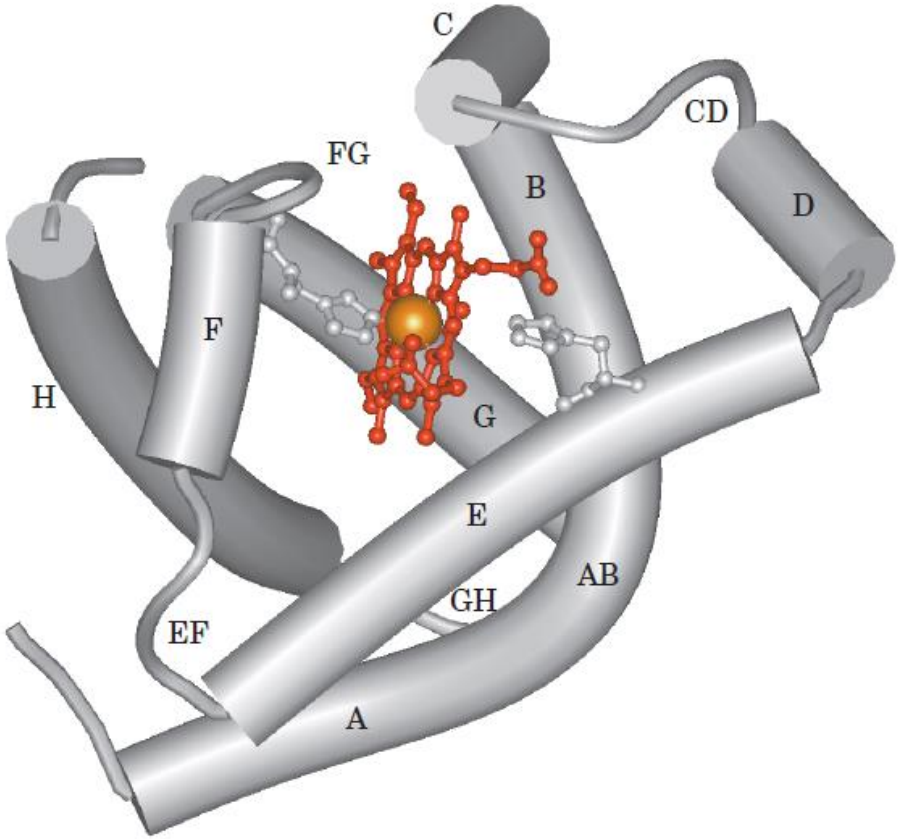
❖The monomeric **neuroglobin** is expressed largely in neurons and helps to protect the brain from hypoxia (low oxygen) or ischemia (restricted blood supply).

❖**Cytoglobin**, another monomeric globin, is found at high levels in a range of tissues, but its function is unknown.

❖Myoglobin (Mr 16,700) is a single polypeptide of 153 amino acid residues with one molecule of heme.

❖As is typical for a globin polypeptide, myoglobin is made up of eight α-helical segments connected by bends.





❖The function of myoglobin depends on the protein's ability not only to bind oxygen but also to release it when and where it is needed.

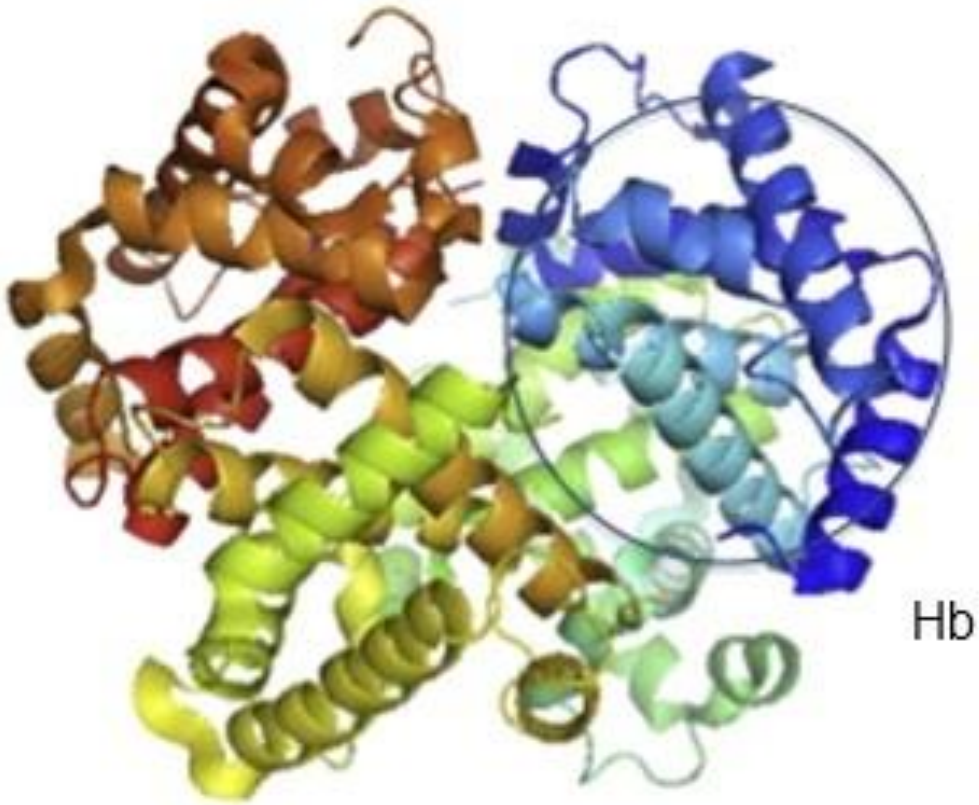
❖Nearly all the oxygen carried by whole blood in animals is bound and transported by hemoglobin in erythrocytes (red blood cells).

❖Normal human erythrocytes are formed from precursor stem cells called hemocytoblasts.

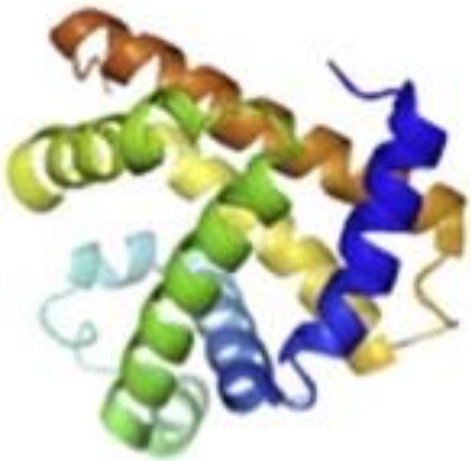
❖In the maturation process, the stem cell produces daughter cells that form large amounts of hemoglobin and then lose their intracellular organelles—nucleus, mitochondria, and endoplasmic reticulum.

❖Erythrocytes are thus incomplete, vestigial cells, unable to reproduce and, their main function is to carry hemoglobin, which is dissolved in the cytosol at a very high concentration (,34% by weight).

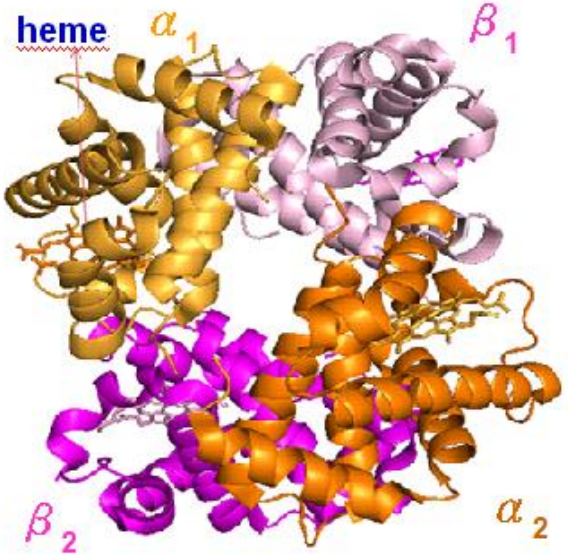
❖In arterial blood passing from the lungs through the heart to the peripheral tissues, hemoglobin is about 96% saturated with oxygen. In the venous blood returning to the heart, hemoglobin is only about 64% saturated.

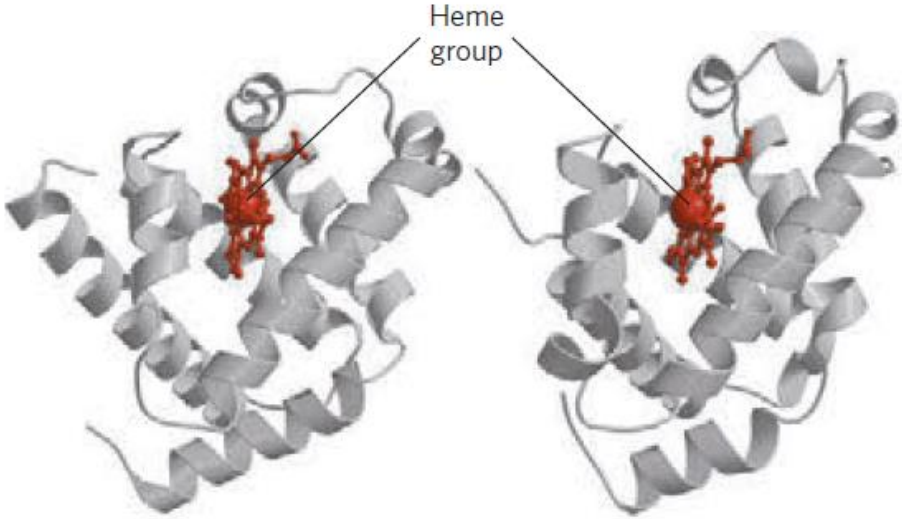


Hb



Mb





Myoglobin

β subunit of
hemoglobin

❖Hemoglobin subunits are structurally similar to myoglobin

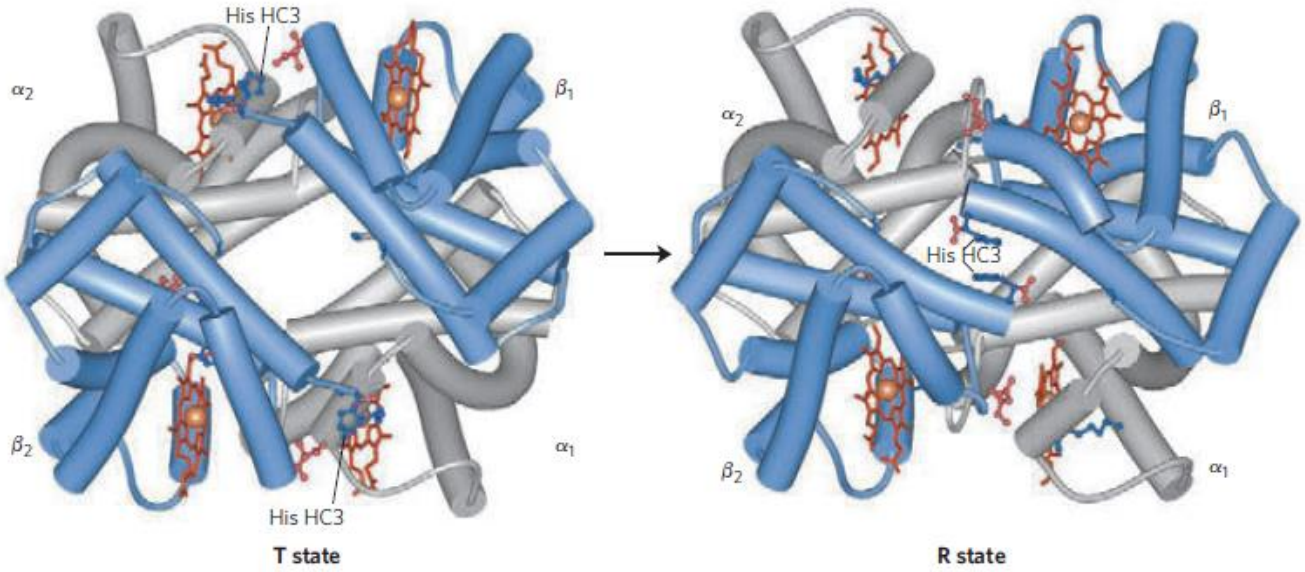
❖Hemoglobin (Mr 64,500; abbreviated Hb) is a **tetrameric** protein containing four heme prosthetic groups, one associated with each polypeptide chain.

❖Adult hemoglobin contains two types of globin, two α chains and two β chains. Hydrophobic interactions predominate at all the αβ dimer interfaces, but there are also many hydrogen bonds and a few ion pairs (or salt bridges).

❖Although fewer than half of the amino acid residues are identical in the polypeptide sequences of the α and βsubunits, the three-dimensional structures of the two types of subunits are very similar.

❖Furthermore, their structures are very similar to that of myoglobin, even though the amino acid sequences of the three polypeptides are identical at only 27 positions.

❖All three polypeptides are members of the globin family of proteins.



❖Hemoglobin has two major conformations: the "relaxed" **R state** and the "tense" **T state**.

❖Although oxygen binds to hemoglobin in either state, it has a significantly higher affinity for hemoglobin in the R state. Oxygen binding stabilizes the R state.

❖When oxygen is absent experimentally, the T state is more stable and is thus the predominant conformation of deoxyhemoglobin.

❖The binding of O_2 to a hemoglobin subunit in the T state triggers a change in conformation to the R state.

❖Hemoglobin must bind oxygen efficiently in the lungs, where the pO_2 is about 13.3 kPa, and release oxygen in the tissues, where the pO_2 is about 4 kPa.

❖A protein that bound O_2 with high affinity would bind it efficiently in the lungs but would not release much of it in the tissues.

❖If the protein bound oxygen with a sufficiently low affinity to release it in the tissues, it would not pick up much oxygen in the lungs.

❖Hemoglobin solves the problem by undergoing a transition from a low-affinity state (the T state) to a high affinity state (the R state) as more O_2 molecules are bound.

❖Binding of the O2 to deoxyhemoglobin leads to conformational changes that making it easier for additional molecules of O_2 to bind.

❖Proteins that have “other shapes,” or conformations, induced by the binding of ligands (or modulators) called as **allosteric** proteins..

❖In addition to carrying nearly all the oxygen required by cells from the lungs to the tissues, hemoglobin carries two end products of cellular respiration—H^+ and CO_2— from the tissues to the lungs and the kidneys, where they are excreted.

❖The CO_2, produced by oxidation of organic fuels in mitochondria, is hydrated to form bicarbonate by the action of **carbonic anhydrase:**

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

❖Carbon dioxide is not very soluble in aqueous solution, and bubbles of CO_2 would form in the tissues and blood if it were not converted to bicarbonate.

❖As you can see from the reaction catalyzed by carbonic anhydrase, the hydration of CO_2 results in an increase in the H1 concentration (a decrease in pH) in the tissues.

❖The binding of oxygen by hemoglobin is profoundly influenced by pH and CO_2 concentration,

❖The binding of H^+ and CO_2 is inversely related to the binding of oxygen.

❖At the relatively low pH and high CO_2 concentration of peripheral tissues, the affinity of hemoglobin for oxygen decreases as H^+ and CO_2 are bound, and O_2 is released to the tissues.

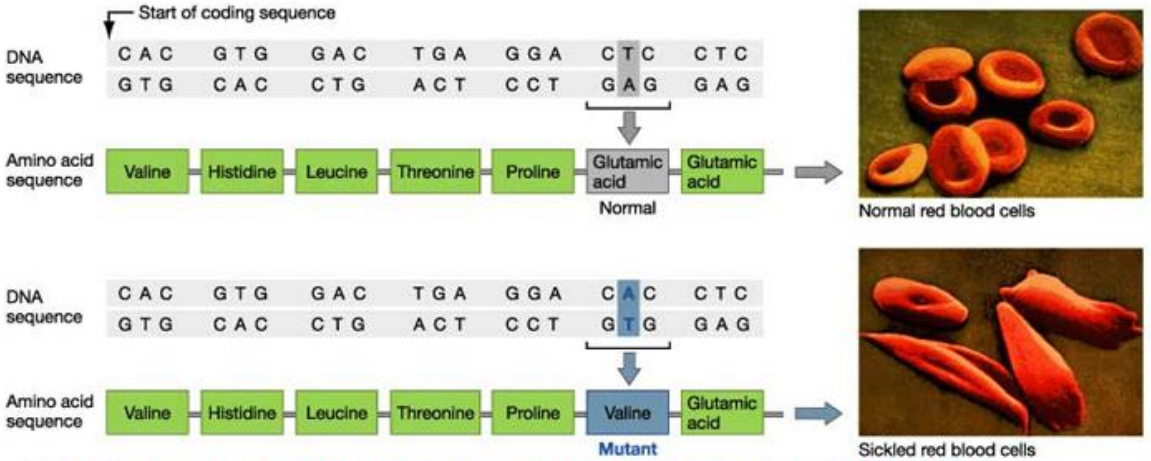
❖Conversely, in the capillaries of the lung, as CO_2 is excreted and the blood pH consequently rises, the affinity of hemoglobin for oxygen increases and the protein binds more O_2 for transport to the peripheral tissues.

❖This effect of pH and CO_2 concentration on the binding and release of oxygen by hemoglobin is called the Bohr effect.

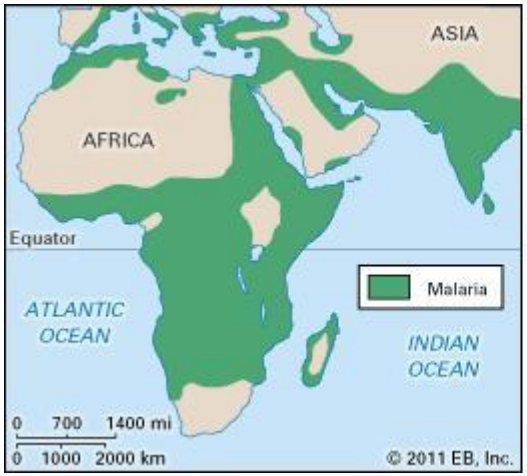
$$HHb^+ + O_2 \rightleftharpoons HbO_2 + H^+$$

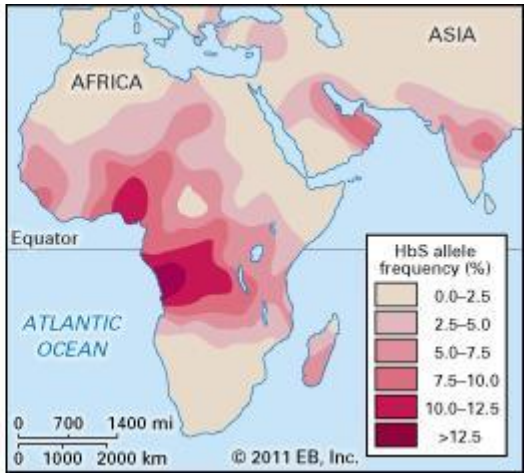
❖Oxygen and H^+ are not bound at the same sites in hemoglobin. Oxygen binds to the iron atoms of the hemes, whereas H^+ binds to any of several amino acid residues in the protein.

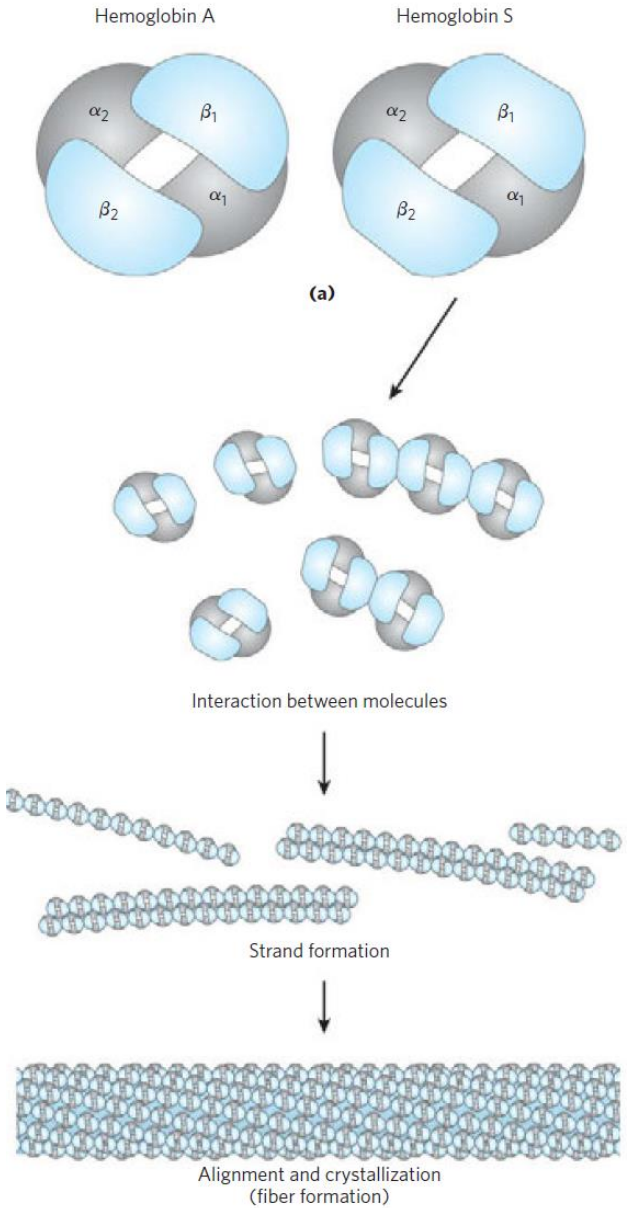
❖Oxygen binding to hemoglobin is also regulated by 2,3-bisphosphoglycerate which is known to greatly reduce the affinity of hemoglobin for oxygen

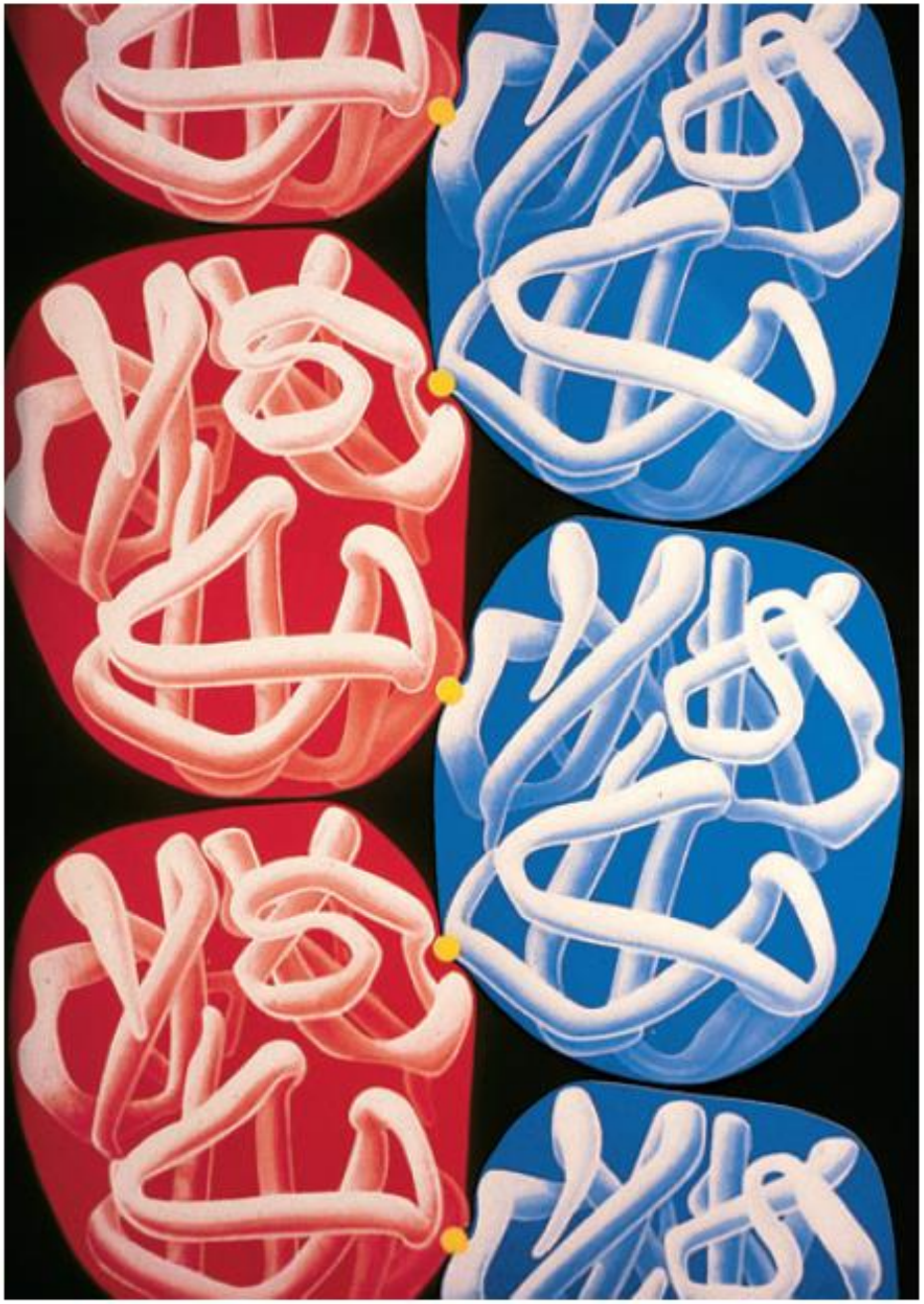


The change in amino acid sequence causes hemoglobin molecules to crystallize when oxygen levels in the blood are low. As a result, red blood cells sickle and get stuck in small blood vessels.

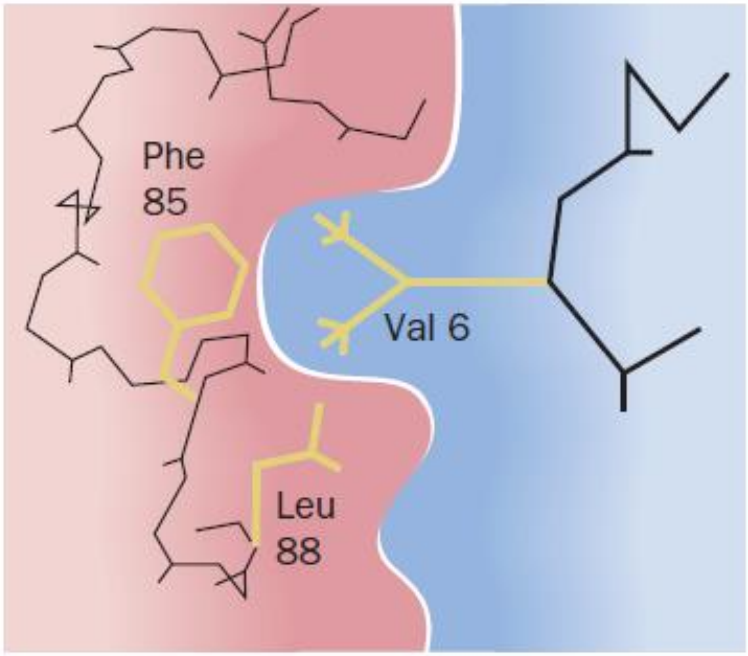








(a)



(b)

The Immune System and Immunoglobulins

❖All vertebrates have an immune system capable of distinguishing molecular “self” from “nonself” and then destroying what is identified as nonself.

❖In this way, the immune system eliminates viruses, bacteria, and other pathogens and molecules that may pose a threat to the organism.

❖On a physiological level, the immune response is an intricate and coordinated set of interactions among many classes of proteins, molecules, and cell types.

❖At the level of individual proteins, the immune response demonstrates how an acutely sensitive and specific biochemical system is built upon the reversible binding of ligands to proteins.

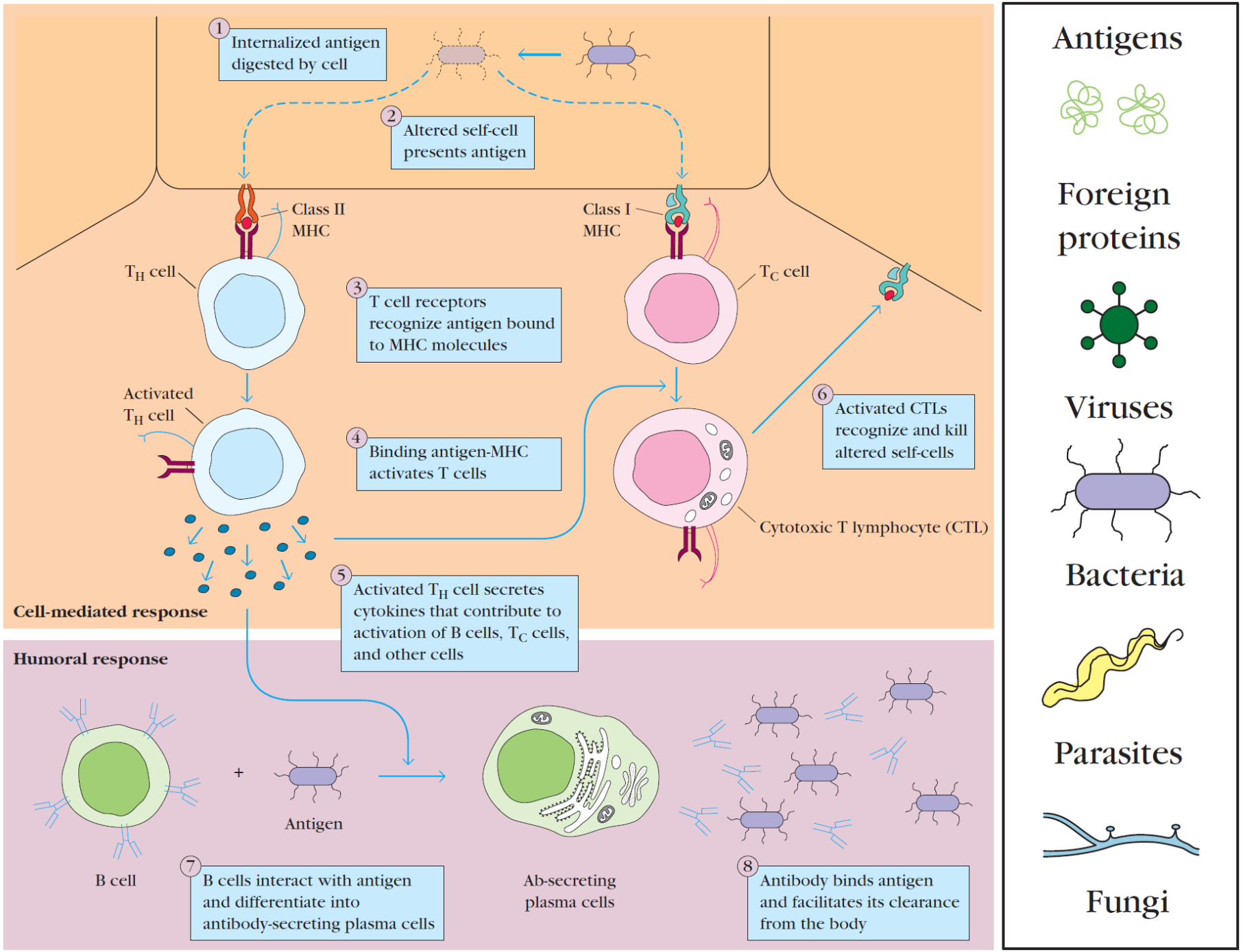
❖Immunity is brought about by a variety of **leukocytes** (white blood cells), including **macrophages** and **lymphocytes**, all of which develop from undifferentiated stem cells in the bone marrow.

❖Leukocytes can leave the bloodstream and patrol the tissues, each cell producing one or more proteins capable of recognizing and binding to molecules that might signal an infection.

❖The immune response consists of two complementary systems, the humoral and cellular immune systems.

❖The **humoral immune system** (Latin humor, “fluid”) is directed at bacterial infections and extracellular viruses (those found in the body fluids), but can also respond to individual foreign proteins.

❖The **cellular immune system** destroys host cells infected by viruses and also destroys some parasites and foreign tissues.


1 Internalized antigen digested by cell
2 Altered self-cell presents antigen
Class II MHC
Class I MHC
T_H cell
T_C cell
3 T cell receptors recognize antigen bound to MHC molecules
Activated T_H cell
4 Binding antigen-MHC activates T cells
6 Activated CTLs recognize and kill altered self-cells
Cytotoxic T lymphocyte (CTL)
5 Activated T_H cell secretes cytokines that contribute to activation of B cells, T_C cells, and other cells
Cell-mediated response
Humoral response
+
Antigen
B cell
7 B cells interact with antigen and differentiate into antibody-secreting plasma cells
Ab-secreting plasma cells
8 Antibody binds antigen and facilitates its clearance from the body
Antigens
Foreign proteins
Viruses
Bacteria
Parasites
Fungi

❖At the heart of the humoral immune response are soluble proteins called **antibodies** or **immunoglobulins**, often abbreviated **Ig**. Immunoglobulins bind bacteria, viruses, or large molecules identified as foreign and target them for destruction.

❖Making up 20% of blood protein, the immunoglobulins are produced by **B lymphocytes**, or **B cells**, so named because they complete their development in the bone marrow.

❖The agents at the heart of the cellular immune response are a class of **T lymphocytes**, or **T cells** (so called because the latter stages of their development occur in the thymus), known as **cytotoxic T cells** (T_C cells, also called killer T cells).

❖Recognition of infected cells or parasites involves proteins called T-cell receptors on the surface of T_C cells.

❖In addition to cytotoxic T cells, there are **helper T cells** (T_H cells), whose function it is to produce soluble signaling proteins called cytokines, which include the interleukins. TH cells interact with macrophages.

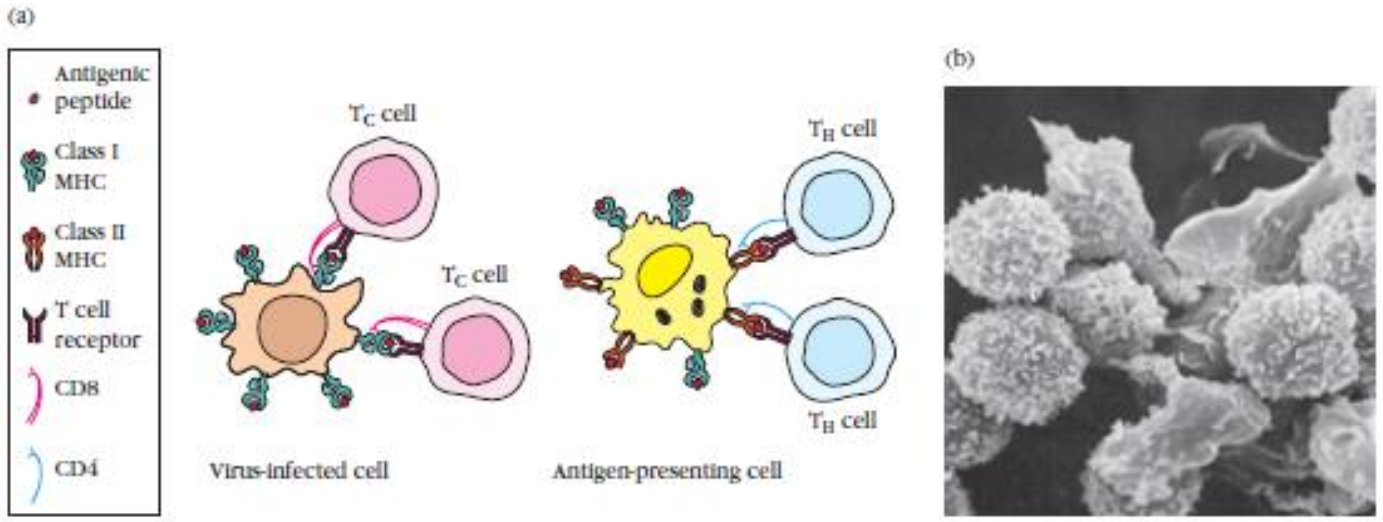


TABLE 5–2 Some Types of Leukocytes Associated with the Immune System

Cell type	Function
Macrophages	Ingest large particles and cells by phagocytosis
B lymphocytes (B cells)	Produce and secrete antibodies
T lymphocytes (T cells)	
Cytotoxic (killer) T cells (T_C)	Interact with infected host cells through receptors on T-cell surface
Helper T cells (T_H)	Interact with macrophages and secrete cytokines (interleukins) that stimulate T_C, T_H, and B cells to proliferate.

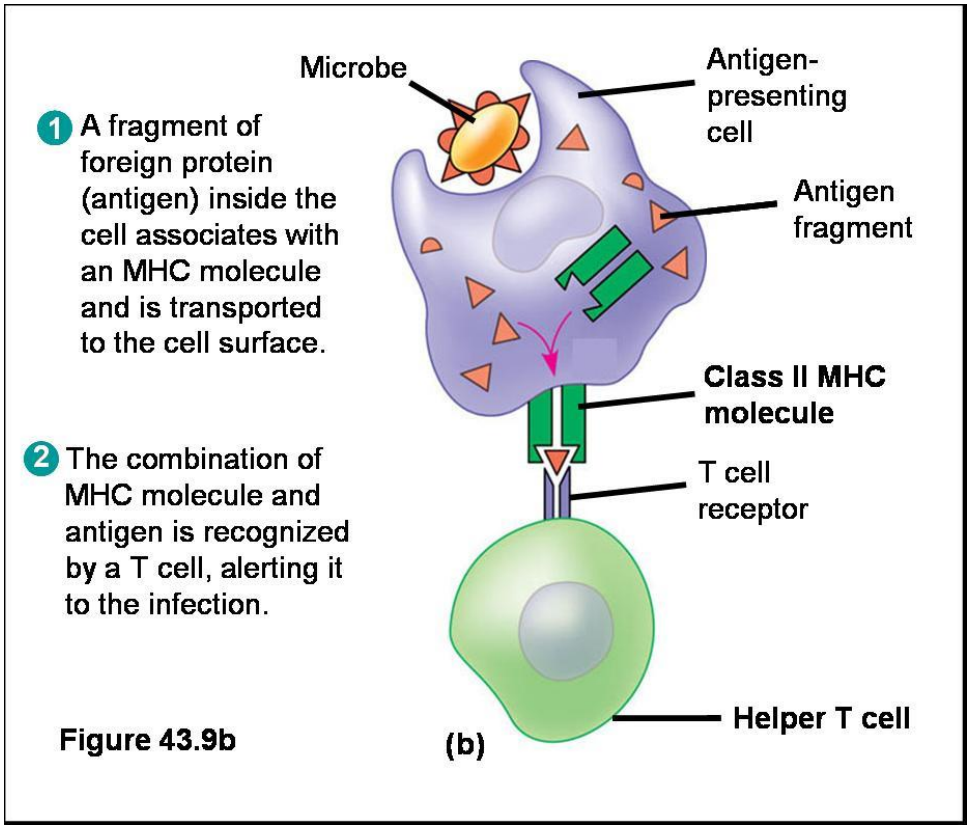


Figure 43.9b

❖Any molecule or pathogen capable of eliciting an immune response is called an **antigen**.

❖An antigen may be a virus, a bacterial cell wall, or an individual protein or other macromolecule.

❖An individual antibody or T-cell receptor binds only a particular molecular structure within the antigen, called its **antigenic determinant** or **epitope**.

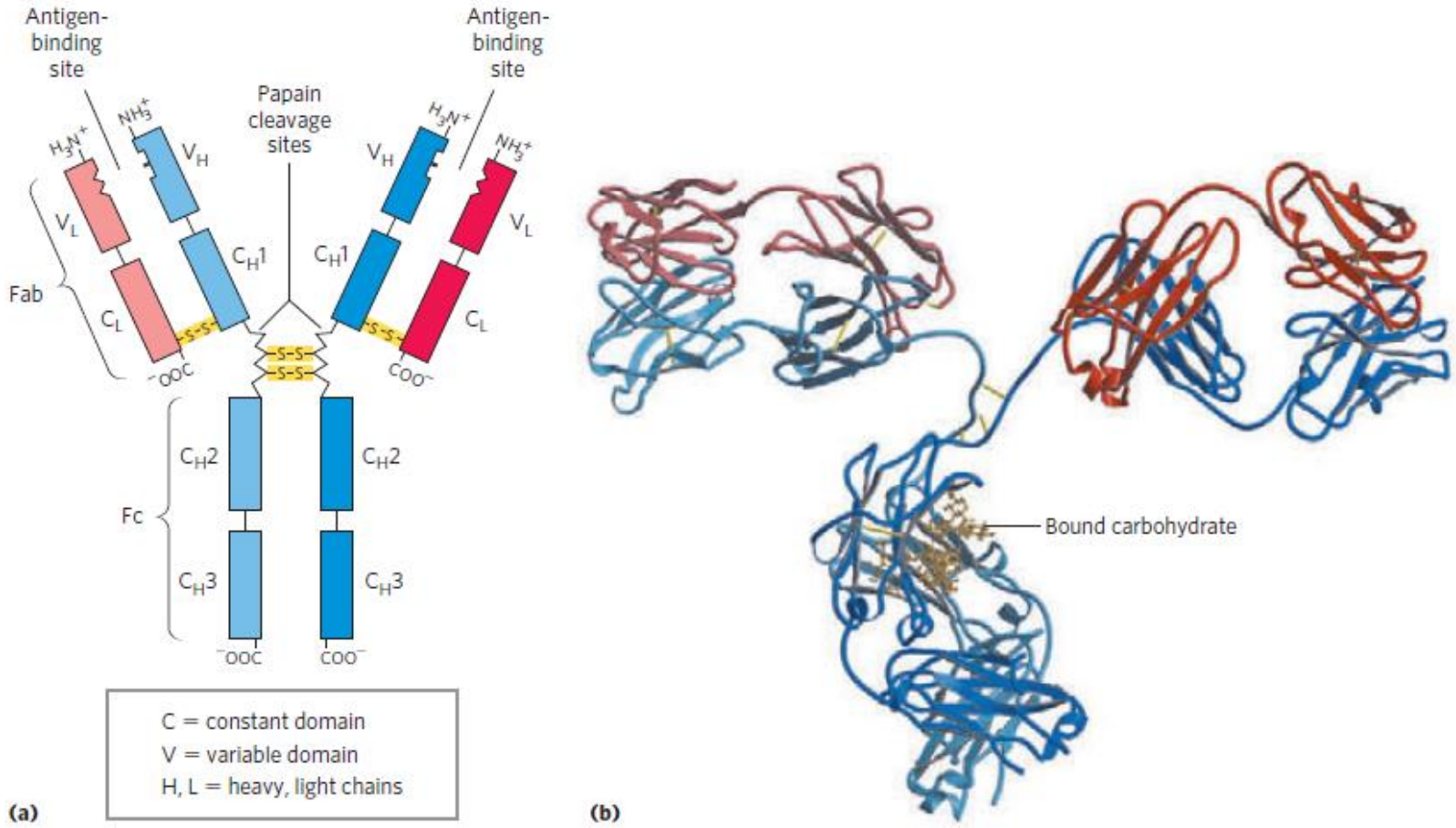
❖It would be unproductive for the immune system to respond to small molecules that are common intermediates and products of cellular metabolism.

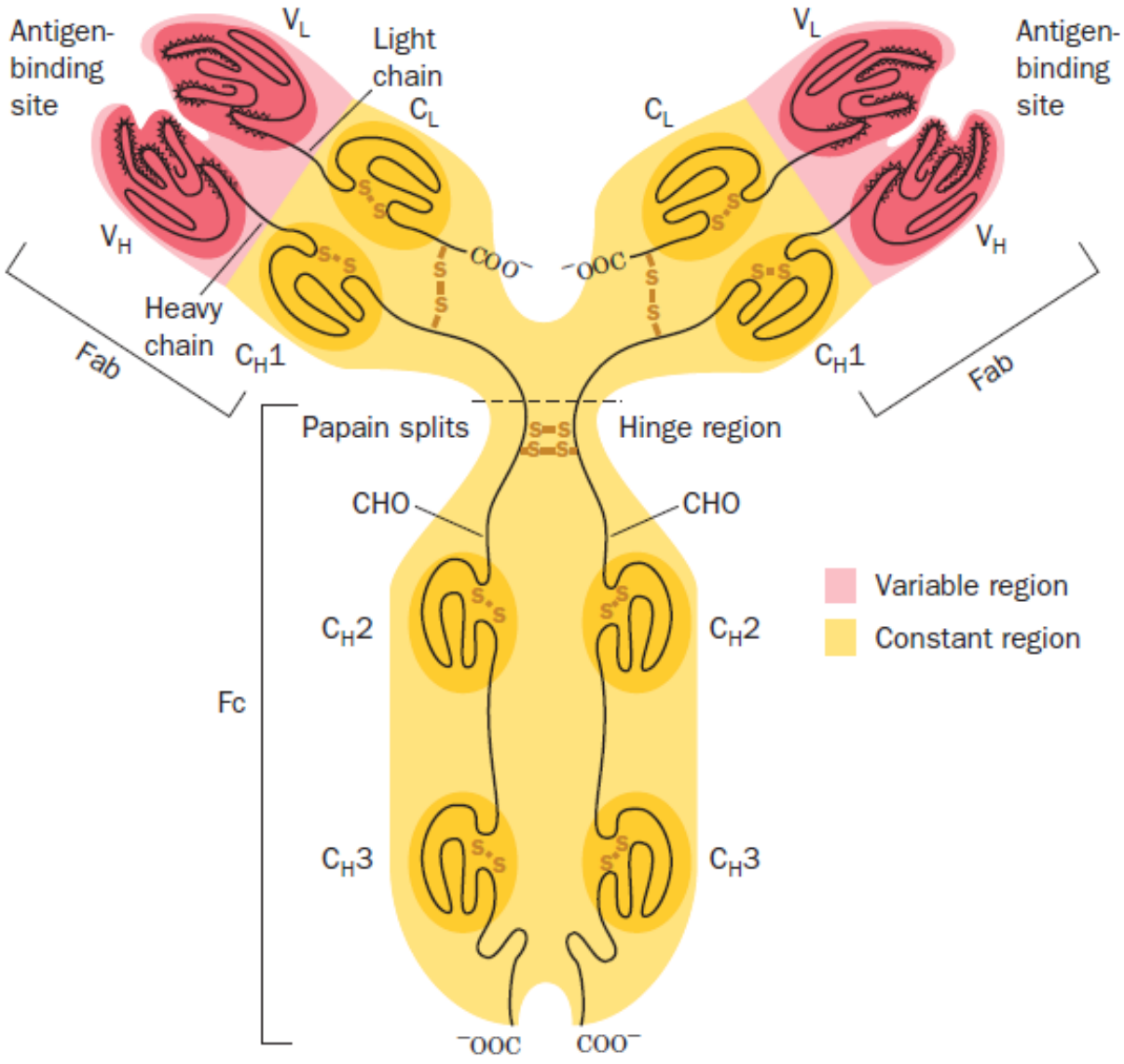
❖Molecules of Mr ,5,000 are generally not antigenic. However, when small molecules are covalently attached to large proteins in the laboratory, they can be used to elicit an immune response.

❖These small molecules are called **haptens**. The antibodies produced in response to protein-linked haptens will then bind to the same small molecules in their free form.

❖**Immunoglobulin G (IgG)** is the major class of antibody molecule and one of the most abundant proteins in the blood serum.

❖IgG has four polypeptide chains: two large ones, called heavy chains, and two light chains, linked by noncovalent and disulfide bonds into a complex of Mr 150,000.




Antigen-binding site
V_L
Light chain
C_L
V_H
Heavy chain
C_H1
Fab
Papain splits
Hinge region
CHO
CHO
C_H2
C_H2
Fc
C_H3
C_H3
^-OOC
COO^-
Variable region
Constant region

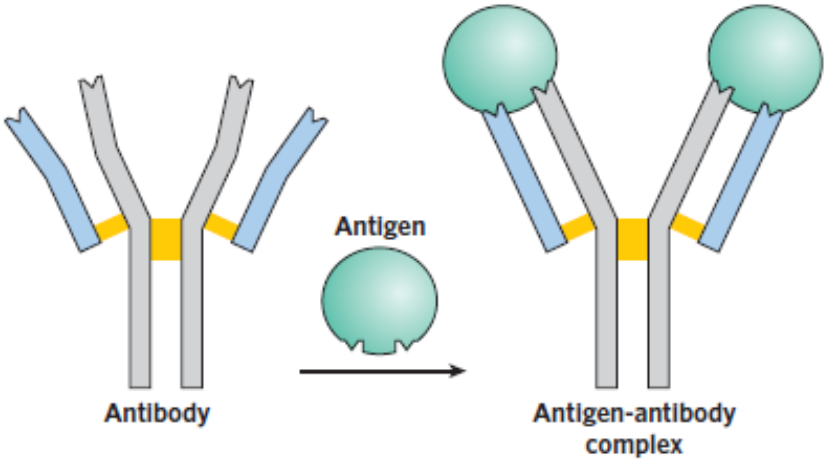
❖In many vertebrates, IgG is but one of five classes of immunoglobulins.

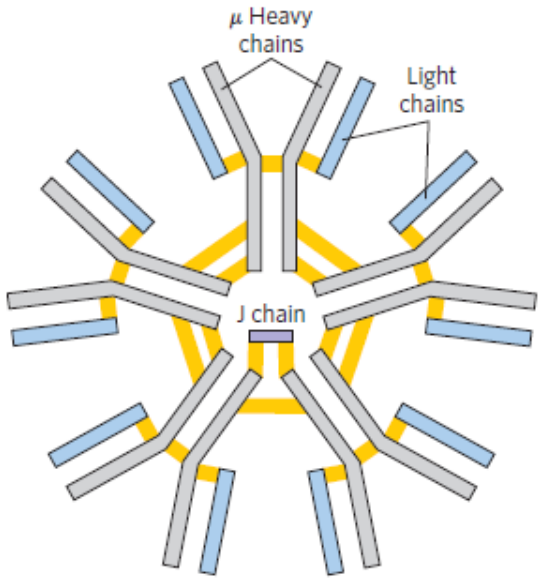
❖Each class, IgA, IgD, IgE, IgG, and IgM, has a characteristic type of heavy chain.

❖The overall structures of IgD and IgE are similar to that of IgG.

❖IgM occurs either in a monomeric, membrane-bound form or in a secreted form that is a cross-linked pentamer of this basic structure.

❖IgA, found principally in secretions such as saliva, tears, and milk, can be a monomer, dimer, or trimer.





❖The extraordinary binding affinity and specificity of antibodies make them valuable analytical reagents.

❖Two types of antibody preparations are in use: polyclonal and monoclonal.

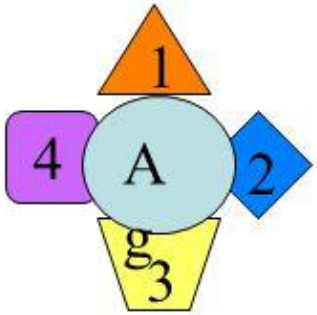
❖Polyclonal antibodies are those produced by many different B lymphocytes responding to one antigen, such as a protein injected into an animal.

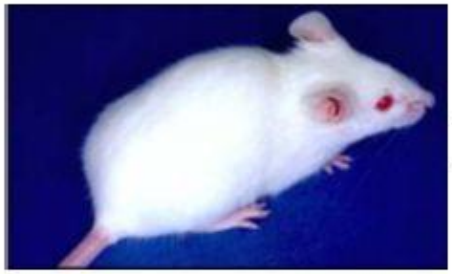
❖Polyclonal preparations contain a mixture of antibodies that recognize different parts of the protein.

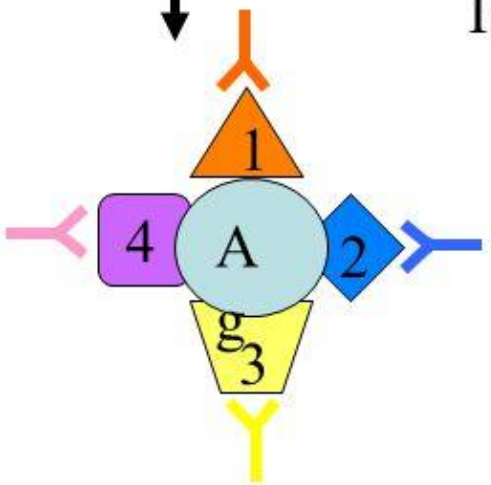
❖Monoclonal antibodies, in contrast, are synthesized by a population of identical B cells (a clone) grown in cell culture.

❖These antibodies are homogeneous, all recognizing the same epitope.

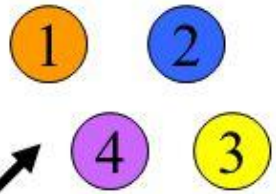
Polyclonal vs Monoclonal Antibodies



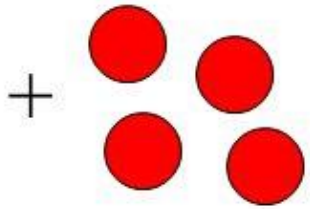




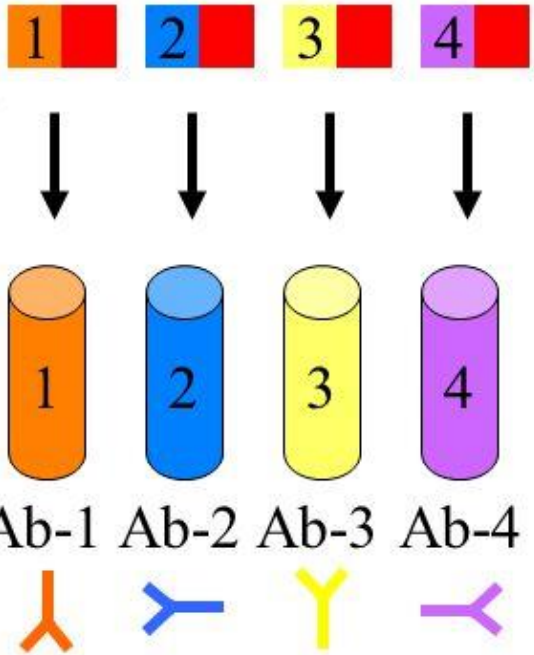
Isolate B-cells
Spleen



Myeloma
cells

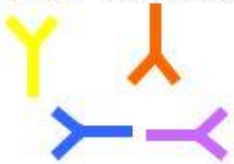


Hybridoma
Cells



Monoclonal
antibodies

Isolate serum

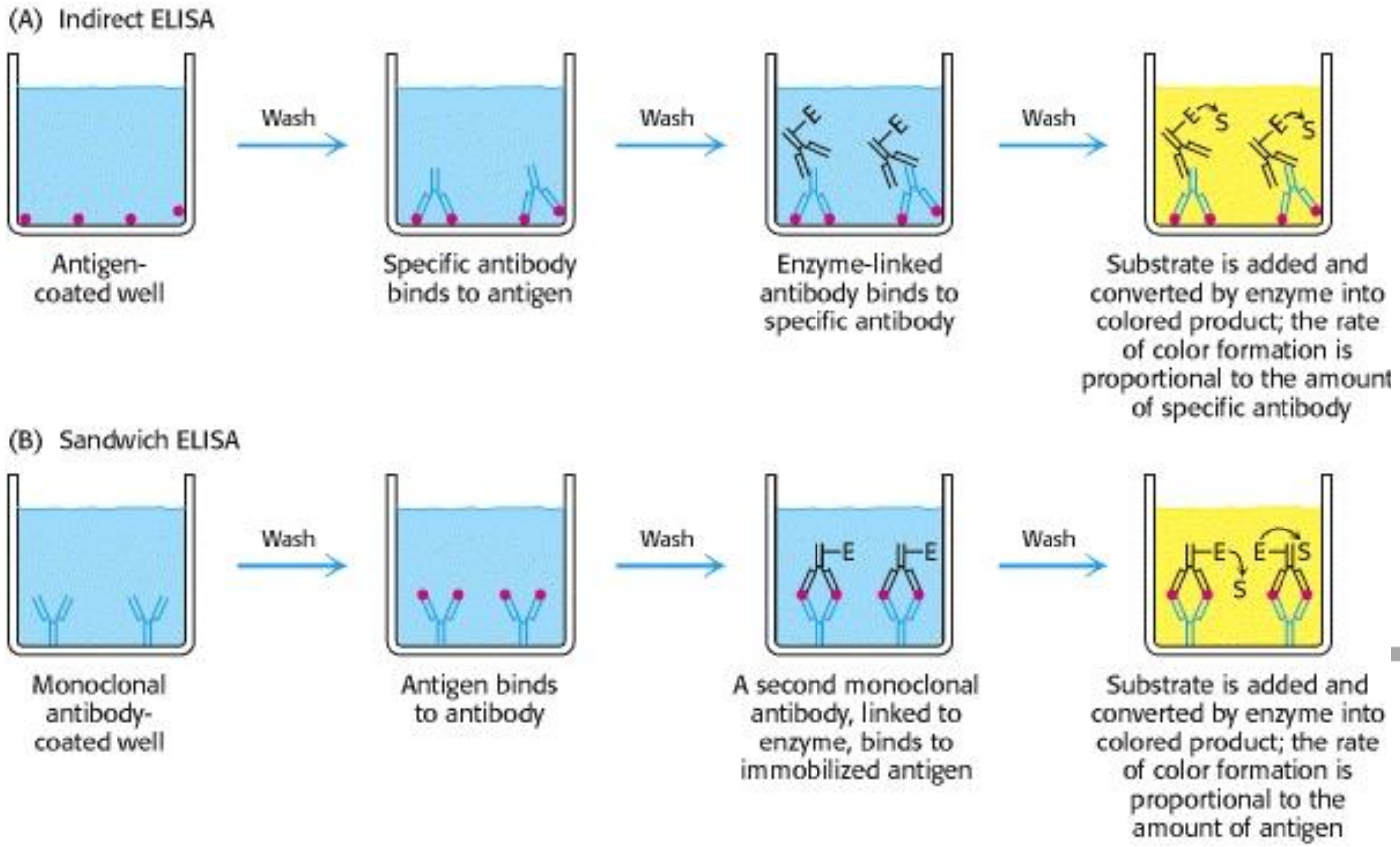


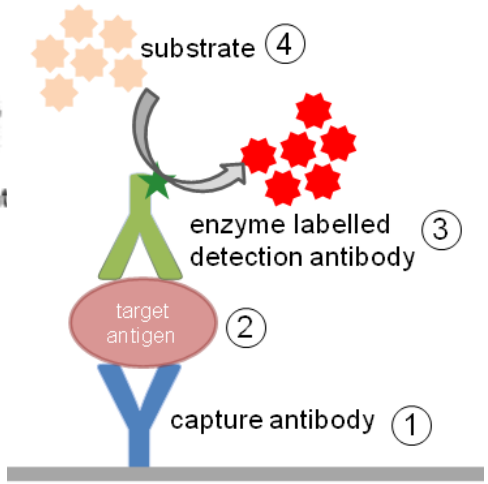
Polyclonal
antibodies

Ab-1, Ab-2
Ab-3, Ab-4

❖An ELISA (enzyme-linked immunosorbent assay) can be used to rapidly screen for and quantify an antigen in a sample

❖ELISA assay has many clinical and scientific applications.





Actin, Myosin, and Molecular Motors

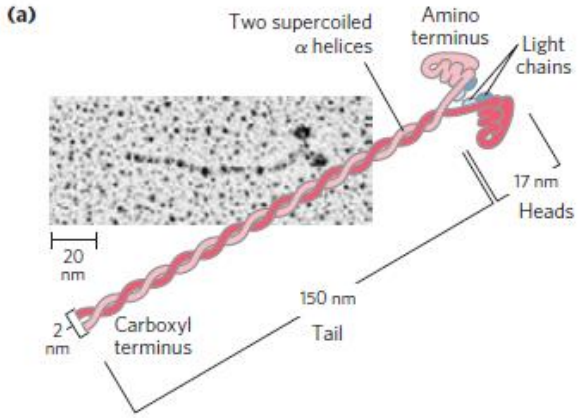
❖Organisms move, cells move, organelles and macromolecules within cells move.

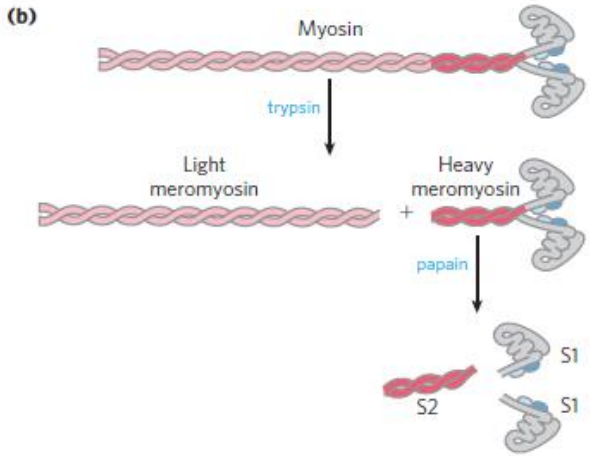
❖Most of these movements arise from the activity of a fascinating class of protein-based molecular motors.

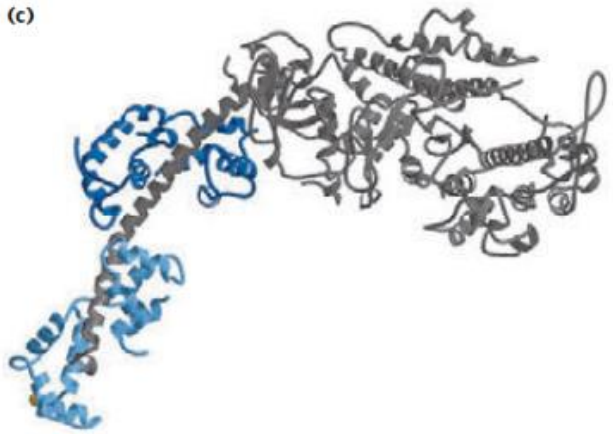
❖Fueled by chemical energy, usually derived from ATP, large aggregates of motor proteins undergo cyclic conformational changes that accumulate into a unified, directional force.

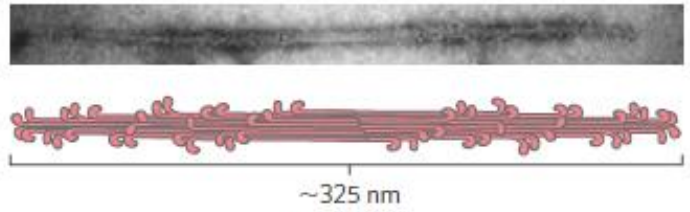
❖The interactions among motor proteins, as you might predict, feature complementary arrangements of ionic, hydrogen-bonding, hydrophobic, and van der Waals interactions at protein binding sites.

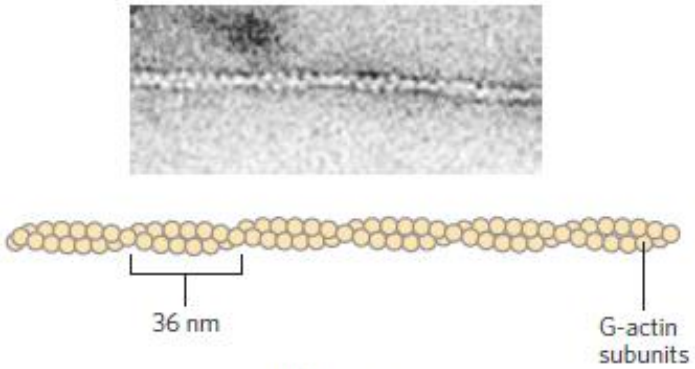
❖The contractile force of muscle is generated by the interaction of two proteins, **myosin** and **actin**. These proteins are arranged in filaments that undergo transient interactions and slide past each other to bring about contraction.

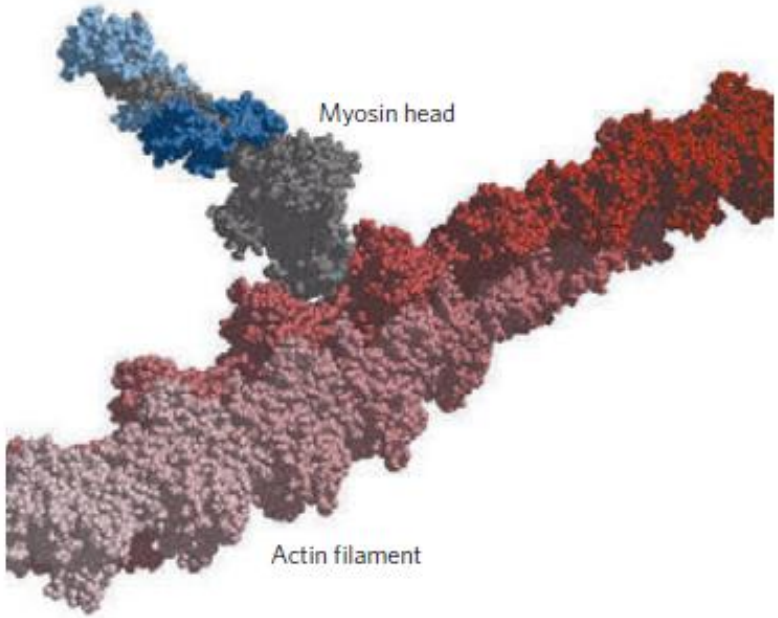

(a)
Two supercoiled α helices
Amino terminus
Light chains
17 nm
Heads
20 nm
2 nm
Carboxyl terminus
150 nm
Tail

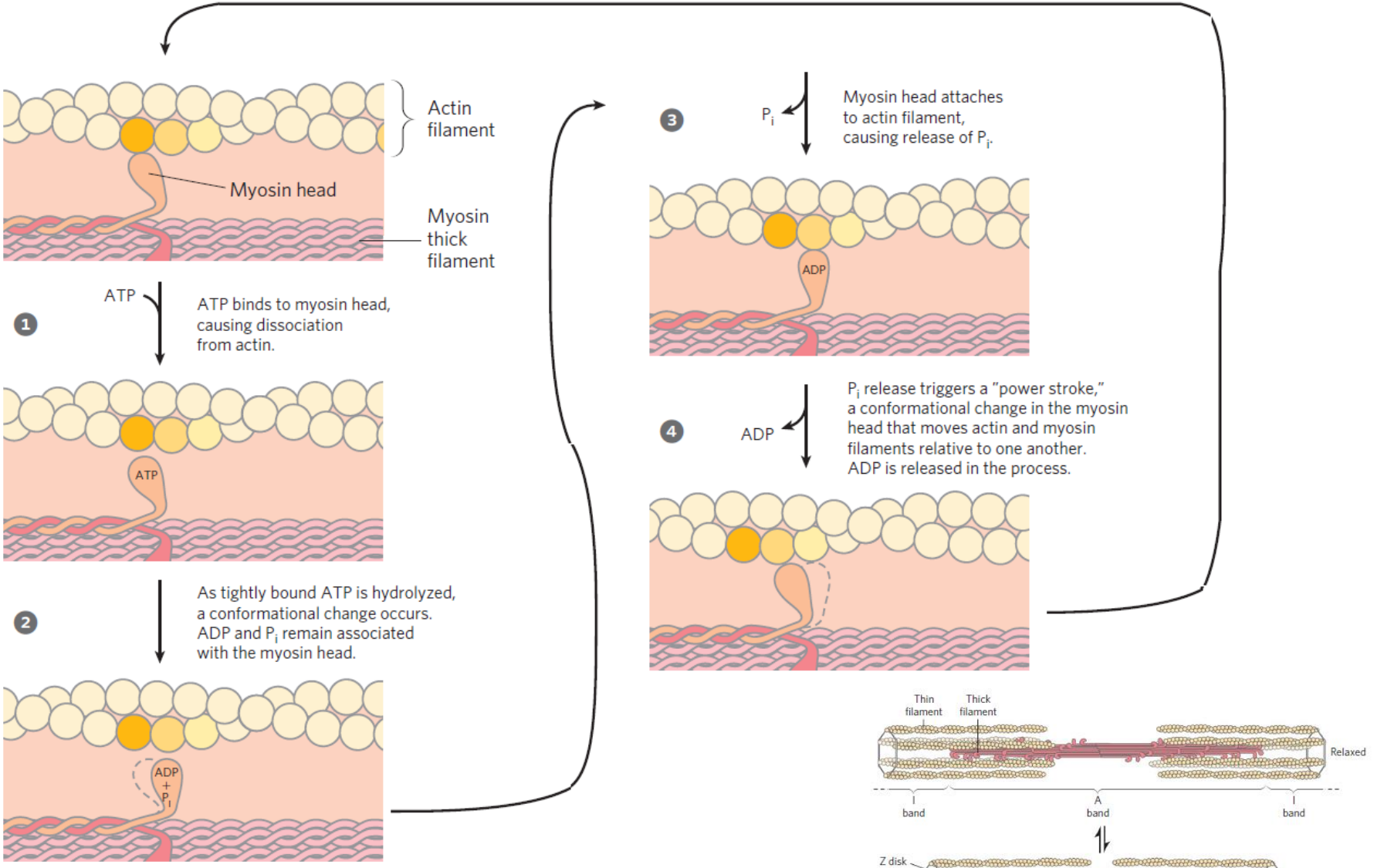

(b)
Myosin
trypsin
Light meromyosin
Heavy meromyosin
+
papain
S1
S2
S1


(c)

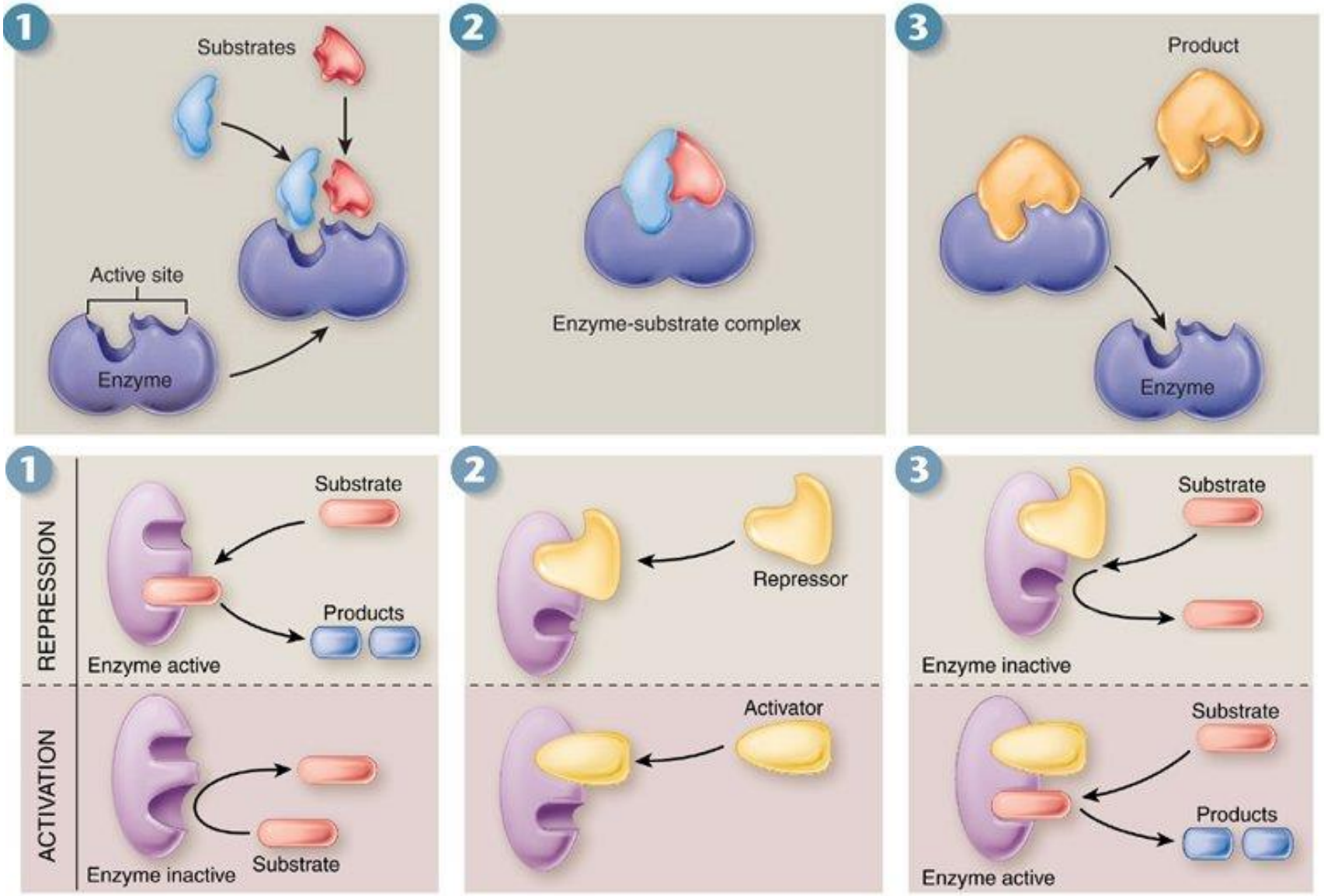

~325 nm
(a) Myosin


36 nm
G-actin subunits
(b) F-actin


Myosin head
Actin filament
(c)


Actin filament
Myosin head
Myosin thick filament
ATP
1
ATP binds to myosin head, causing dissociation from actin.
ATP
2
As tightly bound ATP is hydrolyzed, a conformational change occurs. ADP and P_i remain associated with the myosin head.
ADP + P_i
3
P_i
Myosin head attaches to actin filament, causing release of P_i.
ADP
4
ADP
P_i release triggers a "power stroke," a conformational change in the myosin head that moves actin and myosin filaments relative to one another. ADP is released in the process.
Thin filament
Thick filament
Relaxed
I band
A band
I band
Z disk
Contracted
(a)
(b)

AMINO ACIDS PEPTIDES AND PROTEINS IV: ENZYMES



❖Living systems make use of energy from the environment.

❖Many chemical reactions that occurs in living organisms is thermodynamically favorable, but they are very slow.

❖Therefore in the course of evolution enzymes have been selected to catalyze these reactions.

❖Enzymes are catalysts of biological systems and except a small group of ribozymes they are consist of proteins.

❖Enzymes have extraordinary catalytic power, often far greater than that of synthetic or inorganic catalysts.

❖They have a high degree of specificity for their substrates, they accelerate chemical reactions tremendously, and they function in aqueous solutions under very mild conditions of temperature and pH.

❖Catalytic activity of the enzymes depends on the integrity of their native protein conformation.

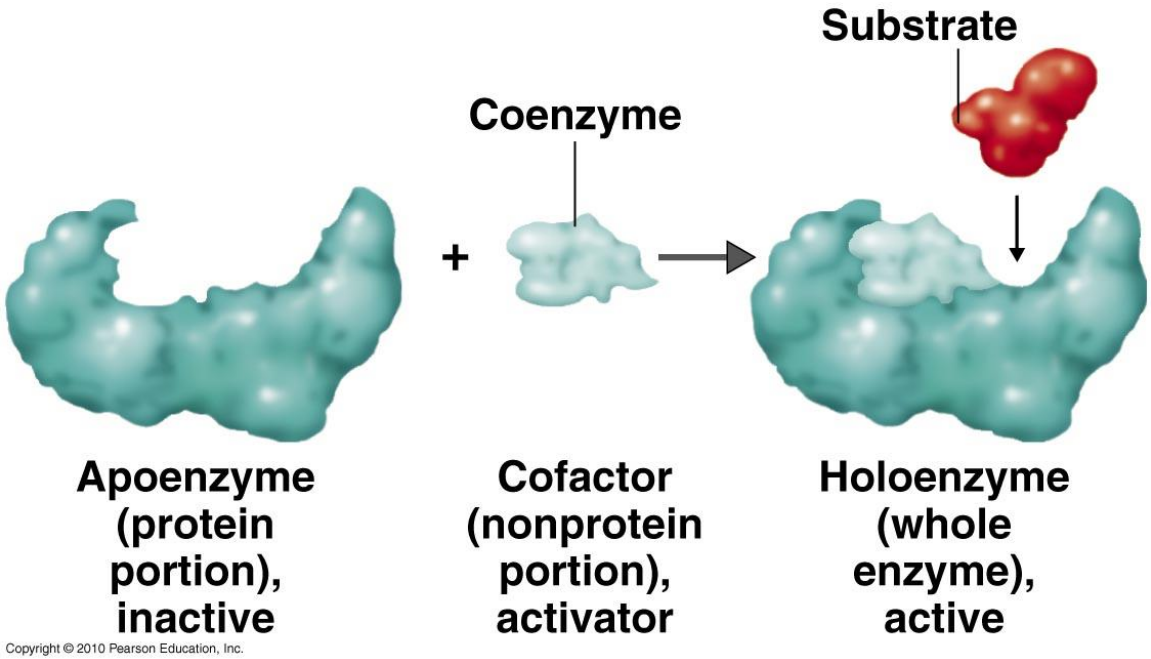
❖If an enzyme is denatured or dissociated into its subunits, catalytic activity is usually lost.

❖Thus the primary, secondary, tertiary, and quaternary structures of protein enzymes are essential to their catalytic activity.

❖Some enzymes require no chemical groups for activity other than their amino acid residues; such as trypsin.

❖Others require an additional chemical component called a **cofactor**—either one or more inorganic ions, such as Fe^{2+}, Mg^{2+}, Mn^{2+}, or Zn^{2+}, or a complex organic or metalloorganic molecule called a **coenzyme**.

❖ A coenzyme or metal ion that is very tightly or even covalently bound to the enzyme protein is called a **prosthetic group**.



❖A complete, catalytically active enzyme together with its bound coenzyme and/or metal ions is called a **holoenzyme.**

❖The protein part of such an enzyme is called the **apoenzyme** or **apoprotein**.

❖Finally, some enzyme proteins are modified covalently by phosphorylation, glycosylation, and other processes.

❖Many enzymes have been named by adding the suffix **"-ase"** to the name of their substrate or to a word or phrase describing their activity.

❖Thus urease catalyzes hydrolysis of urea, and DNA polymerase catalyzes the polymerization of nucleotides to form DNA.

❖Other enzymes were named by their discoverers for a broad function, before the specific reaction catalyzed was known. For example, lysozyme was named for its ability to lyse (break down) bacterial cell walls.

❖Sometimes the same enzyme has two or more names, or two different enzymes have the same name.

❖Because of such ambiguities, and the ever-increasing number of newly discovered enzymes, biochemists, by international agreement, have adopted a system for naming and classifying enzymes.

❖This system divides enzymes into six classes, each with subclasses, based on the type of reaction catalyzed

$$\text{ATP} + \text{D-glucose} \longrightarrow \text{ADP} + \text{D-glucose 6-phosphate}$$

❖The formal systematic name of the enzyme catalyzing the reaction is ATP:glucose phosphotransferase, which indicates that it catalyzes the transfer of a phosphoryl group from ATP to glucose.

❖Its Enzyme Commission (E.C.) number is 2.7.1.1. The first number (2) denotes the class name (transferase); the second number (7), the subclass (phosphotransferase); the third number (1), hydroxyl group as acceptor; and the fourth number (1), shows discovery number.

❖For many enzymes, a trivial name is more frequently used—in this case, hexokinase.

TABLE 6–3 International Classification of Enzymes

Class no.	Class name	Type of reaction catalyzed
1	Oxidoreductases	Transfer of electrons (hydride ions or H atoms)
2	Transferases	Group transfer reactions
3	Hydrolases	Hydrolysis reactions (transfer of functional groups to water)
4	Lyases	Cleavage of C—C, C—O, C—N, or other bonds by elimination, leaving double bonds or rings, or addition of groups to double bonds
5	Isomerases	Transfer of groups within molecules to yield isomeric forms
6	Ligases	Formation of C—C, C—S, C—O, and C—N bonds by condensation reactions coupled to cleavage of ATP or similar cofactor

EC 2 Transferases

Number	Name
EC 2.1	Transferring one-carbon groups
EC 2.1.1	Methyltransferases
EC 2.1.2	Hydroxymethyl-, Formyl- and Related Transferases
EC 2.1.3	Carboxy- and Carbamoyltransferases
EC 2.1.4	Amidinotransferases
EC 2.2	Transferring aldehyde or ketonic groups
EC 2.2.1	Transketolases and Transaldolases
EC 2.3	Acyltransferases
EC 2.3.1	Transferring groups other than amino-acyl groups
EC 2.3.2	Aminoacyltransferases
EC 2.3.3	Acyl groups converted into alkyl on transfer
EC 2.4	Glycosyltransferases
EC 2.4.1	Hexosyltransferases
EC 2.4.2	Pentosyltransferases
EC 2.4.99	Transferring other glycosyl groups
EC 2.5	Transferring alkyl or aryl groups, other than methyl groups
EC 2.5.1	Transferring Alkyl or Aryl Groups, Other than Methyl Groups
EC 2.6	Transferring nitrogenous groups
EC 2.6.1	Transaminases
EC 2.6.2	Amidinotransferases
EC 2.6.3	Oximinotransferases
EC 2.6.99	Transferring Other Nitrogenous Groups

EC 2.7	Transferring phosphorus-containing groups
EC 2.7.1	Phosphotransferases with an alcohol group as acceptor
EC 2.7.2	Phosphotransferases with a carboxy group as acceptor
EC 2.7.3	Phosphotransferases with a nitrogenous group as acceptor
EC 2.7.4	Phosphotransferases with a phosphate group as acceptor
EC 2.7.5	Phosphotransferases with regeneration of donors, apparently catalysing intramolecular transfers
EC 2.7.6	Diphosphotransferases
EC 2.7.7	Nucleotidyltransferases
EC 2.7.8	Transferases for other substituted phosphate groups
EC 2.7.9	Phosphotransferases with paired acceptors
EC 2.7.10	Protein-tyrosine kinases
EC 2.7.11	Protein-serine/threonine kinases
EC 2.7.12	Dual-specificity kinases (those acting on Ser/Thr and Tyr residues)
EC 2.7.13	Protein-histidine kinases
EC 2.7.99	Other protein kinases
EC 2.8	Transferring sulfur-containing groups
EC 2.8.1	Sulfurtransferases
EC 2.8.2	Sulfotransferases
EC 2.8.3	CoA-transferases
EC 2.8.4	Transferring alkylthio groups
EC 2.10	Transferring molybdenum- or tungsten-containing groups
EC 2.10.1	Molybdenumtransferases or tungstentransferases with sulfide groups as acceptors

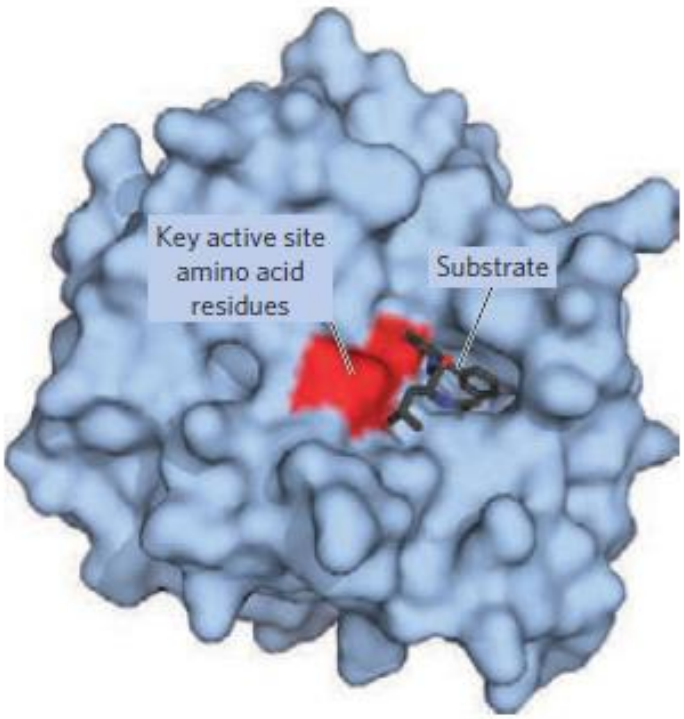
EC 2.7.1.1 hexokinase
EC 2.7.1.2 glucokinase
EC 2.7.1.3 ketohexokinase
EC 2.7.1.4 fructokinase
EC 2.7.1.5 rhamnulokinase
EC 2.7.1.6 galactokinase
EC 2.7.1.7 mannokinase
EC 2.7.1.8 glucosamine kinase
EC 2.7.1.9 deleted
EC 2.7.1.10 phosphoglucokinase
EC 2.7.1.11 6-phosphofructokinase
EC 2.7.1.12 gluconokinase
EC 2.7.1.13 dehydrogluconokinase
EC 2.7.1.14 sedoheptulokinase
EC 2.7.1.15 ribokinase

❖The distinguishing feature of an enzyme-catalyzed reaction is that it takes place within the confines of a pocket on the enzyme called the **active site**.

❖The molecule that is bound in the active site and acted upon by the enzyme is called the **substrate**.

❖The surface of the active site is lined with amino acid residues with substituent groups that bind the substrate and catalyze its chemical transformation.

❖The enzyme-substrate complex, is central to the action of enzymes.



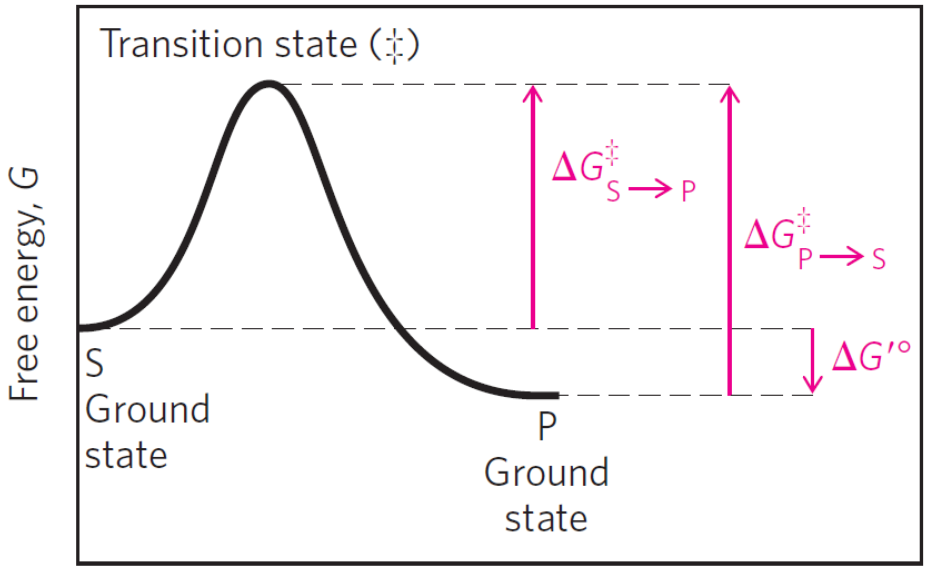
❖A simple enzymatic reaction might be written as

$$E + S \rightleftharpoons ES \rightleftharpoons EP \rightleftharpoons E + P$$

where E, S, and P represent the enzyme, substrate, and product; ES and EP are transient complexes of the enzyme with the substrate and with the product.

❖The function of a catalyst is to increase the rate of a reaction. Catalysts do not affect reaction equilibria.

❖Any reaction, such as S→P, can be described by a reaction coordinate diagram.



❖There is an energy barrier between S and P: the energy required for alignment of reacting groups, formation of transient unstable charges, bond rearrangements, and other transformations required for the reaction to proceed in either direction.

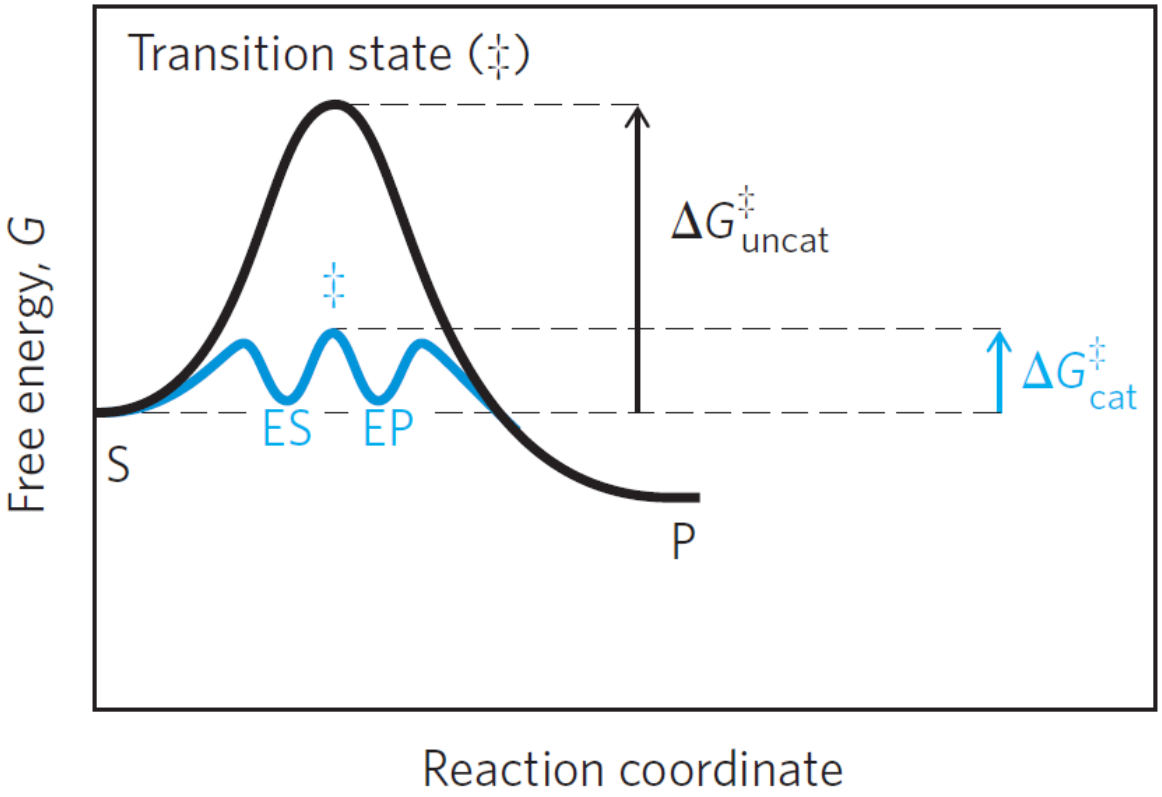
❖To undergo reaction, the molecules must overcome this barrier and therefore must be raised to a higher energy level.

❖At the top of the energy hill is a point at which decay to the S or P state is equally probable (it is downhill either way).

❖This is called the **transition state**.

❖The transition state is not a chemical species with any significant stability and should not be confused with a reaction intermediate (such as ES or EP).

❖The difference between the energy levels of the ground state and the transition state is the **activation energy**, $\Delta G^{\ddagger}$.



❖Reaction rates can be increased by raising the temperature and/or pressure, thereby increasing the number of molecules with sufficient energy to overcome the energy barrier.

❖Alternatively, the activation energy can be lowered by adding a catalyst .

❖Catalysts enhance reaction rates by lowering activation energies.

- The rate of any reaction is determined by the concentration of the reactant (or reactants) and by a rate constant, usually denoted by *k*.

- For the unimolecular reaction S→P, the rate (or velocity) of the reaction, *V*—representing the amount of S that reacts per unit time—is expressed by a rate equati $V = k[\mathrm{S}]$

- In this reaction, the rate depends only on the concentration of S. This is called a first-order reaction.

- If a reaction rate depends on the concentration of two different compounds, the reaction is second order and k is a second order rate constant:

$$V = k[\mathrm{S}_1][\mathrm{S}_2]$$

- From transition-state theory we can derive an expression that relates the magnitude of a rate constant to the activation energy, where k is the Boltzmann constant and h is Planck's constant.

$$k = \frac{\mathbf{k}T}{h} e^{-\Delta G^{\ddagger}/RT}$$

- Much of the catalytic power of enzymes is ultimately derived from the free energy released in forming many weak bonds and interactions between an enzyme and its substrate.

- This binding energy contributes to specificity as well as to catalysis.

- Weak interactions are optimized in the reaction transition state.

- Enzyme active sites are complementary not to the substrates per se but to the transition states through which substrates pass as they are converted to products during an enzymatic reaction.

- How does an enzyme use binding energy to lower the activation energy for a reaction?

- Emil Fischer propose, in 1894, that enzymes were structurally complementary to their substrates, so that they fit together like a lock and key.

- However, the “lock and key” hypothesis can be misleading when applied to enzymatic catalysis.

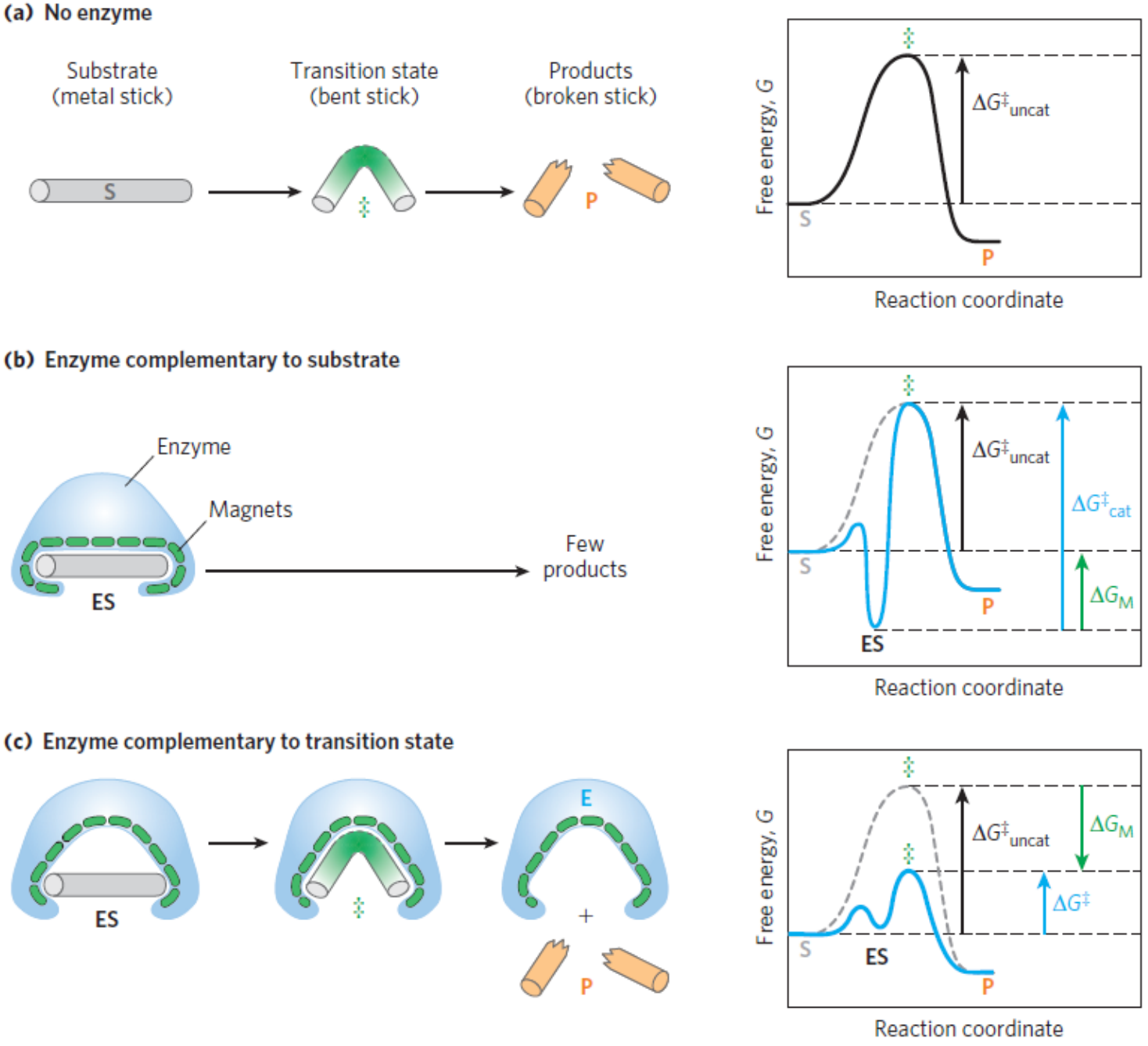
- An enzyme completely complementary to its substrate would be a very poor enzyme,

- The modern notion of enzymatic catalysis is; in order to catalyze reactions, an enzyme must be complementary to the reaction transition state.

- This means that optimal interactions between substrate and enzyme occur only in the transition state.

- Real enzymes work on an analogous principle.

- Some weak interactions are formed in the ES complex, but the full complement of such interactions between substrate and enzyme is formed only when the substrate reaches the transition state.


(a) No enzyme
Substrate
(metal stick)
Transition state
(bent stick)
Products
(broken stick)
S
P
Free energy, G
$\Delta G^{\ddagger}_{uncat}$
Reaction coordinate
(b) Enzyme complementary to substrate
Enzyme
Magnets
ES
Few
products
$\Delta G^{\ddagger}_{cat}$
ΔG_M
(c) Enzyme complementary to transition state
E
+

- Prominent physical and thermodynamic factors contributing to $\Delta G^{\ddagger}$, the barrier to reaction, might include:

- (1) the entropy (freedom of motion) of molecules in solution, which reduces the possibility that they will react together,

- (2) the solvation shell of hydrogenbonded water that surrounds and helps to stabilize most biomolecules in aqueous solution,

- (3) the distortion of substrates that must occur in many reactions, and

- (4) the need for proper alignment of catalytic functional groups on the enzyme.

- Binding energy can be used to overcome all these barriers.

- First, a large restriction in the relative motions of two substrates that are to react, or entropy reduction, is one obvious benefit of binding them to an enzyme.

- Binding energy holds the substrates in the proper orientation to react—a substantial contribution to catalysis, because productive collisions between molecules in solution can be exceedingly rare.

- Second, formation of weak bonds between substrate and enzyme results in desolvation of the substrate. Enzyme-substrate interactions replace most or all of the hydrogen bonds between the substrate and water that would otherwise impede reaction.

- Third, binding energy involving weak interactions formed only in the reaction transition state helps to compensate thermodynamically for any distortion, primarily electron redistribution, that the substrate must undergo to react.

- Finally, the enzyme itself usually undergoes a change in conformation when the substrate binds, induced by multiple weak interactions with the substrate.

- This is referred to as **induced fit**.

- Induced fit serves to bring specific functional groups on the enzyme into the proper position to catalyze the reaction.

- The conformational change also permits formation of additional weak bonding interactions in the transition state.

- In either case, the new enzyme conformation has enhanced catalytic properties.

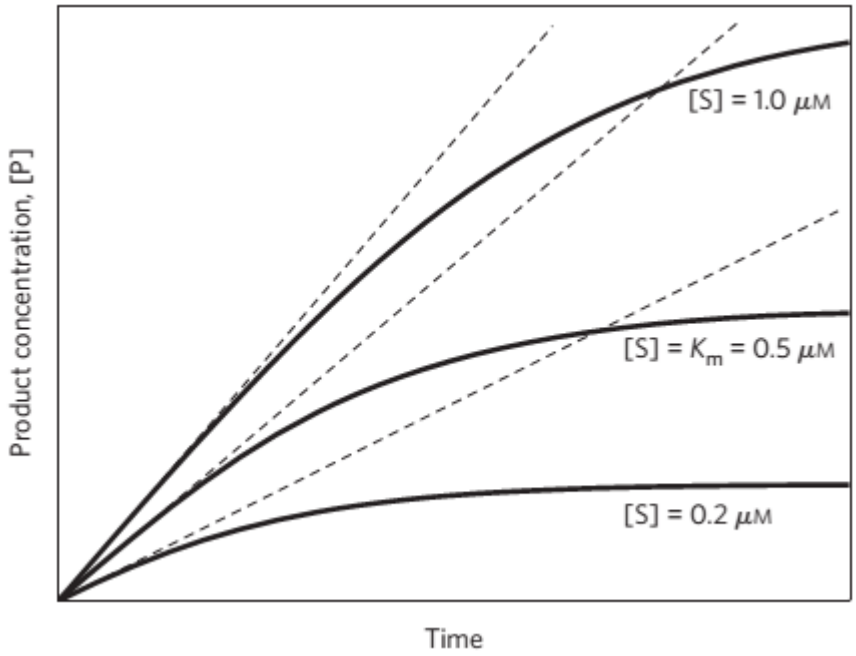
- Biochemists commonly use several approaches to study the mechanism of action of purified enzymes.

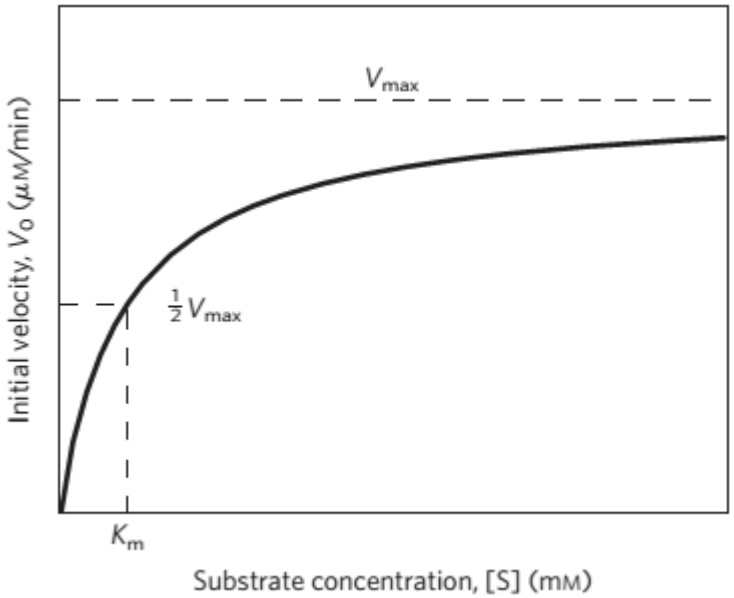
- However, the oldest approach to understanding enzyme mechanisms, and the one that remains most important, is to determine the rate of a reaction and how it changes in response to changes in experimental parameters, a discipline known as **enzyme kinetics**.

- A key factor affecting the rate of a reaction catalyzed by an enzyme is the concentration of substrate, [S].

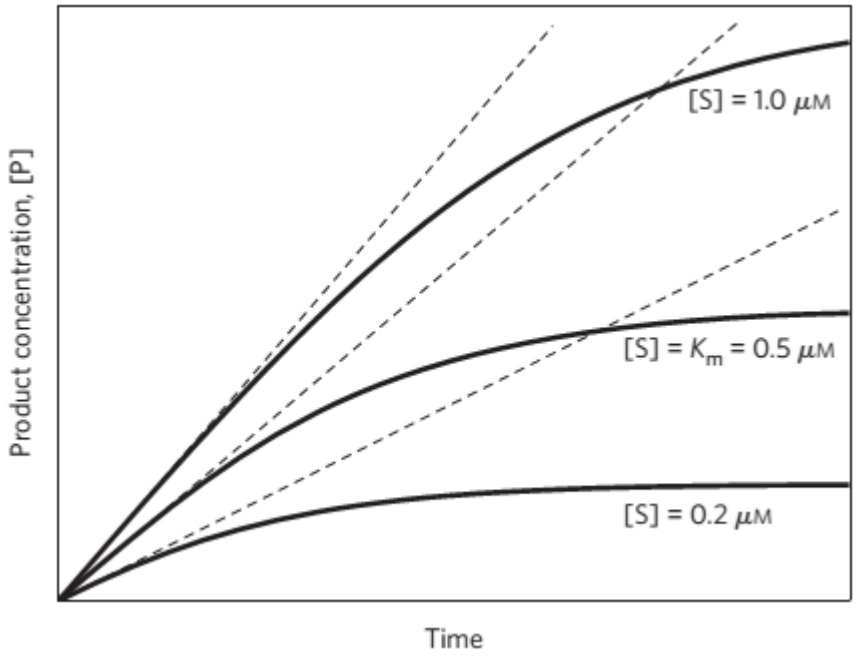
- However, studying the effects of substrate concentration is complicated by the fact that [S] changes during the course of an *in vitro* reaction

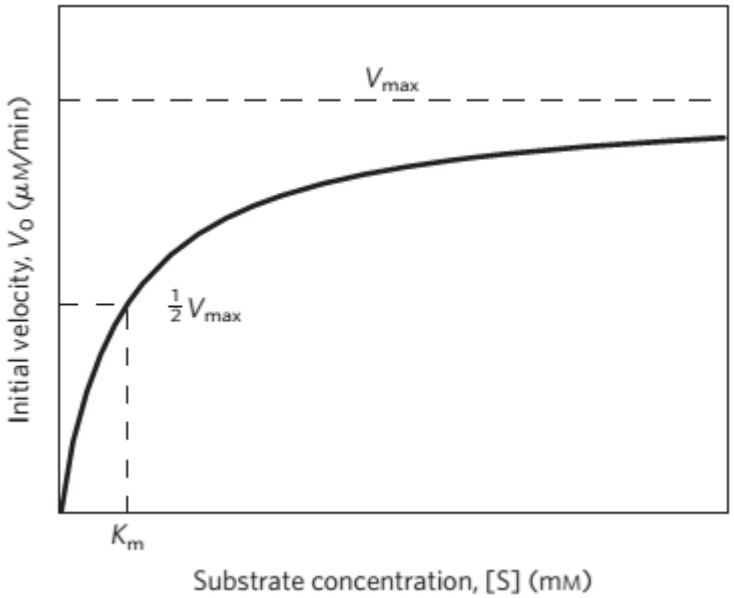
- One simplifying approach in kinetics experiments is to measure the **initial rate (or initial velocity)**, designated $\mathbf{V_0}$.





- At relatively low concentrations of substrate, V_0 increases almost linearly with an increase in [S].
- At higher substrate concentrations, V_0 increases by smaller and smaller amounts in response to increases in [S].
- Finally, a point is reached beyond which increases in V_0 are vanishingly small as [S] increases.
- This plateau-like V_0 region is close to the maximum velocity, V_{max}.





- In the enzymatic reaction, the enzyme first combines reversibly with its substrate to form an enzvme-substrate complex in a relatively fast reversible step.

$$\mathrm{E} + \mathrm{S} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{ES}$$

- The ES complex then breaks down in a slower second step to yield the free enzyme and the reaction product P.

$$\mathrm{ES} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \mathrm{E} + \mathrm{P}$$

- Because the slower second reaction must limit the rate of the overall reaction, the overall rate must be proportional to ES.

- When the enzyme is first mixed with a large excess of substrate, there is an initial period, the **pre–steady state**, during which the concentration of ES builds up.

- Then the reaction quickly achieves a steady state in which [ES] (and the concentrations of any other intermediates) remains approximately constant over time.

- The measured V_0 generally reflects the steady state, even though V_0 is limited to the early part of the reaction, and analysis of these initial rates is referred to as **steady-state kinetics**.

- The curve expressing the relationship between [S] and V_0 has the same general shape for most enzymes (it approaches a rectangular hyperbola), which can be expressed algebraically by the Michaelis-Menten equation.

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

$$V_0 = k_2[ES]$$

$$[E_t] = [E] + [ES]$$

total enzyme — Free enzyme — substrate-bound enzyme

$$[E] = [E_t] - [ES]$$

$$[S] >> [E_t]$$

Rate of ES formation $k_1([E_t] - [ES])[S]$

Rate of ES breakdown $= k_{-1}[ES] + k_2[ES]$

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

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

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

$$[ES] = \frac{k_1[E_t][S]}{k_1[S] + k_{-1} + k_2}$$

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

$(k_{-1} + k_2)/k_1$ **Michaelis constant, K_m.**

$$[ES] = \frac{[E_t][S]}{K_m + [S]}$$

$$V_0 = k_2[ES]$$

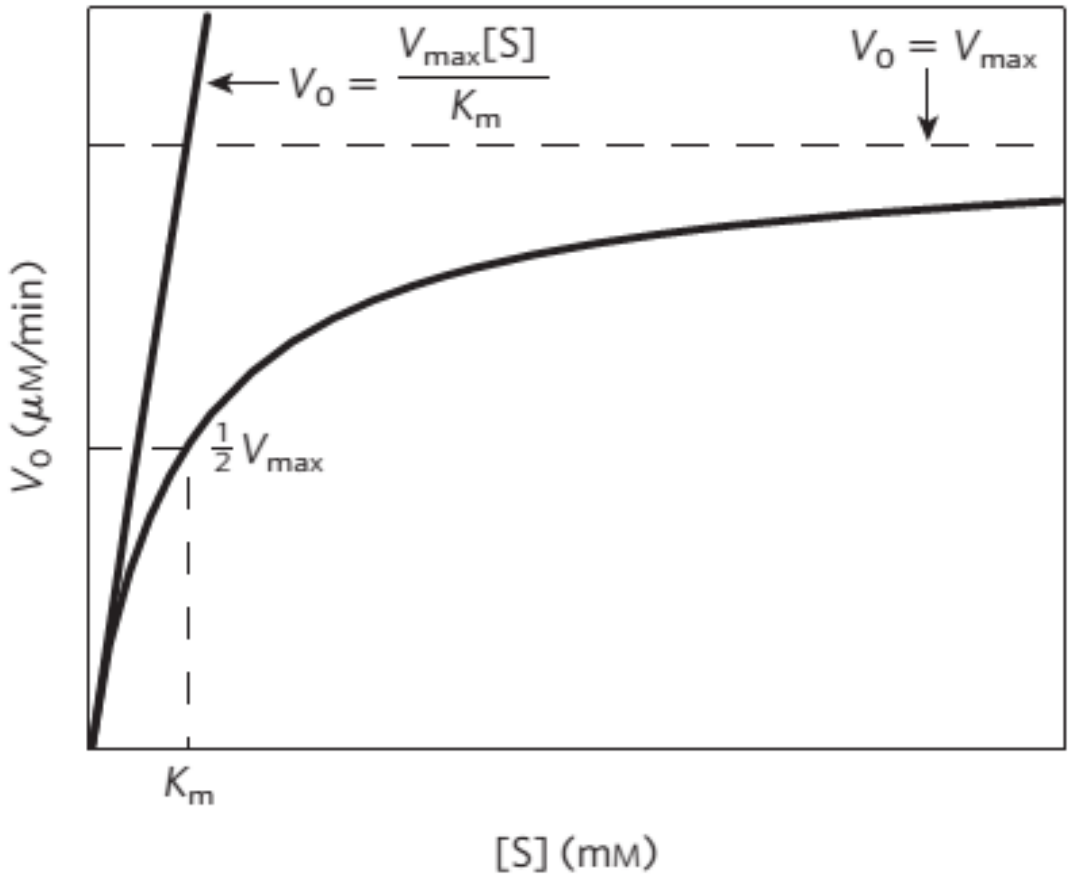
$$[ES] = \frac{[E_t][S]}{K_m + [S]} \qquad V_0 = \frac{k_2[E_t][S]}{K_m + [S]}$$

$[ES] = [E_t]$ $V_0 = V_{max}$ $V_0 = k_2[E_t]$

$$V_0 = \frac{V_{max}[S]}{K_m + [S]}$$

Michaelis-Menten equation

❖ Many enzymes follow Michealis Menten kinetics.



$$V_0 = \frac{V_{\max}[\mathrm{S}]}{K_\mathrm{m} + [\mathrm{S}]}$$

$$\frac{V_{\max}}{2} = \frac{V_{\max}[\mathrm{S}]}{K_\mathrm{m} + [\mathrm{S}]}$$

$$\frac{1}{2} = \frac{[\mathrm{S}]}{K_\mathrm{m} + [\mathrm{S}]}$$

$$K_\mathrm{m} + [\mathrm{S}] = 2[\mathrm{S}]$$

$K_\mathrm{m} = [\mathrm{S}]$, when $V_0 = ½V_{\max}$

❖The Michaelis-Menten equation can be algebraically transformed into versions that are useful in the practical determination of Km and Vmax.

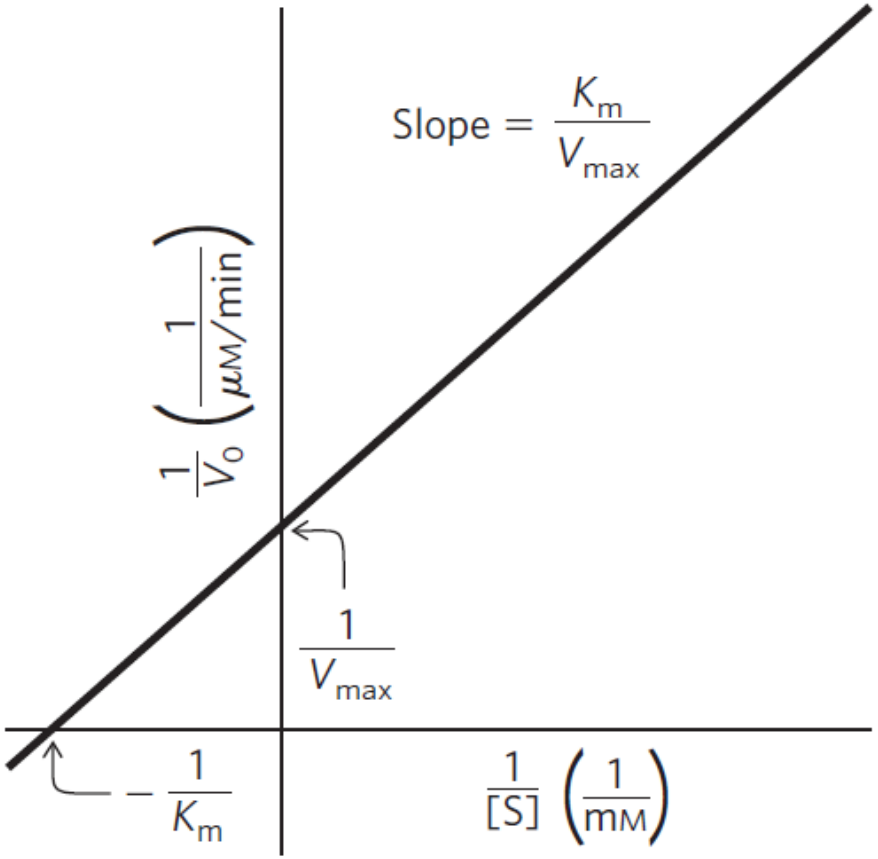
The Michaelis-Menten equation

$$V_0 = \frac{V_{\max}[S]}{K_m + [S]}$$

$$\frac{1}{V_0} = \frac{K_m + [S]}{V_{\max}[S]}$$

$$\frac{1}{V_0} = \frac{K_m}{V_{\max}[S]} + \frac{[S]}{V_{\max}[S]}$$

$$\frac{1}{V_0} = \frac{K_m}{V_{\max}[S]} + \frac{1}{V_{\max}}$$



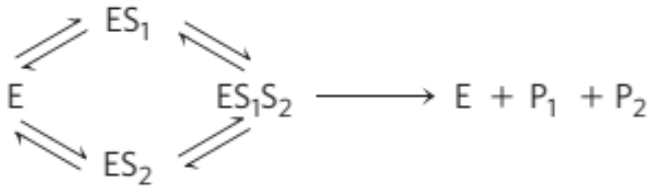
❖This form of the Michaelis-Menten equation is called the **Lineweaver-Burk equation**.

❖We have seen how [S] affects the rate of a simple enzymatic reaction (S → P) with only one substrate molecule. In most enzymatic reactions, however, two (and sometimes more) different substrate molecules bind to the enzyme and participate in the reaction.

❖Enzymatic reactions with two substrates usually involve transfer of an atom or a functional group from one substrate to the other.

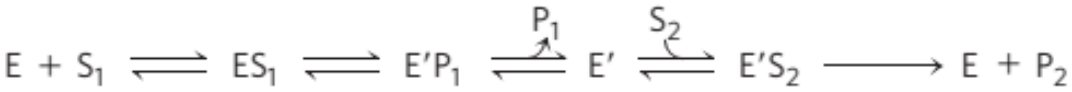
(a) Enzyme reaction involving a ternary complex

Random order



Ordered

(b) Enzyme reaction in which no ternary complex is formed



❖ **Turnover number**, **kcat** is a first-order rate constant and it is equivalent to the number of substrate molecules converted to product in a given unit of time on a single enzyme molecule when the enzyme is saturated with substrate.

❖Kinetic parameters are used to compare enzyme activities.

❖ Reaction rate can be defined from the product formation or loss of substrate.

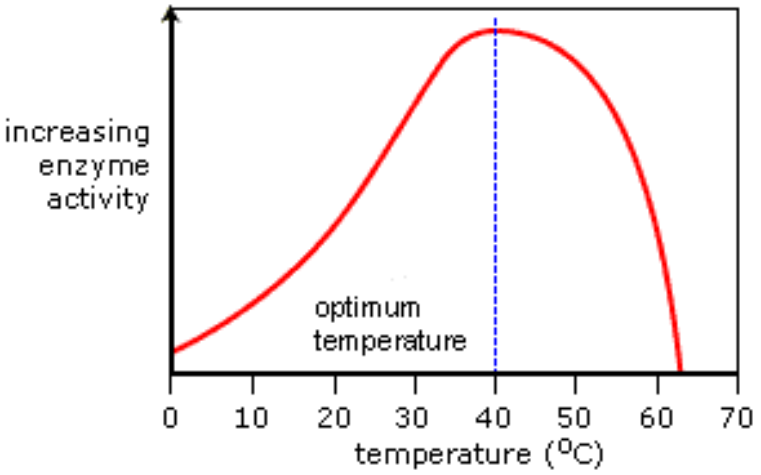
❖ Several factors affect the rate at which enzymatic reactions proceed –

- ❖ temperature,
- ❖ pH,
- ❖ enzyme concentration,
- ❖ substrate concentration,
- ❖ and the presence of any inhibitors or activators.

- Enzymes catalyse reactions by randomly colliding with substrate molecules, increasing temperature increases the rate of reaction, forming more product.

- However, as temperature increases, more bonds, especially the weaker Hydrogen and Ionic bonds, will break as a result of this strain. Eventually, the enzyme will become denatured and will no longer function.

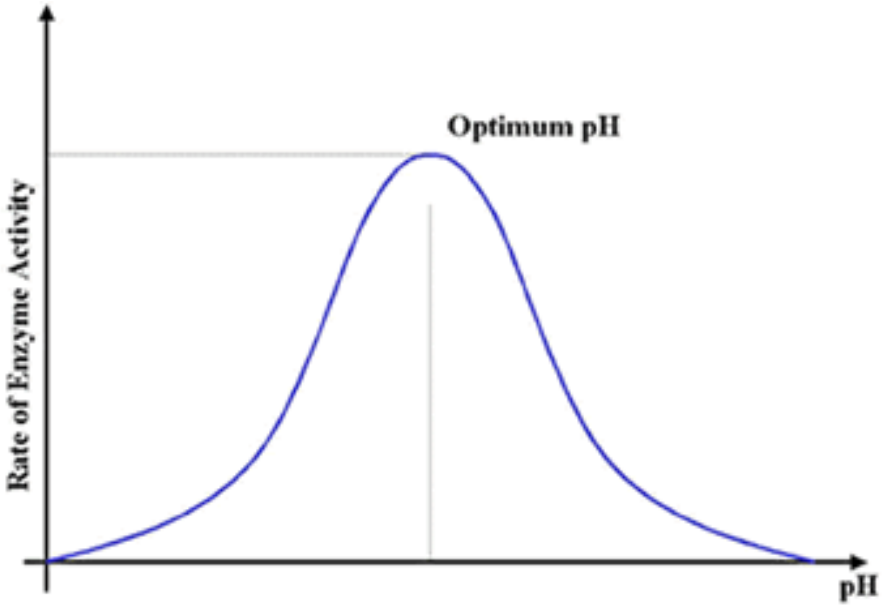
- In summary, as temperature increases, initially the rate of reaction will increase, because of increased Kinetic Energy. However, the effect of bond breaking will become greater and greater, and the rate of reaction will begin to decrease.



❖H^+ and OH^- Ions are charged and therefore interfere with Hydrogen and Ionic bonds that hold together an enzyme, since they will be attracted or repelled by the charges created by the bonds.

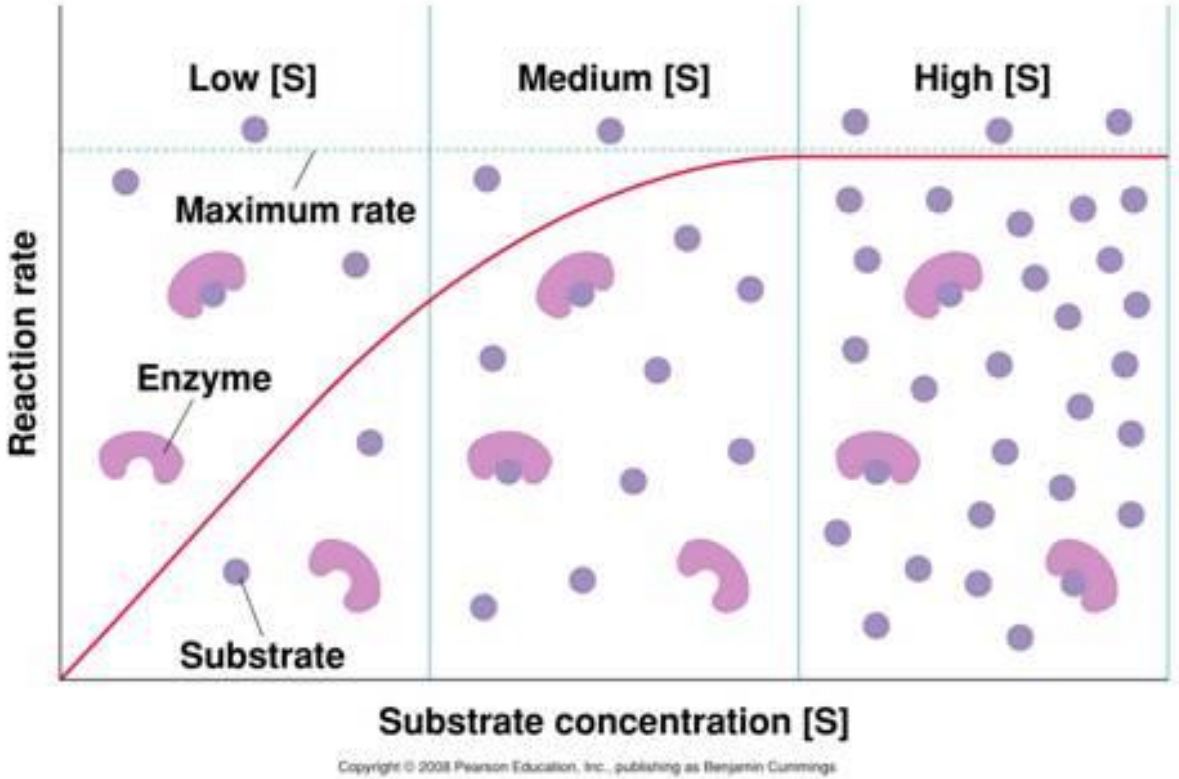
❖Different enzymes have different Optimum pH values. This is the pH value at which the bonds within them are influenced by H^+ and OH^- Ions in such a way that the shape of their active site is the most complementary to the shape of their substrate.

❖ Any change in pH above or below the optimum will quickly cause a decrease in the rate of reaction.



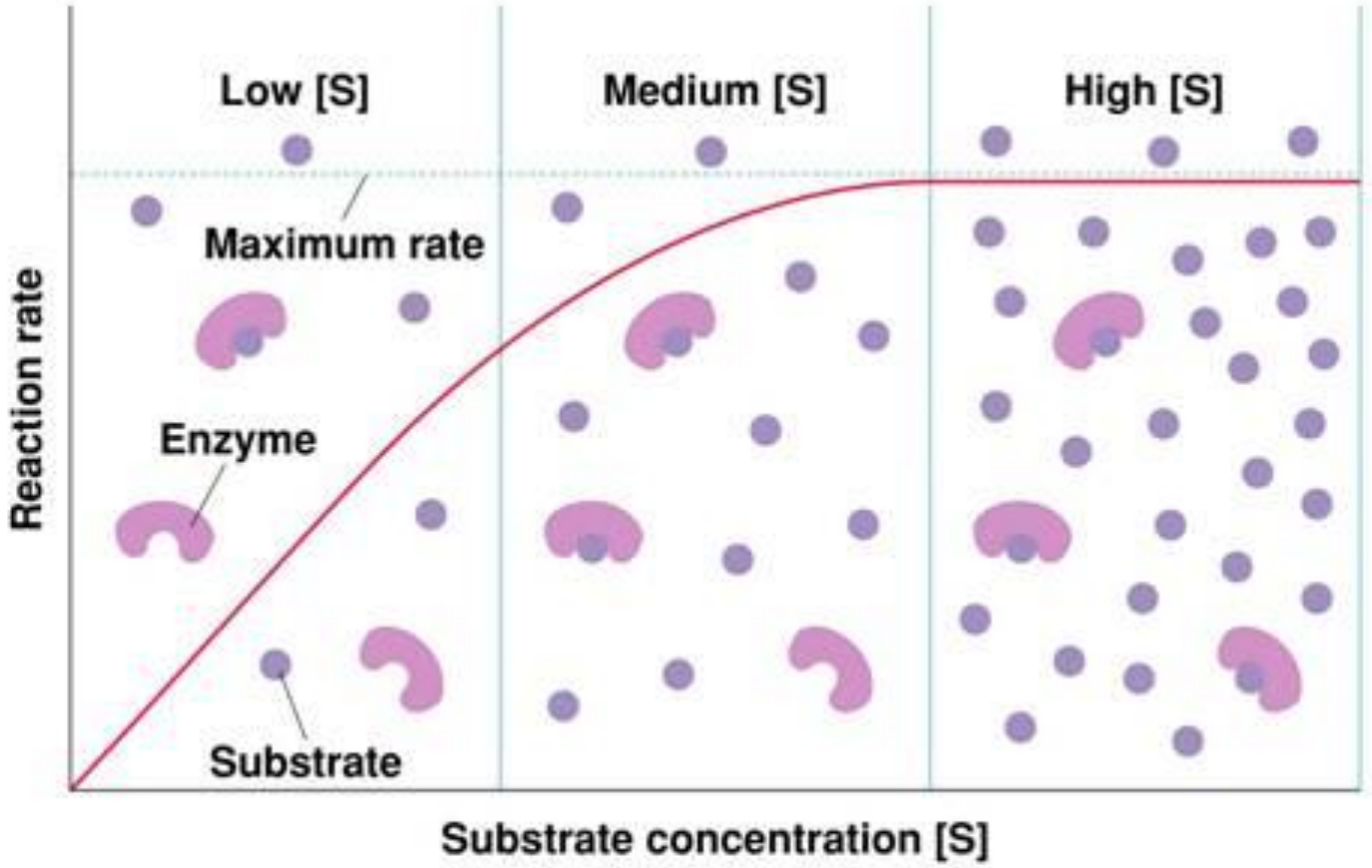
❖ Increasing substrate concentration increases the rate of reaction. This is because more substrate molecules will be colliding with enzyme molecules, so more product will be formed.

❖ However, after a certain concentration, any increase will have no effect on the rate of reaction, since substrate concentration will no longer be the limiting factor. The enzymes will effectively become saturated, and will be working at their maximum possible rate.



❖ Increasing enzyme concentration will increase the rate of reaction, as more enzymes will be colliding with substrate molecules.

❖However, this too will only have an effect up to a certain concentration, where the enzyme concentration is no longer the limiting factor.



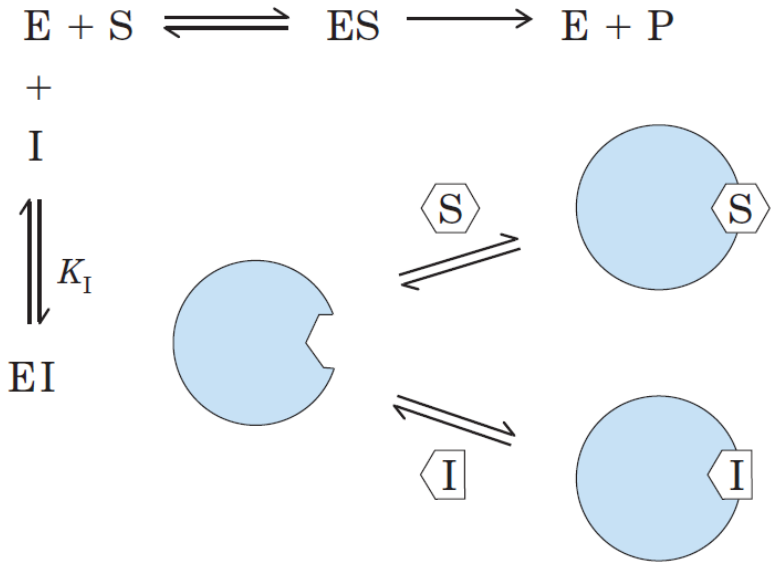
❖Enzyme inhibitors are molecules that interfere with catalysis, slowing or halting enzymatic reactions. This process is known as enzyme inhibition.

❖Enzymes catalyze virtually all cellular processes, so it should not be surprising that enzyme inhibitors are among the most important pharmaceutical agents known.

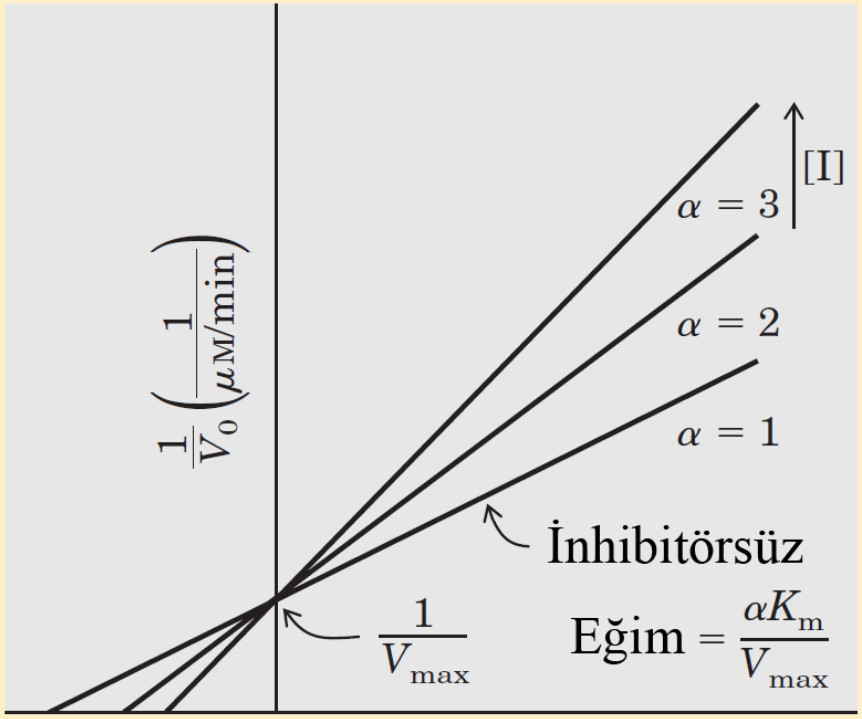
❖For example, aspirin (acetylsalicylate) inhibits the enzyme that catalyzes the first step in the synthesis of prostaglandins, compounds involved in many processes, including some that produce pain.

❖There are two broad classes of enzyme inhibitors: reversible and irreversible.

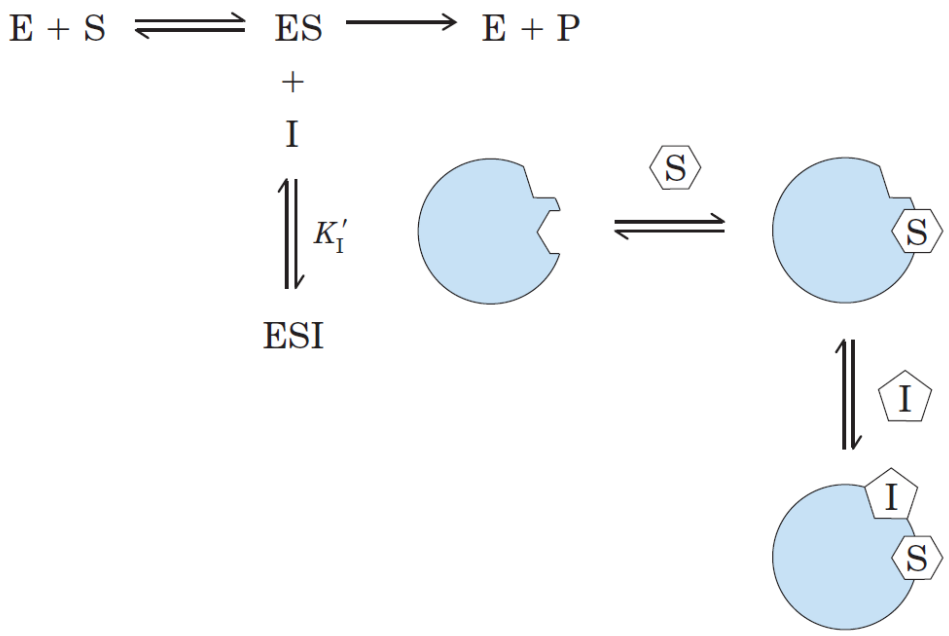
❖A **competitive inhibitor** competes with the substrate for the active site of an enzyme. While the inhibitor (I) occupies the active site, it prevents binding of the substrate to the enzyme.

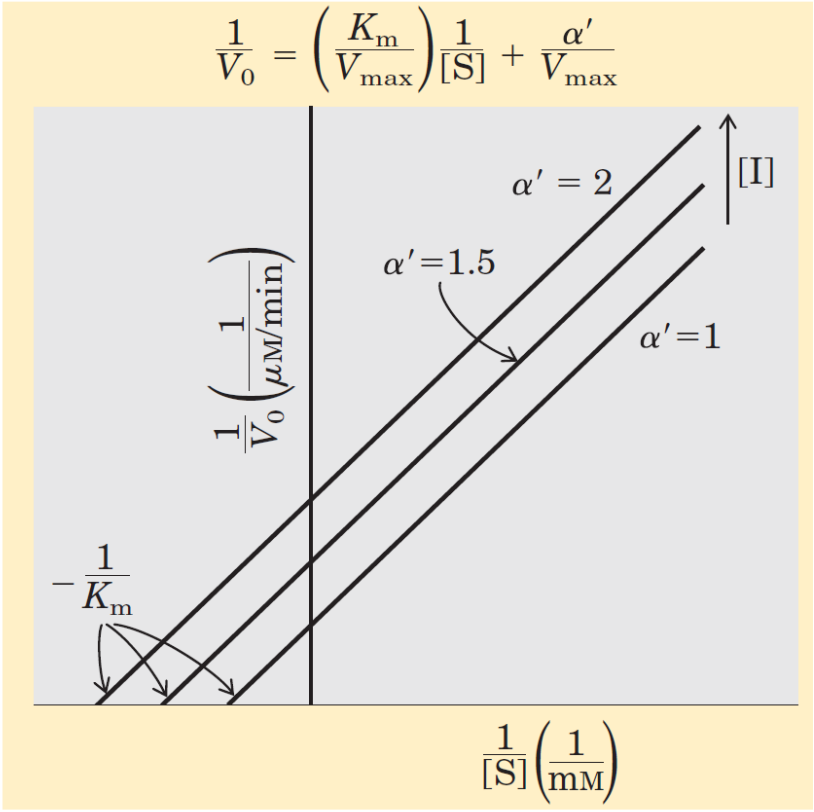


$$\frac{1}{V_0} = \left(\frac{\alpha K_m}{V_{max}}\right)\frac{1}{[S]} + \frac{1}{V_{max}}$$

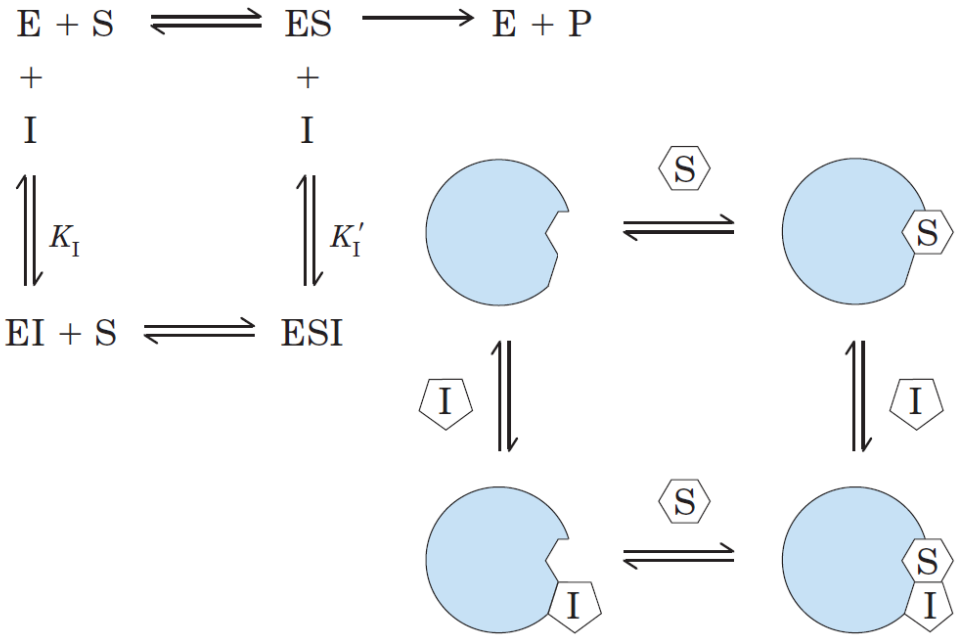


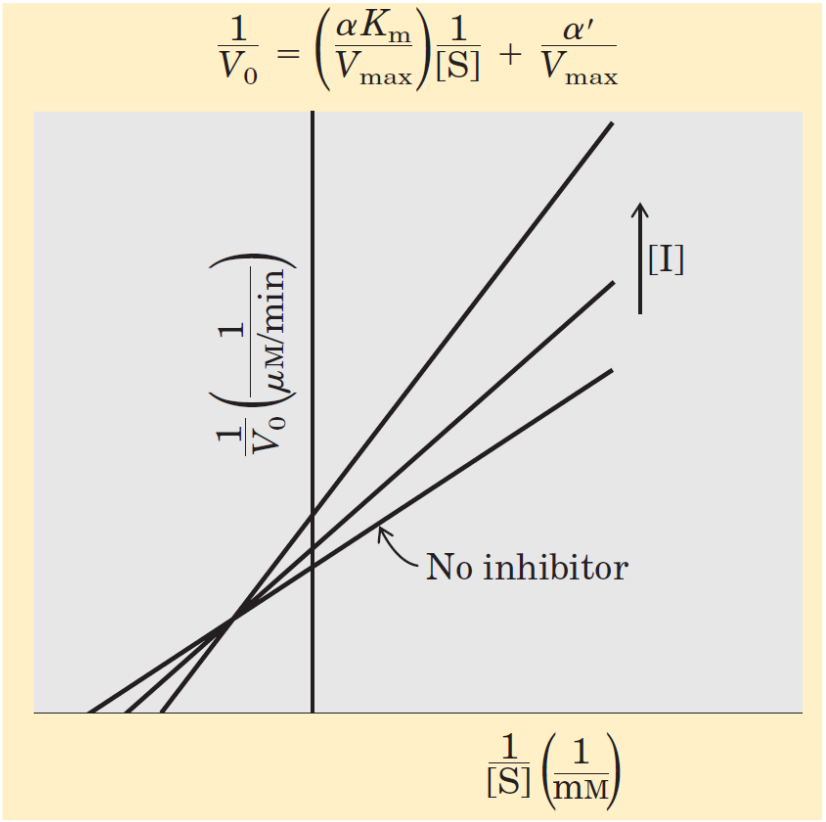
❖An **uncompetitive inhibitor** binds at a site distinct from the substrate active site and, unlike a competitive inhibitor, binds only to the ES complex.



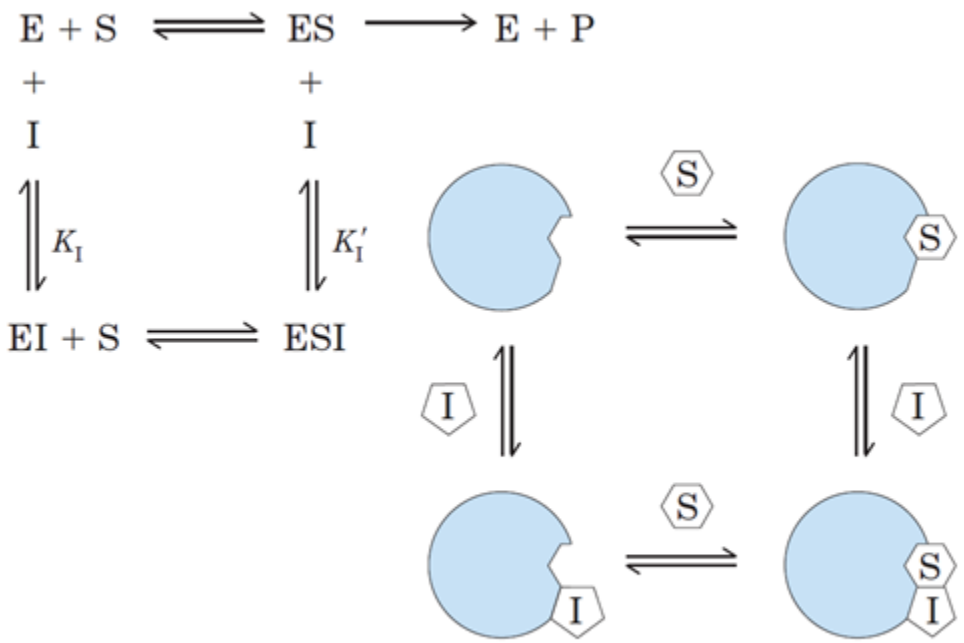


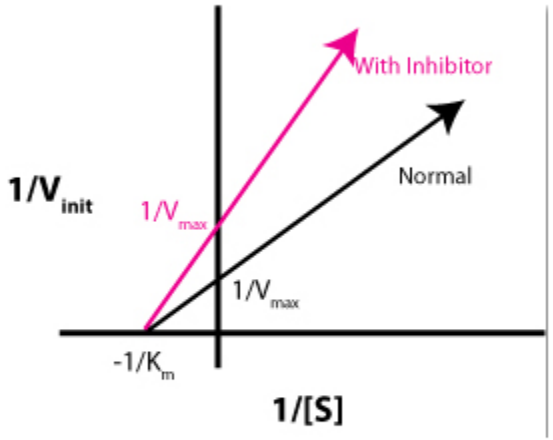
❖A **mixed inhibitor** also binds at a site distinct from the substrate active site, but it binds to either E or ES.



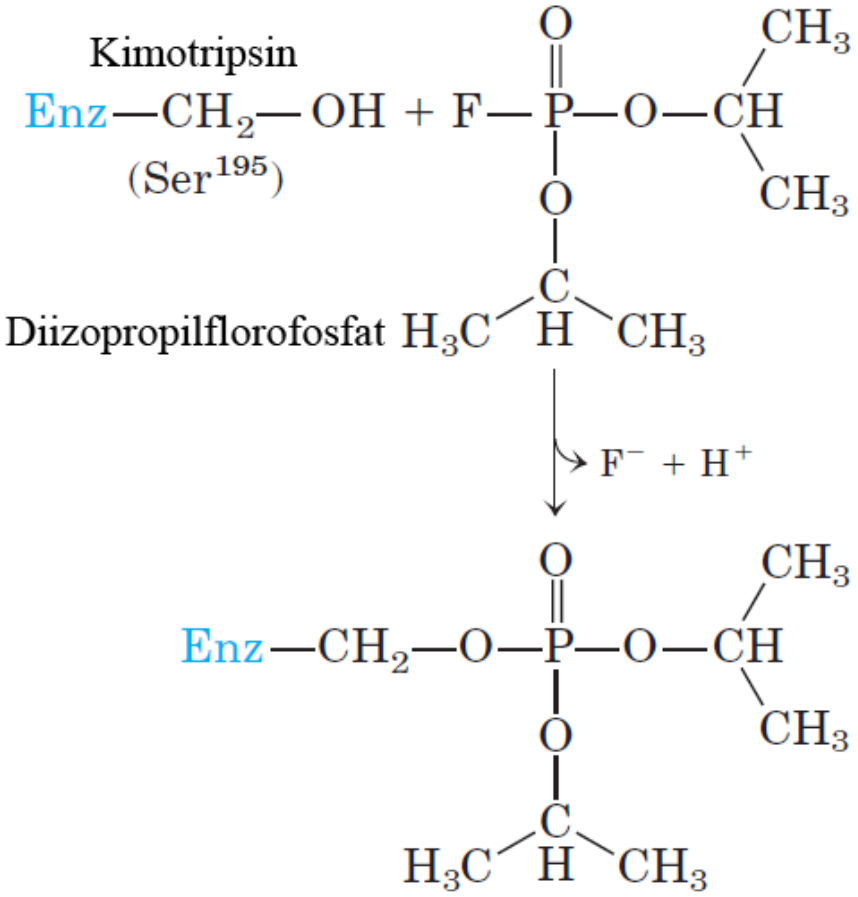


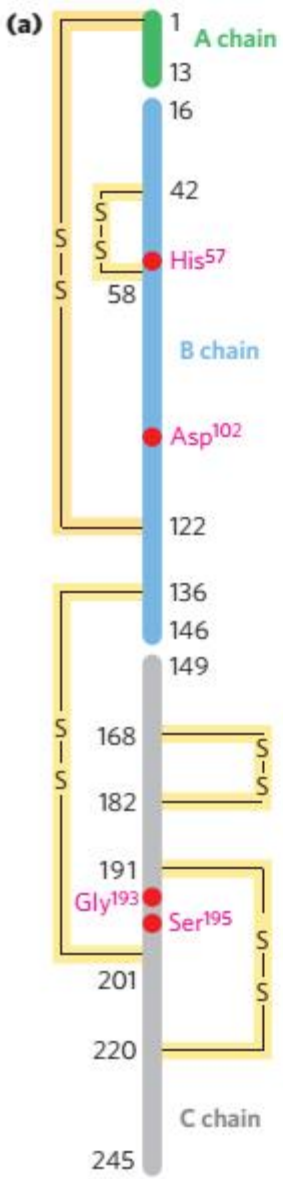
❖A mixed inhibitor usually affects both Km and Vmax. The special case of inhibition, rarely encountered in experiments, classically has been defined as **noncompetitive inhibition**.

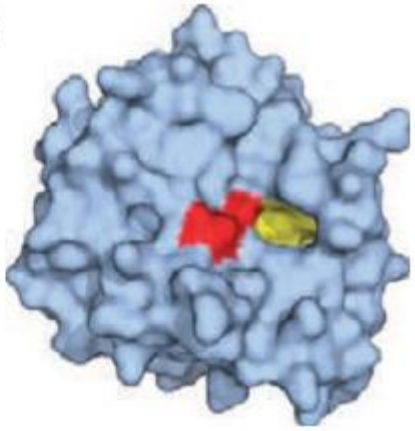


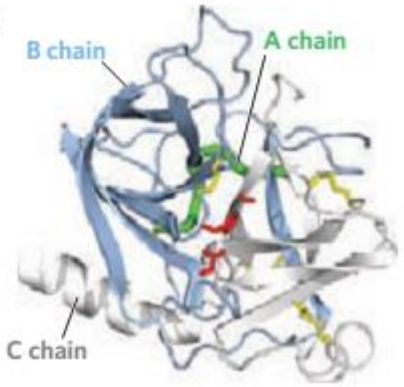


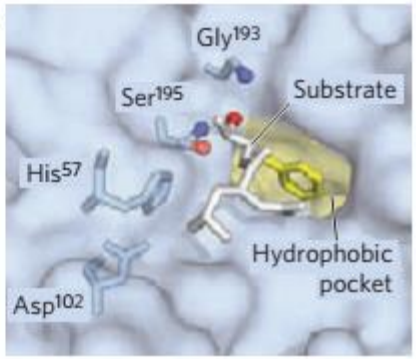
❖The irreversible inhibitors bind covalently with or destroy a functional group on an enzyme that is essential for the enzyme's activity, or form a particularly stable noncovalent association.

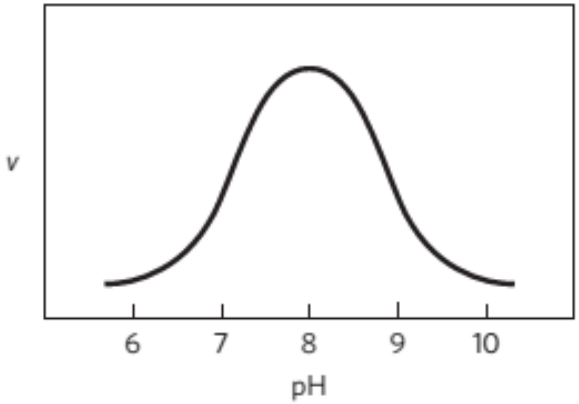


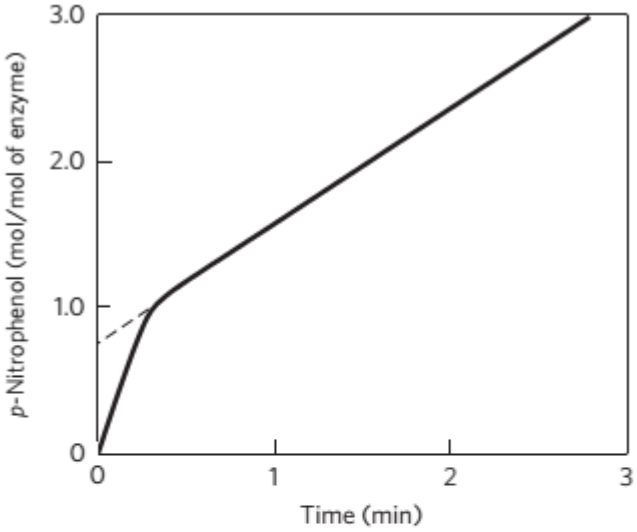

(a)
1
A chain
13
16
42
S
S
S
S
His57
58
B chain
Asp102
122
136
146
149
168
182
191
Gly193
Ser195
201
220
C chain
245

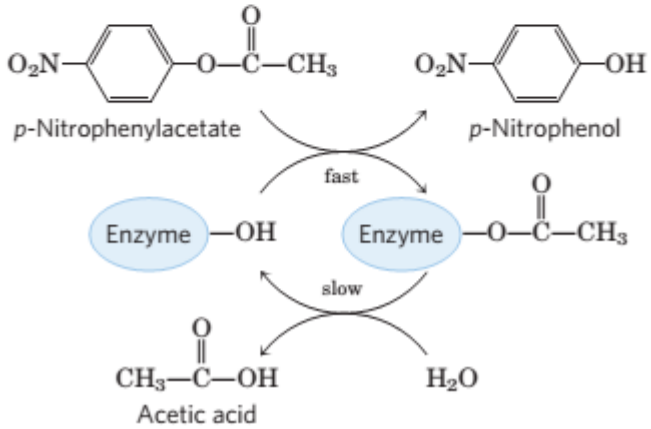

(b)


(c)
B chain
A chain
C chain

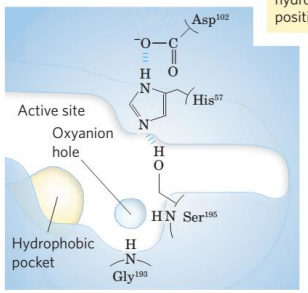

(d)
Gly193
Ser195
Substrate
His57
Hydrophobic pocket
Asp102


v
6
7
8
9
10
pH


3.0
2.0
1.0
0
p-Nitrophenol (mol/mol of enzyme)
0
1
2
3
Time (min)


O2N
O
O—C—CH3
p-Nitrophenylacetate
O2N
OH
p-Nitrophenol
fast
Enzyme
OH
Enzyme
O
O—C—CH3
slow
CH3—C—OH
O
H2O
Acetic acid

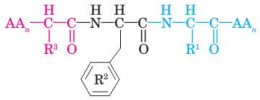
Chymotrypsin (free enzyme)



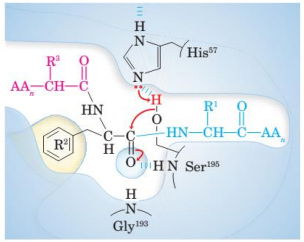
When substrate binds, the side chain of the residue adjacent to the peptide bond to be cleaved nestles in a hydrophobic pocket on the enzyme, positioning the peptide bond for attack.

1

Substrate (a polypeptide)



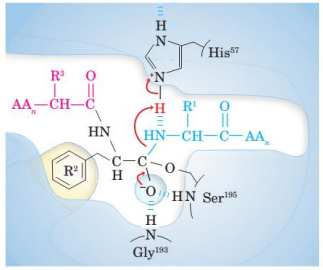
Enzyme-substrate complex



Interaction of Ser195 and His57 generates a strongly nucleophilic alkoxide ion on Ser195; the ion attacks the peptide carbonyl group, forming a tetrahedral acyl-enzyme. This is accompanied by formation of a short-lived negative charge on the carbonyl oxygen of the substrate, which is stabilized by hydrogen bonding in the oxyanion hole.

2

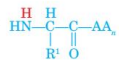
Short-lived intermediate* (acylation)



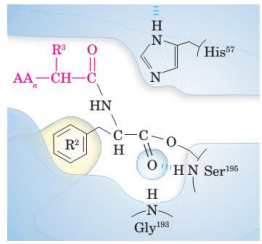
Instability of the negative charge on the substrate carbonyl oxygen leads to collapse of the tetrahedral intermediate; re-formation of a double bond with carbon displaces the bond between carbon and the amino group of the peptide linkage, breaking the peptide bond. The amino leaving group is proton-ated by His57, facilitating its displacement.

3

Product 1

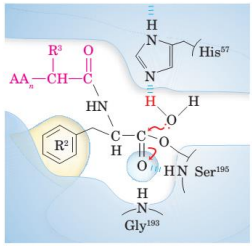


Acyl-enzyme intermediate



4

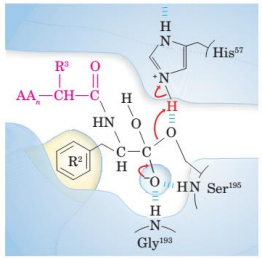
Acyl-enzyme intermediate



An incoming water molecule is deprotonated by general base catalysis, generating a strongly nucleophilic hydroxide ion. Attack of hydroxide on the ester linkage of the acyl-enzyme generates a second tetrahe-dral intermediate, with oxygen in the oxyanion hole again taking on a negative charge.

5

Short-lived intermediate* (deacylation)



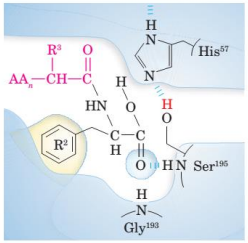
Collapse of the tetrahedral intermediate forms the second product, a carboxylate anion, and displaces Ser195.

and displaces Ser195.

6

Nucleophiles	Electrophiles
Negatively charged oxygen (as in an unprotonated hydroxyl group or an ionized carboxylic acid)	Carbon atom of a carbonyl group (the more electronegative oxygen of the carbonyl group pulls electrons away from the carbon)
Negatively charged sulfhydryl	Protonated imine group (activated for nucleophilic attack at the carbon by protonation of the imine)
Carbanion	Phosphorus of a phosphate group
Uncharged amine group	Proton
Imidazole	
Hydroxide ion	

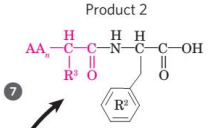
Enzyme-product 2 complex



Dissociation of the second product from the active site regenerates free enzyme.

7

Product 2



❖In cellular metabolism, groups of enzymes work together in sequential pathways to carry out a given metabolic process.

❖In such enzyme systems, the reaction product of one enzyme becomes the substrate of the next.

❖Most of the enzymes in each metabolic pathway follow the kinetic patterns we have already described.

❖Each pathway, however, includes one or more enzymes that have a greater effect on the rate of the overall sequence.

❖These regulatory enzymes exhibit increased or decreased catalytic activity in response to certain signal.

❖The activities of regulatory enzymes are modulated in a variety of ways.

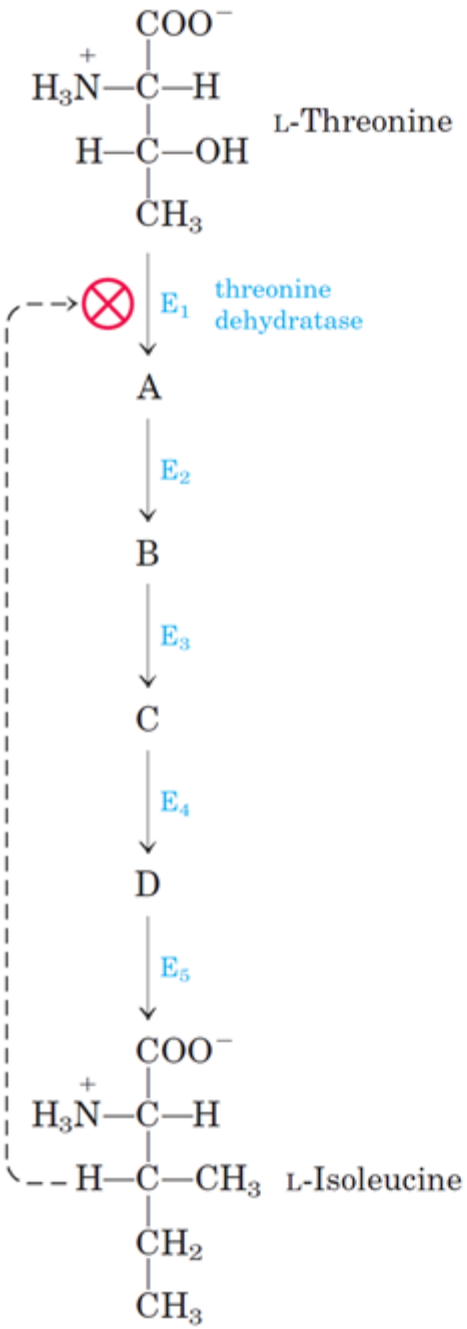
❖Allosteric enzymes function through reversible, noncovalent binding of regulatory compounds called allosteric modulators or allosteric effectors, which are generally small metabolites or cofactors.

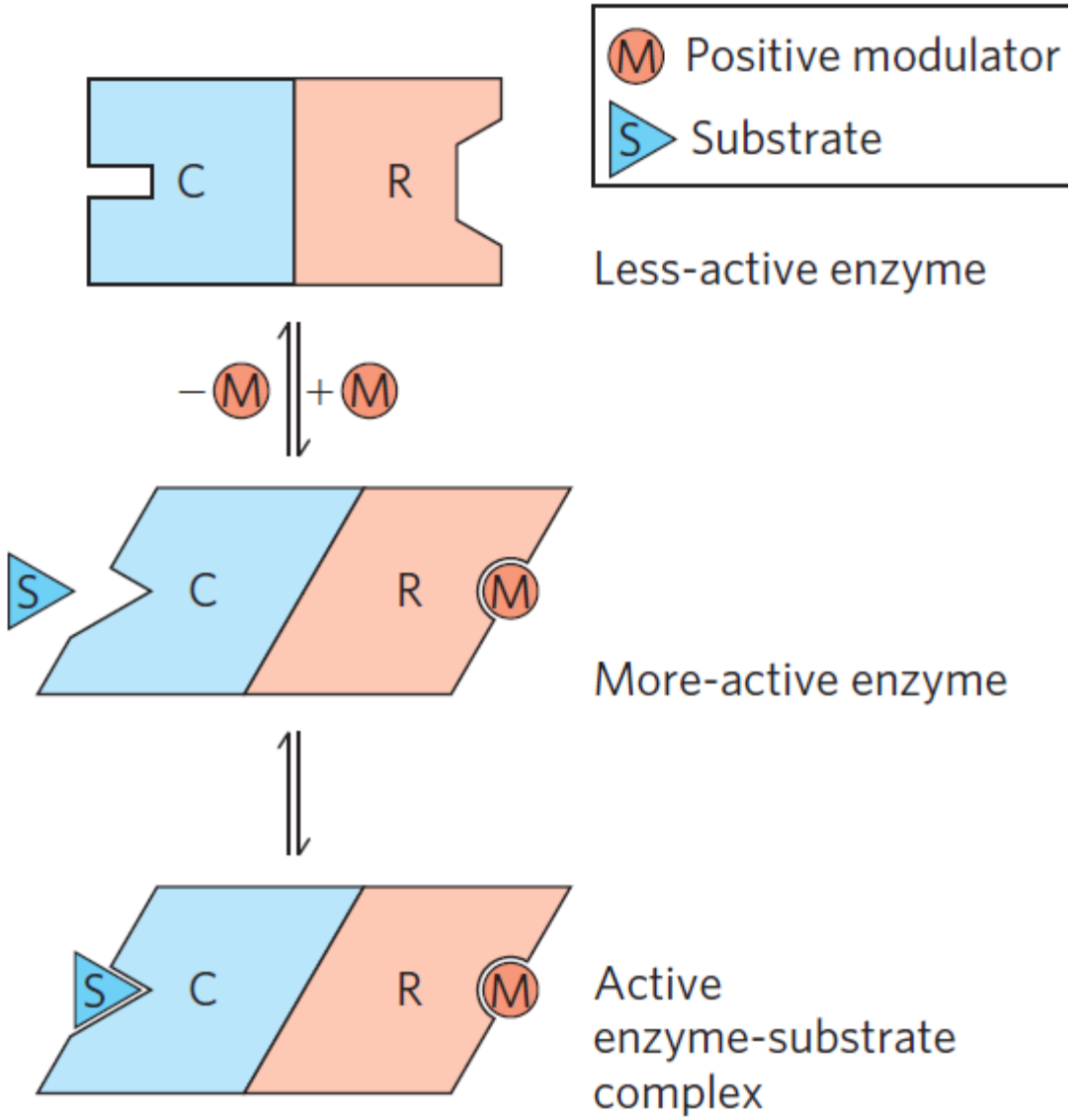
❖Other enzymes are regulated by reversible covalent modification.

❖Metabolic systems have at least two other mechanisms of enzyme regulation.

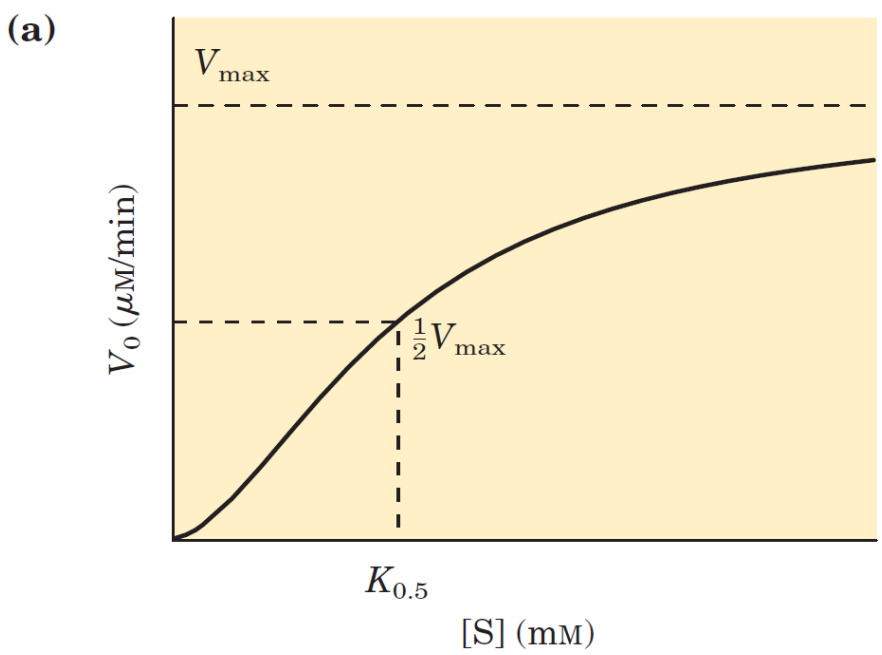
❖Some enzymes are stimulated or inhibited when they are bound by separate regulatory proteins.

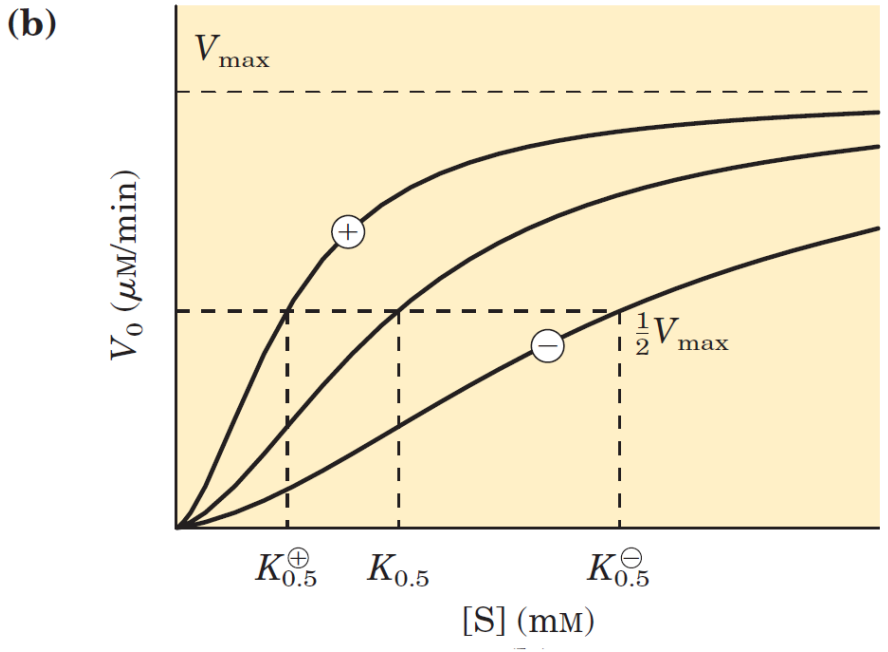
❖Others are activated when peptide segments are removed by proteolytic cleavage; unlike effector-mediated regulation, regulation by proteolytic cleavage is irreversible.


COO⁻
H₃N—C—H
H—C—OH
CH₃
L-Threonine
E1
threonine dehydratase
A
E2
B
E3
C
E4
D
E5
COO⁻
H₃N—C—H
H—C—CH₃
CH₂
CH₃
L-Isoleucine


M Positive modulator
S Substrate
C
R
Less-active enzyme
− M
+ M
S
C
R
M
More-active enzyme
S
C
R
M
Active enzyme-substrate complex

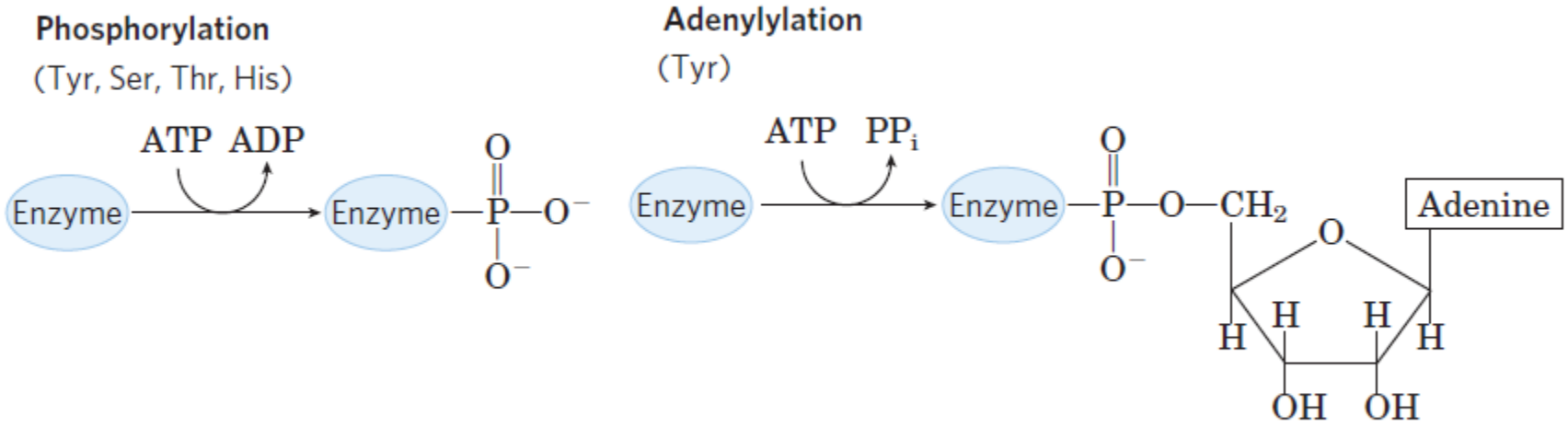
❖Kinetics of allosteric enzymes are different than Michealis-Menten.





❖In another important class of regulatory enzymes, activity is modulated by covalent modification of one or more of the amino acid residues in the enzyme molecule.

❖Common modifying groups include phosphoryl, acetyl, adenylyl, uridylyl, methyl, amide, carboxyl, myristoyl, palmitoyl, prenyl, hydroxyl, sulfate, and adenosine diphosphate ribosyl groups



❖For some enzymes, an inactive precursor called a zymogen is cleaved to form the active enzyme.

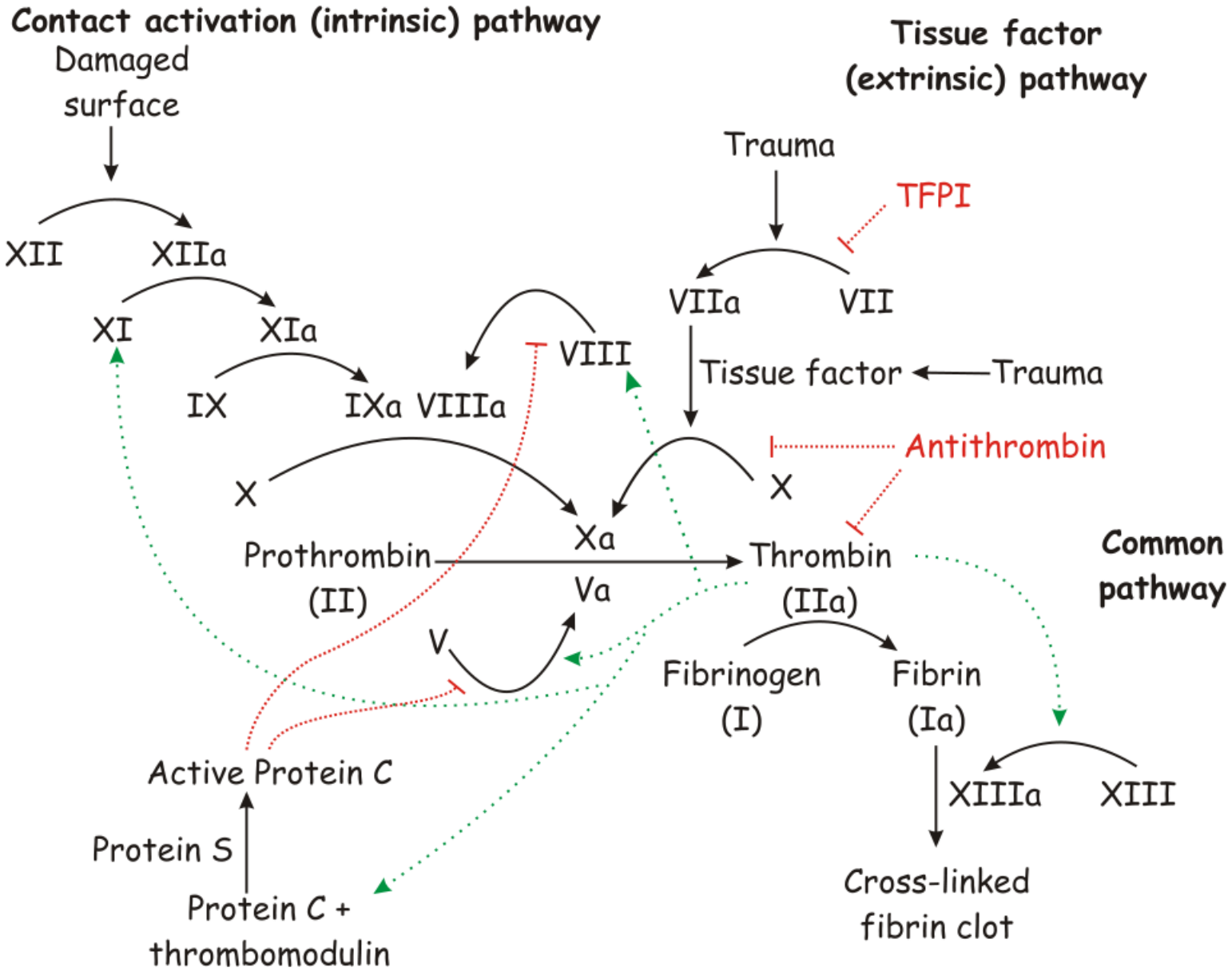
❖Many proteolytic enzymes (proteases) of the stomach and pancreas are regulated in this way.

❖Chymotrypsin and trypsin are initially synthesized as chymotrypsinogen and trypsinogen.

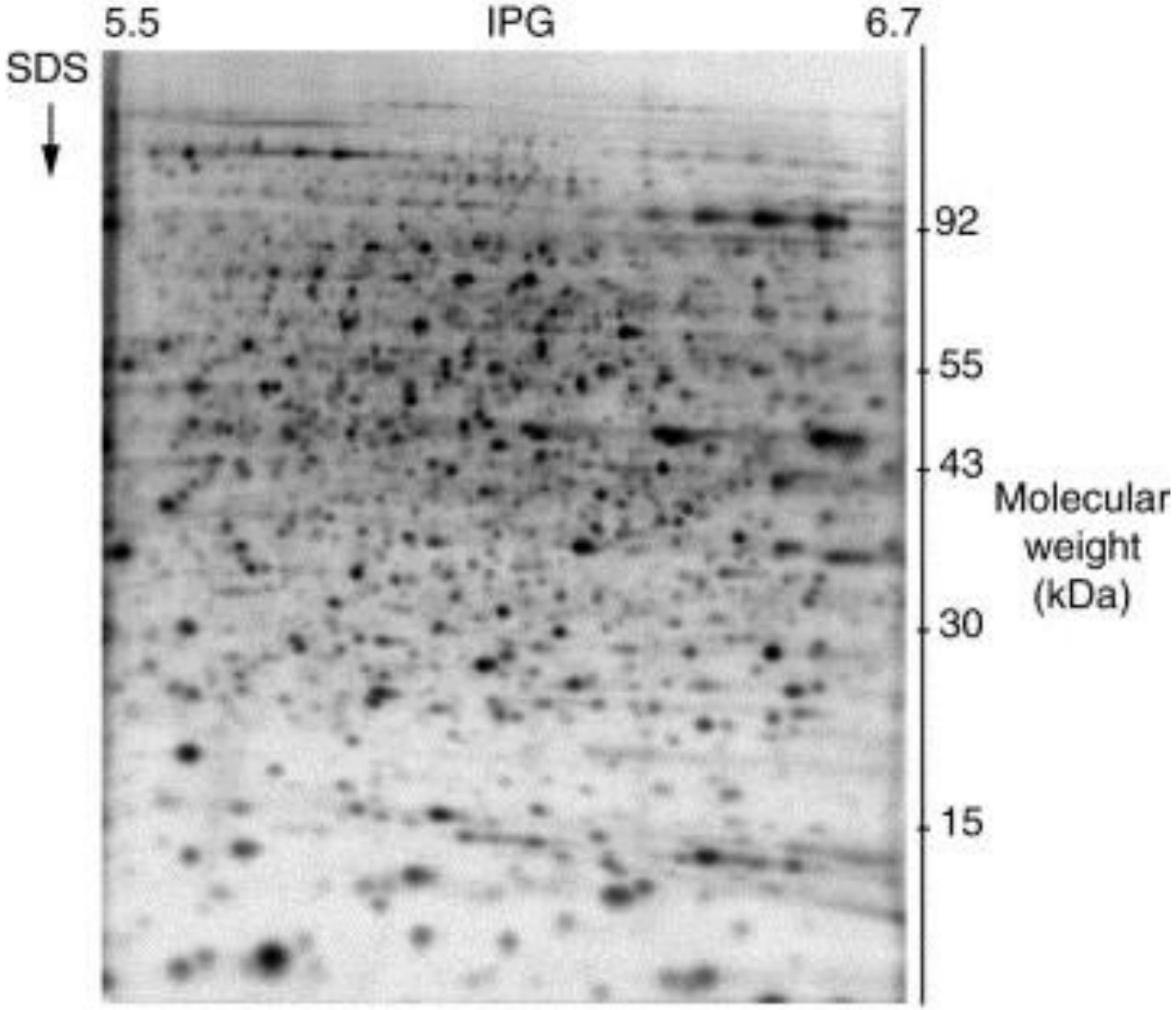
❖Specific cleavage causes conformational changes that expose the enzyme active site.

❖Fibrin is derived from a soluble zymogen called fibrinogen.

❖Because this type of activation is irreversible, other mechanisms are needed to inactivate these enzymes.


Contact activation (intrinsic) pathway
Damaged surface
XII
XIIa
XI
XIa
IX
IXa
VIII
VIIIa
X
Xa
Prothrombin (II)
Thrombin (IIa)
Va
V
Active Protein C
Protein S
Protein C + thrombomodulin
Tissue factor (extrinsic) pathway
Trauma
TFPI
VII
VIIa
Tissue factor
Trauma
Antithrombin
X
Common pathway
Fibrinogen (I)
Fibrin (Ia)
XIII
XIIIa
Cross-linked fibrin clot

AMINO ACIDS, PEPTIDES AND PROTEINS V: Protein Purification and Characterization



Current Opinion in Chemical Biology

❖Our understanding of protein structure and function has been derived from the study of many individual proteins.

❖To study a protein in detail, the researcher must be able to separate it from other proteins and must have the techniques to determine its properties.

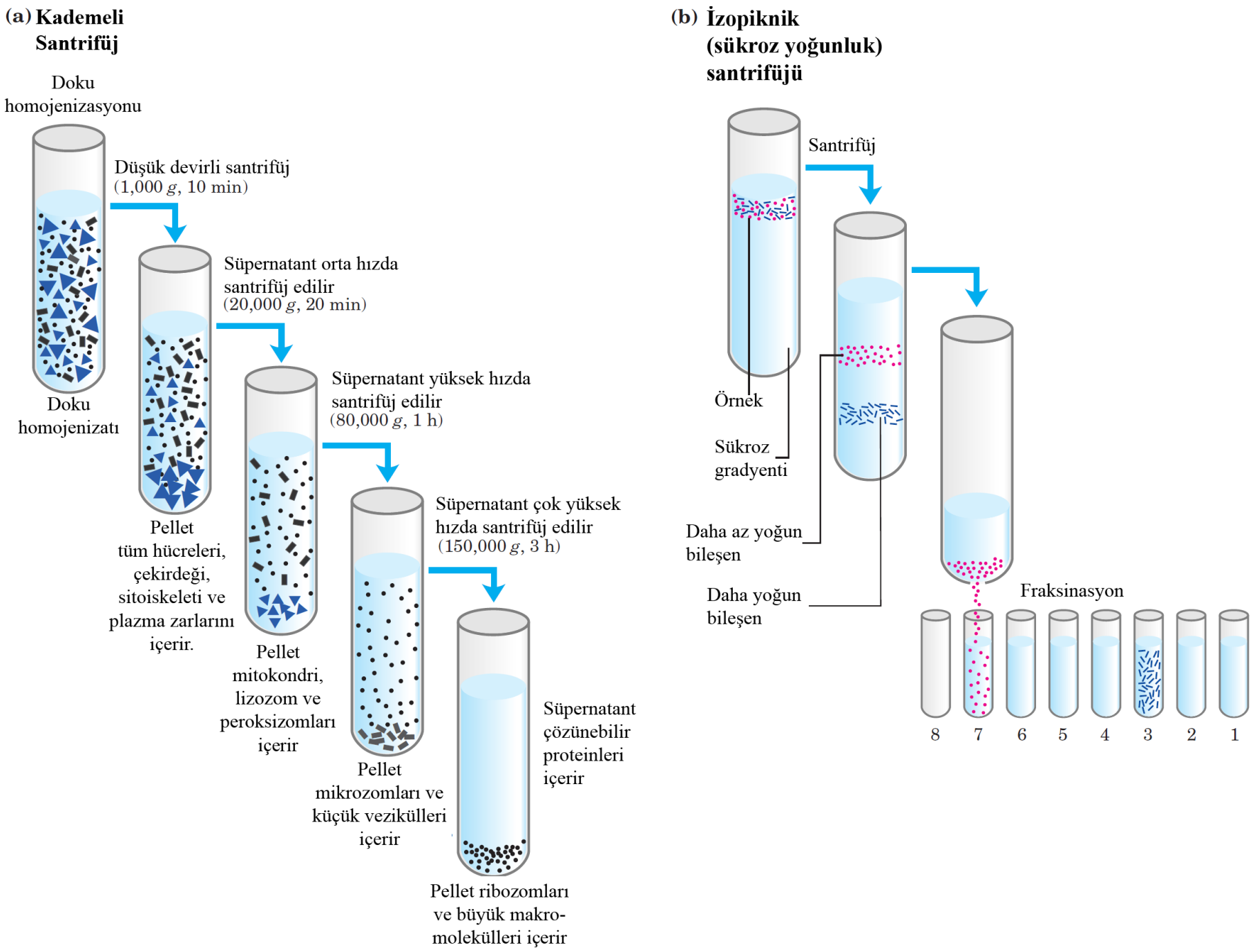
❖ Methods for separating proteins take advantage of properties that vary from one protein to the next, including size, charge, and binding properties

❖The source of a protein is generally tissue or microbial cells.

❖The first step in any protein purification procedure is to break open these cells, releasing their proteins into a solution called a crude extract.

❖If necessary, differential centrifugation can be used to prepare subcellular fractions or to isolate specific organelleas.

❖Commonly, the extract is subjected to treatments that separate the proteins into different fractions based on a property such as size or charge, a process referred to as fractionation.


(a) Kademeli Santrifüj
Doku homojenizasyonu
Düşük devirli santrifüj (1,000 g, 10 min)
Doku homojenizatı
Süpernatant orta hızda santrifüj edilir (20,000 g, 20 min)
Pellet tüm hücreleri, çekirdeği, sitoiskeleti ve plazma zarlarını içerir.
Süpernatant yüksek hızda santrifüj edilir (80,000 g, 1 h)
Pellet mitokondri, lizozom ve peroksizomları içerir
Süpernatant çok yüksek hızda santrifüj edilir (150,000 g, 3 h)
Pellet mikrozomları ve küçük vezikülleri içerir
Süpernatant çözünebilir proteinleri içerir
Pellet ribozomları ve büyük makro-molekülleri içerir
(b) İzopiknik (sükroz yoğunluk) santrifüjü
Santrifüj
Örnek
Sükroz gradyenti
Daha az yoğun bileşen
Daha yoğun bileşen
Fraksinasyon
8
7
6
5
4
3
2
1

❖Once the extract or organelle preparation is ready, various methods are available for purifying one or more of the proteins it contains.

❖Commonly, the extract is subjected to treatments that separate the proteins into different fractions based on a property such as size or charge, a process referred to as fractionation.

❖Early fractionation steps in a purification utilize differences in protein solubility, which is a complex function of pH, temperature, salt concentration, and other factors.

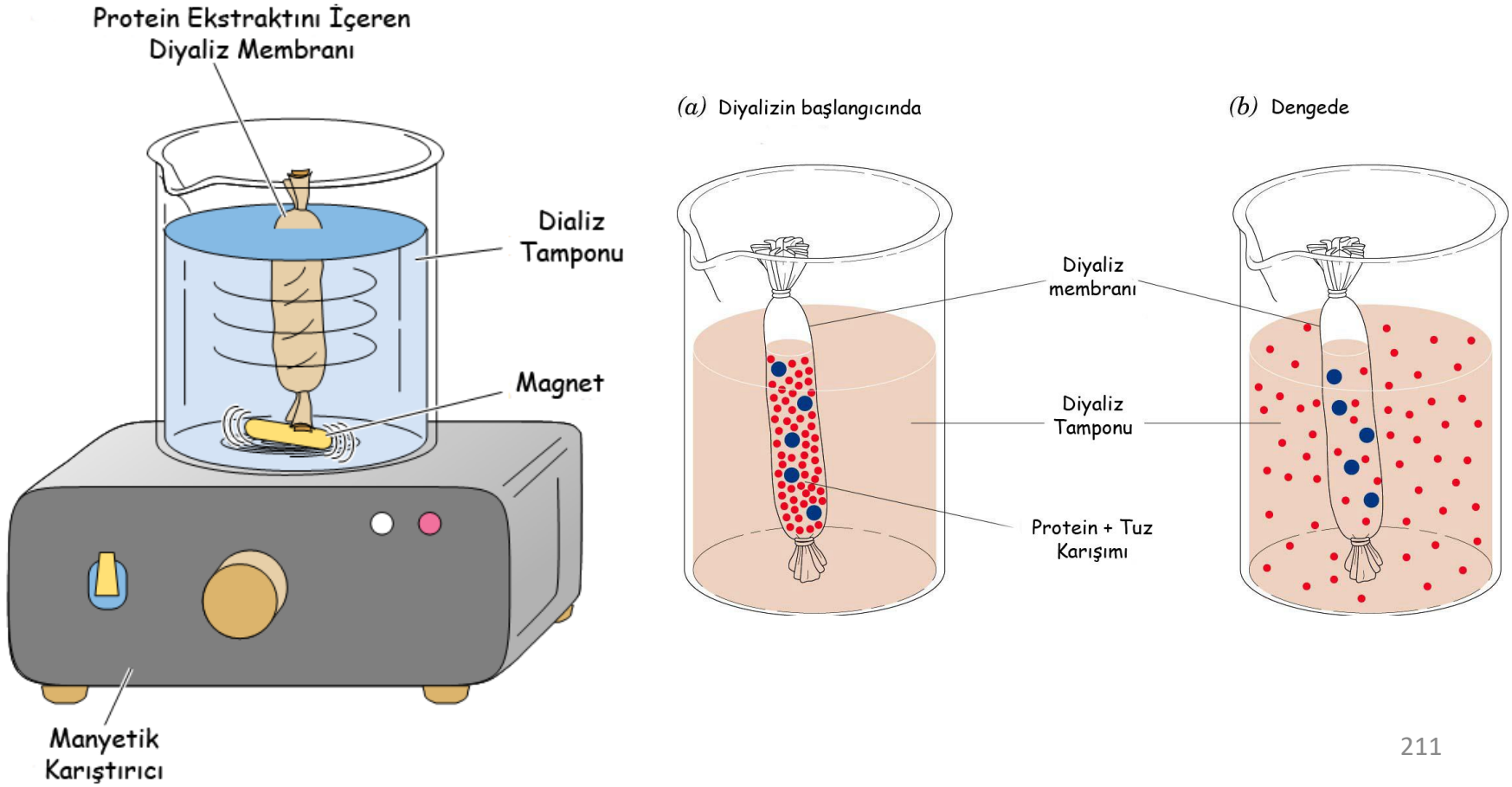
❖The solubility of proteins is generally lowered at high salt concentrations, an effect called “salting out.”

❖The addition of a salt in the right amount can selectively precipitate some proteins, while others remain in solution.

❖Ammonium sulfate ($(NH_4)_2SO_4$) is often used for this purpose because of its high solubility in water.

❖A solution containing the protein of interest often must be further altered before subsequent purification steps are possible.

❖For example, dialysis is a procedure that separates proteins from solvents by taking advantage of the proteins' larger size.

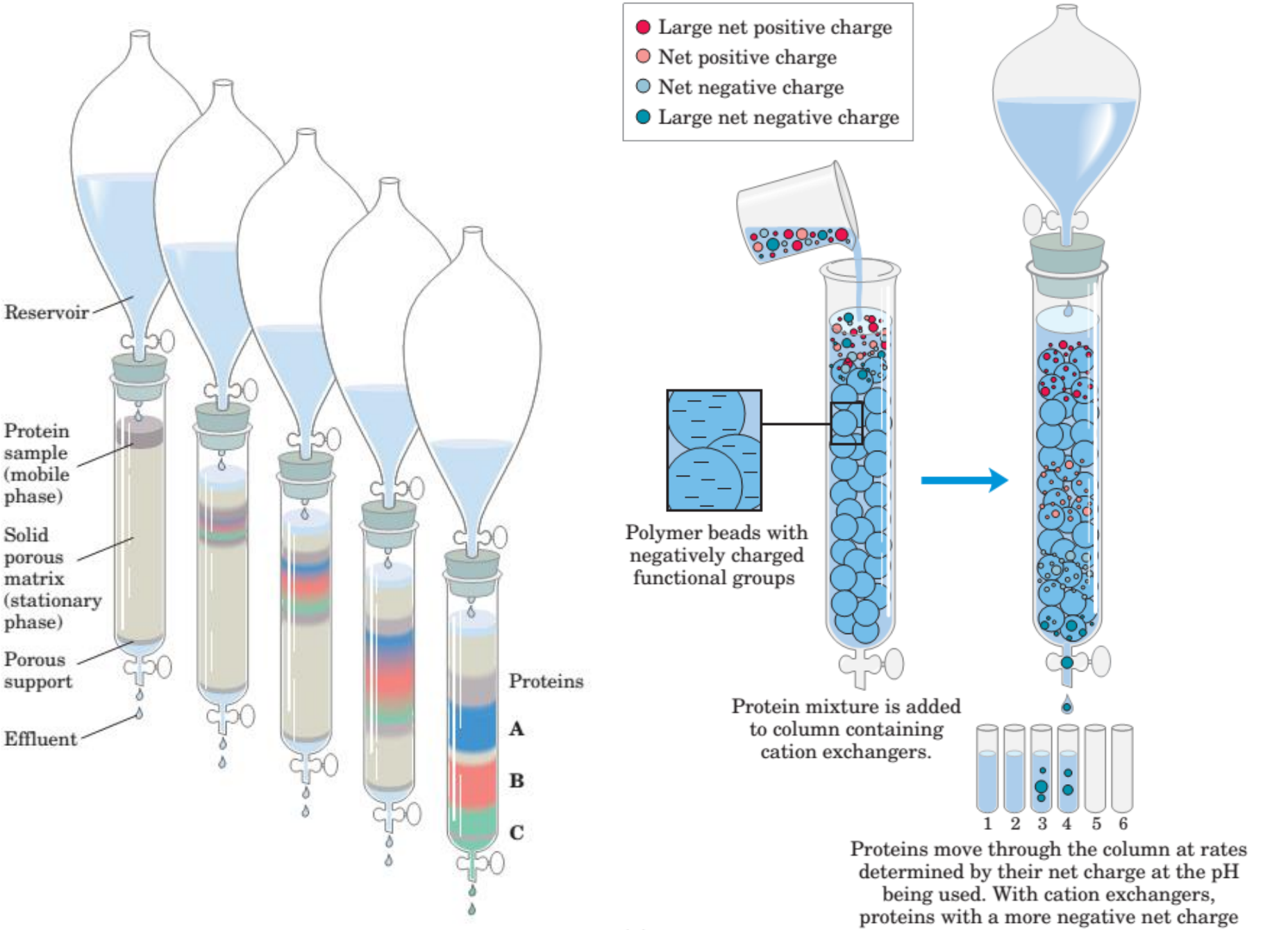


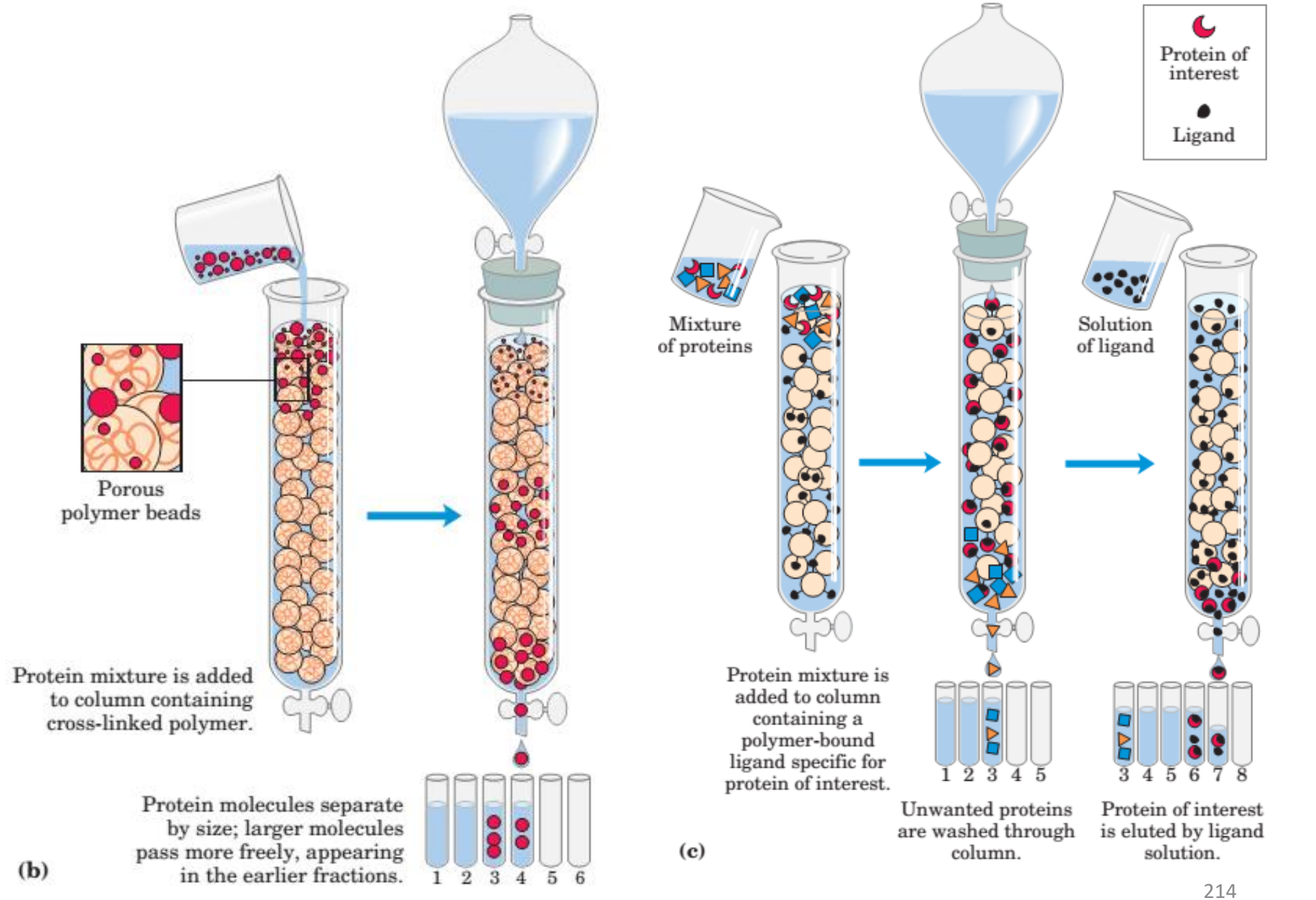
❖The most powerful methods for fractionating proteins make use of column chromatography, which takes advantage of differences in protein charge, size, binding affinity, and other properties.

❖ A porous solid material with appropriate chemical properties (the stationary phase) is held in a column, and a buffered solution (the mobile phase) percolates through it.

❖The protein-containing solution, layered on the top of the column, percolates through the solid matrix as an ever-expanding band within the larger mobile phase.

❖Individual proteins migrate faster or more slowly through the column depending on their properties.


Large net positive charge
Net positive charge
Net negative charge
Large net negative charge
Reservoir
Protein sample (mobile phase)
Solid porous matrix (stationary phase)
Porous support
Effluent
Proteins
A
B
C
Polymer beads with negatively charged functional groups
Protein mixture is added to column containing cation exchangers.
1 2 3 4 5 6
Proteins move through the column at rates determined by their net charge at the pH being used. With cation exchangers, proteins with a more negative net charge move faster and elute earlier.
(a)


Protein of interest
Ligand
Porous polymer beads
Protein mixture is added to column containing cross-linked polymer.
Protein molecules separate by size; larger molecules pass more freely, appearing in the earlier fractions.
1 2 3 4 5 6
(b)
Mixture of proteins
Protein mixture is added to column containing a polymer-bound ligand specific for protein of interest.
1 2 3 4 5
Unwanted proteins are washed through column.
Solution of ligand
3 4 5 6 7 8
Protein of interest is eluted by ligand solution.
(c)

❖ Ion-exchange chromatography exploits differences in the sign and magnitude of the net electric charges of proteins at a given pH.

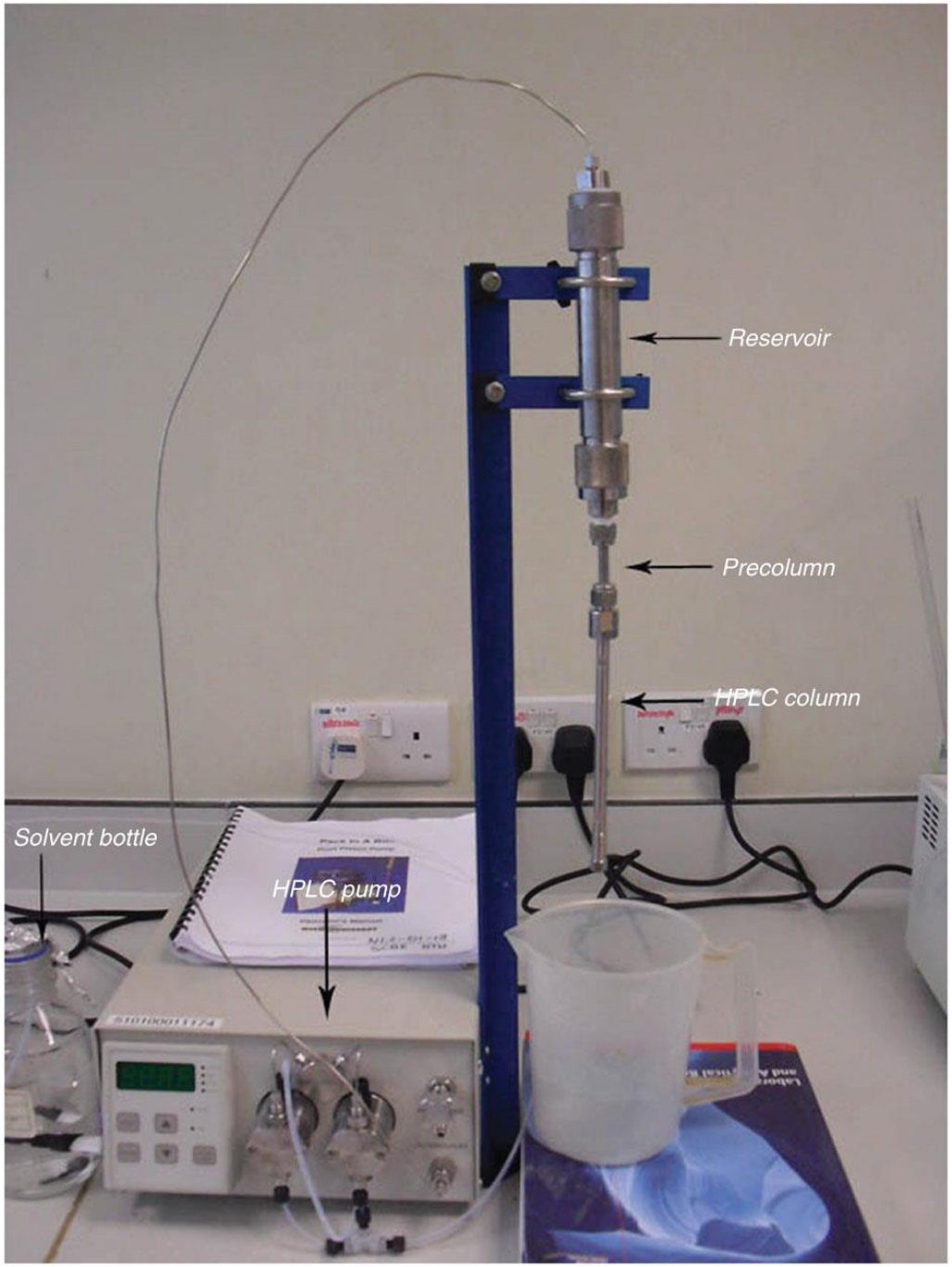
❖Separation can be optimized by gradually changing the pH and/or salt concentration of the mobile phase so as to create a pH or salt gradient.

❖Size-exclusion chromatography, also called gel filtration, separates proteins according to size. The column matrix is a cross-linked polymer with pores of selected size.

❖Affinity chromatography separates proteins by their binding specificities.

❖The proteins retained on the column are those that bind specifically to a ligand cross-linked to the beads.

❖One of the most advance chromatografic tehnique is high performance liquid chromatography; HPLC.

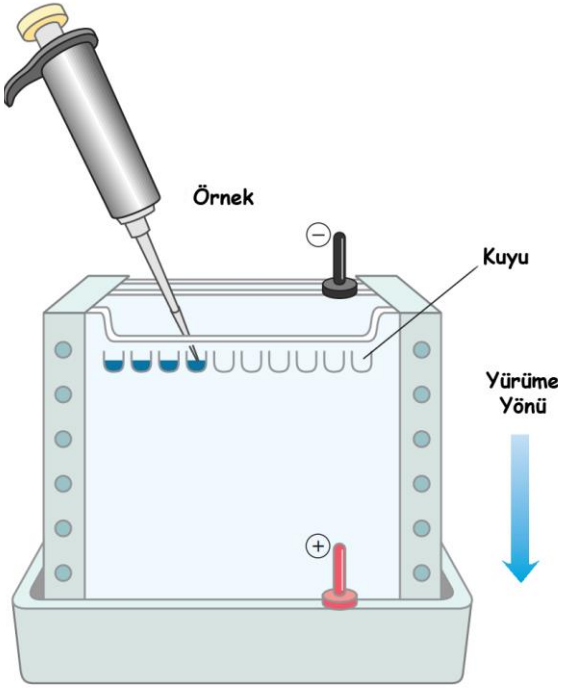


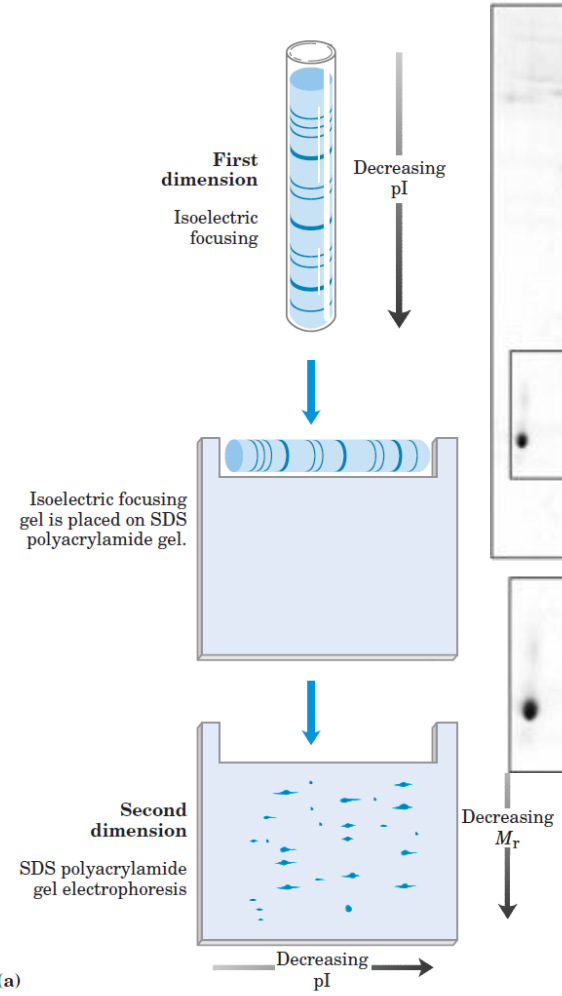
❖Another important technique for the separation of proteins is based on the migration of charged proteins in an electric field, a process called electrophoresis.

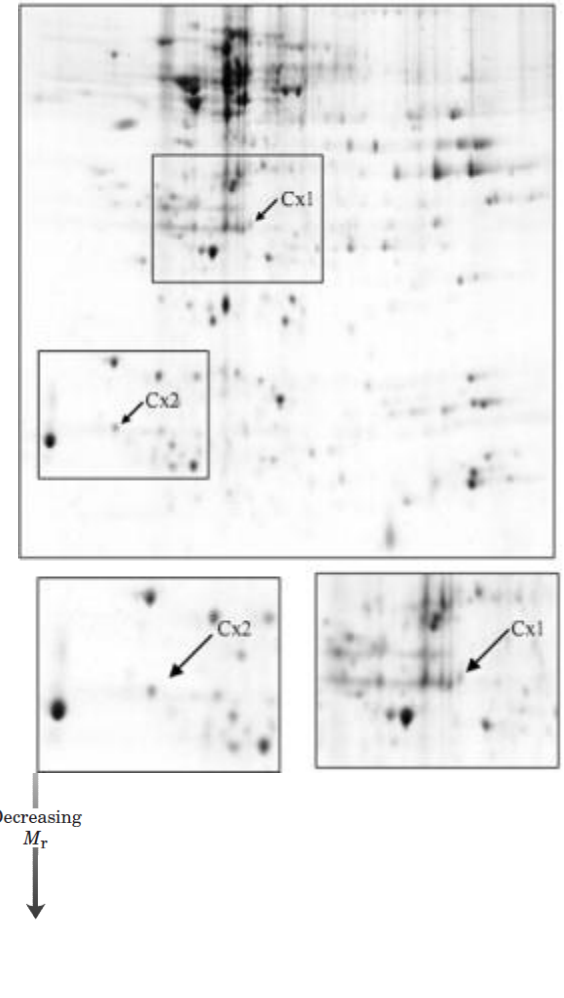
❖These procedures are not generally used to purify proteins in large amounts, because simpler alternatives are usually available and electrophoretic methods often adversely affect the structure and thus the function of proteins.

❖Electrophoresis is, however, especially useful as an analytical method. Its advantage is that proteins can be visualized as well as separated, permitting a researcher to estimate quickly the number of different proteins in a mixture or the degree of purity of a particular protein preparation.

❖An electrophoretic method commonly employed for estimation of purity and molecular weight makes use of the detergent sodium dodecyl sulfate (SDS).


Örnek
Kuyu
Yürüme Yönü


First dimension
Isoelectric focusing
Decreasing pI
Isoelectric focusing gel is placed on SDS polyacrylamide gel.
Second dimension
SDS polyacrylamide gel electrophoresis
Decreasing M_r
Decreasing pI
(a)


Cx1
Cx2
Cx2
Cx1

- Correct Tissue Selection

❖First step in protein purification is the selection of an appropriate tissue.

❖ If a protein is targeted beforehand, an easy to obtain, relatively less complex, protease free source which contains high amount of desired protein should be selected.

❖As a general rule, all purification processes done in +4°C. If long period of time needed, samples should be kept -80°C.

- Homogenisation

❖In order to isolate an intracellular protein, cells must be destroyed (or lysed). Certain chemical and mechanical processes could be use for this purpose.

❖Success of cell lyses generally depended upon various factors such as buffer selection, use of protease inhibitors and osmolarity of lysis buffer.

TABLE 1.1
Various Cell Lysis Techniques

Technique	Principle	Time of Lysis	Example
Enzyme digestion	Digestion of cell wall leading to osmotic disruption of cell membrane	15–30 min	Gram positive bacteria
Osmotic shock lysis	Osmotic disruption of cell membrane	<5 min	Red blood cells
Hand homogenization	Cells are forced through narrow gap leading to disruption of cell membrane	10–15 min	Liver tissue
Blade homogenizer	Large cells are broken by chopping action	5–10 min	Muscle tissue, animal tissue, plant tissue
Grinding with alumina or sand	Cell walls are ripped off by micro roughness	5–15 min	Bacteria
Grinding with glass beads	Cell walls are ripped off by rapid vibration of glass	10–20 min	Bacteria
French press	Cells are forced through small orifice at very high pressure. Shear forces disrupt cells.	10–30 min	Bacteria, plant cells
Sonication	Cell disruption by shear forces and cavitation caused by high-pressure sound waves	5–10 min	Bacteria

❖It is not possible to apply one protein isolation protocol to all tissue and cell samples. Best method to isolate the targeted protein should be defined with literature search and trial and error approach.

❖Proteins have very different properties and they are greatly affected by hydrogen ion concentration, therefore pH should be controlled extensively.

❖Various buffer solutions could be use for keeping the pH of the solution a constant value.

❖To choose the proper buffer, you should consider various factors such as pKa, temperature effect, interference with other compounds, UV absorption, suitability with other purification techniques and price.

❖After required buffer selected, buffer solution could be prepared at the working temperature (generally +4°C) with the help of a calibrated pH meter.

❖Proteolysis is an important problem for protein extraction. Protease inhibitiors should be used to counter that problem.

TABLE 1.3
Limitations of Buffers Commonly Used in Extractions

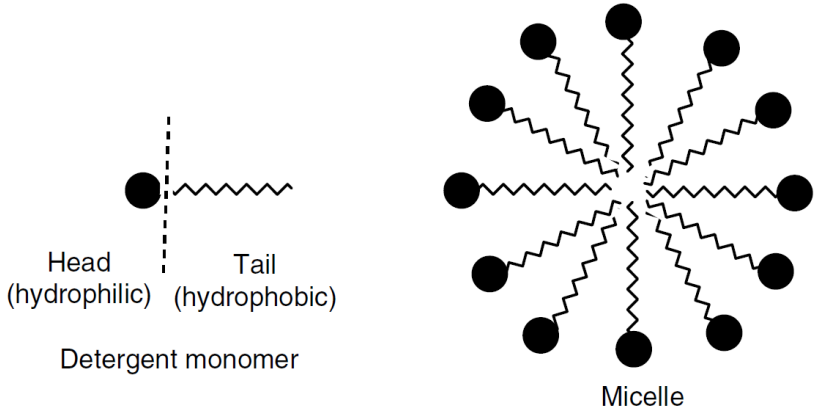
Buffer	Advantage	Disadvantage
Phosphate	Compatible with gel permeation and cation-exchange chromatography Compatible with most cross-linking reagents Inexpensive	Weak buffering capacity in the pH 8–11 Precipitates in the presence of polyvalent cations Inhibits a wide variety of enzymes, including kinase, phosphatase, dehydrogenase Not suitable in anion-exchange chromatography
Tris	Suitable in gel permeation and anion-exchange chromatography Inexpensive	Poor buffer below pH 7 Passes through biological membrane Contains a reactive primary amine and thus forms Schiff base adduct with aldehyde and inhibits protein conjugation by amine-based cross-linkers Sensitive to temperature
Borate	Inexpensive	Forms complexes with the ribose moieties of nucleic acid, and other mono- and oligosaccharides
Citrate	Inexpensive	Binds to some proteins and forms metal complexes
Carbonate	Inexpensive	Limited solubility Since carbonate is in equilibrium with CO_2, studies should be carried out in a closed system
Good buffers (e.g., MES, MOPS, HEPES)	Relatively free of side effects Low UV absorbance Effect of temperature and ionic strength is minimum	Most Good buffers interfere with the Lowry protein assay All piperazine-based Good buffers (HEPES, PIPES, etc.) form radicals under various conditions and are thus not suitable for the systems where redox processes are being studied Expensive

TABLE 1.8
Common protease inhibitors

Inhibitors	Solubility	Recommended Stock	Concentration Used
Serine protease inhibitors			
Phenylmethylsulfonyl fluoride (PMSF)	2-propanol, ethanol	100 mM	0.1–1 mM
Benzamide	2-propanol, ethanol	100 mM	1 mM
Benzamidine-HCl	Water	100 mM	1 mM
ε-Amino n-caproic acid	Water	500 mM	5 mM
Aprotinin	Water	1 mg/ml	1–5 μg/ml
Cysteine (thiol) protease inhibitors			
Sodium p-hydroxymercuribenzoate (PHMB)	Water	100 mM	1 mM
Antipain	Water	1 mg/ml	1 μg/ml
Leupeptin	Water	1 mg/ml	1 μg/ml
Aspartate (acidic) protease inhibitors			
Pepstatin	Methanol	10 mg/ml	0.1 mg/ml
Diazoacetylnorleucine methyl ester (DAN)	Methanol	100 mM	5 mM
Metalloprotease inhibitor			
EGTA [ethylene glycol bis(β-aminoethyl ether) N, N, N′, N′-tetraacetic acid]	Water	500 mM	10 mM

❖Detergents are usually employ to the purification of membrane proteins.

❖Detergents are amphiphilic molecules; their “tail” portion is hydrophobic and “head” portion is hydrophilic.

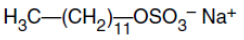


❖Detergents could be classified as ionic, nonionic and zwitterionic.

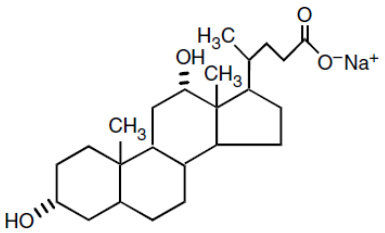
❖Another important feature of detergents is their capability of forming micelles.

Detergent	MW (Da)	CMC	N	HLB	Micelle MW (kDa)	Conc. for Solubilization	Specific Use
Ionic							
Sodium dodecyl sulfate	288.5	7–10 mM, 0.23%	60	40	18		Electrophoresis and electrofocusing Micellar chromatography
Sodium cholate	431	3–10 mM, 0.2%			1.8		Liposome preparation
Sodium deoxycholate	433	2–4 mM, 0.1%	20				Liposome preparation
Non-ionic							
Triton X-100	Approx. 628	0.3 mM, 0.02%	140	13.5	90	0.2–0.6 mg/mg protein	Selective solubilization of membranes Enzyme immunoassay
Triton X-114	Approx. 543	0.17 mM, 0.01%		12.4			Protein fractionation by phase separation Liposome preparation
Octylglucoside	292.4	15–25 mM, 0.5%			8	20–45 mM	Protein crystallization
Tween 20	1,230	50 µM, 0.006%	62	16.7	76		Enzyme immunoassay
Zwitterionic							
CHAPS	614.9	6.5 mM, 0.4%	10		6	6–10 mM	*In situ* hybridization
Zwittergent 3-14	364	0.3 mM			30		

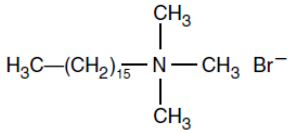
Sodium dodecylsulfate (SDS)



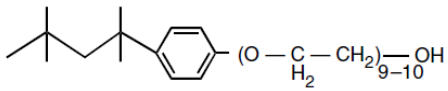
Sodium deoxycholate



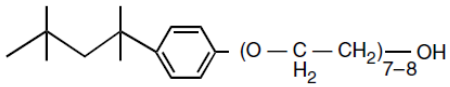
Cetyltrimethylammonium bromide (CTAB)



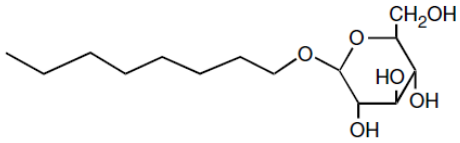
Triton X-100



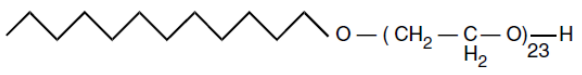
Triton X-114



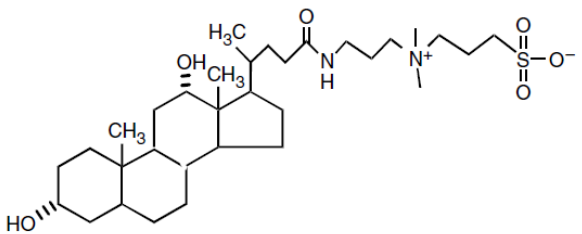
Octyl glucoside



Brij 35

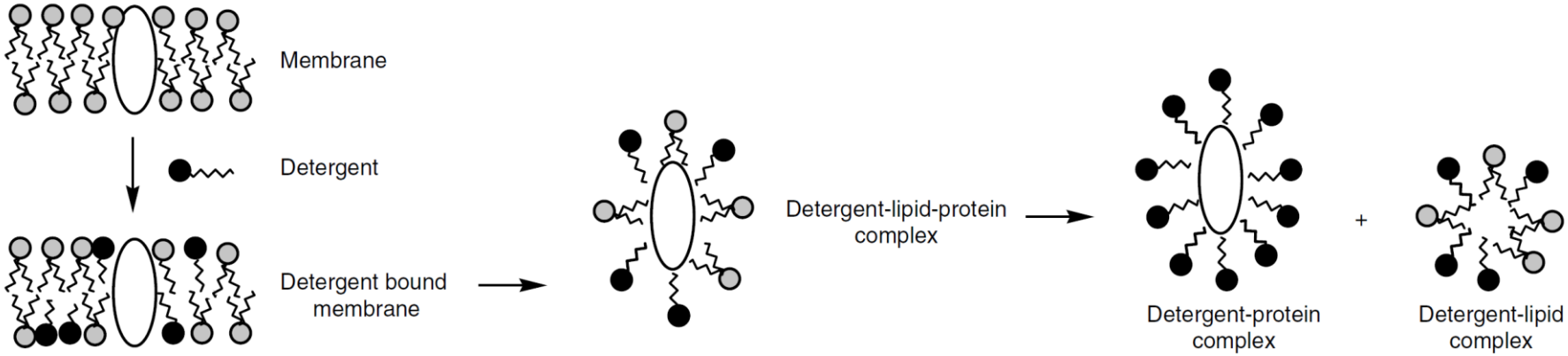


CHAPS



Zwittergent 3-14

❖Nonionic or zwitterionic detergents used for solubilization of membrane proteins.



❖For protein extraction, as mentioned before, mechanical and chemical lysis methods could be used.

❖Chemical lysis generally consist of alkaline, enzymatic or detergent application to the cells.

❖For Gram positive bacteria, lysozyme enzymes could be used.

❖Osmotic shock lysis with a hypotonic solution could be used for cells that do not protected by a cell wall.

❖Mechanic methods such as sonication, high pressure spraying, blade homogenizer and spherical mills also useful for the homogenisation of tissues.

❖However mechanic methods produce heat and therefore temperature of the solution must be controlled regularly.

❖Certain aggregates could be form during homogenization and for poorly soluble proteins caotropic agents such as urea and guanidine hydrochloride should be used.

➢ Solubilization

- ❖ Proteins should be precipitated after crude extract obtained from tissue.
- ❖ For this purpose salts such as ammonium sulfate and sodium chloride and organic solvents such as cold ethanol and acetone could be used.
- ❖ Ammonium sulfate and other salts should be removed from the solution using dialysis, ultrafiltration and gel exclusion chromatography.
- ❖ Organic solvent also removed from the solution before further purification.

- Determination of protein concentration.

❖After proteins solubilize protein concentration determined and according to the amount of proteins further steps of purification planned.

❖There are certain advantages and disadvantage of all protein determination techniques. Therefore most suitable one for the sample should be choose.

❖A_{280} (Simple Absorbance Method): Aromatic amino acids have maximum absorbance at 280 nm. (+) Fast, simple. (-) Low sensitivity (0.2-2 mg/mL), differences from protein to protein.

❖A_{205} : Peptide bonds have maximum absorption at 205 nm. (+) Higher sensitiviy. (-) many chemical used in purification interfere with this method.

❖Various colorimetric methods also developed for protein determination. They are much more sensitive than simple absorbance method.

❖This methods required using a standard protein (generally BSA).

❖Biuret: peptide bonds form complex with alkaline copper ions. Blue color read at 550 nm. (+) Protein to protein differences is low.
(-)Lower sensitivity (1-6 mg/mL) and interference with reducing agents.

❖Lowry: Similar to Biuret reaction however sensitivity is increased with the use of phosphomolibdotungstate. (+) High sensitivity (0.01-1mg/mL)
(-) interference with reducing agents, ammonium sulfate, chelating agents and detergents.

❖Bicinchoninic Acid (BCA) Similar to biuret reaction however interferences are very low due to the selectivity of the reagent. (+) Interference from other chemical is low and high sensitivity (0.5-10µg/mL). (-) Protein to protein differences and interference with phospholipids

❖Bradford: Basic and aromatic amino acids react with Coomassie Brilliant Blue G-250. (+) High sensitivity (5-40 µg/mL) and low interference with detergents. (-) Medium protein to protein differences.

❖Colloidal gold. (+) Highest sensitivity (2-20µg/mL) and low interference. (-) High protein to protein differences.

❖Also certain fluorometric methods could be used for protein determination and they are generally sensitive.

TABLE 2.1
Advantages and Disadvantages of Some Common Methods for Protein Determination

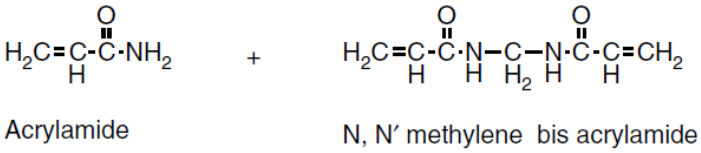
Method	Range of Sensitivity	Minimum Vol./Amount	Destructive	Time of Assay	Protein-to-Protein Variation	Comments
A_{280}	0.2–2 mg/ml	100 µl/20 µg	No	Instant	Large	Interference by UV absorbing materials
A_{205}	10–50 µg/ml	100 µl/1 µg	No	Instant	Low	Interference by UV absorbing materials
Biuret	1–5 mg/ml	2 ml/2 mg	Yes	1 h	Low	Rapid color; interference by caustic reagents, NH_4^+
Lowry	0.01–1 mg/ml	0.5 ml/5 µg	Yes	40 min	Moderate	Many interfering substances; slow color development
BCA	10–1200 µg/ml (standard) 0.5–10 µg/ml (micro)	0.5 ml/250 ng	Yes	40 min – 2 h	Moderate	Warming required, interference by reducing agents
Bradford	25–200 µg/ml (standard) 5–40 µg/ml (micro)	100 µl/0.5 µg	Yes	10 min	Moderate	Interference by detergents and strong bases
Colloidal gold	20–200 ng 1–10 ng (Modified assay)	10 µl/1 ng	Yes	30 min	Moderate	Assay pH should be acidic (around pH 3)
Ninhydrin	0.1–1 mg/ml	10 µl/1 µg	Yes	>24 h	Low	Protein must be hydrolyzed
Fluorescamine	0.5–50 µg/ml (standard) 50 ng–25 µg/ml (micro)	100 µl/5 ng	Yes	<10 min	Moderate	Not compatible with buffers containing primary amines
o-phthalaldehyde	10–500 µg/ml (standard) 0.05–2 µg/ml (micro)	50 µl/2.5 ng	Yes	<10 min	Moderate	Not compatible with buffers containing primary amines
CBQCA	0.12–1.8 µg/ml	80 µl/10 ng	Yes	90 min	Moderate	Buffers containing primary amines and reducing agents are not compatible
Nano-Orange^R	0.01–10 µg/ml	2 ml/20 ng	Yes	10 min at 95°C	Low	Not compatible with detergents

➢ Electrophoresis

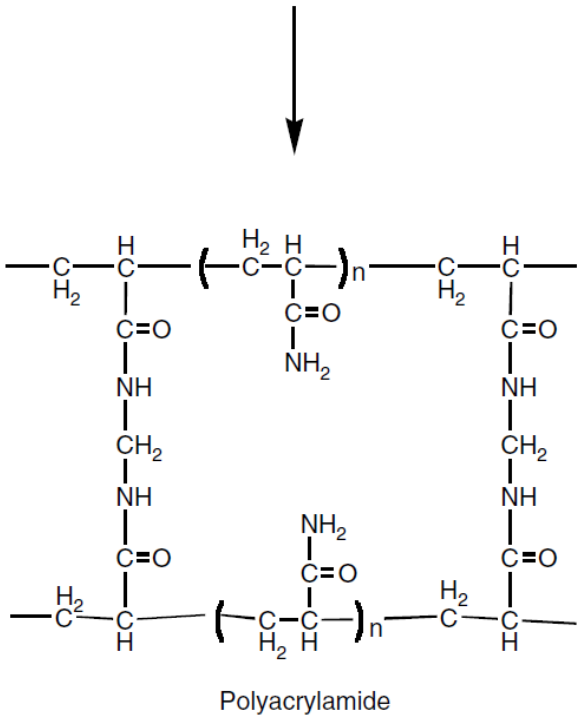
❖Proteins could be separated using electrophoresis.

❖Conventional electrophoresis uses a single electrical field to cause biomolecules to migrate through a matrix according to its mass-to-charge ratio

❖Most frequently used method is Poliacrylamide gel electrophoresis (PAGE).



Acrylamide + N, N′ methylene bis acrylamide



Polyacrylamide

TABLE 3.1
Various PAGE Systems and Their Applications

PAGE System	Application	Comment
SDS-PAGE (Laemmli)	Determination of subunit molecular weight Homogeneity test of a purified protein	Native protein activity is lost Not suitable for low molecular weight proteins/peptides (<10 kDa)
SDS-urea PAGE	Separation of membrane proteins Suitable for low molecular weight proteins	—
Non-denaturing PAGE	Homogeneity test of a purified protein	Native protein activity usually retained Not reliable for molecular weight estimation
Tricine PAGE	Separation of low molecular weight proteins/peptides (1 – 40 kDa range)	Protein band in the gel can be excised for amino acid sequencing without significant interference.
Non-urea SDS-PAGE (modified Laemmli)	Separation of low molecular weight proteins/peptides (as low as 5 kDa)	—
Acid-urea PAGE	Separation of basic proteins such as histones	Long run Proteins move toward cathode Electrode connection is opposite to the SDS-PAGE configuration
CTAB-PAGE	Determination of native molecular weight Native activity assay	Proteins move toward cathode Electrode connection is opposite to the SDS-PAGE configuration

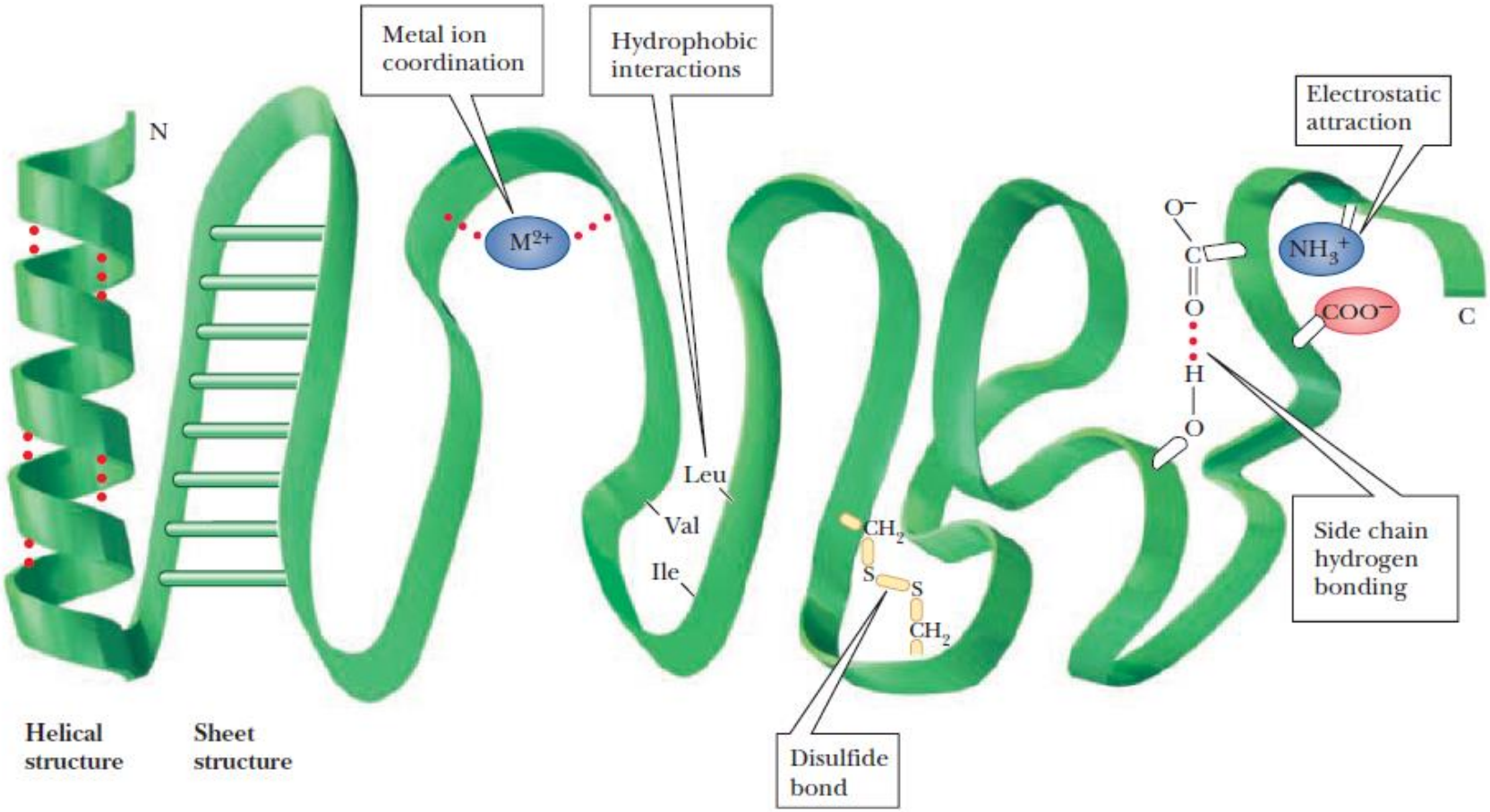
❖Under denaturating condition sodium dodesyl sulfate cover proteins with a negative charge and there fore proteins move only according to their size.

❖Two-dimensional gel electrophoresis (2-D electrophoresis) is a powerful and widely used method for the analysis of complex protein mixtures extracted from cells, tissues, or other biological samples.

❖This technique separate proteins in two steps, according to two independent properties: the first-dimension is isoelectric focusing (IEF), which separates proteins according to their isoelectric points (pI); the second-dimension is SDS-polyacrylamide gel electrophoresis (SDS-PAGE), which separates proteins according to their molecular weights (MW).

❖In this way, complex mixtures consisted of thousands of different proteins can be resolved and the relative amount of each protein can be determined.

Technique	Property Required	Remarks	When to Use (Recommended Application)
Membrane filtration	Molecular size	Fractionation as well as concentration. Loss of protein by non-specific adsorption.	At the beginning of a purification procedure. Particularly useful for concentrating large volumes of culture medium.
Centrifugation	Molecular size, shape, density	Commonly used for cellular fractionation	
Preparative isoelectric focusing	pI	Proteins precipitate in the rotofor chamber	
Preparative electrophoresis	Charge		
Size exclusion	Molecular size	Usually low resolution. Provide information about protein molecular weight.	At the end of a purification procedure
Ion-exchange	Charge	Protein binding capacity usually high	At the beginning of a purification procedure
Reversed phase	Hydrophobicity	Resolution varies according to gel size. Commonly used for peptide separation.	Used for separation of peptides, digested purified proteins, and other applications where loss of protein's biological activity is not a concern.
Hydrophobic	Hydrophobicity		After ammonium sulfate fractionation, but before ion-exchange chromatography
Hydroxyapatite	Charge		
Affinity	Binding ligand	Usually specific separation. Limited by availability of immobilized ligand. Expensive to scale up.	At the beginning of a purification procedure
Covalent	Thiol groups	Limited to thiol-containing proteins. Specific separation.	For thiol-containing proteins
Chromatofocusing	Charge, pI		Useful to separate isoforms of closely spaced pIs. Use after affinity chromatography.


Metal ion coordination
Hydrophobic interactions
Electrostatic attraction
N
M^{2+}
Leu
Val
Ile
CH_2
S
S
CH_2
O^-
C
O
H
O
NH_3^+
COO^-
C
Side chain hydrogen bonding
Helical structure
Sheet structure
Disulfide bond