جامعة الملك سعود
King Saud University
Department of Chemistry
College of Science

Biopolymers

Chem. 563

Dr. Mohamed El-Newehy
http://fac.ksu.edu.sa/melnewehy

PROTEINS

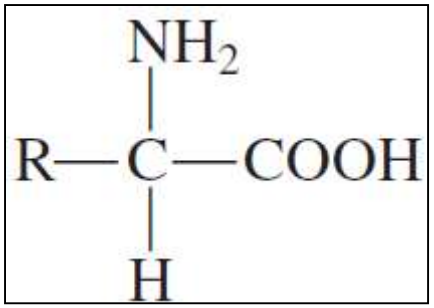
- ***Proteins*** are naturally occurring polymers composed of amino acid units joined one to another by amide (or peptide) bonds.
- This word is derived from the Greek porteios, "of chief importance."

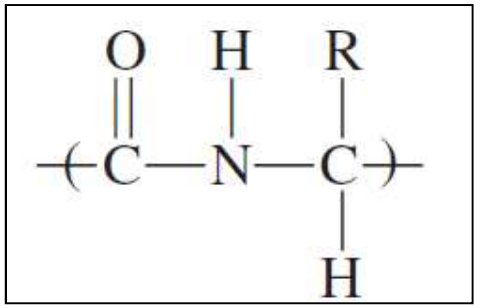
  Examples; *Spider webs, animal hair and muscle, egg whites, and hemoglobin.*

- ***Peptides*** are oligomers of amino acids that play important roles in many biological processes (*oligomers or relatively low molecular weight proteins*)

  Example*, the peptide hormone insulin controls our blood sugar levels, bradykinin controls our blood pressure, and oxytocin regulates uterine contraction and lactation.*

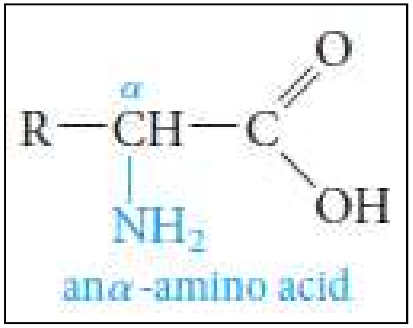
- Thus, ***proteins***, ***peptides***, and ***amino acids*** are essential to the structure, function, and reproduction of living matter.
- All ***α-amino acids*** found in proteins except glycine contain a chiral carbon atom and are L-amino acids





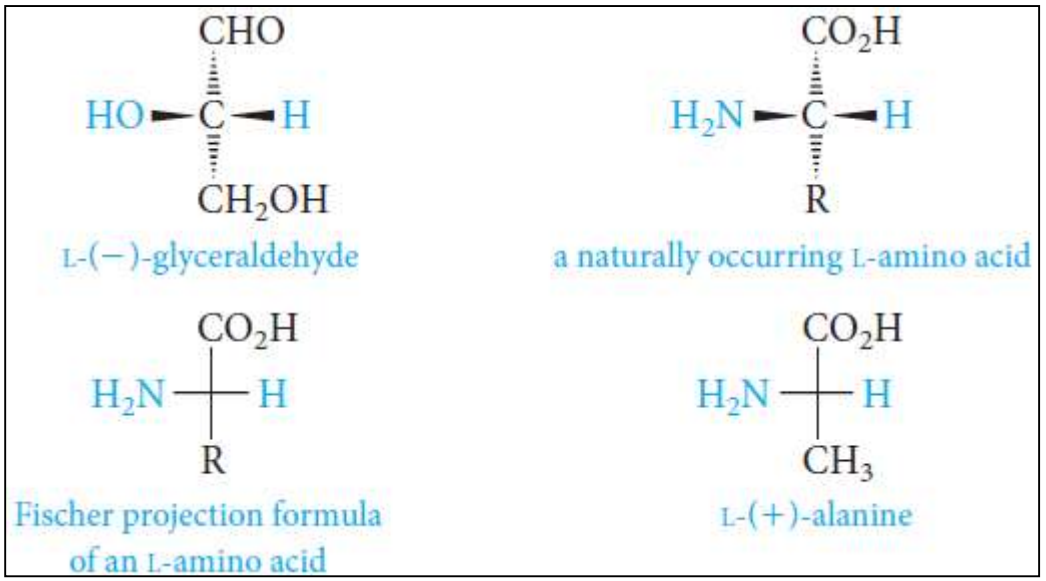
Naturally Occurring Amino Acids

- The ***amino acids*** obtained from protein hydrolysis are α-amino acids.



an α-amino acid

- All except glycine are therefore optically active.
- They have the L configuration relative to glyceraldehyde.



L-(−)-glyceraldehyde

a naturally occurring L-amino acid

Fischer projection formula of an L-amino acid

L-(+)-alanine

Naturally Occurring Amino Acids

Table 17.1 Names and Formulas of the Common Amino Acids

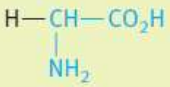
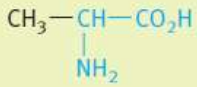
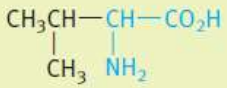
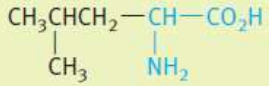
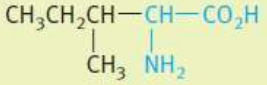
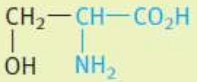
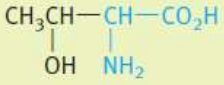
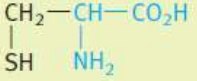
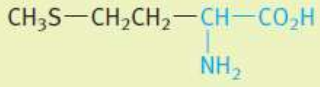
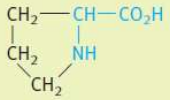
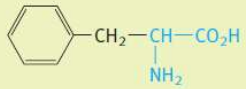
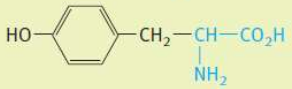
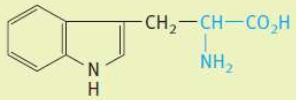
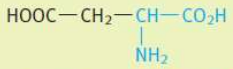
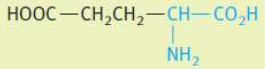
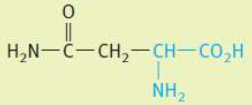
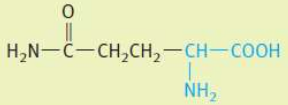
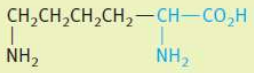
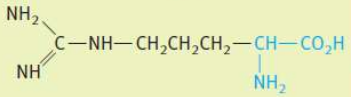
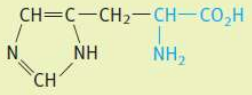
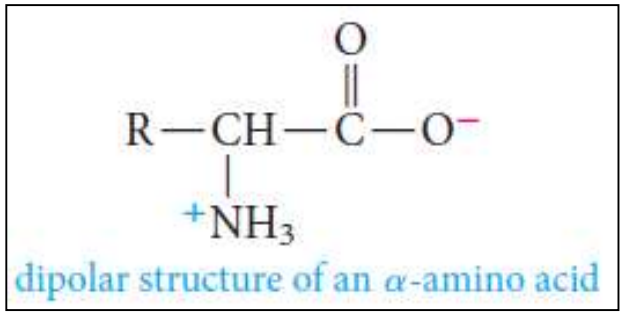
Name	Three-letter abbreviation (isoelectric point) one-letter abbreviation	Formula	R
A. One amino group and one carboxyl group			
1. glycine	Gly (6.0) G		
2. alanine	Ala (6.0) A		
3. valine	Val (6.0) V		R is hydrogen or an alkyl group.
4. leucine	Leu (6.0) L		
5. isoleucine	Ile (6.0) I		
6. serine	Ser (5.7) S		
7. threonine	Thr (5.6) T		R contains an alcohol function.
8. cysteine	Cys (5.0) C		
9. methionine	Met (5.7) M		R contains sulfur.

Table 17.1 (continued)

Name	Three-letter abbreviation (isoelectric point) one-letter abbreviation	Formula	R
10. proline	Pro (6.3) P		The amino group is secondary and part of a ring.
11. phenylalanine	Phe (5.5) F		One hydrogen in alanine is replaced by an aromatic or heteroaromatic (indole) ring.
12. tyrosine	Tyr (5.7) Y		
13. tryptophan	Trp (5.9) W		
B. One amino group and two carboxyl groups			
14. aspartic acid	Asp (3.0) D		
15. glutamic acid	Glu (3.2) E		
16. asparagine	Asn (5.4) N		
17. glutamine	Gln (5.7) Q		
C. One carboxyl group and two basic groups			
18. lysine	Lys (9.7) K		The second basic group is a primary amine, a guanidine, or an imidazole.
19. arginine	Arg (10.8) R		
20. histidine	His (7.6) H		

The Acid–Base Properties of Amino Acids

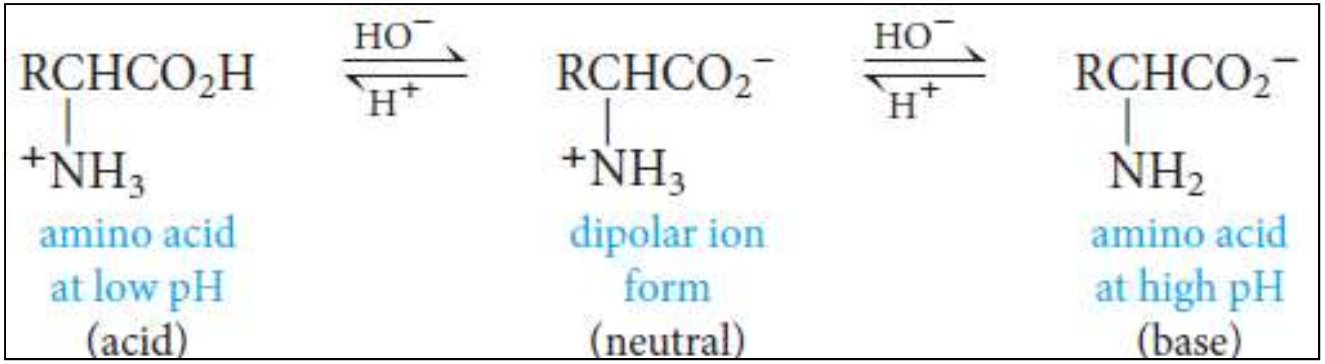
- The carboxylic acid and amine functional groups are *simultaneously* present in amino acids.
- *Amino acids* with one amino group and one carboxyl group are better represented by a ***dipolar ion structure***.



dipolar structure of an α-amino acid

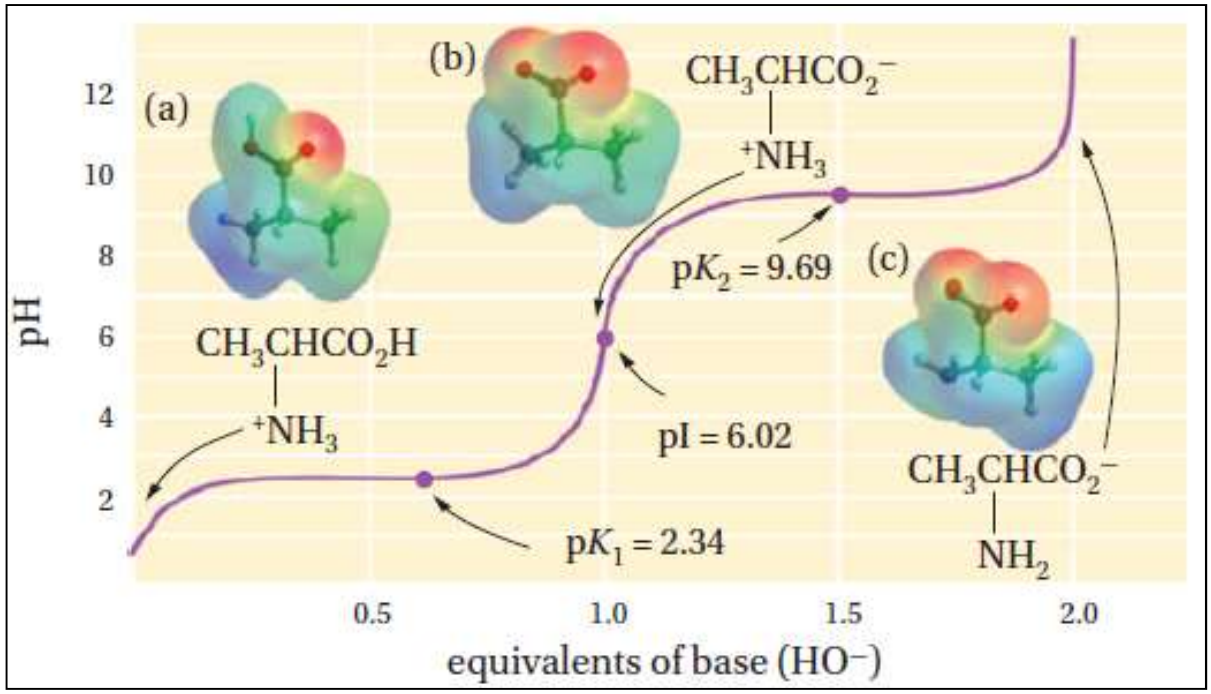
- **Amino acids are *amphoteric*** .

  They can behave as acids and donate a proton to a strong base, or they can behave as bases and accept a proton from a strong acid.



The Acid–Base Properties of Amino Acids

- **Titration curve for alanine, a typical amino acid of this kind.**
  - **At low pH (acidic solution), the amino acid is in the form of a substituted ammonium ion.**
  - **At high pH (basic solution), it is present as a substituted carboxylate ion.**
  - **At some intermediate pH (for alanine, pH 6.02), the amino acid is present as the dipolar ion with an ammonium ($-NH_3^+$) and a carboxylate ($-CO_2^-$) unit.**



Isoelectric point (pI),

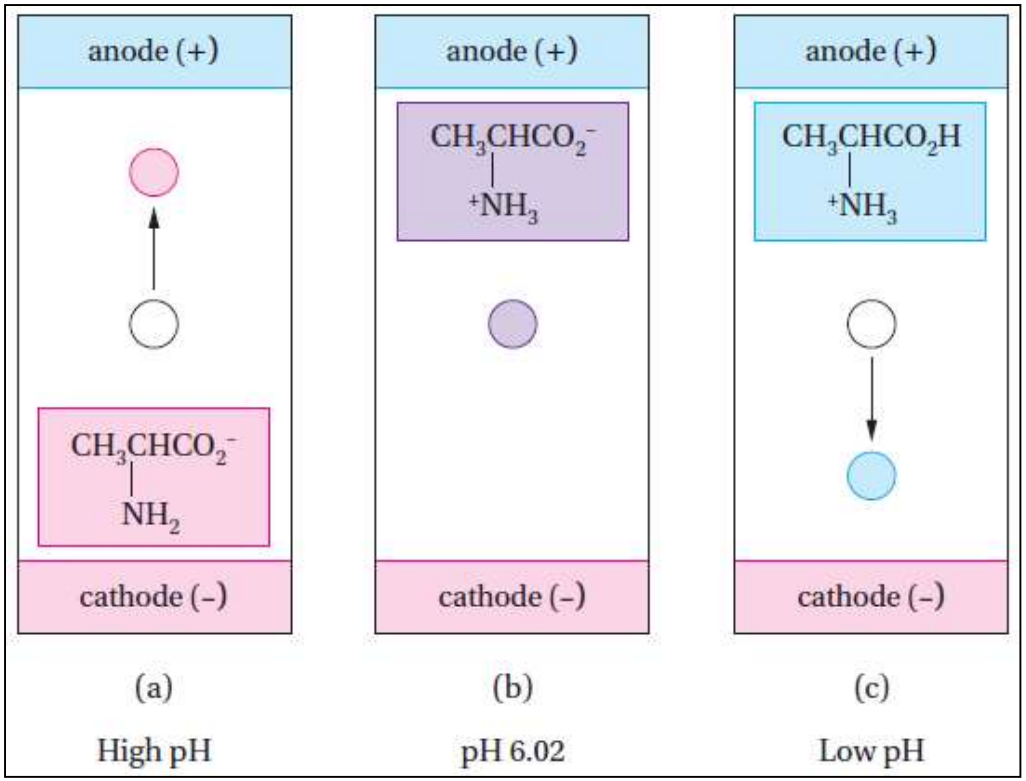
At some intermediate pH, called the isoelectric point (pI), the amino acid will be dipolar and have a net charge of zero. It will be unable to move toward either electrode.

Electrophoresis

- ***Electrophoresis*** is an important method for separating amino acids that takes advantage of these charge differences.

In a typical electrophoresis experiment,

- A mixture of amino acids is placed on a solid support (e.g., paper), and the support is bathed in an aqueous solution at a controlled pH.
- An electrical field is then applied across the paper.



The migration of an amino acid (such as alanine) in an electric field depends on pH:

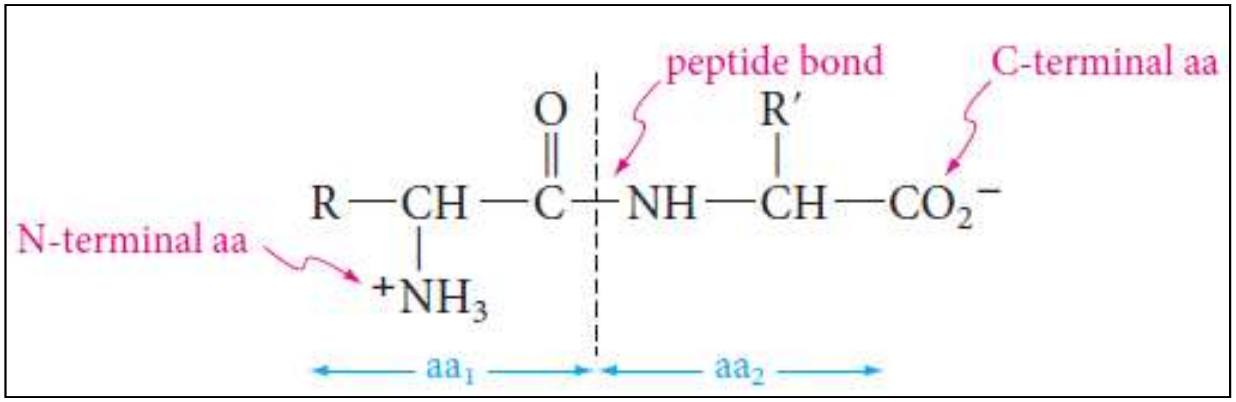
(a) At high pH, the alanine is negatively charged and migrates toward the positive anode.

(b) At the isoelectric point (pH 6.02), the alanine is neutral and does not migrate.

(c) At low pH, the alanine is positively charged and migrates toward the negatively charged cathode.

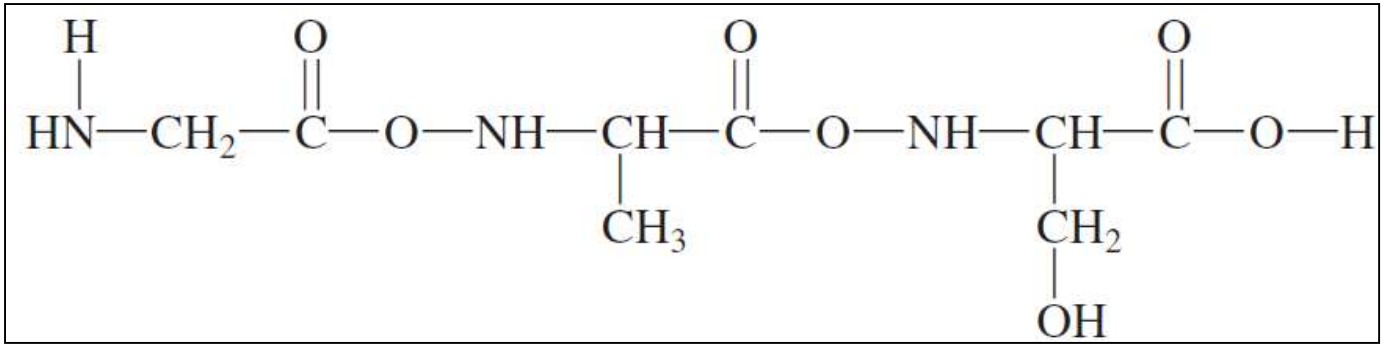
Polypeptides

- ***Amino acids*** are linked in peptides and proteins by an amide bond between the *carboxyl group of one amino acid* and the *α-amino group of another amino acid*.
- Emil Fischer, who first proposed this structure, called this *amide bond* (peptide bond).
- A molecule containing only two amino acids (the shorthand aa is used for amino acid) joined in this way is a dipeptide:



Polypeptides

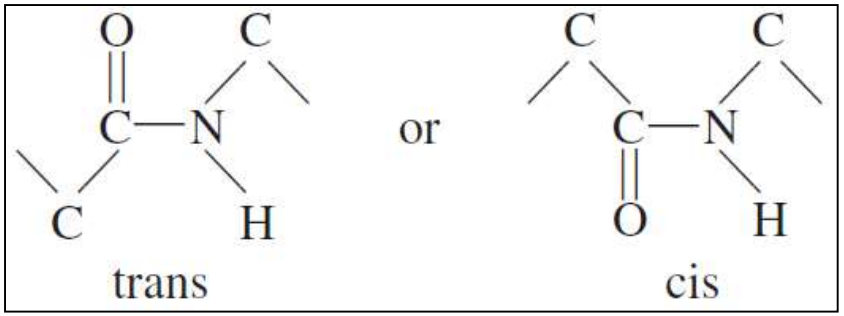
- **In writing out sequences for polypeptides** it is usual to use a three-letter abbreviation or a one-letter abbreviation starting with the N terminus to the left and going to the C-O terminus to the right.
- Thus the trimer becomes Gly-Ala-Ser or GlyAlaSer or EGAS where the E signals the N terminus.



- *If there are a number of acidic and/or basic groups on the polypeptide* the molecule is said to be a **polyampholyte**
- *If there are only positive or negative charges* it is called a **polyelectrolyte**.

Polypeptides

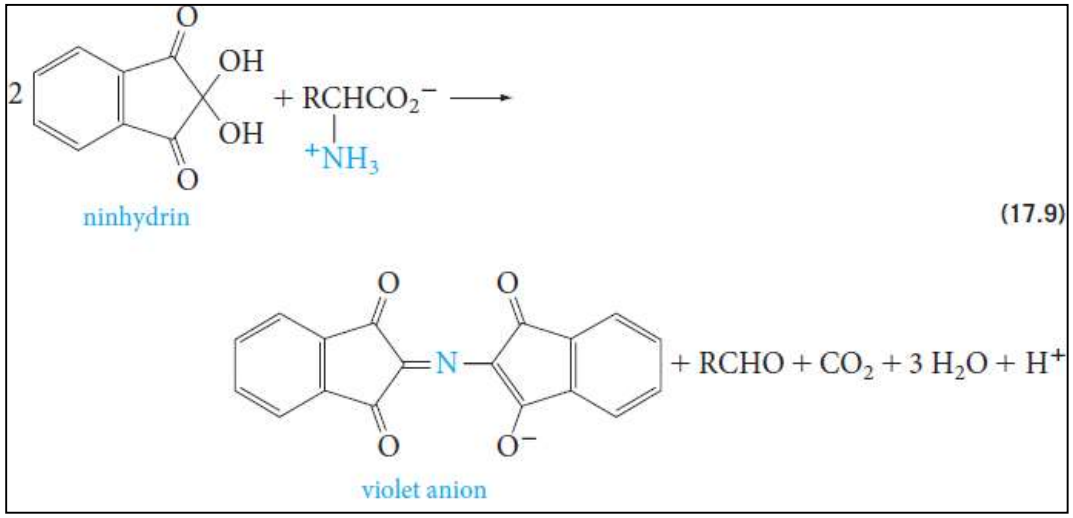
- Even though the atoms within a peptide bond are coplanar, they can exist in two possible configurations; *cis* and *trans*.



- The ***trans* form is usually favored** whenever there is a bulky group on the adjacent carbon(s) because the groups will interfere more in the *cis* structure.
- ***Proteins*** may be hydrolyzed by dilute acids, and the mixture of amino acids or residues produced may be separated and identified by paper chromatography.
- The reagent ninhydrin yields characteristic colored products with amino acids.

Polypeptides

- NOTE: **The Ninhydrin Reaction**
  - Ninhydrin is a useful reagent for detecting amino acids and determining the concentrations of their solutions.
  - It is the hydrate of a cyclic triketone, and when it reacts with an amino acid, a violet dye is produced.



- Only the nitrogen atom of the violet dye comes from the amino acid; the rest of the amino acid is converted to an aldehyde and carbon dioxide.
- Therefore, the same violet dye is produced from all α-amino acids with a primary amino group, and the intensity of its color is directly proportional to the concentration of the amino acid present.
- Only proline, which has a secondary amino group, reacts differently to give a yellow dye, but this, too, can be used for analysis.

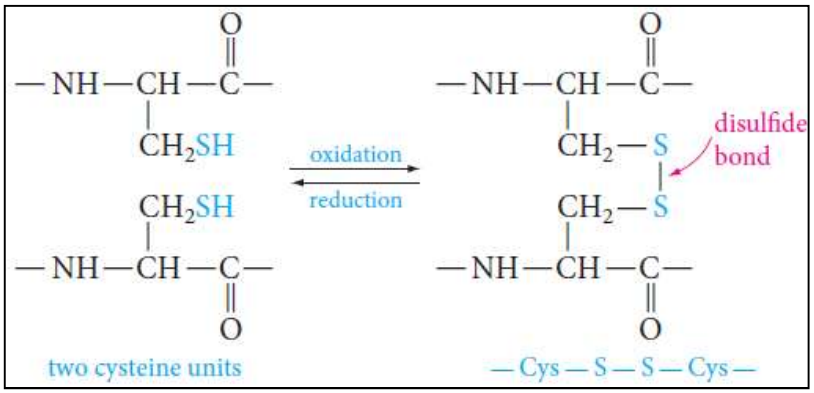
Polypeptides

- While humans synthesize about a dozen of the 20 amino acids needed for good health,
- The other 8 are obtained from outside our bodies, generally from eating foods that supply these ***essential amino acids*** ***(****phenylalanine, valine, threonine, tryptophan, methionine, leucine, isoleucine, lysine, and histidine****)***.
- Different foods are good sources of different amino acids.

Polypeptides

The Disulfide Bond

- Aside from the peptide bond, the only other type of covalent bond between amino acids in peptides and proteins is the **disulfide bond**.
- Two cysteine units can be linked by a disulfide bond.



- If the two cysteine units are in different parts of the *same* chain of a peptide or protein, a disulfide bond between them will form a “loop,” or large ring.
- If the two units are on different chains, the disulfide bond will cross-link the two chains.
- Disulfide bonds can easily be broken by mild reducing Agents.

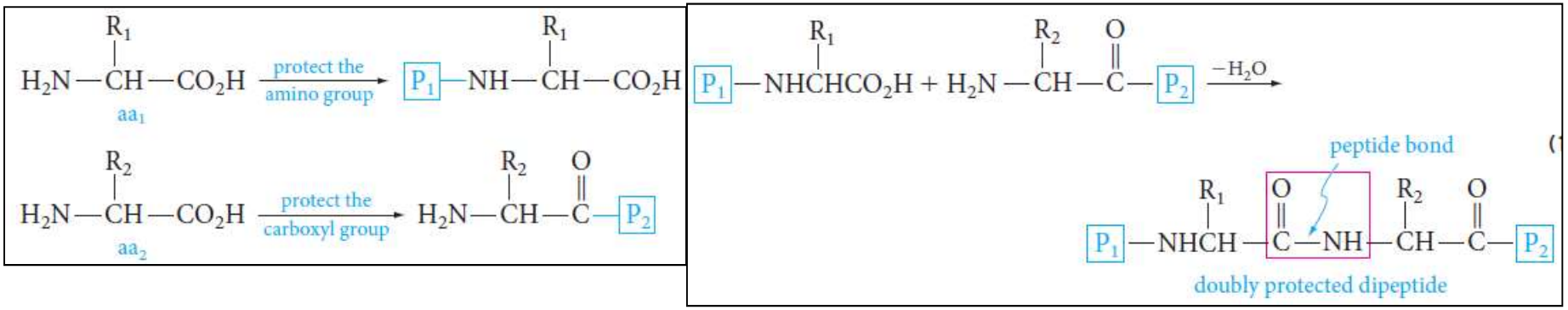
Peptide Synthesis

- Many methods have been developed to link amino acids in a controlled manner and they require careful strategy.
- Amino acids are bifunctional.

  To link the carboxyl group of one amino acid to the amino group of a second amino acid, we must first prepare each compound by **protecting the amino group of the first and the carboxyl group of the second**.

- **After the peptide bond is formed,** we must be able to *remove the protecting groups under conditions that do not hydrolyze the peptide bond.*

These methods were used by Vincent du Vigneaud and his colleagues



Peptide Synthesis

Solid-Phase Technique

- In 1965, solid phase technique developed by Bruce Merrifield.
- In this way, excess reagents and by-products can be removed simply by washing and filtering the solid.
- The growing peptide chain does not need to be purified at any intermediate stage.
- When the peptide is fully constructed, it is cleaved chemically from the solid support.
- ***Typically,*** the *solid phase is a cross-linked polystyrene* in which some (usually 1% to 10%) of the aromatic rings contain chloromethyl ($ClCH_2$-) groups.

Peptide Synthesis

Solid-Phase Technique

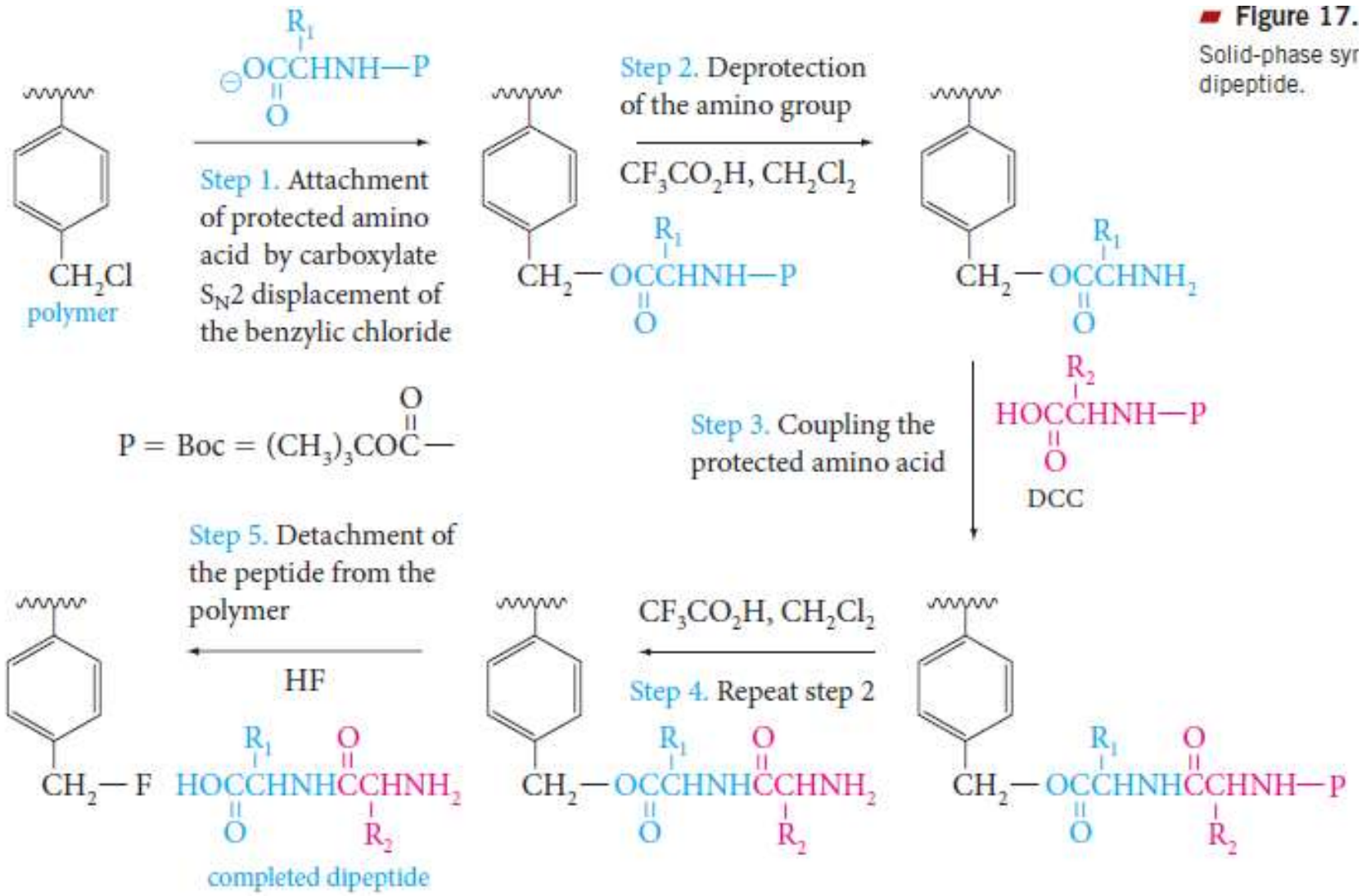


Figure 17.
Solid-phase syr
dipeptide.

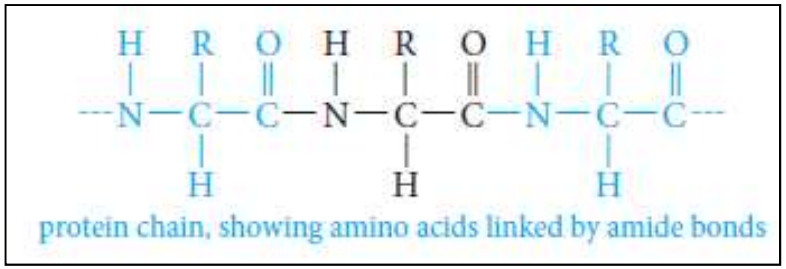
Peptide Synthesis

Solid-Phase Technique

- The process begins with the attachment of the C-terminal amino acid to the chloromethylated polymer.
- Nucleophilic substitution by the carboxylate anion of an N-blocked protected C-terminal amino acid displaces chloride from the chloromethyl group forming an ester, protecting the C site while attaching it to the solid support.
- The blocking group is removed by addition of acid and the polymer containing the nonprotected N terminus is washed to remove unwanted byproducts.
- A peptide bond is formed by condensation to an N-blocked protected amino acid.
- Again, the solid phase system is washed to remove byproducts.
- The block group is removed by acid treatment and the site is ready for attachment by another amino acid.
- This cycle is repeated eventually producing the desired polypeptide without isolation of intermediate products.

Proteins

- ***Proteins*** are biopolymers composed of many amino acids connected to one another through amide (peptide) bonds.
- The backbone of proteins is a repeating sequence of one nitrogen and two carbon atoms.
- They play numerous roles in biological systems.
- Some proteins are major components of structural tissue (muscle, skin, nails, and hair).
- Others transport molecules from one part of a living system to another.
- The shapes of proteins as macromolecules; the *primary structure*, and three-dimensional aspects of peptide and protein structure, usually referred to as their *secondary, tertiary,* and *quaternary structures*.



protein chain, showing amino acids linked by amide bonds

Proteins

The Primary Structure of Proteins

- The term ***primary structure*** is used to describe the sequence of amino acid units (configuration) in a polypeptide chain i.e.
  1) Which amino acids are present and how many of each there are
  2) The sequence of the amino acids in the chain.
- The **sequence for N-terminal amino acids** in a chain may be determined by use of a technique developed by Nobel Laureate Sanger,
  The amino end group reacted with 2,4-dinitrofluorobenzene and characterized the yellow aromatic amino acid produced by hydrolysis.
- This process is repeated after the end amino acid has been hydrolyzed off.
- The **C-terminal amino acids** may be determined by using hydrazine to form hydrazides from the cleaved amino groups.
- Since the free carboxyl end group is not affected by hydrazine, the terminal amino acid is readily identified.

Proteins

The Primary Structure of Proteins

A) Amino Acid Analysis

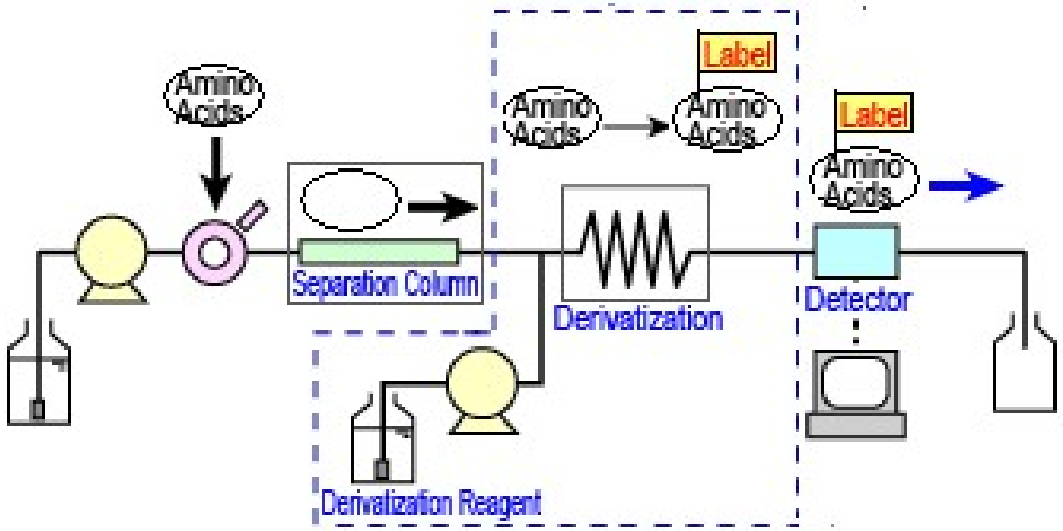
- Since peptides and proteins consist of amino acids held together by amide bonds, they can be hydrolyzed to their amino acid components.
- This hydrolysis is typically accomplished by heating the peptide or protein with 6 M HCl at 110°C for 24 h.
- Analysis of the resulting amino acid mixture requires a procedure that separates the amino acids from one another, identifies each amino acid present, and determines its amount.

Proteins

The Primary Structure of Proteins

A) Amino Acid Analysis

- An instrument called an *amino acid analyzer* performs these tasks automatically in the following way.



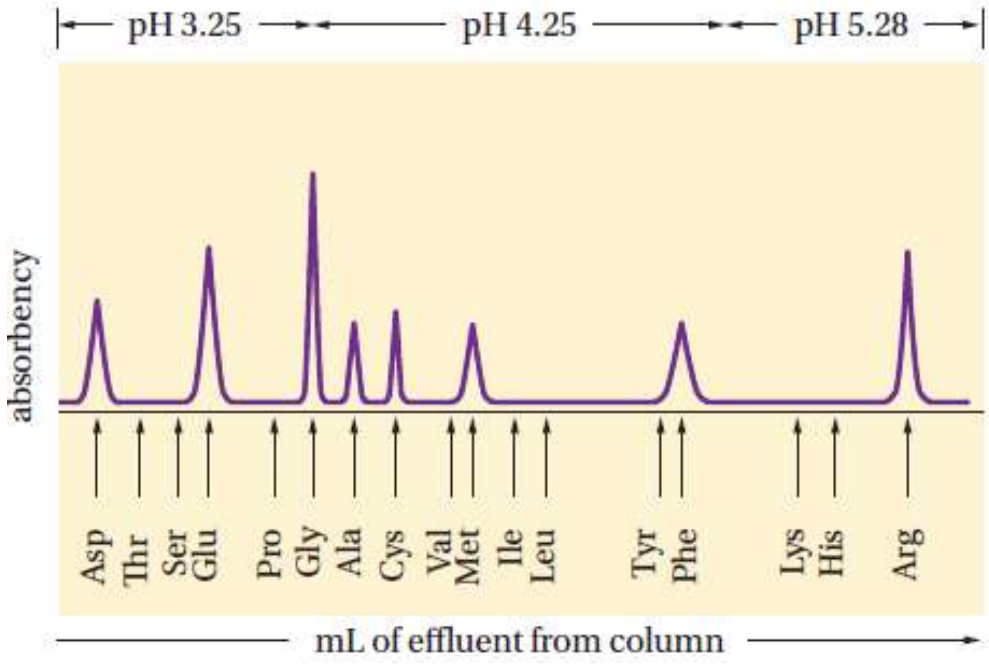
- The amino acid mixture from the complete hydrolysis of a few milligrams of the peptide or protein is placed at the top of a column packed with material that selectively absorbs amino acids (*an insoluble resin that contains strongly acidic groups*).
- These groups protonate the amino acids.
- Next, a buffer solution of known pH is pumped through the column.
- The amino acids pass through the column at different rates, depending on their structure and basicity, and are thus separated.

Proteins

The Primary Structure of Proteins

A) Amino Acid Analysis

- The column effluent is met by a stream of ninhydrin reagent.
- Therefore, the effluent is alternately violet or colorless, depending on whether or not an amino acid is being eluted from the column.
- The intensity of the color is automatically recorded as a function of the volume of effluent. Calibration with known amino acid mixtures allows each amino acid to be identified by the appearance time of its peak.
- Furthermore, the intensity of each peak gives a quantitative measure of the amount of each amino acid that is present.



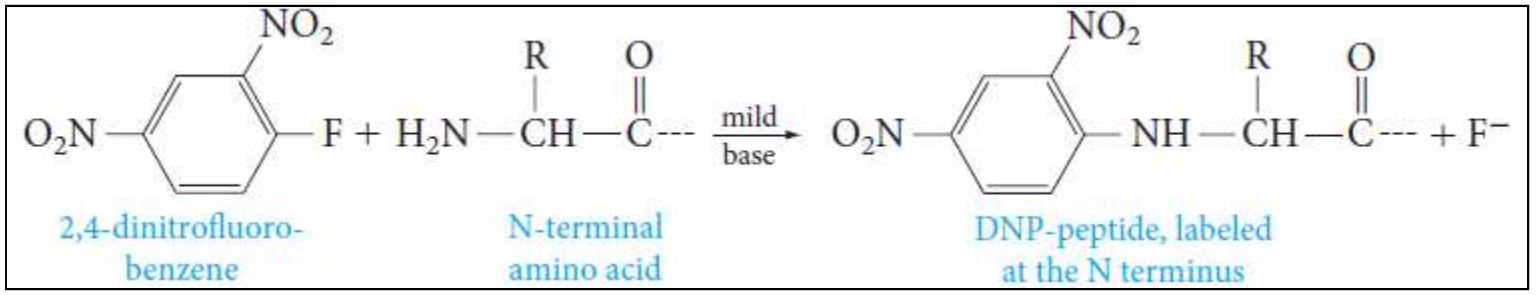
- This sample contains eight amino acids: aspartic acid, glutamic acid, glycine, alanine, cysteine, methionine, phenylalanine, and arginine.

Proteins

The Primary Structure of Proteins

B) Sequence Determination

- *Frederick Sanger devised a method for sequencing peptides based on the observation that the N-terminal amino acid differs from all others in the chain by having a free amino group.*
  - If that amino group were to react with some reagent prior to hydrolysis, then after hydrolysis, that amino acid would be labeled and could be identified.
  - Sanger's reagent is 2,4-dinitrofluorobenzene, which reacts with the NH_2 group of amino acids and peptides to give yellow 2,4-dinitrophenyl (DNP) derivatives.



  - Hydrolysis of a peptide treated this way would give the DNP derivative of the N-terminal amino acid; other amino acids in the chain would be unlabeled.
  - In this way, the N-terminal amino acid could be identified.

Proteins

The Primary Structure of Proteins

B) Sequence Determination

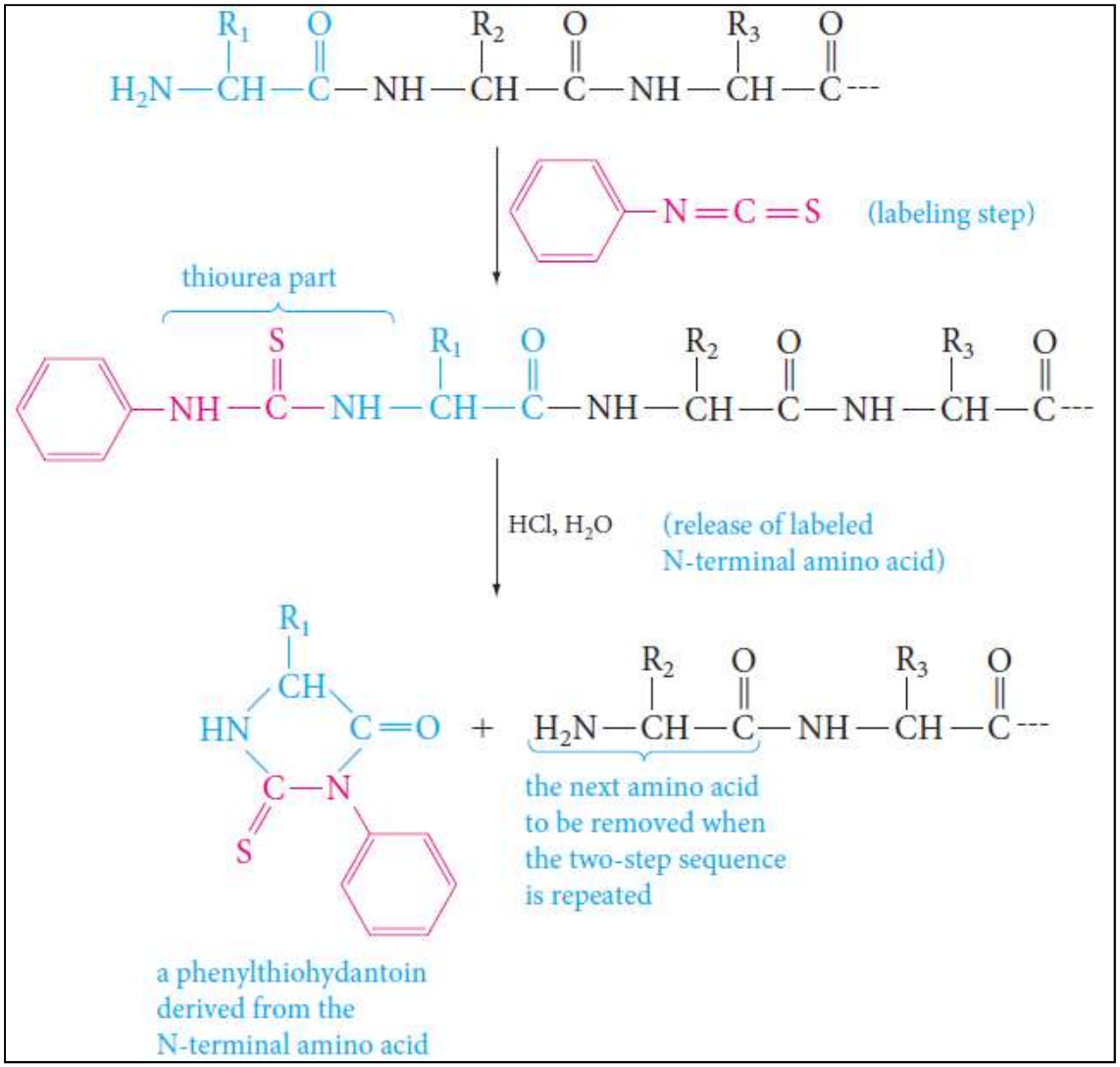
- *Pehr Edman*, *an ideal method for sequencing a peptide or protein would have a reagent that clips off just one amino acid at a time from the end of the chain, and identifies it.*

- **Edman's reagent** is phenyl isothiocyanate, $C_6H_5N = C = S$.
- The steps in selectively labeling and releasing the N-terminal amino acid are:
  - **In the first step**, the N-terminal amino acid acts as a nucleophile toward the C = S bond of the reagent to form a thiourea derivative.
  - **In the second step**, the N-terminal amino acid is removed in the form of a heterocyclic compound, a phenylthiohydantoin.
- The specific phenylthiohydantoin that is formed can be identified by comparison with reference compounds separately prepared from the known amino acids.
- Then the two steps are repeated to identify the next amino acid, and so on.
- The method has been automated, so currently amino acid "sequenators" can easily determine, in a day, the sequence of the first 50 or so amino acids in a peptide, starting at the N-terminal end.
- But the Edman method cannot be used indefinitely, due to the gradual buildup of impurities.

Proteins

The Primary Structure of Proteins

B) Sequence Determination



The Edman degradation of peptides.

Proteins

The Primary Structure of Proteins

C) Cleavage of Selected Peptide Bonds

- If a protein contains several hundred amino acid units, it is best to first partially hydrolyze the chain to smaller fragments that can be separated and subsequently sequenced by the Edman method.
- Certain chemicals or enzymes are used to cleave proteins at *particular* peptide bonds.
- For example, the enzyme *trypsin* (an intestinal digestive enzyme) specifically hydrolyzes polypeptides only at the carboxy end of arginine and lysine.

Table 17.3 Reagents for Specific Cleavage of Polypeptides

Reagent	Cleavage site
trypsin	carboxyl side of Lys, Arg
chymotrypsin	carboxyl side of Phe, Tyr, Trp
cyanogen bromide (CNBr)	carboxyl side of Met
carboxypeptidase	the C-terminal amino acid

Proteins

The Secondary Structure of Proteins

- The term secondary structure is used to describe the molecular shape or conformation of a molecule.
- The most important factor in determining the secondary structure of materials is
  - its precise or primary structure (For proteins, it is then the *amino acid sequence*).
  - Hydrogen bonding is also an important factor in determining the secondary structures of natural materials and those synthetic materials that can hydrogen-bond.
- In fact, for proteins, secondary structures are generally those that allow a maximum amount of hydrogen bonding.
- This hydrogen bonding also acts to stabilize the secondary structure while crosslinking acts to lock in a structure.

Proteins

The Secondary Structure of Proteins

- Because proteins consist of long chains of amino acids strung together, one might think that their shapes are rather amorphous, or “floppy” and ill-defined. NO
- Many proteins have been isolated in pure crystalline form and are polymers with well-defined shapes.
- Indeed, even in solution, the shapes seem to be quite regular.

A) Geometry of the Peptide Bond

- Simple amides have a planar geometry, that the amide C-N bond is shorter than usual, and that rotation around that bond is restricted.
- Bond planarity and restricted rotation, which are consequences of resonance,
- are also important in peptide bonds. The rather rigid geometry and restricted rotation of the peptide bond help to impart a definite shape to proteins.

Proteins

The Secondary Structure of Proteins

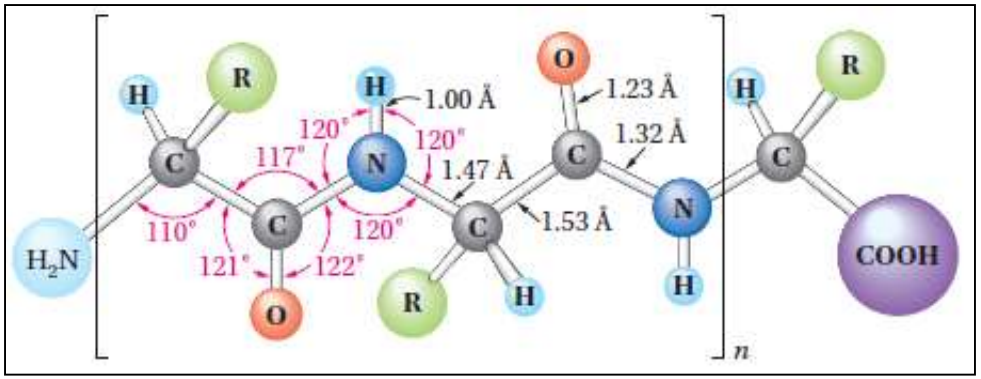
A) Geometry of the Peptide Bond

- The characteristic dimensions, which are common to all peptides and proteins:

(1) The amide group is flat; the carbonyl carbon, the nitrogen, and the four atoms connected to them all lie in a single plane.

(2) The short amide C-N distance (1.32 Å, compared with 1.47 Å for the other C-N bond) and the 120° bond angles around that nitrogen show that it is essentially sp2-hybridized and that the bond between it and the carbonyl carbon is like a double bond.

(3) Although each amide group is planar, two adjacent amide groups need not be coplanar because of rotation about the other single bonds in the chain; that is, rotation can occur around the two single bonds to the –CHR- group.

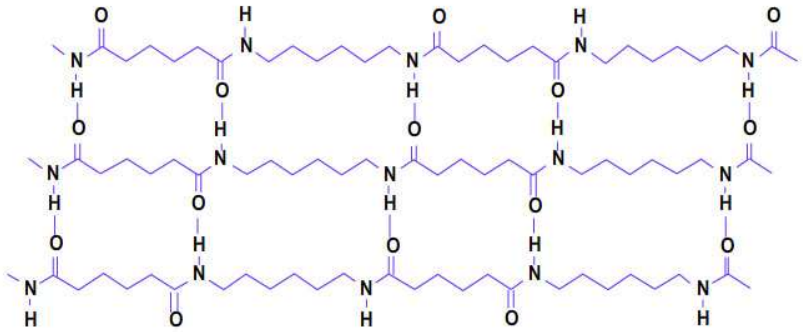


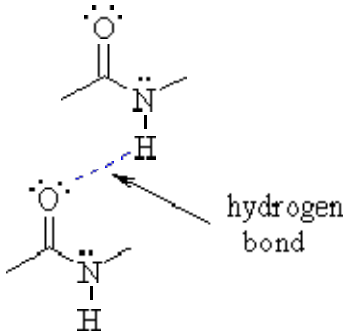
Proteins

The Secondary Structure of Proteins

B) Hydrogen Bonding

- Amides readily form intermolecular hydrogen bonds between the carbonyl group and the N-H group, bonds of the type C = O · · · H-N.
- The chain may coil in such a way that the N-H of one peptide bond can hydrogen-bond with a carbonyl group of another peptide bond farther down the same chain, thus rigidifying the coiled structure.
- Although a single hydrogen bond is relatively weak (perhaps only 5 kcal/mol of energy), the possibility of forming multiple intrachain or inter-chain hydrogen bonds makes this a very important factor in protein structure.



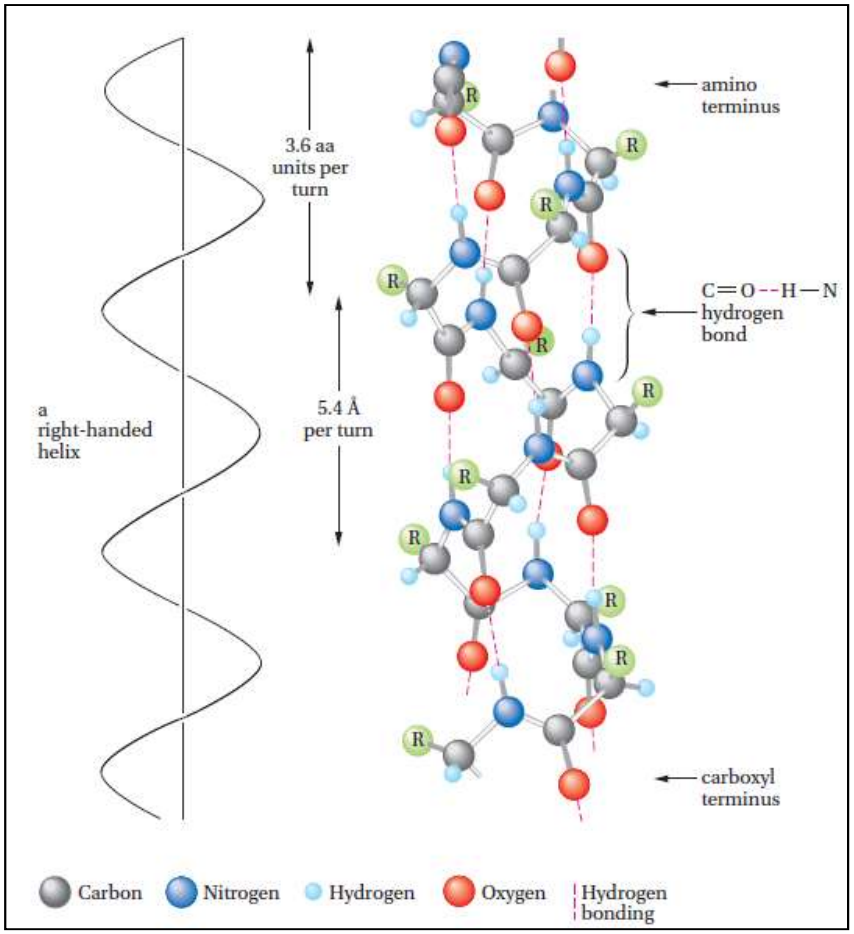


Proteins

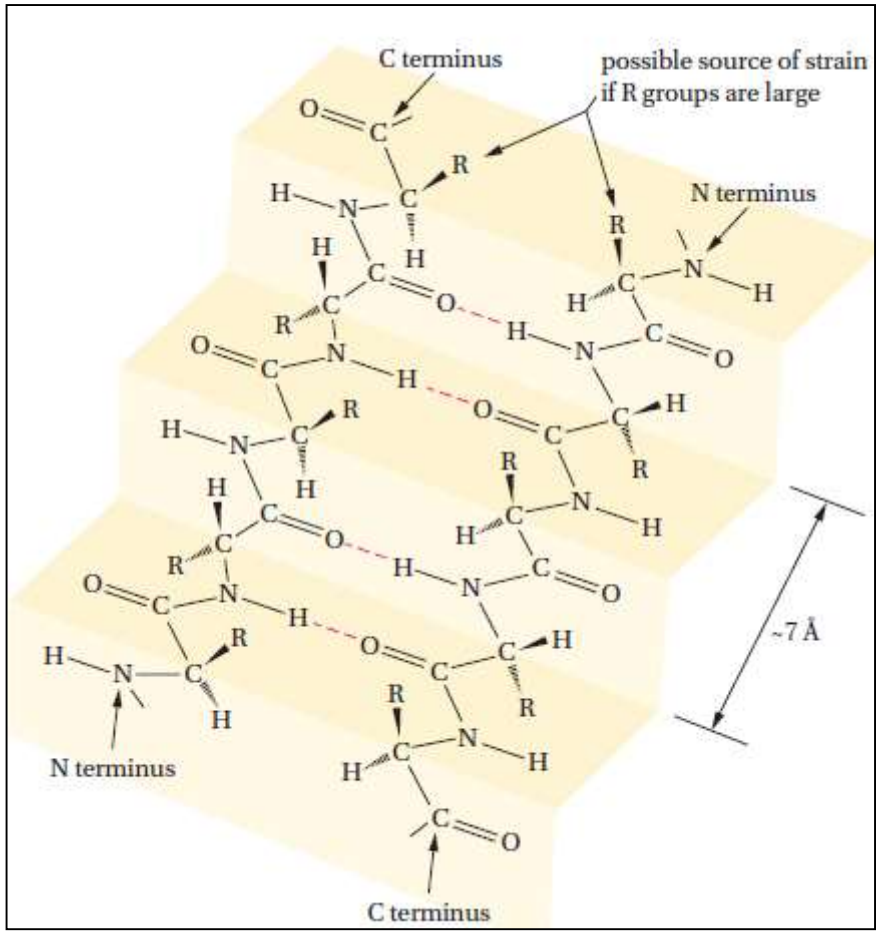
The Secondary Structure of Proteins

C) The α Helix and the Pleated Sheet

- In nature, the two most common secondary structures are helical and sheets.



Segment of an α helix, showing three turns of the helix, with 3.6 amino acid units per turn. Hydrogen bonds are shown as dashed colored lines.



A segment of the pleated-sheet structure of β-keratin. Adjacent chains run in opposite directions and are held together by hydrogen bonds (shown in color). R groups project above or below the mean plane of the sheet.

Proteins

The Secondary Structure of Proteins

C) The α Helix and the Pleated Sheet

keratin

- **X-ray crystallographic studies of α-keratin**, a structural protein present in hair, wool, horns, and nails, showed that some feature of the structure repeats itself every 5.4 Å.
- *Linus Pauling* was able to suggest a structure that explains this and other features of the x-ray studies.
  - Pauling proposed that the polypeptide chain coils about itself in a spiral manner to form a helix, held rigid by intra-chain hydrogen bonds.
  - The a helix, as it is called, is right-handed and has a pitch of 5.4 Å, or 3.6 amino acid units.

Proteins

The Secondary Structure of Proteins

C) The α Helix and the Pleated Sheet

- Note several features of the α helix.
  - Proceeding from the N terminus (at the top of the structure as drawn in the figure), each carbonyl group points ahead or down toward the C terminus and is hydrogen-bonded to an N-H bond farther down the chain.
  - The N-H bonds all point back toward the N terminus.
  - All of the hydrogen bonds are roughly aligned with the long axis of the helix.
  - The very large number of hydrogen bonds (one for each amino acid unit) strengthens the helical structure.
  - The R groups of the individual amino acid units are directed outward and do not disrupt the central core of the helix.
- It turns out that the α helix is a natural pattern into which many proteins or segments of proteins fold.

Proteins

The Secondary Structure of Proteins

C) The α Helix and the Pleated Sheet

- Helices are often described in terms of a repeat distance.
- Helices generally do not have an integral number of repeat units or residues per turn.
- The α helix repeats after 18 amino acid residues taking five turns to repeat. Thus, the number of residues per turn is 18/5 = 3.6 residues/turn.

Proteins

The Secondary Structure of Proteins

C) The α Helix and the Pleated Sheet

- **The structural protein β-keratin**, obtained from silk fibroin, shows a different repeating pattern (7 Å) in its x-ray crystal structure.
- Pauling suggested a pleated-sheet arrangement of the peptide chain.
  - In the pleated sheet, peptide chains lie side by side and are held together by inter-chain hydrogen bonds.
  - Adjacent chains run in opposite directions.
  - The repeating unit in each chain, which is stretched out compared with the a helix, is about 7 Å.
  - In the pleated-sheet structure, the R groups of amino acid units in any one chain alternate above and below the mean plane of the sheet.
  - If the R groups are large, there will be appreciable steric repulsion between them on adjacent chains.

Proteins

The Secondary Structure of Proteins

C) The α Helix and the Pleated Sheet

- For this reason, the pleated-sheet structure is important only in proteins that have a high percentage of amino acid units with small R groups.
- In the β-keratin of silk fibroin, for example, 36% of the amino acid units are glycine (R = H) and another 22% are alanine (R = CH_3).

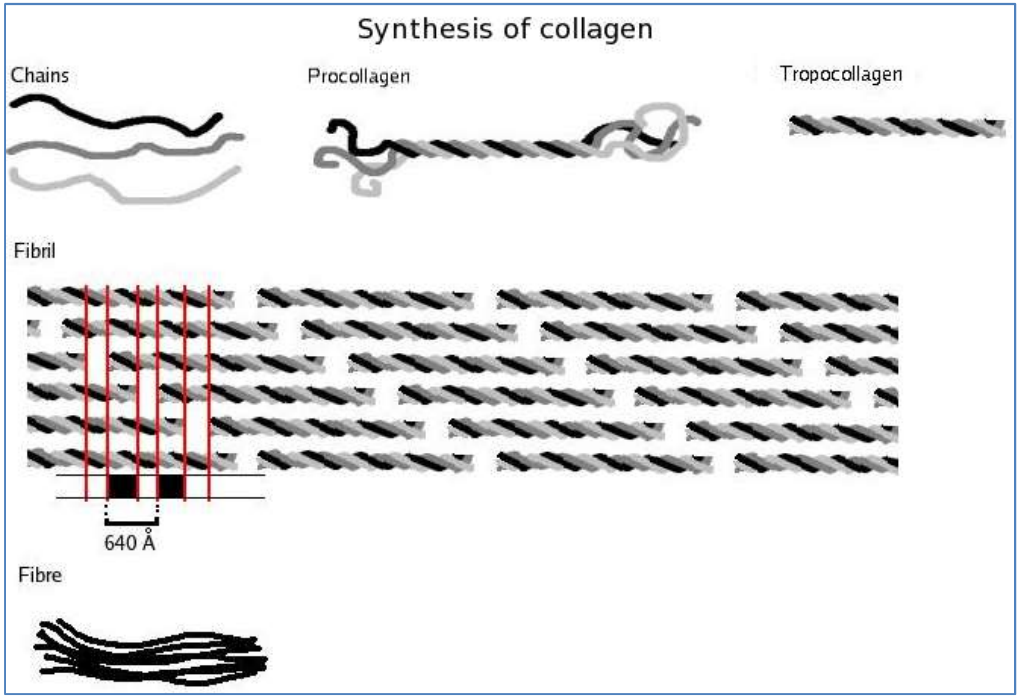
Proteins

The Secondary Structure of Proteins

C) The α Helix and the Pleated Sheet

Collagen

- The basic building block of collagen is a triple helix of three polypeptide chains called the tropocollagen unit.
- Each chain is about 1000 residues long. The individual collagen chains form left-handed helices with about 3.3 residues per turn.



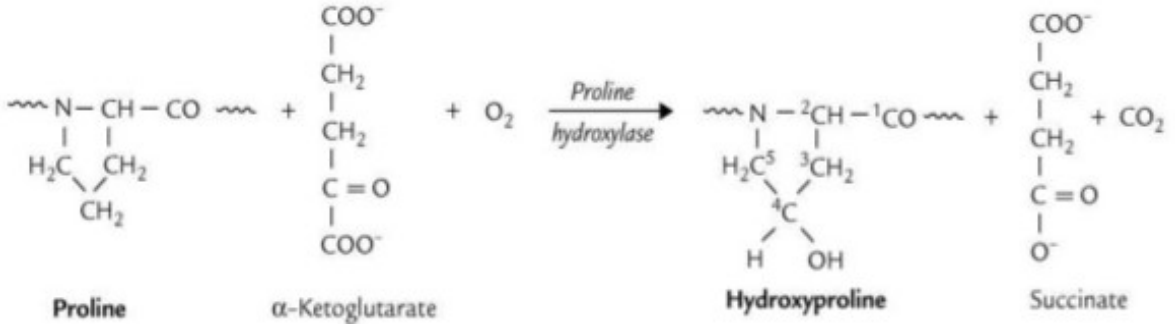
Proteins

The Secondary Structure of Proteins

C) The α Helix and the Pleated Sheet

Collagen

- In order to form this triple-stranded helix, every third residue must be glycine because glycine offers a minimum of bulk.
- Another interesting theme in collagen is the additional hydrogen bonding that occurs because of the presence of hydroxyproline derived from the conversion of proline to hydroxproline.
- Collagen fibers are strong.
- In tendons, the collagen fibers have a strength similar to that of hard drawn copper wire.



Proteins

The Secondary Structure of Proteins

C) The α Helix and the Pleated Sheet

Collagen

- Much of the toughness of collagen is the result of the crosslinking of the tropocollagen units to one another through a reaction involving lysine side chains.

  Lysine side chains are oxidized to aldehydes that react with either a lysine residue or with one another through an aldol condensation and dehydration resulting in a crosslink.

- This process continues throughout our life, resulting in our bones and tendons becoming less elastic and more brittle.
- Again, a little crosslinking is essential, but more crosslinking leads to increased fracture and brittleness.

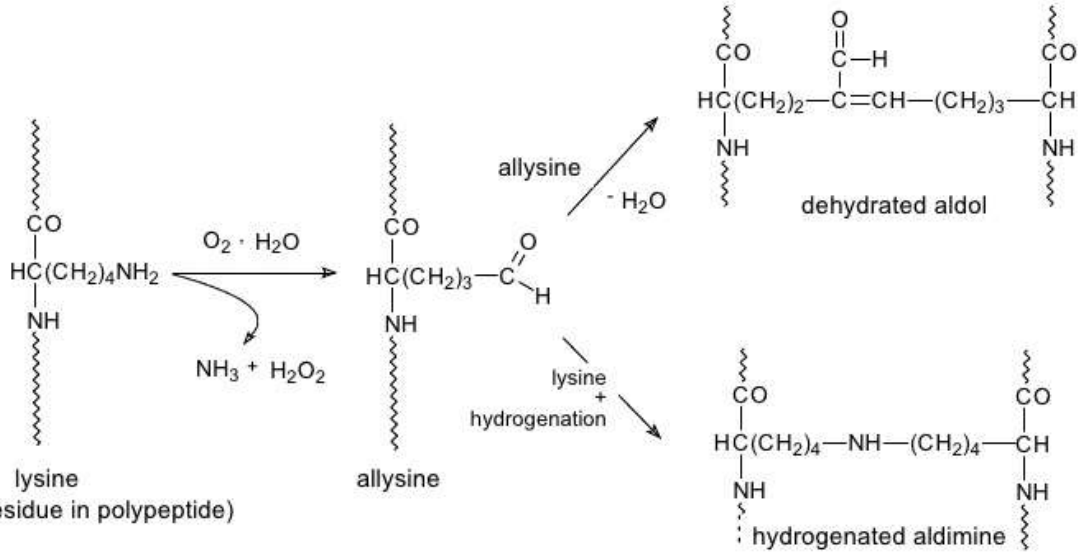
Proteins

The Secondary Structure of Proteins

C) The α Helix and the Pleated Sheet

Cross-links in collagen

products of reaction between the amino groups in side chains of **lysine** with the modified lysine side chains comprising the aldehyde group (the result of oxidation of lysine to **allysine**) – aldol type or aldimine type of cross-links.



- Collagen is a major ingredient in some "gelation" materials.
- Here, collagen forms a triple helix for some of its structure while other parts are more randomly flowing single collagen chain segments.
- The bundled triple-helical structure acts as the rigid part of the polymer while the less ordered amorphous chains act as a soft part of the chain.
- The triple helix also acts as a noncovalently bonded crosslink.

Proteins

The Secondary Structure of Proteins

C) The α Helix and the Pleated Sheet

Elastin

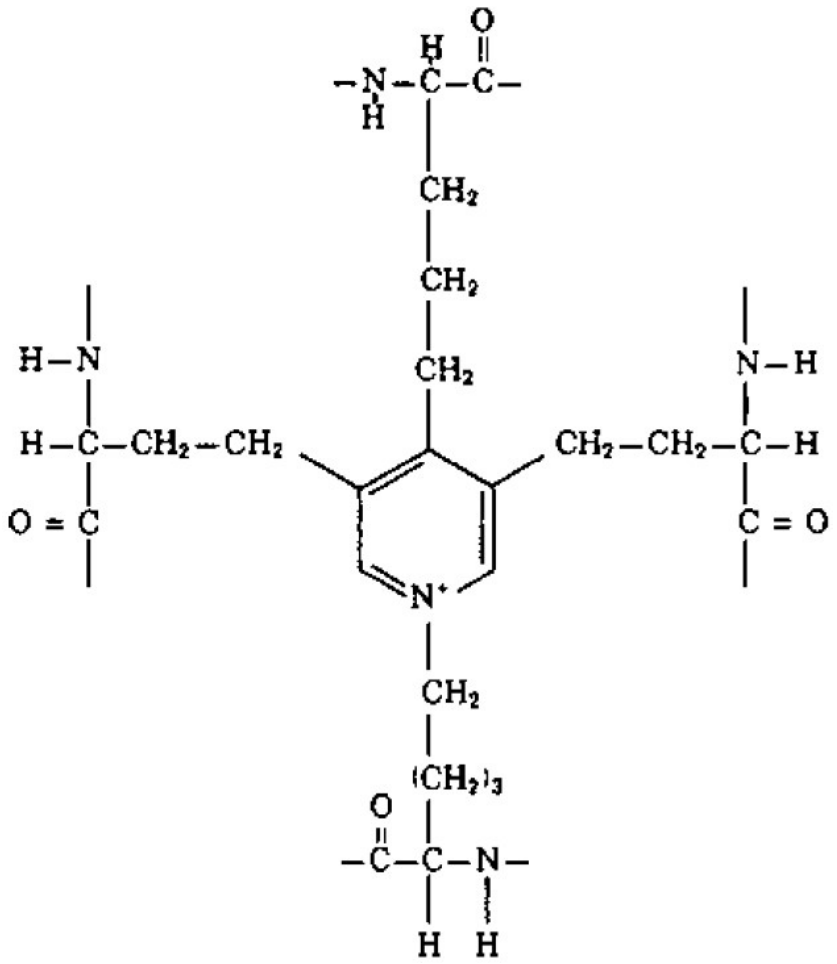
- Collagen is found where strength is needed, but some tissues, such as arterial blood vessels and ligaments, need materials that are elastic.
- Elastin is the protein of choice for such applications.
- Elastin is rich in glycine, alanine, and valine, and it is easily extended and flexible.
- Its conformation approaches that of a random coil so that secondary forces are relatively weak allowing elastin to be readily extended as tension is applied.
- The structure also contains some lysine side chains that are involved in crosslinking.
- The crosslinking is accomplished when four lysine side chains are combined to form a desmosine crosslink

Proteins

The Secondary Structure of Proteins

C) The α Helix and the Pleated Sheet

Elastin



Central structure of desmosine

Proteins

Tertiary Structure: Fibrous and Globular Proteins

- The term *tertiary structure* is used to describe the shaping or folding of macromolecules.
- These larger structures generally contain elements of the secondary structures.
- Often hydrogen bonding and crosslinking lock in such structures.
- Proteinsn can be divided into two broad groups:
  - Fiberous or fibrillar proteins,
    - Fibrous proteins are long macromolecules that are attached through either inter- or intra-hydrogen bonding of the individual residues within the chain.
    - Solubility, partial or total, occurs when these hydrogen bonds are broken.
  - globular proteins
    - They are generally soluble in acidic, basic, or neutral aqueous solutions.
- The key lies mainly in the amino acid makeup itself.
- But what about the diverse R groups of the various amino acids?
- Some amino acids have nonpolar R groups, simple alkyl or aromatic groups. Others have highly polar R groups, with carboxylate or ammonium ions and hydroxyl or other polar groups.

  Different R groups affect the gross properties of a protein.

Proteins

Tertiary Structure: Fibrous Proteins

- ***Fibrous proteins*** are animal structural materials and hence are water insoluble.
- They fall into three general categories:
  - the keratins, which make up protective tissue, such as skin, hair, feathers, claws, and nails;
  - the collagens, which form connective tissue, such as cartilage, tendons, and blood vessels; and
  - the silks, such as the fibroin of spider webs and cocoons.
- Keratins and collagens have helical structures, whereas silks have pleated-sheet structures.
- A large fraction of the R groups attached to these frameworks are nonpolar, accounting for the insolubility of these proteins in water.

Proteins

Tertiary Structure: Fibrous Proteins

- In hair, three a helices are braided to form a rope, the helices being held together by disulfide cross-links.
- The ropes are further packed side by side in bundles that ultimately form the hair fiber.
- The α-keratin of more rigid structures, such as nails and claws, is similar to that of hair, except that there is a higher percentage of cysteine amino acid units in the polypeptide chain.
- Therefore, there are more disulfide cross-links, giving a firmer, less flexible overall structure.
- To summarize, nonpolar R groups and disulfide cross-links, together with helical or sheet-like backbones, tend to give fibrous proteins their rather rigid, insoluble structures.

Proteins

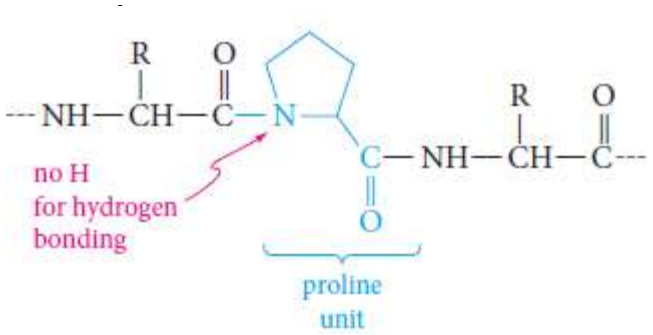
Tertiary Structure: Globular Proteins

- ***Globular proteins*** are very different from fibrous proteins.
- They tend to be water soluble and have roughly spherical shapes, as their name suggests.
- Instead of being structural, globular proteins perform various other biological functions. **They may be**
  - enzymes (biological catalysts),
  - hormones (chemical messengers that regulate biological processes),
  - transport proteins (carriers of small molecules from one part of the body to another, such as hemoglobin, which transports oxygen in the blood),
  - storage proteins (which act as food stores; ovalbumin of egg white is an example).

Proteins

Tertiary Structure: Globular Proteins

- ***Globular proteins*** have more amino acids with polar or ionic side chains than the water-insoluble fibrous proteins.
- An enzyme or other globular protein that carries out its function mainly in the aqueous medium of the cell will adopt a structure in which the nonpolar, hydrophobic R groups point in toward the center and the polar or ionic R groups point out toward the water.
- ***Globular proteins*** are mainly helical, but they have folds that permit the overall shape to be globular.
- One of the 20 amino acids, proline has a second
- Wherever a proline unit occurs in the primary pe
  no N-H group available for intra-chain hydrogen

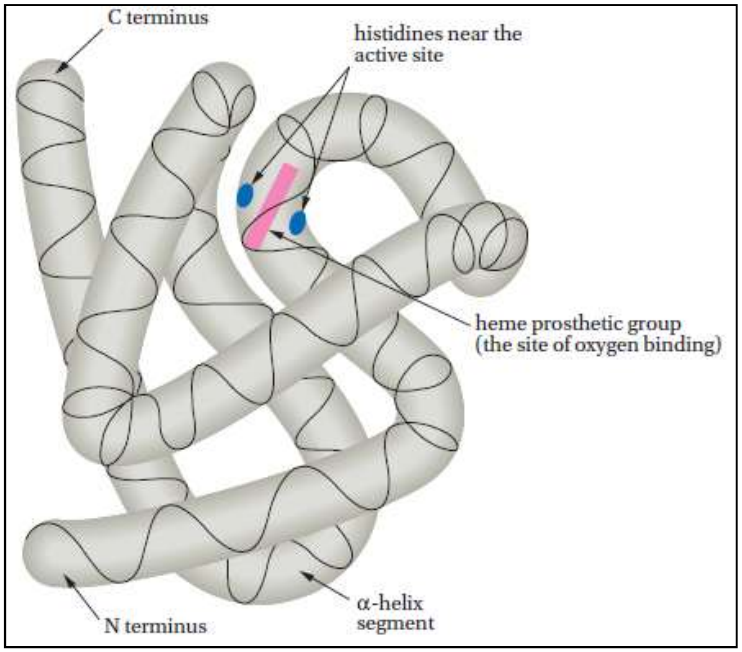


Proteins

Tertiary Structure: Globular Proteins

- *Myoglobin*, the oxygen-transport protein of muscle, is a good example of a globular protein.

- It contains 153 amino acid units, yet is extremely compact, with very little empty space in its interior.
- Approximately 75% of the amino acid units in myoglobin are part of eight right-handed α-helical sections.
- The interior of myoglobin consists almost entirely of nonpolar R groups, such as those of leucine, valine, phenylalanine, and methionine.
- The only interior polar groups are two histidines. These perform a necessary function at the active site of the protein, where the nonprotein portion, a molecule of the porphyrin heme, binds the oxygen.
- The outer surface of the protein includes many highly polar amino acid residues (lysine, arginine, glutamic acid, etc.).



Schematic drawing of myoglobin. Each of the tubular sections is a segment of α-helix, but the overall shape is globular.

Proteins

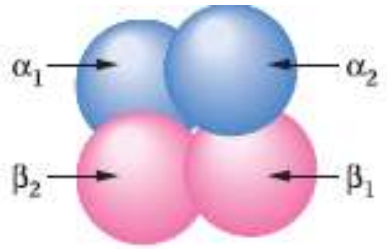
Summary

- Amino acid content of a peptide or protein influences its shape.
- These interactions are mainly a consequence of disulfide bonds and of the polarity or nonpolarity of the R groups, their shape, and their ability to form hydrogen bonds.
- When we refer to the tertiary structure of a protein, we refer to all contributions of these factors to its three-dimensional structure.

Proteins

Quaternary Protein Structure

- The term quaternary structure is employed to describe the overall shape of groups of chains of proteins.
- ***Hemoglobin*** is composed of four distinct but similar protein macromolecules, each with its own tertiary structure that comes together to give the quaternary hemoglobin structure

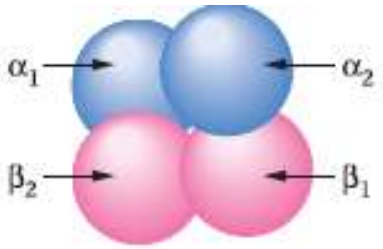


Schematic drawing of the four hemoglobin subunits.

Proteins

Quaternary Protein Structure

- Some high-molecular-weight proteins exist as aggregates of several subunits.
- These aggregates are referred to as the quaternary structure of the protein.
- Aggregation helps to keep nonpolar portions of the protein surface from being exposed to the aqueous cellular environment.
- ***Hemoglobin***, the oxygen-transport protein of red cells, provides an example of such aggregation.
- It consists of four almost spherical units, two α units with 141 amino acids and two β units with 146 amino acids.
- The four units come together in a tetrahedral array.



Schematic drawing of the four hemoglobin subunits.