Amino acids are detected as they emerge from the column by reaction with ninhydrin (Section 27.2D) followed by absorption spectroscopy. Current procedures for hydrolysis of polypeptides and analysis of amino acid mixtures have been refined to the point where it is possible to obtain amino acid composition from as little as 50 nanomoles (50×10^{-9} mole) of polypeptide. Figure 27.6 shows the analysis of a polypeptide hydrolysate by ion-exchange chromatography. Note that during hydrolysis, the side-chain amide groups of asparagine and glutamine are hydrolyzed, and these amino acids are detected as aspartic acid and glutamic acid. For each glutamine or asparagine hydrolyzed, an equivalent amount of ammonium chloride is formed.

B. Sequence Analysis

After the amino acid composition of a polypeptide has been determined, the next step is to determine the order in which the amino acids are joined in the polypeptide chain. The most common sequencing strategy is to cleave the polypeptide at specific peptide bonds (using, for example, cyanogen bromide or certain proteolytic enzymes), determine the sequence of each fragment (using, for example, the Edman degradation), and then match overlapping fragments to arrive at the sequence of the polypeptide.

Cyanogen Bromide

Cyanogen bromide (BrCN) is specific for cleavage of peptide bonds formed by the carboxyl group of methionine (Figure 27.7). The products of this cleavage are a substituted γ-lactone (Section 18.1C) derived from the *N*-terminal portion of the polypeptide and a second fragment containing the *C*-terminal portion of the polypeptide.

A three-step mechanism can be written for this reaction. The strategy for cyanogen bromide cleavage depends on chemical manipulation of the leaving ability of the sulfur atom of methionine. Because CH_3S^- is the anion of a weak acid, it is a very poor leaving group, just as OH^- is a poor leaving group (Section 9.4F). Yet, just as the oxygen atom of an alcohol can be transformed into a better leaving group by converting it into an oxonium ion (by protonation), so too can the sulfur atom of methionine be transformed into a better leaving group by converting it into a sulfonium ion.

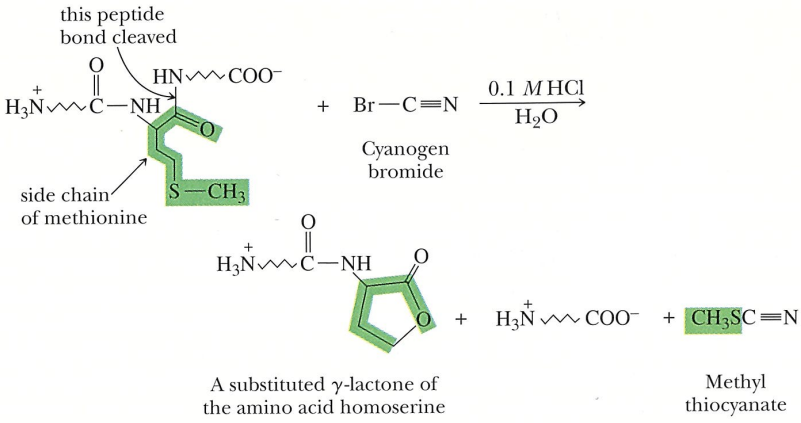
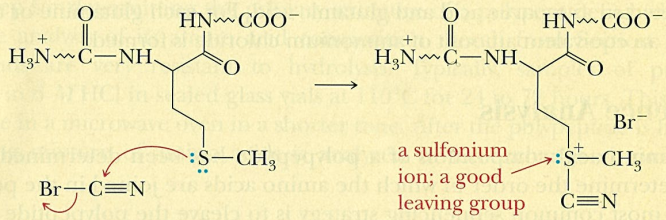


Figure 27.7
Cleavage by cyanogen bromide, BrCN, of a peptide bond formed by the carboxyl group of methionine.

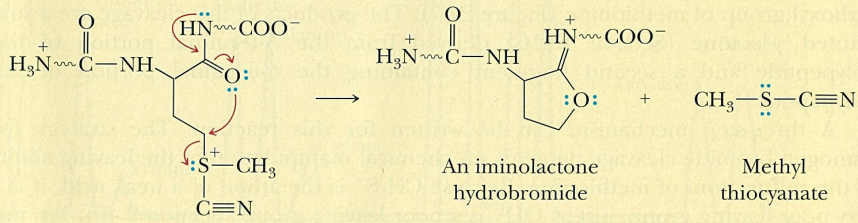
Mechanism *Cleavage of a Peptide Bond at Methionine by Cyanogen Bromide*

Step 1: Reaction is initiated by nucleophilic attack of the divalent sulfur atom of methionine on the carbon of cyanogen bromide displacing bromide ion. The product of this nucleophilic displacement is a sulfonium ion.



Cyanogen bromide

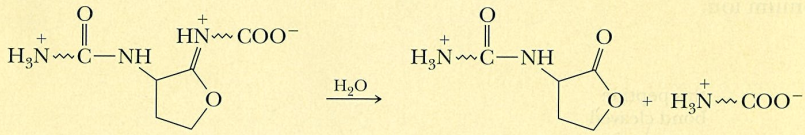
Step 2: An internal S_N2 reaction in which the oxygen of the methionine carbonyl group attacks the γ-carbon and displaces methyl thiocyanate gives a five-membered ring. Note that the oxygen of a carbonyl group is at best a weak nucleophile. This displacement is facilitated, however, because the sulfonium ion is a very good leaving group and because of the ease with which a five-membered ring is formed.



An iminolactone hydrobromide

Methyl thiocyanate

Step 3: Hydrolysis of the imino group gives a γ-lactone derived from the *N*-terminal end of the original polypeptide.



A substituted γ-lactone of the amino acid homoserine

Enzyme-Catalyzed Hydrolysis of Peptide Bonds

A group of proteolytic enzymes, among them trypsin and chymotrypsin, can be used to catalyze the hydrolysis of specific peptide bonds. Trypsin catalyzes the hydrolysis of peptide bonds formed by the carboxyl groups of arginine and lysine; chymotrypsin catalyzes the hydrolysis of peptide bonds formed by the carboxyl groups of phenylalanine, tyrosine, and tryptophan (Table 27.3).

Table 27.3 Cleavage of Specific Peptide Bonds Catalyzed by Trypsin and Chymotrypsin

Enzyme	Catalyzes Hydrolysis of Peptide Bond Formed by Carboxyl Group of
Trypsin	Arginine, lysine
Chymotrypsin	Phenylalanine, tyrosine, tryptophan

Example 27.5

Which of these dipeptides are hydrolyzed by trypsin? By chymotrypsin?
(a) Arg-Glu-Ser **(b)** Phe-Gly-Lys

Solution

(a) Trypsin catalyzes the hydrolysis of peptide bonds formed by the carboxyl groups of lysine and arginine. Therefore, the peptide bond between arginine and glutamic acid is hydrolyzed in the presence of trypsin.

$$\text{Arg-Glu-Ser} + H_2O \xrightarrow{\text{trypsin}} \text{Arg} + \text{Glu-Ser}$$

Chymotrypsin catalyzes the hydrolysis of peptide bonds formed by the carboxyl groups of phenylalanine, tyrosine, and tryptophan. Because none of these three aromatic amino acids is present, tripeptide (a) is not affected by chymotrypsin.

(b) Tripeptide (b) is not affected by trypsin. Although lysine is present, its carboxyl group is at the *C*-terminal end and is not involved in peptide bond formation. Tripeptide (b) is hydrolyzed in the presence of chymotrypsin.

$$\text{Phe-Gly-Lys} + H_2O \xrightarrow{\text{chymotrypsin}} \text{Phe} + \text{Gly-Lys}$$

Problem 27.5

Which of these tripeptides are hydrolyzed by trypsin? By chymotrypsin?
(a) Tyr-Gln-Val **(b)** Thr-Phe-Ser **(c)** Thr-Ser-Phe

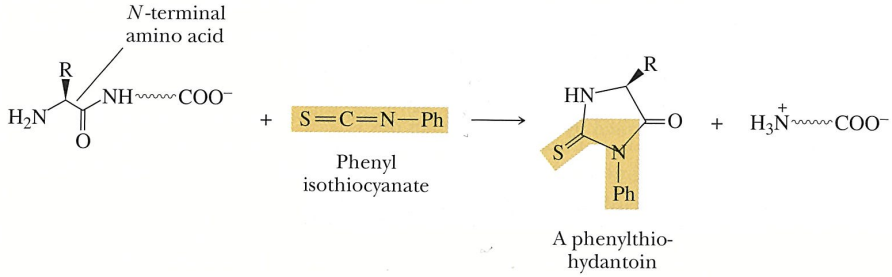
Edman Degradation

Edman degradation A method for selectively cleaving and identifying the *N*-terminal amino acid of a polypeptide chain.

Of the various chemical methods developed for determining the amino acid sequence of a polypeptide, the one most widely used today is the **Edman degradation,** introduced in 1950 by Pehr Edman of the University of Lund, Sweden. In this procedure, a polypeptide is treated with phenyl isothiocyanate, $C_6H_5N{=}C{=}S$, and then with acid. The effect of Edman degradation is to remove the *N*-terminal amino acid selectively as a substituted phenylthiohydantoin (Figure 27.8), which is then separated and identified.

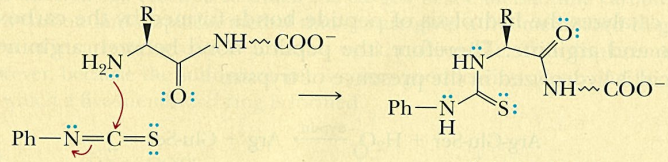
The key feature of the Edman degradation is successive C=N, C=O, and C=O addition reactions.

Figure 27.8
Edman degradation. Treatment of a polypeptide with phenyl isothiocyanate followed by acid selectively cleaves the *N*-terminal amino acid as a substituted phenylthiohydantoin.

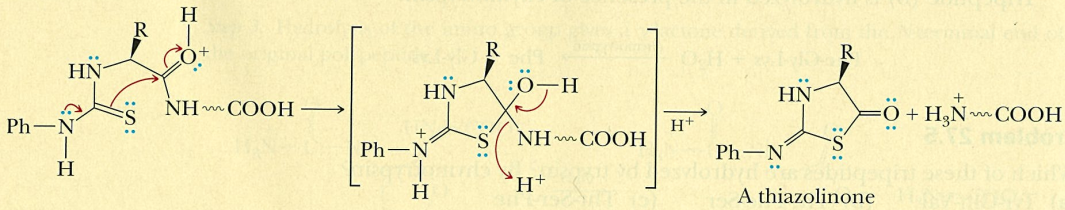


Mechanism *Edman Degradation—Cleavage of an N-Terminal Amino Acid*

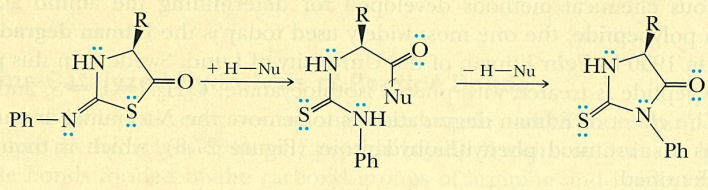
Step 1: Nucleophilic addition of the *N*-terminal amino group to the C=N bond of phenyl isothiocyanate gives a derivative of *N*-phenylthiourea.



Step 2: Heating the derivatized polypeptide in HCl at 100°C results in nucleophilic addition of sulfur to the carbonyl of the adjacent amide group to give a tetrahedral carbonyl addition intermediate, which collapses to give a thiazolinone ring derived from the *N*-terminal amino acid.



Step 3: The thiazolinone ring undergoes isomerization by ring opening followed by reclosure to give a more stable phenylthiohydantoin, which is separated and identified by chromatography.



The special value of the Edman degradation is that it cleaves the *N*-terminal amino acid from a polypeptide without affecting any other bonds in the chain. Furthermore, Edman degradation can be repeated on the shortened polypeptide, causing the next amino acid in the sequence to be cleaved and identified. In practice, it is now possible to sequence as many as the first 20 to 30 amino acids in a polypeptide by this method using only a few milligrams of material.

Most polypeptides in nature are longer than 20 to 30 amino acids, the practical limit to the number of amino acids that can be sequenced by repetitive Edman degradation. The special value of cleavage with cyanogen bromide, trypsin, and chymotrypsin is that a long polypeptide chain can be cleaved at specific peptide bonds into smaller polypeptide fragments, and each fragment can then be sequenced separately. The entire procedure, from the chemical cleavage step to the separation of the product can be automated. Samples of a pure protein can be added to a machine and the first several *N*-terminal amino acids determined automatically.

Example 27.6

Deduce the amino acid sequence of a pentapeptide from the following experimental results. Note that under the column labeled "Amino Acid Composition," the amino acids are listed in alphabetical order. In no way does this listing give any information about primary structure.

Experimental Procedure	Amino Acid Composition
Pentapeptide	Arg, Glu, His, Phe, Ser
Edman degradation	Glu
Hydrolysis catalyzed by chymotrypsin	
Fragment A	Glu, His, Phe
Fragment B	Arg, Ser
Hydrolysis catalyzed by trypsin	
Fragment C	Arg, Glu, His, Phe
Fragment D	Ser

Solution

Edman degradation cleaves Glu from the pentapeptide; therefore, glutamic acid must be the *N*-terminal amino acid.

Glu-(Arg, His, Phe, Ser)

Fragment A from chymotrypsin-catalyzed hydrolysis contains Phe. Because of the specificity of chymotrypsin, Phe must be the *C*-terminal amino acid of fragment A. Fragment A also contains Glu, which we already know is the *N*-terminal amino acid. From these observations, conclude that the first three amino acids in the chain must be Glu-His-Phe and then write the following partial sequence:

Glu-His-Phe-(Arg, Ser)

The fact that trypsin cleaves the pentapeptide means that Arg must be within the pentapeptide chain; it cannot be the *C*-terminal amino acid. Therefore, the complete sequence must be

Glu-His-Phe-Arg-Ser

Problem 27.6

Deduce the amino acid sequence of an undecapeptide (11 amino acids) from the experimental results shown in the table.

Experimental Procedure	Amino Acid Composition
Undecapeptide	Ala, Arg, Glu, Lys_2, Met, Phe, Ser, Thr, Trp, Val
Edman degradation	Ala
Trypsin-catalyzed hydrolysis	
Fragment E	Ala, Glu, Arg
Fragment F	Thr, Phe, Lys
Fragment G	Lys
Fragment H	Met, Ser, Trp, Val
Chymotrypsin-catalyzed hydrolysis	
Fragment I	Ala, Arg, Glu, Phe, Thr
Fragment J	Lys_2, Met, Ser, Trp, Val
Treatment with cyanogen bromide	
Fragment K	Ala, Arg, Glu, Lys_2, Met, Phe, Thr, Val
Fragment L	Trp, Ser

Sequencing by Mass Spectrometry

As mentioned in Chapter 14, mass spectrometry is increasingly being used for the direct sequencing of proteins in very small quantities. This topic was fully addressed in Connections to Biological Chemistry: "Mass Spectra of Biological Macromolecules" in Section 14.3 and will not be further developed here.

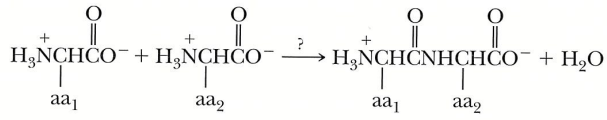
Sequencing Proteins from the Coding Nucleotide Sequence

As it has become easier to determine nucleotide sequences (see Section 28.5), it is now often easier to sequence the nucleotide that codes for a protein than to sequence the protein itself. In some cases, this has lead to discovery of new proteins of unknown function. Often comparison of the revealed protein sequences with those of known proteins from simpler organisms discloses sequence homologies that suggest the function of the new proteins. Comparisons with yeast, whose genome was one of the first to be sequenced, and for which most of the proteins coded for have known functions, have been particularly fruitful.

27.5 Synthesis of Polypeptides

A. The Problems

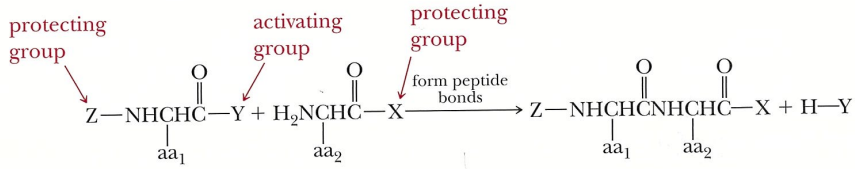
The problem in peptide synthesis is to join the carboxyl group of amino acid 1 (aa_1) by an amide (peptide) bond to the amino group of amino acid 2 (aa_2).



B. The Strategy

A rational strategy for the synthesis of peptide bonds and polypeptides requires three steps.

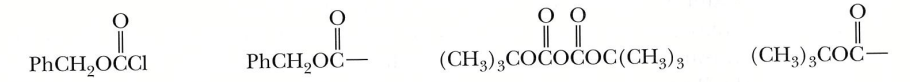
1. Protect the α-amino group of amino acid aa_1 to reduce its nucleophilicity so that it does not participate in nucleophilic addition to the carboxyl group of either aa_1 or aa_2.
2. Protect the α-carboxyl group of amino acid aa_2 so that it is not susceptible to nucleophilic attack by the α-amino group of another molecule of aa_2.
3. Activate the α-carboxyl group of amino acid aa_1 so that it is susceptible to nucleophilic attack by the α-amino group of aa_2.



After dipeptide aa_1—aa_2 has been formed, the protecting group Z can be removed, and chain growth can be continued from the *N*-terminal end of the dipeptide. Alternatively, the protecting group X can be removed and chain growth can be continued from the *C*-terminal end. The range of protecting groups and activating groups is large, and experimental conditions have been found to attach and remove them as desired.

C. Amino-Protecting Groups

The most common strategy for protecting amino groups and reducing their nucleophilicity is to react them with carbonyl derivatives. The reagents most commonly used for this purpose are benzyloxycarbonyl chloride and di-*tert*-butyl dicarbonate. In the terminology adopted by the IUPAC, the benzyloxycarbonyl group is given the symbol Z, and the *tert*-butoxycarbonyl group is given the symbol BOC—.



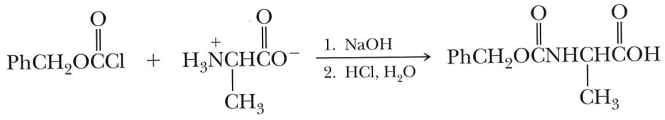
Benzyloxycarbonyl chloride

Benzyloxycarbonyl (Z—) group

Di-*tert*-butyl dicarbonate

tert-Butoxycarbonyl (BOC—) group

Treatment of an amino group with either of these reagents forms a new functional group called a carbamate. A carbamate is an ester of carbamic acid; that is, it is an ester of the monoamide of carbonic acid.

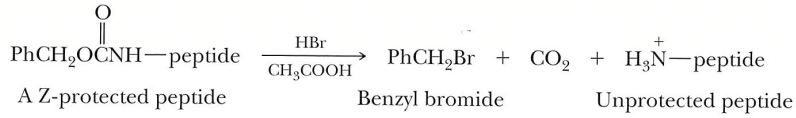


Benzyloxycarbonyl chloride (Z—Cl)

Alanine

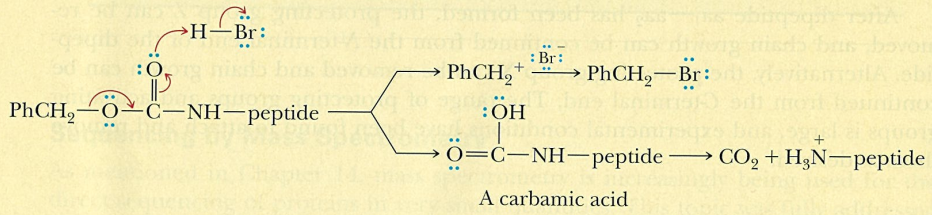
N-Benzyloxycarbonylalanine (Z—Ala) A carbamate

The special advantage of the carbamate group is that it is stable to dilute base but can be removed by treatment with HBr in acetic acid.



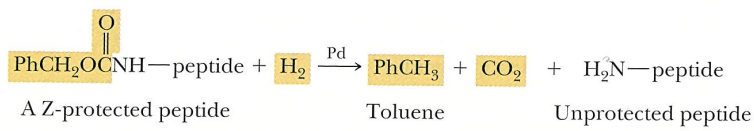
A study of the mechanism for removal of this protecting group has shown that the reaction is first order in $[H^+]$ and involves formation of a carbocation and a carbamic acid. A carbamic acid spontaneously loses carbon dioxide to form the free amine. The carbocation reacts with an available nucleophile such as halide ion to form an alkyl halide.

Mechanism *Acid-Catalyzed Removal of a Benzyloxycarbonyl Protecting Group*



Note that because acid-catalyzed removal of this protecting group is carried out in nonaqueous media, there is no danger of simultaneous acid-catalyzed hydrolysis of peptide (amide) bonds within the newly synthesized polypeptide. This relationship exists because water is required for hydrolysis of a peptide bond.

The benzyloxycarbonyl group can also be removed by treatment with H_2 in the presence of a transition metal catalyst (hydrogenolysis, Section 21.5C). In hydrogenolysis of a Z-protecting group, one product is toluene. The other is a carbamic acid, which undergoes spontaneous decarboxylation to give carbon dioxide and the unprotected peptide.



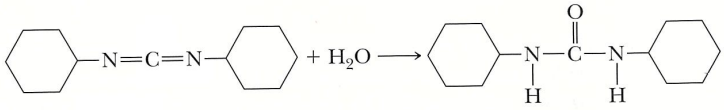
D. Carboxyl-Protecting Groups

Carboxyl groups are most often protected by conversion to methyl, ethyl, or benzyl esters. Methyl and ethyl esters are prepared by Fischer esterification (Section 17.7A)

and are removed by hydrolysis in aqueous base (Section 18.4C) under mild conditions. Benzyl esters are conveniently removed by hydrogenolysis with H_2 over a palladium or platinum catalyst (Section 21.5C). Benzyl groups can also be removed by treatment with HBr in acetic acid.

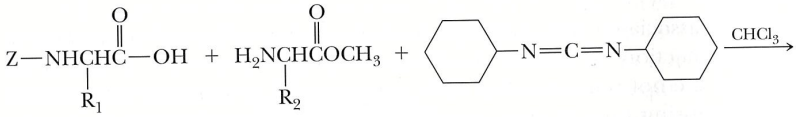
E. Peptide-Bond Forming Reactions

The reagent most commonly used to bring about peptide bond formation is 1,3-dicyclohexylcarbodiimide (DCC). This reagent is the anhydride of a disubstituted urea, and, when treated with water, it is converted to *N,N'*-dicyclohexylurea (DCU).

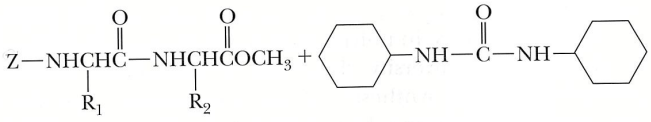


1,3-Dicyclohexylcarbodiimide (DCC) — *N,N'*-Dicyclohexylurea (DCU)

When an amino-protected aa_1 and a carboxyl-protected aa_2 are treated with DCC, this reagent acts as a dehydrating agent; it removes —OH from the carboxyl group and —H from the amino group to form an amide bond. More specifically, DCC activates the α-carboxyl group of aa_1 toward nucleophilic acyl substitution by converting its —OH group into a better leaving group.



Amino-protected aa_1 — Carboxyl-protected aa_2 — 1,3-Dicyclohexylcarbodiimide (DCC)



Amino and carboxyl protected dipeptide — *N,N'*-Dicyclohexylurea (DCU)

An abbreviated mechanism for this intermolecular dehydration is shown in Figure 27.9. An acid-base reaction in Step 1 between the carboxyl group of aa_1 and a nitrogen of DCC followed in Step 2 by addition of the carboxylate anion to the C=N double bond results in electrophilic addition to a C=N double bond. The O-acylisourea formed is the nitrogen analog of a mixed anhydride. Nucleophilic addition of the amino group of aa_2 to the carbonyl group of the O-acylisourea in Step 3 generates a tetrahedral carbonyl addition intermediate that collapses in Step 4 to give a dipeptide and DCU.

Mechanism *Dicyclohexylcarbodiimide and Formation of a Peptide Bond*

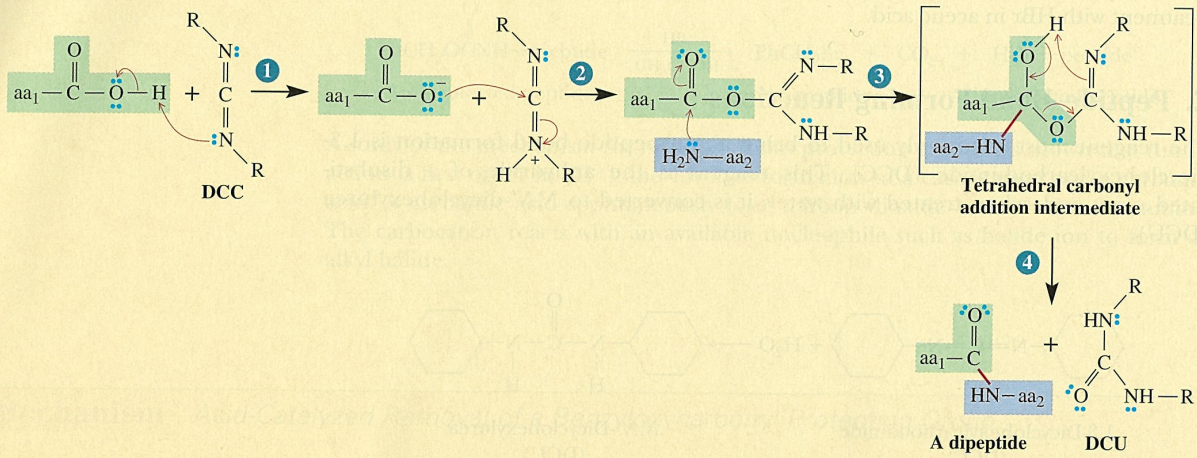


Figure 27.9
The role of 1,3-dicyclohexylcarbodiimide (DCC) in the formation of a peptide bond between an amino-protected amino acid (aa_1) and a carboxyl-protected amino acid (aa_2).

F. Solid-Phase Synthesis

A major problem associated with polypeptide synthesis is purification of intermediates after each protection, activation, coupling, and deprotection step. If unreacted starting materials are not removed after each step, the final product is contaminated by polypeptides missing one or more amino acids. The required purification steps are not only laborious and time consuming, but they also inevitably result in some loss of the desired product. These losses become especially severe in the synthesis of larger polypeptides.

A major advance in polypeptide synthesis came in 1962 when R. Bruce Merrifield of Rockefeller University described a solid-phase synthesis (alternatively called polymer-supported synthesis) of the tetrapeptide, Leu-Ala-Gly-Ala, by a technique that now bears his name. Merrifield was awarded the 1984 Nobel Prize for chemistry for his work in developing the solid-phase method for peptide synthesis.

The solid support used by Merrifield was a type of polystyrene in which about 5% of the phenyl groups carry a chloromethyl ($—CH_2Cl$) group in their para positions (Figure 27.10). These chloromethyl groups, like all benzylic halides, are particularly reactive in nucleophilic substitution reactions.

In the Merrifield method, the *C*-terminal amino acid is joined as a benzyl ester to the solid polymer support, and then the polypeptide chain is extended one amino acid at a time from the *N*-terminal end. The advantage of polypeptide synthesis on a solid support is that the polymer beads with the peptide chains anchored on them are completely insoluble in the solvents used in the synthesis. Furthermore, excess reagents (for example, DCC) and byproducts (for example, DCU) are removed after each step simply by washing the polymer beads. When synthesis is completed, the

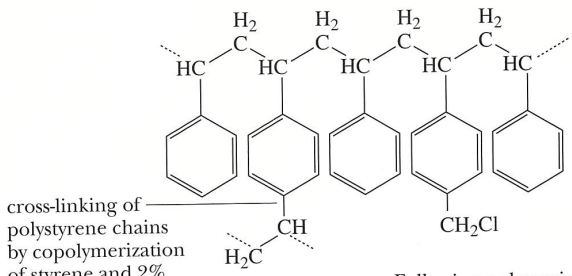


Figure 27.10
The support used for the Merrifield solid-phase synthesis is a chloromethylated polystyrene resin.

polypeptide is released from the polymer beads by cleavage of the benzyl ester. The steps in solid-phase synthesis of a polypeptide are summarized in Figure 27.11.

Thanks to automation, the synthesis of polypeptides is now a routine procedure in chemical research. It is common for researchers to order several peptides at a time for use in fields as diverse as medicine, biology, material science, and biomedical engineering.

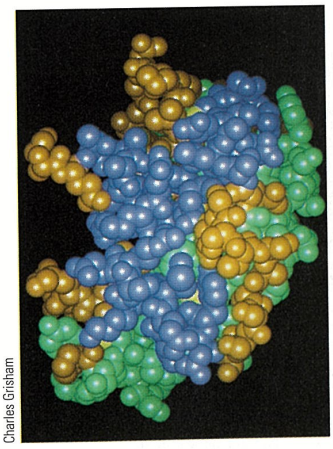
A dramatic illustration of the power of the solid-phase method was the synthesis of the enzyme ribonuclease by Merrifield in 1969. The synthesis involved 369 chemical reactions and 11,931 operations, all of which were performed by an automated machine and without any intermediate isolation stages. Each of the 124 amino acids was added as an *N*-*tert*-butoxycarbonyl derivative and coupled using DCC. Cleavage from the resin and removal of all protective groups gave a mixture that was purified by ion-exchange chromatography. The specific activity of the synthetic enzyme was 13–24% of that of the natural enzyme. The fact that the specific activity of the synthetic enzyme was lower than that of the natural enzyme was probably attributable to the presence of polypeptide byproducts closely related to but not identical to the natural enzyme. Synthesizing ribonuclease (124 amino acids) requires forming 123 peptide bonds. If each peptide bond is formed in 99% yield, the yield of homogeneous polypeptide is $0.99^{123} = 29\%$. If each peptide bond is formed in 98% yield, the yield is 8%. Thus, even with yields as high as 99% in each peptide bond-forming step, a large portion of the synthetic polypeptides have one or more sequence defects. Many of these, nonetheless, may be fully or partially active.

27.6 Three-Dimensional Shapes of Polypeptides and Proteins

A. Geometry of a Peptide Bond

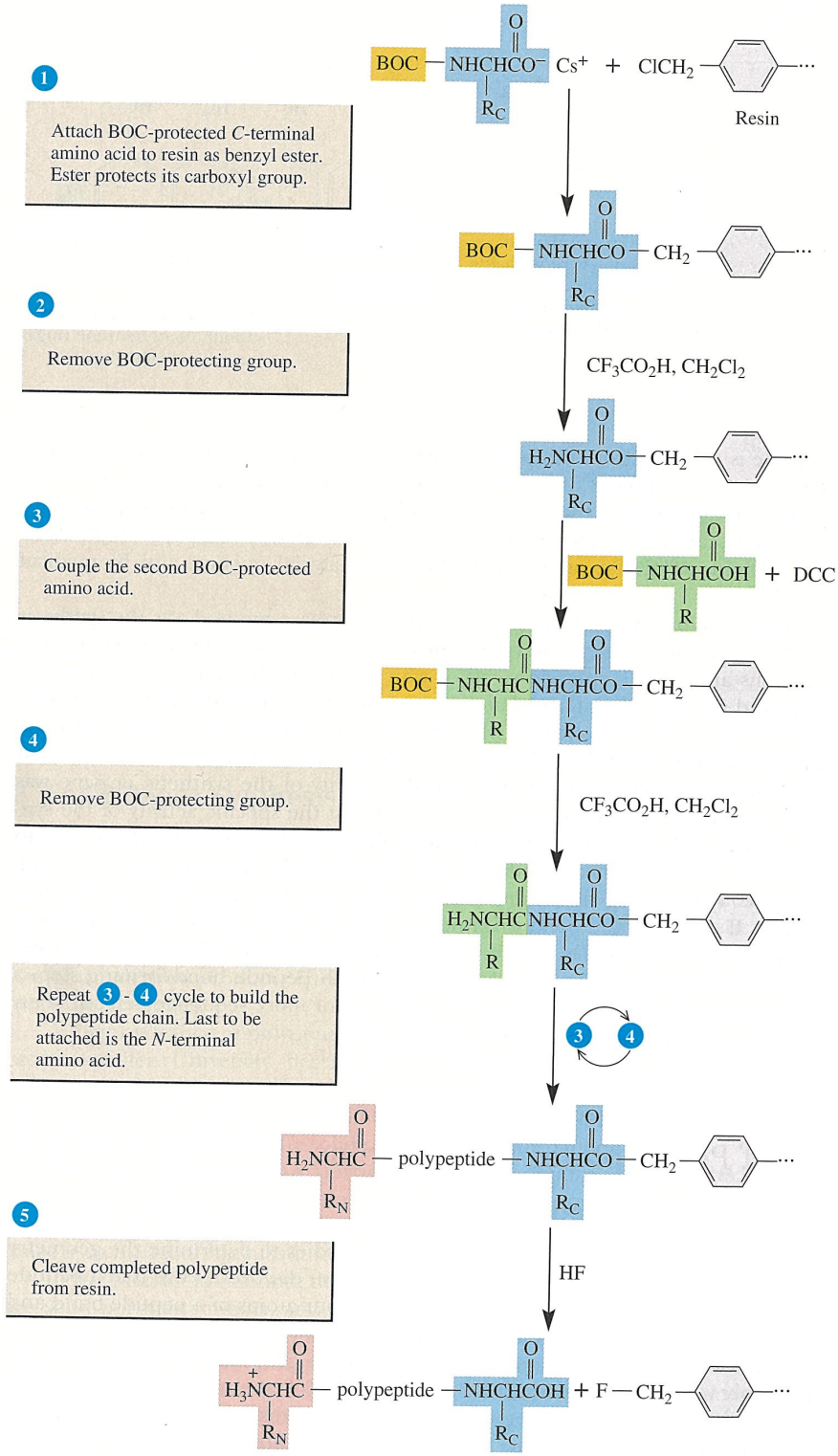
In the late 1930s, Linus Pauling began a series of studies to determine the geometry of a peptide bond. One of his first and most important discoveries was that a peptide bond itself is planar. As shown in Figure 27.12, the four atoms of a peptide bond and the two α-carbons joined to it all lie in the same plane.

Had you been asked in Chapter 1 to describe the geometry of a peptide bond, you probably would have predicted bond angles of 120° about the carbonyl carbon and 109.5° about the amide nitrogen. However, as fully discussed in Connections to Biological Chemistry: "The Unique Structure of Amide Bonds" in Chapter 18, both


Charles Grisham

A model of the protein ribonuclease A. The purple segments are regions of α-helix and the yellow segments are regions of β-pleated sheet, both of which are described in Section 27.6. Other colors represent loop regions.

Figure 27.11
Steps in the Merrifield solid-phase polypeptide synthesis.



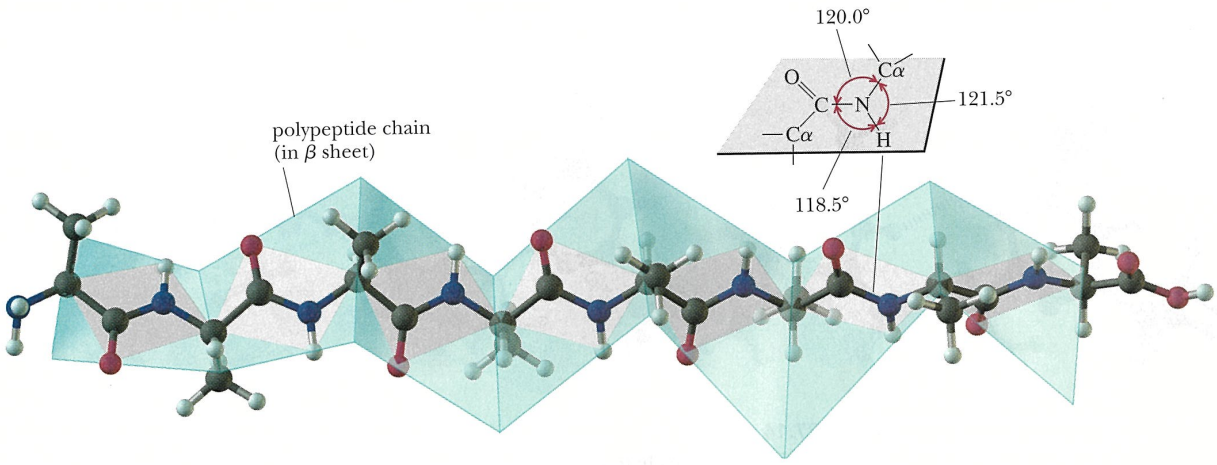
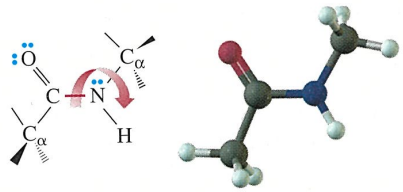
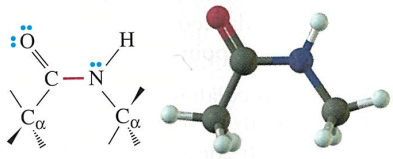


Figure 27.12
Planarity of a peptide bond. Bond angles about the carbonyl carbon and the amide nitrogen are approximately 120°.

atoms are actually planar with approximately 120° bond angles about each because of resonance of the nitrogen lone pair with the carbonyl. Two configurations are possible for the atoms of a planar peptide bond. In one, the two α-carbons are *cis* to each other; in the other, they are *trans* to each other. The *trans* configuration is more favorable because the α-carbons with the bulky groups bonded to them are farther from each other than they are in the *cis* configuration. Almost all peptide bonds in naturally occurring proteins studied to date have the *trans* configuration. Proline is found *cis* most of the time, and there are some well-known examples of other *cis* peptide bonds as well.



s-*trans* configuration



s-*cis* configuration

Secondary structure of proteins
The ordered arrangements (conformations) of amino acids in localized regions of a polypeptide or protein.

B. Secondary Structure

Secondary (2°) structure refers to ordered arrangements (conformations) of amino acids in localized regions of a polypeptide or protein molecule. The first studies of polypeptide conformations were carried out by Linus Pauling and Robert Corey beginning in 1939. They assumed that in conformations of greatest stability, all atoms in a peptide bond lie in the same plane, and there is hydrogen bonding between the N—H of one peptide bond and the C=O of another, as shown in Figure 27.13.

On the basis of model building, Pauling proposed that two types of secondary structure should be particularly stable: the α-helix and the antiparallel β-pleated sheet. X-ray crystallography has validated this prediction completely.

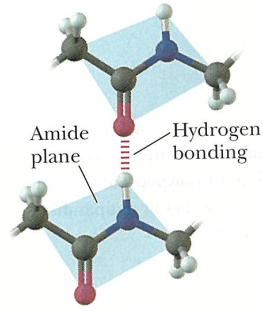


Figure 27.13
Hydrogen bonding between amide groups.

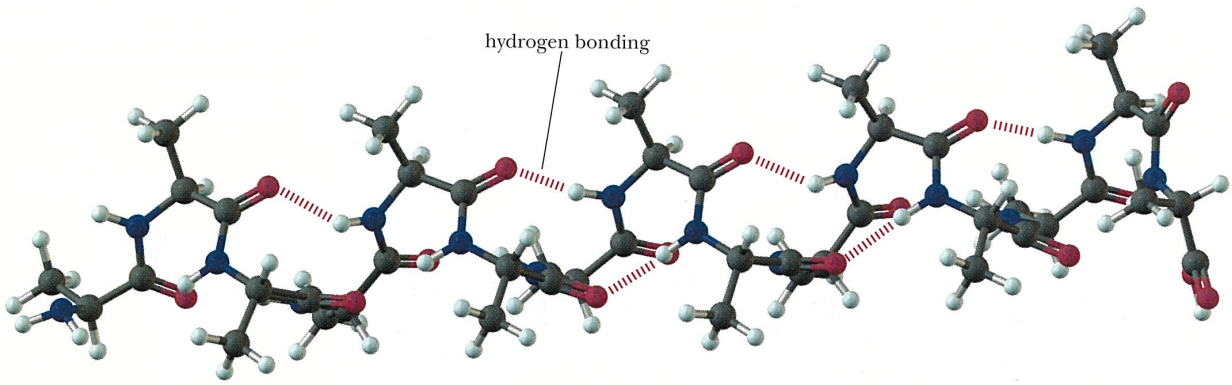


Figure 27.14
An α-helix. The peptide chain is repeating units of L-alanine.

The α-Helix

α-Helix A type of secondary structure in which a section of polypeptide chain coils into a spiral, most commonly a right-handed spiral.

In an **α-helix** pattern shown in Figure 27.14, a polypeptide chain is coiled in a spiral. As you study this section of α-helix, note the following.

1. The helix is coiled in a clockwise, or right-handed, manner. Right-handed means that if you turn the helix clockwise, it twists away from you. In this sense, a right-handed helix is analogous to the right-handed thread of a common wood or machine screw.
2. There are 3.6 amino acids per turn of the helix.
3. Each peptide bond is *trans* and planar.
4. The N—H group of each peptide bond points roughly downward, parallel to the axis of the helix, and the C=O of each peptide bond points roughly upward, also parallel to the axis of the helix.
5. The carbonyl group of each peptide bond is hydrogen-bonded to the N—H group of the peptide bond four amino acid units away from it. Hydrogen bonds are shown as dashed lines.
6. All R— groups point outward from the helix.

Almost immediately after Pauling proposed the α-helix conformation, other researchers proved the presence of α-helix conformations in keratin, the protein of hair and wool. It soon became obvious that the α-helix is one of the fundamental folding patterns of polypeptide chains.

The β-Pleated Sheet

β-Pleated sheet A type of secondary structure in which sections of polypeptide chains are aligned parallel or antiparallel to one another.

An antiparallel **β-pleated sheet** consists of extended polypeptide chains with neighboring chains running in opposite (antiparallel) directions. In a parallel β-pleated sheet, the polypeptide chains run in the same direction. Unlike the α-helix arrangement, N—H and C=O groups lie in the plane of the sheet and are roughly perpendicular to the long axis of the sheet. The C=O group of each peptide bond is hydrogen-bonded to the N—H group of a peptide bond of a neighboring chain (Figure 27.15).

As you study this section of β-pleated sheet, note the following.

1. The three polypeptide chains lie adjacent to each other and run in opposite (antiparallel) directions.
2. Each peptide bond is planar, and the α-carbons are *trans* to each other.

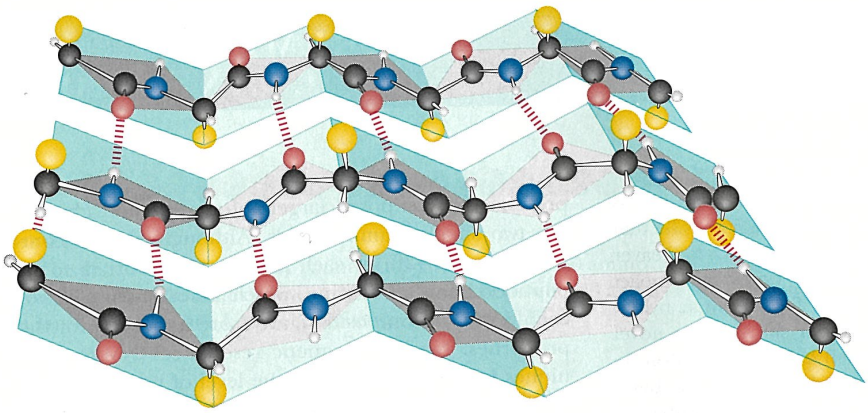


Figure 27.15
β-Pleated sheet conformation with three polypeptide chains running in opposite (antiparallel) directions. Hydrogen bonding between chains is indicated by dashed lines.

3. The C=O and N—H groups of peptide bonds from adjacent chains point at each other and are in the same plane so that hydrogen bonding is possible between adjacent polypeptide chains.
4. The R-groups on any one chain alternate, first above and then below the plane of the sheet, and so on.

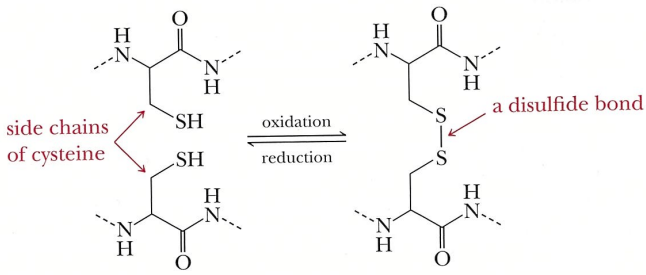
The β-pleated sheet conformation is stabilized by hydrogen bonding between N—H groups of one chain and C=O groups of an adjacent chain. By comparison, the α-helix is stabilized by hydrogen bonding between N—H and C=O groups within the same polypeptide chain.

C. Tertiary Structure

Tertiary structure of proteins
The three-dimensional arrangement in space of all atoms in a single polypeptide chain.

Tertiary (3°) structure refers to the overall folding pattern and arrangement in space of all atoms in a single polypeptide chain. No sharp dividing line exists between secondary and tertiary structures. Secondary structure refers to the spatial arrangement of amino acids close to one another on a polypeptide chain, whereas tertiary structure refers to the three-dimensional arrangement of all atoms of a polypeptide chain. Among the most important factors in maintaining 3° structure are disulfide bonds, hydrophobic interactions, hydrogen bonding, and salt linkages.

Disulfide bonds (Section 10.9G) play an important role in maintaining tertiary structure. Disulfide bonds are formed between side chains of two cysteine units by oxidation of their thiol groups (—SH) to form a disulfide bond. Treatment of a disulfide bond with a reducing agent regenerates the thiol groups.



CHEMICAL CONNECTIONS

Spider Silk

Spider silk has some remarkable properties. Research is currently concentrated on the strong dragline silk that forms the spokes of a web of the Golden Orb Weaver (*Nephila clavipes*). This silk has three times the impact strength of Kevlar and is 30% more flexible than nylon. The commercial application of spider silk is not a novel concept. Eighteenth-century French entrepreneur Bon de Saint-Hilaire attempted to mass-produce silk in his high-density spider farms but failed because of cannibalism among his territorial arachnid workers. In contrast, native New Guineans continue to successfully collect and utilize spider silk for a wide range of applications including bags and fishing nets. Today, the only way to obtain large amounts of silk is to extract it from the abdomens of immobilized spiders, but scientific advances make the mass production and industrial application of spider silk increasingly possible.

Biologically produced dragline silk is a combination of two liquid proteins, Spidroin 1 and 2, which become oriented and solidify as they travel through a complex duct system in the spider's abdomen. These proteins are composed largely of alanine and glycine, the two smallest amino acids. Although glycine comprises almost 42% of each protein, the short, 5 to 10 peptide chains of alanine, which account for 25% of each protein's composition, are more important for the properties. Nuclear magnetic resonance (NMR) techniques have vastly improved the level of understanding of spider silk's structure, which was originally determined by x-ray crystallography. NMR data of spidroins containing deuterium-tagged alanine have shown that all alanines are configured into β-pleated sheets. Furthermore, the NMR data suggest that 40% of the alanine β-sheets are highly structured while the other 60% are less oriented, forming fingers that reach out from each individual strand. These fingers are believed to join the oriented alanine β-sheets and the glycine-rich, amorphous "background" sectors of the polypeptide.

Currently, genetically modified *Escherichia coli* is used to mass-produce Spidroin 1 and 2. However, DNA redundancy initially caused synthesis problems when the spider genes were transposed into the bacteria. The *E. coli* did not transcribe some of the codons in the same way that spider cells would, forcing scientists to modify the DNA. When the proteins could be synthesized, it was necessary to develop a system to mimic the natural production of spider silk while preventing the silk from contacting the air and subsequently hardening. After the two proteins are separated from the *E. coli*, they are drawn together into methanol through separate needles. Another approach is to dissolve the silk in formic acid or to add codons for hydrophilic amino acids, in this case histidine and arginine, to keep the artificial silk pliable. The industrial and practical applications of spider silk will not be fully known until it can be abiotically synthesized in large quantities.



Tom Bean/Stone/Getty

Golden orb weaver.

Based on a Chem 30H honors paper by Paul Celestre, UCLA.

Figure 27.16 shows the amino acid sequence of human insulin. This protein consists of two polypeptide chains: an A chain of 21 amino acids and a B chain of 30 amino acids. The A chain is bonded to the B chain by two interchain disulfide bonds. An intrachain disulfide bond also connects the cysteine units at positions 6 and 11 of the A chain.



Figure 27.16
Human insulin. The A chain of 21 amino acids and B chain of 30 amino acids are connected by interchain disulfide bonds between A7 and B7 and between A20 and B19. In addition, a single intrachain disulfide bond occurs between A6 and A11.

As an example of 2° and 3° structure, let us look at the three-dimensional structure of myoglobin—a protein found in skeletal muscle and particularly abundant in diving mammals, such as seals, whales, and porpoises. Myoglobin and its structural relative, hemoglobin, are the oxygen storage and transport molecules of vertebrates. Hemoglobin binds molecular oxygen in the lungs and transports it to myoglobin in muscles. Myoglobin stores molecular oxygen until it is required for metabolic oxidation.

Myoglobin consists of a single polypeptide chain of 153 amino acids. Myoglobin also contains a single heme unit. Heme consists of one Fe^{2+} ion coordinated in a square planar array with the four nitrogen atoms of a molecule of porphyrin (Figure 27.17).

Determination of the three-dimensional structure of myoglobin represented a milestone in the study of molecular architecture. For their contribution to this research, John C. Kendrew and Max F. Perutz, both of Britain, shared the 1962 Nobel Prize for chemistry. The secondary and tertiary structures of myoglobin are shown in Figure 27.18. The single polypeptide chain is folded into a complex, almost boxlike shape.

Following are important structural features of the three-dimensional shape of myoglobin.

1. The backbone consists of eight relatively straight sections of α-helix, each separated by a bend in the polypeptide chain. The longest section of α-helix has 24 amino acids, the shortest has 7. Some 75% of the amino acids are found in these eight regions of α-helix.

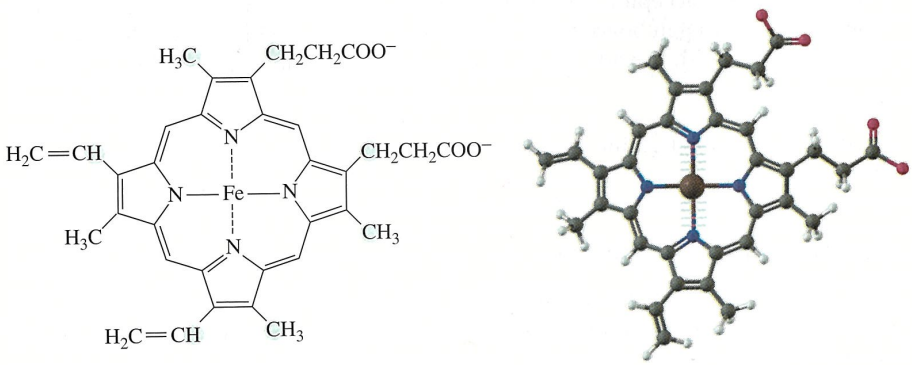
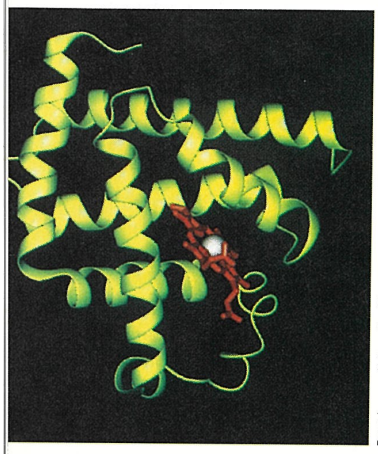


Figure 27.17
The structure of heme, found in myoglobin and hemoglobin.



Brent Iverson, University of Texas

Figure 27.18
Ribbon model of myoglobin. The polypeptide chain is shown in yellow, the heme ligand in red, and the Fe atom as a white sphere.

2. Hydrophobic side chains of phenylalanine, alanine, valine, leucine, isoleucine, and methionine are clustered in the interior of the molecule where they are shielded from contact with water. **Hydrophobic interactions** are a major factor in directing the folding of the polypeptide chain of myoglobin into this compact, three-dimensional shape.
3. The outer surface of myoglobin is coated with hydrophilic side chains, such as those of lysine, arginine, serine, glutamic acid, histidine, and glutamine, which interact with the aqueous environment by **hydrogen bonding.** The only polar side chains that point to the interior of the myoglobin molecule are those of two histidine units, which point inward toward the heme group.
4. Oppositely charged amino acid side chains close to each other in the three-dimensional structure interact by electrostatic attractions called **salt linkages.** An example of a salt linkage is the attraction of the side chains of lysine ($—NH_3^+$) and glutamic acid ($—COO^-$).

The tertiary structures of hundreds of proteins have also been determined. It is clear that proteins contain α-helix and β-pleated sheet structures, but that wide variations exist in the relative amounts of each. Lysozyme, with 129 amino acids in a single polypeptide chain, has only 25% of its amino acids in α-helix regions. Cytochrome, with 104 amino acids in a single polypeptide chain, has no α-helix structure but does contain several regions of β-pleated sheet. Yet, whatever the proportions of α-helix, β-pleated sheet, or other periodic structure, most nonpolar side chains of water-soluble proteins are directed toward the interior of the molecule, whereas polar side chains are on the surface of the molecule and in contact with the aqueous environment. Note that this arrangement of polar and nonpolar groups in water-soluble proteins very much resembles the arrangement of polar and nonpolar groups of soap molecules in micelles (Figure 26.3). It also resembles the arrangement of phospholipids in lipid bilayers (Figure 26.13).

Example 27.7

With which of the following amino acid side chains can the side chain of threonine form hydrogen bonds?

(a) Valine **(b)** Asparagine **(c)** Phenylalanine
(d) Histidine **(e)** Tyrosine **(f)** Alanine

Solution

The side chain of threonine contains a hydroxyl group that can participate in hydrogen bonding in two ways: Its oxygen has a partial negative charge and can function as a hydrogen bond acceptor, and its hydrogen has a partial positive charge and can function as a hydrogen bond donor. Therefore, the side chain of threonine can form hydrogen bonds with the side chains of tyrosine, asparagine, and histidine.

Problem 27.7

At pH 7.4, with what amino acid side chains can the side chain of lysine form salt linkages?

D. Quaternary Structure

Quaternary structure The arrangement of polypeptide monomers into a noncovalently bonded aggregate.

Most proteins of molecular weight greater than 50,000 consist of two or more noncovalently linked polypeptide chains. The arrangement of protein monomers into an aggregation is known as **quaternary (4°) structure.** A good example is hemoglobin, a

protein that consists of four separate polypeptide chains: two α-chains of 141 amino acids each and two β-chains of 146 amino acids each. The quaternary structure of hemoglobin is shown in Figure 27.19.

A major factor stabilizing the aggregation of protein subunits is the **hydrophobic effect.** When separate polypeptide chains fold into compact three-dimensional shapes to expose polar side chains to the aqueous environment and shield nonpolar side chains from water, hydrophobic "patches" still appear on the surface, in contact with water. These patches can be shielded from water if two or more monomers assemble so that their hydrophobic patches are in contact. The numbers of subunits of several proteins of known quaternary structure are shown in Table 27.4. Other important factors include correctly located complementary hydrogen bonding and charged sites on different subunits. The formation of aggregates of well-defined structure based on specific structural units on the subunits is being explored in the new field of molecular recognition.

Hydrophobic effect The tendency of nonpolar groups to cluster to shield themselves from contact with an aqueous environment.

Table 27.4 Quaternary Structure of Selected Proteins

Protein	Number of Subunits
Alcohol dehydrogenase	2
Aldolase	4
Hemoglobin	4
Lactate dehydrogenase	4
Insulin	6
Glutamine synthetase	12
Tobacco mosaic virus protein disc	17



Brent Iverson, University of Texas

Figure 27.19
Ribbon model of hemoglobin. The α-chains are shown in purple, the β-chains in yellow, the heme ligands in red, and the Fe atoms as white spheres.

Summary

Amino acids are compounds that contain both an amino group and a carboxyl group (Section 27.1A). A **zwitterion** is an internal salt of an amino acid. With the exception of glycine, all protein-derived amino acids are chiral (Section 27.1B). In the D,L convention, all are L-amino acids. In the *R*,*S* convention, 18 are (*S*)-amino acids. Although cysteine has the same absolute configuration, it is an (*R*)-amino acid because of the manner in which priorities are assigned about the tetrahedral chiral center. Isoleucine and threonine contain a second chiral center. The 20 protein-derived amino acids are commonly divided into four categories (Section 27.1C): nine with nonpolar side chains, four with polar but un-ionized side chains, four with acidic side chains, and three with basic side chains.

The **isoelectric point, pI,** of an amino acid, polypeptide, or protein is the pH at which it has no net charge (Section 27.2C). **Electrophoresis** is the process of separating compounds on the basis of their electric charge (Section 27.2D). Compounds having a high charge density move more rapidly than those with a lower charge density. Any amino acid or protein in a solution with a pH that equals the pI of the compound remains at the origin.

A **peptide bond** is the special name given to the amide bond formed between α-amino acids (Section 27.3). A **polypeptide** is a biological macromolecule containing many amino acids, each joined to the next by a peptide bond. By convention, the sequence of amino acids in a polypeptide is written beginning with the ***N*-terminal amino acid** toward the ***C*-terminal amino acid. Primary (1°) structure** of a polypeptide is the sequence of amino acids in the polypeptide chain (Section 27.4).

In solid-phase synthesis (Section 27.5F), or polymer-supported synthesis of polypeptides, the *C*-terminal amino acid is joined to a chloromethylated polystyrene resin as a benzyl ester. The polypeptide chain is then extended one amino acid at a time from the *N*-terminal end. When synthesis is completed,

the polypeptide chain is released from the solid support by cleavage of the benzyl ester.

A peptide bond is planar (Section 27.6A), that is, the four atoms of the amide and the two α-carbons of a peptide bond lie in the same plane. Bond angles about the amide nitrogen and the amide carbonyl carbon are 120°. **Secondary (2°) structure** (Section 26.7B) refers to the ordered arrangement (conformations) of amino acids in localized regions of a polypeptide or protein. Two types of secondary structure are the α-helix and the β-pleated sheet. **Tertiary (3°) structure** (Section 27.6C) refers to the overall folding pattern and arrangement in space of all atoms in a single polypeptide chain. **Quaternary (4°) structure** (Section 27.6D) is the arrangement of polypeptide monomers into a noncovalently bonded aggregate.

Key Reactions

1. Acidity of an α-Carboxyl Group (Section 27.2A)

An α-COOH (pK_a approximately 2.19) of a protonated amino acid is a considerably stronger acid than acetic acid (pK_a 4.76) or other low-molecular-weight aliphatic carboxylic acid owing to the electron-withdrawing inductive effect of the α-NH_3^+ group.

$$\underset{\displaystyle NH_3^+}{\overset{\displaystyle |}{RCHCOOH}} + H_2O \rightleftharpoons \underset{\displaystyle NH_3^+}{\overset{\displaystyle |}{RCHCOO^-}} + H_3O^+ \qquad pK_a = 2.19$$

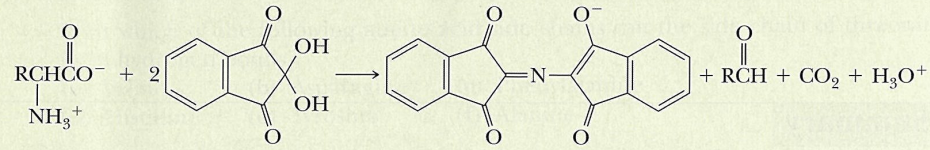
2. Acidity of an α-Ammonium Group (Section 27.2A)

An α-NH_3^+ group (pK_a approximately 9.47) is a slightly stronger acid than a primary aliphatic ammonium ion (pK_a approximately 10.76).

$$\underset{\displaystyle NH_3^+}{\overset{\displaystyle |}{RCHCOO^-}} + H_2O \rightleftharpoons \underset{\displaystyle NH_2}{\overset{\displaystyle |}{RCHCOO^-}} + H_3O^+ \qquad pK_a = 9.47$$

3. Reaction of an α-Amino Acid with Ninhydrin (Section 27.2D)

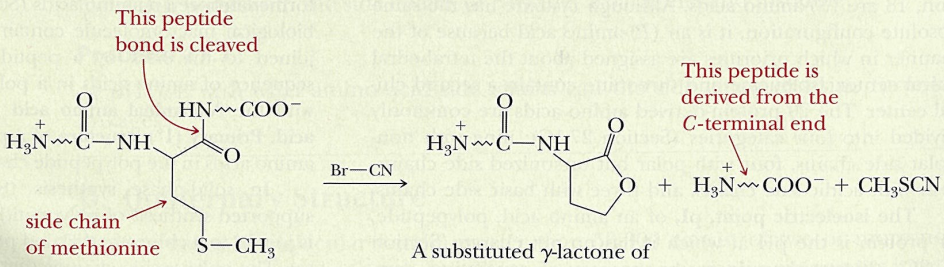
Treatment of an α-amino acid with ninhydrin gives a purple-colored solution. Treatment of proline with ninhydrin gives an orange-colored solution.



An α-amino acid — Ninhydrin — Purple-colored anion

4. Cleavage of a Peptide Bond by Cyanogen Bromide (Section 27.4B)

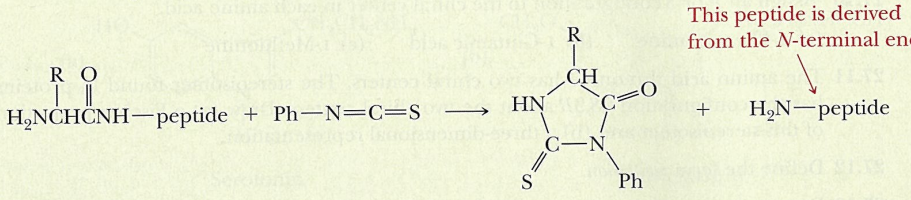
Cleavage is regioselective for a peptide bond formed by the carboxyl group of methionine.



A substituted γ-lactone of the amino acid homoserine

5. Edman Degradation (Section 27.4B)

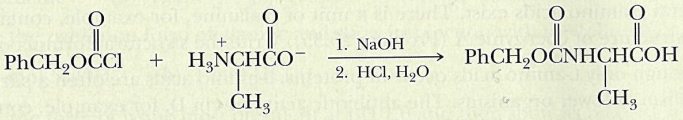
Treatment with phenyl isothiocyanate followed by acid removes the *N*-terminal amino acid as a substituted phenylthiohydantoin, which is then separated and identified.



Phenyl isothiocyanate — A phenylthiohydantoin

6. The Benzyloxycarbonyl (*Z*—) Protecting Group (Section 27.5C)

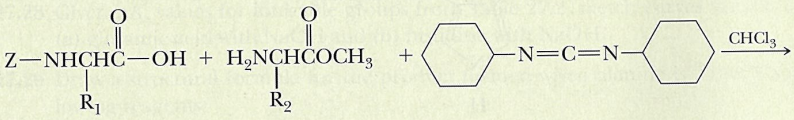
The benzyloxycarbonyl protecting group is prepared by treatment of an unprotected α-NH_2 group with benzyloxycarbonyl chloride. It is removed by treatment with HBr in acetic acid or by hydrogenolysis.



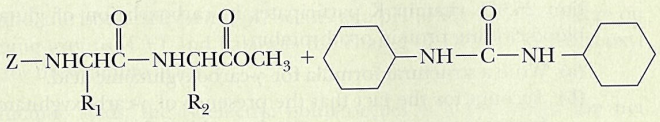
Benzyloxycarbonyl chloride (Z—Cl) — Alanine — *N*-Benzyloxycarbonylalanine (Z—Ala)

7. Peptide Bond Formation Using 1,3-Dicyclohexylcarbodiimide (Section 27.5E)

This substituted carbodiimide is a dehydrating agent and is converted to a disubstituted urea. The reaction is efficient and yields are generally very high.



Amino protected aa_1 — Carboxyl protected aa_2 — 1,3-Dicyclohexylcarbodiimide (DCC)



Amino and carboxyl protected dipeptide — *N*,*N*′-Dicyclohexylurea (DCU)

Problems

Amino Acids

27.8 What amino acid does each abbreviation stand for?

(a) Phe **(b)** Ser **(c)** Asp **(d)** Gln
(e) His **(f)** Gly **(g)** Tyr

ORGANIC Chemistry Now™
Assess your understanding of this chapter's topics with additional quizzing and conceptual-based problems at **http://org.brookscole.com/bfi4**

27.9 The configuration of the chiral center in α-amino acids is most commonly specified using the D,L convention. It can also be identified using the *R,S* convention (Section 3.3). Does the chiral center in L-serine have the *R* or the *S* configuration?

27.10 Assign an *R* or *S* configuration to the chiral center in each amino acid.

(a) L-Phenylalanine **(b)** L-Glutamic acid **(c)** L-Methionine

27.11 The amino acid threonine has two chiral centers. The stereoisomer found in proteins has the configuration 2*S*,3*R* about the two chiral centers. Draw (a) a Fischer projection of this stereoisomer and (b) a three-dimensional representation.

27.12 Define the term *zwitterion*.

27.13 Draw zwitterion forms of these amino acids.

(a) Valine **(b)** Phenylalanine **(c)** Glutamine

27.14 Why are Glu and Asp often referred to as acidic amino acids?

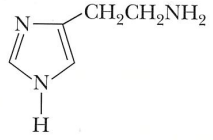
27.15 Why is Arg often referred to as a basic amino acid? Which two other amino acids are also basic amino acids?

27.16 What is the meaning of the alpha as it is used in α-amino acid?

27.17 Several β-amino acids exist. There is a unit of β-alanine, for example, contained within the structure of coenzyme A (Problem 25.35). Write the structural formula of β-alanine.

27.18 Although only L-amino acids occur in proteins, D-amino acids are often a part of the metabolism of lower organisms. The antibiotic actinomycin D, for example, contains a unit of D-valine, and the antibiotic bacitracin A contains units of D-asparagine and D-glutamic acid. Draw Fischer projections and three-dimensional representations for these three D-amino acids.

27.19 Histamine is synthesized from one of the 20 protein-derived amino acids. Suggest which amino acid is its biochemical precursor and the type of organic reaction(s) involved in its biosynthesis (for example, oxidation, reduction, decarboxylation, nucleophilic substitution).



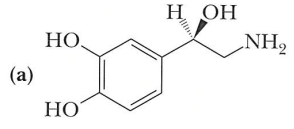
Histamine

27.20 As discussed in Chemical Connections: "Vitamin K, Blood Clotting, and Basicity" in Section 26.6D, vitamin K participates in carboxylation of glutamic acid residues of the blood-clotting protein prothrombin.

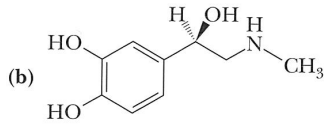
(a) Write a structural formula for γ-carboxyglutamic acid.

(b) Account for the fact that the presence of γ-carboxyglutamic acid escaped detection for many years; on routine amino acid analyses, only glutamic acid was detected.

27.21 Both norepinephrine and epinephrine are synthesized from the same protein-derived amino acid. From which amino acid are they synthesized, and what types of reactions are involved in their biosynthesis?

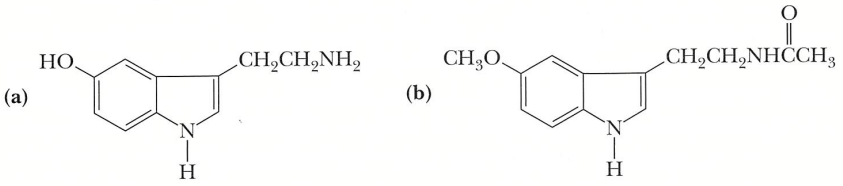


Norepinephrine



Epinephrine
(Adrenaline)

27.22 From which amino acid are serotonin and melatonin synthesized, and what types of reactions are involved in their biosynthesis?



Serotonin Melatonin

Acid-Base Behavior of Amino Acids

27.23 Draw a structural formula for the form of each amino acid most prevalent at pH 1.0.

(a) Threonine **(b)** Arginine **(c)** Methionine **(d)** Tyrosine

27.24 Draw a structural formula for the form of each amino most prevalent at pH 10.0.

(a) Leucine **(b)** Valine **(c)** Proline **(d)** Aspartic acid

27.25 Write the zwitterion form of alanine and show its reaction with the following.

(a) 1 mol NaOH **(b)** 1 mol HCl

27.26 Write the form of lysine most prevalent at pH 1.0 and then show its reaction with the following. Consult Table 27.2 for pK_a values of the ionizable groups in lysine.

(a) 1 mol NaOH **(b)** 2 mol NaOH **(c)** 3 mol NaOH

27.27 Write the form of aspartic acid most prevalent at pH 1.0 and then show its reaction with the following. Consult Table 27.2 for pK_a values of the ionizable groups in aspartic acid.

(a) 1 mol NaOH **(b)** 2 mol NaOH **(c)** 3 mol NaOH

27.28 Given pK_a values for ionizable groups from Table 27.2, sketch curves for the titration of **(a)** glutamic acid with NaOH and **(b)** histidine with NaOH.

27.29 Draw a structural formula for the product formed when alanine is treated with the following reagents.

(a) Aqueous NaOH **(b)** Aqueous HCl
(b) CH_3CH_2OH, H_2SO_4 **(b)** $(CH_3CO)_2O$, CH_3COONa

27.30 For lysine and arginine, the isoelectric point, pI, occurs at a pH where the net charge on the nitrogen-containing groups is +1 and balances the charge of −1 on the α-carboxyl group. Calculate pI for these amino acids.

27.31 For aspartic and glutamic acids, the isoelectric point occurs at a pH where the net charge on the two carboxyl groups is −1 and balances the charge of +1 on the α-amino group. Calculate pI for these amino acids.

27.32 Account for the fact that the isoelectric point of glutamine (pI 5.65) is higher than the isoelectric point of glutamic acid (pI 3.08).

27.33 Enzyme-catalyzed decarboxylation of glutamic acid gives 4-aminobutanoic acid (Section 27.1D). Estimate the pI of 4-aminobutanoic acid.

27.34 Guanidine and the guanidino group present in arginine are two of the strongest organic bases known. Account for their basicity.

27.35 At pH 7.4, the pH of blood plasma, do the majority of protein-derived amino acids bear a net negative charge or a net positive charge? Explain.

27.36 Do the following compounds migrate to the cathode or to the anode on electrophoresis at the specified pH?

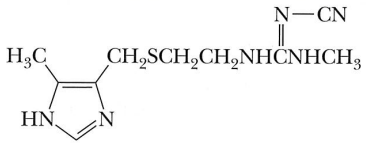
(a) Histidine at pH 6.8 **(b)** Lysine at pH 6.8
(c) Glutamic acid at pH 4.0 **(d)** Glutamine at pH 4.0
(e) Glu-Ile-Val at pH 6.0 **(f)** Lys-Gln-Tyr at pH 6.0

27.37 At what pH would you carry out an electrophoresis to separate the amino acids in each mixture?

(a) Ala, His, Lys **(b)** Glu, Gln, Asp **(c)** Lys, Leu, Tyr

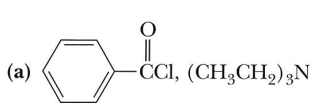
27.38 Examine the amino acid sequence of human insulin (Figure 27.16), and list each Asp, Glu, His, Lys, and Arg in this molecule. Do you expect human insulin to have an isoelectric point nearer that of the acidic amino acids (pI 2.0–3.0), the neutral amino acids (pI 5.5–6.5), or the basic amino acids (pI 9.5–11.0)?

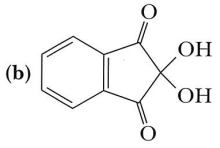
27.39 A chemically modified guanidino group is present in cimetidine (Tagamet), a widely prescribed drug for the control of gastric acidity and peptic ulcers. Cimetidine reduces gastric acid secretion by inhibiting the interaction of histamine with gastric H_2 receptors. In the development of this drug, a cyano group was added to the substituted guanidino group to alter its basicity. Do you expect this modified guanidino group to be more basic or less basic than the guanidino group of arginine? Explain.

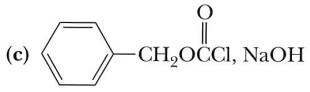


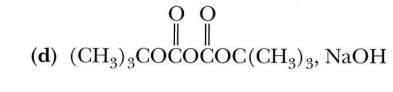
Cimetidine
(Tagamet)

27.40 Draw a structural formula for the product formed when alanine is treated with the following reagents.









(e) Product(c) + L-Alanine ethyl ester + DCC
(f) Product(d) + L-Alanine ethyl ester + DCC

Primary Structure of Polypeptides and Proteins

27.41 If a protein contains four different SH groups, how many different disulfide bonds are possible if only a single disulfide bond is formed? How many different disulfides are possible if two disulfide bonds are formed?

27.42 How many different tetrapeptides can be made under the following conditions?

(a) The tetrapeptide contains one unit each of Asp, Glu, Pro, and Phe.
(b) All 20 amino acids can be used, but each only once.

27.43 A decapeptide has the following amino acid composition.

Ala_2, Arg, Cys, Glu, Gly, Leu, Lys, Phe, Val

Partial hydrolysis yields the following tripeptides.

Cys-Glu-Leu + Gly-Arg-Cys + Leu-Ala-Ala + Lys-Val-Phe + Val-Phe-Gly

One round of Edman degradation yields a lysine phenylthiohydantoin. From this information, deduce the primary structure of this decapeptide.

27.44 Following is the primary structure of glucagon, a polypeptide hormone of 29 amino acids. Glucagon is produced in the α-cells of the pancreas and helps maintain blood glucose levels in a normal concentration range.

1 5 10 15 20 25 29
His-Ser-Glu-Gly-Thr-Phe-Thr-Ser-Asp-Tyr-Ser-Lys-Tyr-Leu-Asp-Ser-Arg-Arg-Ala-Gln-Asp-Phe-Val-Gln-Trp-Leu-Met-Asn-Thr

Which peptide bonds are hydrolyzed when this polypeptide is treated with each reagent?

(a) Phenyl isothiocyanate **(b)** Chymotrypsin **(c)** Trypsin **(d)** Br—CN

27.45 A tetradecapeptide (14 amino acid residues) gives the following peptide fragments on partial hydrolysis. From this information, deduce the primary structure of this polypeptide. Fragments are grouped according to size.

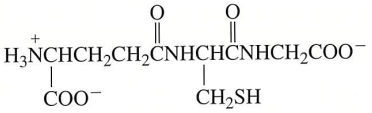
Pentapeptide Fragments	Tetrapeptide Fragments
Phe-Val-Asn-Gln-His	Gln-His-Leu-Cys
His-Leu-Cys-Gly-Ser	His-Leu-Val-Glu
Gly-Ser-His-Leu-Val	Leu-Val-Glu-Ala

27.46 Draw a structural formula of these tripeptides. Mark each peptide bond, the *N*-terminal amino acid, and the *C*-terminal amino acid.

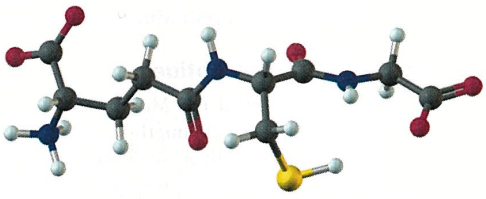
(a) Phe-Val-Asn **(b)** Leu-Val-Gln

27.47 Estimate the pI of each tripeptide on Problem 27.46.

27.48 Glutathione (G-SH), one of the most common tripeptides in animals, plants, and bacteria, is a scavenger of oxidizing agents. In reacting with oxidizing agents, glutathione is converted to G-S-S-G.



Glutathione



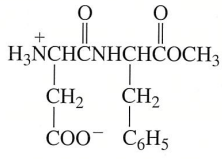
(a) Name the amino acids in this tripeptide.
(b) What is unusual about the peptide bond formed by the *N*-terminal amino acid?
(c) Write a balanced half-reaction for the reaction of two molecules of glutathione to form a disulfide bond. Is glutathione a biological oxidizing agent or a biological reducing agent?
(d) Write a balanced equation for reaction of glutathione with molecular oxygen, O_2, to form G-S-S-G and H_2O. Is molecular oxygen oxidized or reduced in this process?

27.49 Following are a structural formula and a ball-and-stick model for the artificial sweetener aspartame. Each amino acid has the L configuration.

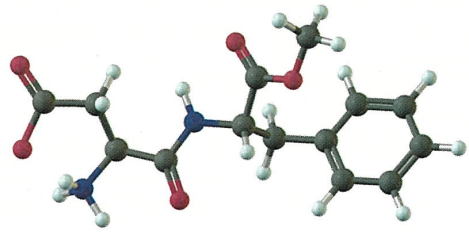


Charles D. Winters

Aspartame is present in many artificially sweetened foods and beverages.

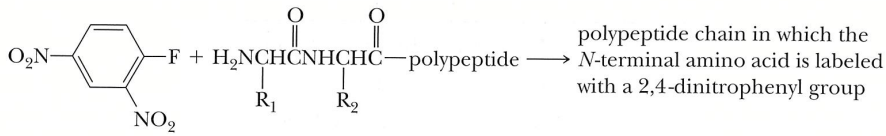


Aspartame



(a) Name the two amino acids in this molecule.
(b) Estimate the isoelectric point of aspartame.
(c) Draw structural formulas for the products of hydrolysis of aspartame in 1 *M* HCl.

27.50 2,4-Dinitrofluorobenzene, very often known as Sanger's reagent after the English chemist Frederick Sanger who popularized its use, reacts selectively with the *N*-terminal amino group of a polypeptide chain. Sanger was awarded the 1958 Nobel Prize for chemistry for his work in determining the primary structure of bovine insulin. One of the few persons to be awarded two Nobel Prizes, he also shared the 1980 award in chemistry with American chemists, Paul Berg and Walter Gilbert, for the development of chemical and biological analyses of DNAs.



2,4-Dinitrofluorobenzene

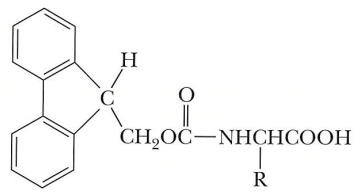
(*N*-Terminal end of a polypeptide chain)

Following reaction with 2,4-dinitrofluorobenzene, all amide bonds of the polypeptide chain are hydrolyzed, and the amino acid labeled with a 2,4-dinitrophenyl group is separated by either paper or column chromatography and identified.

(a) Write a structural formula for the product formed by treatment of the *N*-terminal amino group with Sanger's reagent and propose a mechanism for its formation.
(b) When bovine insulin is treated with Sanger's reagent followed by hydrolysis of all peptide bonds, two labeled amino acids are detected: glycine and phenylalanine. What conclusions can be drawn from this information about the primary structure of bovine insulin?
(c) Compare and contrast the structural information that can be obtained from use of Sanger's reagent with that from use of the Edman degradation.

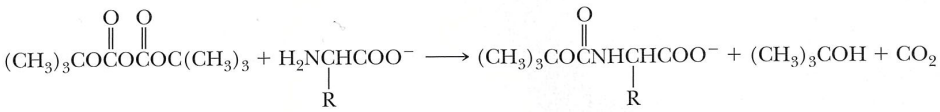
Synthesis of Polypeptides

27.51 In a variation of the Merrifield solid-phase peptide synthesis, the amino group is protected by a fluorenylmethoxycarbonyl (FMOC) group. This protecting group is removed by treatment with a weak base such as the secondary amine, piperidine. Write a balanced equation and propose a mechanism for this deprotection.



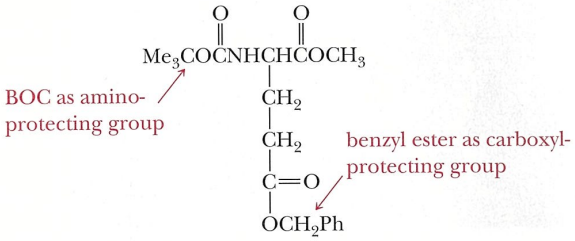
Fluorenylmethoxycarbonyl (FMOC) group

27.52 The BOC-protecting group may be added by treatment of an amino acid with di-*tert*-butyl dicarbonate as shown in the following reaction sequence. Propose a mechanism to account for formation of these products.



Di-*tert*-butyl dicarbonate BOC-amino acid

27.53 The side-chain carboxyl groups of aspartic acid and glutamic acid are often protected as benzyl esters.



(a) Show how to convert the side-chain carboxyl group to a benzyl ester using benzyl chloride as a source of the benzyl group.

(b) How do you deprotect the side-chain carboxyl under mild conditions without removing the BOC-protecting group at the same time?

Three-Dimensional Shapes of Polypeptides and Proteins

27.54 Examine the α-helix conformation. Are amino acid side chains arranged all inside the helix, all outside the helix, or randomly?

27.55 Distinguish between intermolecular and intramolecular hydrogen bonding between the backbone groups on polypeptide chains. In what type of secondary structure do you find intermolecular hydrogen bonds? In what type do you find intramolecular hydrogen bonding?

27.56 Many plasma proteins found in an aqueous environment are globular in shape. Which amino acid side chains would you expect to find on the surface of a globular protein and in contact with the aqueous environment? Which would you expect to find inside, shielded from the aqueous environment? Explain.

(a) Leu **(b)** Arg **(c)** Ser **(d)** Lys **(e)** Phe

27.57 Denaturation of a protein is a physical change, the most readily observable result of which is loss of biological activity. Denaturation stems from changes in secondary, tertiary, and quaternary structure through disruption of noncovalent interactions including hydrogen bonding and hydrophobic interactions. Three common denaturing agents are sodium dodecyl sulfate (SDS), urea, and heat. What kinds of noncovalent interactions might each reagent disrupt?