Chapter 26
Amino Acids, Peptides, and Proteins

Solutions to In-Text Problems

26.1 The peptide is drawn in the form that exists at neutral pH.

(b) The structure of Glu-Gln-Phe-Arg:

26.3 (b) L-Alloisoleucine has the 2S configuration, that is, the same configuration at the α-carbon as L-leucine, but has a different configuration at the β-carbon. D-Alloisoleucine is the enantiomer of L-alloisoleucine.

26.4 (b) In terms of the sequence rules for determining configuration, the two carbons attached to the α-carbons of cysteine or serine can be represented as C(O,O,O) and C(X,H,H), where X = S or O, respectively. When X = S (cysteine), the latter carbon receives priority. When X = O (serine), the former receives priority. Sulfur receives priority over oxygen because it has higher atomic number, and the priority is decided at the first point of difference. (See Sec. 4.2B, text p. 134.)

26.5 (b)

26.6 (b) In the major neutral form of G-D-G-L-F (Gly-Asp-Gly-Leu-Phe), the side-chain carboxy group of Asp is unionized and the carboxy-terminal carboxy group is ionized because it has the lower pKₐ of the two carboxy
groups (Table 26.1, text p. 1268). This form would exist at a pH between the $pK_a$ values of the two carboxy groups. That is, the pH at which this neutral form predominates is relatively acidic. (See Sec. 26.3B.) Notice that the major neutral form of a peptide (such as this one) may not be the form that predominates at neutral pH.

26.7 (a) The amino group of tyrosine can be protonated, and both the carboxy group and the phenolic O—H group can be ionized.

(b) At pH 6, the net charge on tyrosine is zero. A pH value of 6 is below the $pK_a$ of the conjugate acid of the amino group, and the amino group is therefore protonated; a pH value of 6 is above the $pK_a$ of the carboxy group, which is therefore ionized; and a pH value of 6 is below the $pK_a$ of the phenolic O—H group, which is therefore un-ionized.

(c) The structure of tyrosine in aqueous solution at pH 6:

26.9 The general rule of thumb is that if the peptide contains more acidic than basic groups it is an acidic peptide; if it contains more basic than acidic groups it is a basic peptide; and if the number of acidic and basic groups are equal, the peptide is neutral. (An acidic group is a group that is in its conjugate-base form at neutral pH, such as a carboxy group; a basic group is a group that is in its conjugate-acid form at neutral pH, such as an amino group.)

(b) This peptide contains three basic groups: the $\alpha$-amino group of the amino-terminal residue Leu and the side-chain amino groups of the Lys residues. The peptide contains one acidic group: the carboxy group of the carboxy-terminal Lys residue. Because the peptide contains more basic residues than acidic residues; it is a basic peptide. Its net charge at pH 6 is +2.

(d) This peptide contains three acidic groups: the side-chain carboxy groups of the Asp and Glu residues and the carboxy group of the carboxy-terminal Ile residue. The peptide contains two basic groups: the $\alpha$-amino group of the amino-terminal residue Glu and the side-chain amino group of the Lys residue. Because the peptide contains more acidic residues than basic residues, it is an acidic peptide. Its net charge at pH 6 is −1.

26.12 At pH 6 the ion-exchange column is negatively charged because it bears strongly acidic sulfonic acid pendant groups. At this pH, the peptide Ac-Leu-Gly bears a negative charge, is repelled by the column, and will emerge first; Lys-Gly-Leu has a net charge of +1, is attracted to the column, and will emerge next; and Lys-Gly-Arg has a net charge of +2, is most strongly attracted to the column, and will emerge last.

26.13 (b) All three methods would work in principle for leucine. 

Alkylation of ammonia:

\[
\begin{align*}
(CH_3)_2CHCH_2CHCO_2H + NH_3 (excess) \rightarrow (CH_3)_2CHCH_2CHCO_2 + NH_3 \\
2\text{-bromo-4-methylpentanoic acid leucine}
\end{align*}
\]
26.14 (b) Alanine is esterified to give its methyl ester:

26.17 The box on text p. 1283 shows that solid-phase peptide synthesis requires about three synthetic steps per residue. Hence, about 300 steps are required for the synthesis of a protein containing 100 amino acid residues. If the average yield of each step is $Y$, then

$$300 \log Y = \log 0.50 = -0.3010$$

$$\log Y = -0.001003, \text{ or } Y = 0.998$$

Thus, an average yield of 99.8% per step would be required.

If you used 100 steps for your calculation, the average yield is 99.3%. Either way, the yield must be nearly perfect at each step to achieve even a 50% overall yield. Notice that, because there are so many steps, a drop of a small fraction of a percent in average yield per step has a large effect on the overall yield.

26.18 (b) The 20% piperidine step removes both Fmoc protecting groups. Hence, both amino groups react in subsequent acylation reactions. (The Lys residue is drawn in more structural detail for clarity.)

26.19 (a) Compound $A$ results from the $S_N2$ coupling reaction to the resin (analogous to Eq. 26.22 on text p. 1285), and compound $B$ results from deprotection of $A$:
Compound C is the coupling product that results from coupling of the lysine derivative reacting at its carboxylic acid group with the free amino group of B, and compound D results from removal of the Fmoc protecting group, but not the Boc group, from compound C.

\[
\begin{align*}
\text{FmocNH} & - \text{CH} - \text{C} - \text{NHCH}_2\text{C} - \text{O} - \text{CH}_2 \\
\text{NHBoc} & \\
\text{H}_2\text{N} & - \text{CH} - \text{C} - \text{NHCH}_2\text{C} - \text{O} - \text{CH}_2 \\
\text{NHBoc} &
\end{align*}
\]

Compound E results from coupling of the carboxy terminus of Boc-Val with the free amino group of D. Peptide P results from both removal of E from the resin and removal of the Boc group.

\[
\begin{align*}
\text{BocVal} & - \text{NH} - \text{CH} - \text{C} - \text{NHCH}_2\text{C} - \text{O} - \text{CH}_2 \\
\text{NHBoc} & \\
\text{Val-Lys-Gly} & + \ F_3\text{C} - \text{C} - \text{O} - \text{CH}_2 \\
\text{P} & + \text{CO}_2 \\
\text{(from reaction of tert-butyl cation of the Boc group with trifluoroacetic acid)}
\end{align*}
\]

Note that the tripeptide P is in the di-cationic, amino-protonated form as a result of the acidic deprotection conditions. The neutral (zwitterionic) form would be obtained by neutralization.

(b) The Boc group prevents the side-chain amino group of lysine from reacting with the carboxylic acid group of another molecule of itself under the coupling conditions. Also, it remains intact when the amino-terminal Fmoc group is removed with piperidine. The lysine side-chain thus remains protected in the subsequent coupling step as well.

(c) Boc-Val introduces a protecting group that can be removed at the same time as the peptide is released from the resin. Use of Fmoc-Val would require an additional deprotection step with 20% piperidine.

26.20 (b) Glu and Gln are not differentiated by amino acid analysis because the side-chain carboxamide group of Gln, like that of Asn in part (a), is hydrolyzed; that is, Gln is converted into Glu.

26.21 The indole side-chain of tryptophan is very hydrocarbonlike, more so than the side chain of any other amino acid. [Recall (Sec. 25.2, text p. 1224) that pyrrole and indole are neither basic nor acidic.] Hence, we would expect to see AQC-Trp emerge on the far right of the amino acid analysis chromatogram, beyond AQC-Tyr.

26.23 Trypsin catalyzes the hydrolysis of peptides or proteins at the carbonyl group of arginine (R) or lysine (K) residues. Therefore, the peptide E-R-G-A-N-I-K-K-H-E-M would produce the following peptides upon trypsin-catalyzed hydrolysis:

(1) E-R
(2) G-A-N-I-K
(3) K
(4) K-H-E-M
(5) H-E-M

Peptides (2) and (4) originate by hydrolysis at the peptide bond between the two K residues. Peptides (3) that is, K, and (5) originate from hydrolysis at the peptide bond between the K and H residues to give G-A-N-I-K-K followed by hydrolysis at the C-terminal peptide bond of this peptide. However, peptide (5) does not originate from (4), because, when a lysine residue is at the amino terminus of a peptide, the hydrolysis of its peptide bond is not catalyzed by trypsin. (Trypsin is an endopeptidase.)
26.25 (a) b-Type fragmentation of the peptide N-F-E-S-G-K would produce the following \( m/z \) values. In this notation, each amino acid letter represents a residue (that is, \(-\text{NH}--\text{CHR}--\text{CO}--\)). Therefore, the \( H-- \) must be added to indicate the additional hydrogen at the amino terminus \( H_2N-- \), and one mass unit must be added to the corresponding residue mass.

\[
\begin{align*}
H--N & \Rightarrow \quad 1+114.0 = 115.0 \\
H--N-F & \Rightarrow \quad 1+114.0 + 147.1 = 262.1 \\
H--N-F-E & \Rightarrow \quad 1+114.0 + 147.1 + 129.0 = 391.1 \\
H--N-F-E-S & \Rightarrow \quad 1+114.0 + 147.1 + 129.0 + 87.0 = 478.1 \\
H--N-F-E-S-G & \Rightarrow \quad 1+114.0 + 147.1 + 129.0 + 87.0 + 57.0 = 535.1 \\
H--N-F-E-S-G-K--OH & \Rightarrow \quad 1+114.0 + 147.1 + 129.0 + 87.0 + 57.0 + 128.1 + 17.0 = 620.2
\end{align*}
\]

26.26 (b) The \( M + 1 \) ion (shown here protonated on the \(--\text{OH}--\) group of the carboxylic acid) loses water to give an acylium ion, which is the ion of interest. Although carboxylic acids are normally protonated on the carbonyl group, the small amount of the form shown here, once formed, is very unstable.

26.27 (a) The amino group serves as a nucleophile in an addition to the isothiocyanate.

26.28 (a) As Eq. 26.41a on text p. 1303 as well as its mechanism in the solution to Problem 26.27(a) show, the Edman degradation depends on the presence of a free terminal amino group in the peptide. Because this group is blocked as an amide in acetylated peptides, such peptides cannot undergo the Edman degradation.

26.30 The compound is shown in its ionization state at \( pH = 7.4 \) (physiological \( pH \)). (Refer to the configuration of \( L \)-threonine on text p. 1271.) The configuration that is naturally occurring in proteins is assumed.

26.35 Although the sulfur has two unshared pairs, the nitrogen of the thiazole is like the nitrogen of pyridine or imidazole and is the most basic site on the ring; hence, it is likely to be the hydrogen-bond acceptor. An amide backbone \( N--H \) is the only hydrogen bond donor site on a peptide backbone.
Solutions to Additional Problems

26.37 The results for proline:

(a) 

(b) 

(c) 

(d) 

(e) 

(f) 

(g) 

(h) 

(i) 

(j) 

26.38 (b) The most basic amino acid is arginine. It has the highest isoelectric point.

(d) Glycine is not chiral and therefore it cannot be optically active.

26.39 (b) Fmoc-Leu, once attached to the resin, must be deprotected before the subsequent coupling reaction can be expected to succeed. Polly forgot the deprotection step with 20% piperidine, which is essential for removing the Fmoc group.

26.41 Cysteic acid is essentially an aspartic acid in which the side-chain carboxy group has been substituted with a sulfonic acid group. Because sulfonic acids are more acidic than carboxylic acids, cysteic acid is expected to be
(and is) more acidic than aspartic acid. Thus, the correct answer is (1): cysteic acid has a lower isoelectric point than aspartic acid.

26.43 The amino-terminal residue of the peptide P is valine. Because dansyl-valine is obtained, the \( \alpha \)-amino group of valine must have been free in the peptide, and hence, valine must have been the amino-terminal residue. The sequence of the other residues cannot be determined from the data given.

“Dansylation” is a method for determining the amino-terminal residue of a peptide.

26.45 (a) Trypsin should catalyze the cleavage of glucagon at Lys and Arg residues. The following five fragments are expected. (Remember that, by convention, peptides written in this way have their amino ends on the left.) We assume that all cleavages proceed to completion, although in practice intermediate cleavage fragments might be formed as well.

(1) His-Ser-Gln-Gly-Thr-Phe-Thr-Ser-Asp-Tyr-Ser-Lys
(2) Tyr-Leu-Asp-Ser-Arg
(3) Arg
(4) Ala-Gln-Asp-Phe-Val-Gln-Trp-Leu-Met-Asn-Thr
(5) Arg-Ala-Gln-Asp-Phe-Val-Gln-Trp-Leu-Met-Asn-Thr

If cleavage occurs after the first Arg in the Arg-Arg sequence of glucagon, then the peptide (5) is formed, and the second Arg becomes the amino terminus of this peptide. Cleavage does not occur at the peptide bond of this Arg because trypsin is an endopeptidase. (Eq. 26.36b, text p. 1296.) If cleavage of glucagon occurs after the second Arg, then Tyr-Leu-Asp-Ser-Arg-Arg is formed. Subsequent cleavage of this peptide at the Arg-Arg bond gives (2) and (3).

(b) The products of this Edman cleavage would be the PTH derivative of histidine along with a new peptide, which is glucagon without its amino-terminal residue.

26.47 The presence of two amino-terminal residues suggests that either (a) insulin is an approximately equimolar mixture of two proteins, or (b) it is a single protein containing two peptide chains. In fact, the latter is correct; insulin is two polypeptide chains connected by disulfide bonds. Glycine is the amino-terminal residue of one chain, and phenylalanine is the amino-terminal residue of the other.

26.49 This amino acid is really a substituted malonic acid. When heated in acid, it, like most malonic acid derivatives, decarboxylates. In this case, the product is glutamic acid. (See Eq. 20.40, text p. 977.)
26.51 The principle is that a peptide with a given charge migrates to the electrode of opposite charge; a peptide with zero charge does not migrate. Because Gly-Lys is a basic peptide, it has a net positive charge at pH 6, and it therefore migrates to the cathode, the negatively charged electrode. Because Gly-Asp is an acidic peptide, it has a net negative charge at pH 6, and it migrates to the anode, the positively charged electrode. Gly-Ala is a neutral peptide; it has net zero charge at pH 6, and therefore it does not migrate.

26.53 Such a resin should be less effective in partitioning the amino acids with hydrocarbon side chains, because they will be “less soluble” in the resin phase. In effect, there are fewer CH₂— groups on the resin to provide favorable interactions with the amino acid side chains. Furthermore, the amino acid side-chains would be forced to be closer to the more polar and hydrophilic resin surface, with which they might not interact as favorably. The separation should be less effective. The separation should be less effective with C8-silica.

26.55 (b) Hydrogen peroxide oxidizes the disulfide bonds to sulfenic acids, then to sulfinic acids, and finally to sulfonic acids (see Fig. 10.3, text p. 472).

26.58 The mechanism of cleavage in trifluoroacetic acid involves a benzylic carbocation intermediate. This carbocation is resonance-stabilized not only by the para-oxygen, but also by the ortho-oxygen as well. (Draw the resonance structures.) Because the carbocation intermediate is more stable, the cleavage reaction is faster. Hence, milder conditions (lower acid concentration) can be used for the cleavage.

26.61 (b) The side-chain amino group of lysine serves as a nucleophile to open the anhydride. The resulting carboxy group is ionized at the pH of the reaction.

(d) The carbodiimide promotes a condensation between the side-chain carboxy group of aspartic acid and the amino group of the glycine ester.

26.63 Every synthesis has the same final steps:
The challenge in each case is to prepare the alkyl halide $R\text{—}Br$ from the given starting material.

(a) Alternatively, the same alkene could be hydroborated with $\text{BD}_3$, the resulting organoborane oxidized to the primary alcohol by alkaline $\text{H}_2\text{O}_2$, and the alcohol converted into the alkyl bromide with concentrated $\text{HBr}$ and $\text{H}_2\text{SO}_4$ catalyst.

(b) 

(c) Alkylate the acetamidomalonate ion with this halide. Then, in the final hydrolysis step, use concentrated $\text{HBr}$, which will also cleave the methyl ether. Recall that $\alpha$-bromo carbonyl compounds are very reactive in $\text{S}_\text{N}2$ reactions (Sec. 22.3, text p. 1042).

26.64 In the first step, the amine adds to the Edman reagent; for the mechanism of this reaction, see the solution to Problem 26.27(a) on p. 5 of this chapter. The sulfur of the thiourea then serves as a nucleophile to close a six-membered ring and cleave the peptide. The mechanism below begins with the product of the reaction between the Edman reagent and the peptide, which is protonated on the carbonyl oxygen by the catalyzing acid.

26.65 This reaction is an intramolecular ester aminolysis. The mechanism below begins with the conjugate base of the peptide formed by loss of a proton from the terminal amino group.
You may have noticed that diet soft drinks that go on sale in the supermarket are not quite as sweet as really fresh ones. This is because some of the sweetener has degraded by the mechanism just depicted.

26.67 The chemistry is very similar to that of Problem 26.66(i), text p. 1330. Compound $A$ is the hydrazide; compound $B$ is the acyl azide; compound $C$ is the conjugate acid of an $N$-(1-aminoalkyl)amide, which is a nitrogen analog of a hemiacetal. This hydrolyzes in base to the aldehyde shown, ammonia, and acetamide $D$. (If the conditions were harsh enough, acetamide $D$ would be converted into acetate ion and ammonia. If you postulated that compound $D$ is acetate ion, your answer is equally satisfactory.)

26.69 Peptide $I$ results from intramolecular nucleophilic reaction with the conjugate-base anion of the neighboring amide bond. This occurs mostly at Asn-Gly because glycine has no carbon side chain; in residues other than glycine, the carbon side chains can cause rate-retarding van der Waals repulsions in the ring-closure step.

Derivative $I$ is an imide, which is the nitrogen analog of an anhydride. Hydroxide can react with either carbonyl carbon of the imide ring to open the ring, thus generating either peptide $J$ or peptide $K$. The mechanism for the formation of peptide $K$ is as follows:
26.71 (a) In the \( Z \) conformation, the two large groups \( \text{Pep}^N \) and the \( N \)-alkyl group \( \text{Pep}^C \) are trans, and thus van der Waals repulsions between these groups cannot occur. In the \( E \) conformation, these two groups are close enough that van der Waals repulsions can result. These raise the energy of the peptide; hence, the \( Z \) conformation is energetically preferred.

(b) The proline nitrogen bears two \( N \)-alkyl groups, whereas the peptide-bond nitrogens of other residues have only one. Thus, in either the \( E \) or the \( Z \) conformation, an alkyl group is cis to the \( \text{Pep}^N \) group. Although the group that is cis to \( \text{Pep}^N \) in the \( E \) configuration is larger, there is a much smaller difference in energy between \( E \) and \( Z \) conformations for proline residues than there is for other residues. Hence, \( E \) conformations in peptide bonds are sometimes found at proline residues in proteins and peptides.

Notice that the \( E \) conformation of proline in a peptide forms a turn or bend in the peptide chain. Proline residues with \( E \) conformations are often found within turns in the three-dimensional structures of proteins.
26.73 The formula of compound $A$ is the same as that of $N$-acetylaspartic acid minus the elements of water. Since treatment of a dicarboxylic acid with acetic anhydride is a good way to prepare a cyclic anhydride (see Eq. 20.31, text p. 973), it is reasonable to suppose that compound $A$ is the cyclic anhydride of aspartic acid. When this anhydride is treated with L-alanine, the amino group of alanine can react with either of the two carbonyl groups to give either of the two isomeric peptides $B$ and $C$. Both would hydrolyze to give alanine and aspartic acid (as well as acetic acid).

![Diagram of $N$-acetylaspartic anhydride](image)

26.74 (b) (See the enzyme mechanism on p. 720 of the Study Guide and Solutions Manual.) The lysozyme reaction with triethylxonium fluoroborate, $\text{Et}_3O^+ - \text{BF}_4$, results in the formation of the ethyl ester of the carboxylate of Asp52. As shown in Sec. 11.6A on text p. 508, oxonium salts are powerful alkylating agents.

![Enzyme mechanism](image)

The obliteration of enzyme activity could be due to either one or both of the following reasons. First, the ethyl group of the ethyl ester may block the active site between the Asp52 and Glu35 residues and thus prevent the substrate from binding. Second, the mechanism shows that the negatively charged Asp52 residue plays a key role in stabilizing the cationic intermediates, possibly by nucleophilic involvement. Formation of the ethyl ester eliminates the nucleophilic capability of the carboxylate group.