Amino Acids, Peptides, and Proteins
Proteins – Amides from Amino Acids

• Amino acids contain a basic amino group and an acidic carboxyl group
• Joined as amides between the –NH$_2$ of one amino acid and the –CO$_2$H to the next amino acid
• Chains with fewer than 50 units are called peptides
• Protein: large chains that have structural or catalytic functions in biology
Why this Chapter?

• Amino acids are the fundamental building blocks of proteins

• To see how amino acids are incorporated into proteins and the structures of proteins
Structures of Amino Acids

- In neutral solution, the COOH is ionized and the NH$_2$ is protonated
- The resulting structures have “+” and “-” charges (a dipolar ion, or zwitterion)
- They are like ionic salts in solution
The Common Amino Acids

- 20 amino acids form amides in proteins
- All are $\alpha$-amino acids - the amino and carboxyl are connected to the same C
- They differ by the other substituent attached to the $\alpha$ carbon, called the side chain, with H as the fourth substituent
- Proline is a five-membered secondary amine, with N and the $\alpha$ C part of a five-membered ring
- See table 26.1 to examine names, abbreviations, physical properties, and structures of 20 commonly occurring amino acids
Abbreviations and Codes

Alanine \textbf{A}, \textit{Ala}
Arginine \textbf{R}, \textit{Arg}
Asparagine \textbf{N}, \textit{Asn}
Aspartic acid \textbf{D}, \textit{Asp}
Cysteine \textbf{C}, \textit{Cys}
Glutamine \textbf{Q}, \textit{Gln}
Glutamic Acid \textbf{E}, \textit{Glu}
Glycine \textbf{G}, \textit{Gly}
Histidine \textbf{H}, \textit{His}
Isoleucine \textbf{I}, \textit{Ile}

Leucine \textbf{L}, \textit{Leu}
Lysine \textbf{K}, \textit{Lys}
Methionine \textbf{M}, \textit{Met}
Phenylalanine \textbf{F}, \textit{Phe}
Proline \textbf{P}, \textit{Pro}
Serine \textbf{S}, \textit{Ser}
Threonine \textbf{T}, \textit{Thr}
Tryptophan \textbf{W}, \textit{Trp}
Tyrosine \textbf{Y}, \textit{Tyr}
Valine \textbf{V}, \textit{Val}
Chirality of Amino Acids

- Glycine, 2-amino-acetic acid, is achiral
- In all the others, the $\alpha$ carbons of the amino acids are centers of chirality
- The stereochemical reference for amino acids is the Fischer projection of L-serine
- Proteins are derived exclusively from L-amino acids
Types of side chains

• Neutral: Fifteen of the twenty have neutral side chains
• Asp and Glu have a second COOH and are acidic
• Lys, Arg, His have additional basic amino groups side chains (the N in tryptophan is a very weak base)
• Cys, Ser, Tyr (OH and SH) are weak acids that are good nucleophiles
Histidine

- Contains an imidazole ring that is partially protonated in neutral solution
- Only the pyridine-like, doubly bonded nitrogen in histidine is basic. The pyrrole-like singly bonded nitrogen is nonbasic because its lone pair of electrons is part of the $6\pi$ electron aromatic imidazole ring.

\[
\begin{align*}
\text{Basic;} & \\
\text{pyridine-like} & \\
\text{Nonbasic;} & \\
\text{pyrrole-like} & \\
\text{Imidazole ring}
\end{align*}
\]
Essential Amino Acids

• All 20 of the amino acids are necessary for protein synthesis

• Humans can synthesize only 10 of the 20

• The other 10 must be obtained from food
Amino Acids, the Henderson Hasselbalch Equation, and Isoelectric Points

- In acidic solution, the carboxylate and amine are in their conjugate acid forms, an overall cation
- In basic solution, the groups are in their base forms, an overall anion
- In neutral solution cation and anion forms are present
- This pH where the overall charge is 0 is the isoelectric point, pl
If pKa values for an amino acid are known the fractions of each protonation state can be calculated (Henderson-Hasselbach Equation)

\[ \text{pH} = \text{pK}_a - \log \left[ \frac{[A^-]}{[HA]} \right] \]

This permits a titration curve to be calculated or pK$_a$ to be determined from a titration curve.
pl Depends on Side Chain

- The 15 amino acids with thiol, hydroxyl groups or pure hydrocarbon side chains have pl = 5.0 to 6.5 (average of the pKₐ’s)
- D and E have acidic side chains and a lower pl
- H, R, K have basic side chains and higher pl
Electrophoresis

- Proteins have an overall pI that depends on the net acidity/basicity of the side chains
- The differences in pI can be used for separating proteins on a solid phase permeated with liquid
- Different amino acids migrate at different rates, depending on their isoelectric points and on the pH of the aqueous buffer

Strip buffered to pH = 6.00

- Basic pI = 7.50
- Neutral pI = 6.00
- Acidic pI = 4.50
Synthesis of Amino Acids

- Bromination of a carboxylic acid by treatment with Br$_2$ and PBr$_3$ then use NH$_3$ or phthalimide to displace Br
The Amidomalonate Synthesis

- Based on malonic ester synthesis.
- Convert diethyl acetamidomalonate into enolate ion with base, followed by alkylation with a primary alkyl halide.
- Hydrolysis of the amide protecting group and the esters and decarboxylation yields an α-amino acid.

Diethyl acetamidomalonate

(R,S)-Aspartic acid (55%)
Reactive Amination of $\alpha$-Keto Acids

- Reaction of an $\alpha$-keto acid with $\text{NH}_3$ and a reducing agent produces an $\alpha$-amino acid
Enantioselective Synthesis of Amino Acids

- Amino acids (except glycine) are chiral and pure enantiomers are required for any protein or peptide synthesis
- Resolution of racemic mixtures is inherently inefficient since at least half the material is discarded
- An efficient alternative is enantioselective synthesis
Enantioselective Synthesis of Amino Acids (cont’d)

• Chiral reaction catalyst creates diastereomeric transition states that lead to an excess of one enantiomeric product
• Hydrogenation of a Z enamido acid with a chiral hydrogenation catalyst produces S enantiomer selectively

\[
\text{[Rh}(R,R\text{-DiPAMP})(\text{COD})]^+ \text{ BF}_4^- \quad \text{Ph}
\]

\[
\text{H}_3\text{NH}^+\text{CO}_2^- \quad \text{(S)-Phenylalanine}
\]
Peptides and Proteins

- Proteins and peptides are amino acid polymers in which the individual amino acid units, called residues, are linked together by amide bonds, or peptide bonds.
- An amino group from one residue forms an amide bond with the carboxyl of a second residue.
Peptide Linkages

- Two dipeptides can result from reaction between A and S, depending on which COOH reacts with which NH₂ we get AS or SA.
- The long, repetitive sequence of —N—CH—CO— atoms that make up a continuous chain is called the protein’s backbone.
- Peptides are always written with the N-terminal amino acid (the one with the free —NH₂ group) on the left and the C-terminal amino acid (the one with the free —CO₂H group) on the right.
- Alanylserine is abbreviated Ala-Ser (or A-S), and serylalanine is abbreviated Ser-Ala (or S-A).
Amino Acid Analysis of Peptides

- The sequence of amino acids in a pure protein is specified genetically.
- If a protein is isolated it can be analyzed for its sequence.
- The composition of amino acids can be obtained by automated chromatography and quantitative measurement of eluted materials using a reaction with ninhydrin that produces an intense purple color.

\[ \text{Ninhydrin} + \text{α-Amino acid} \xrightarrow{\text{NaOH, H}_2\text{O}} \text{(purple color)} + \text{RCH} + \text{CO}_2 \]
Peptide Sequencing: The Edman Degradation

- The Edman degradation cleaves amino acids one at a time from the N-terminus and forms a detectable, separable derivative for each amino acid.
Peptide Synthesis

- Peptide synthesis requires that different amide bonds must be formed in a desired sequence.
- The growing chain is protected at the carboxyl terminal and added amino acids are N-protected.
- After peptide bond formation, N-protection is removed.
Carboxyl Protecting Groups

- Usually converted into methyl or benzyl esters
- Removed by mild hydrolysis with aqueous NaOH
- Benzyl esters are cleaved by catalytic hydrogenolysis of the weak benzylic C–O bond
Amino Group Protection

- An amide that is less stable than the protein amide is formed and then removed
- The \textit{tert}-butoxycarbonyl amide (BOC) protecting group is introduced with di-\textit{tert}-butyl dicarbonate
- Removed by brief treatment with trifluoroacetic acid
Peptide Coupling

- Amides are formed by treating a mixture of an acid and amine with dicyclohexylcarbodiimide (DCC)

1. The amino group of alanine is protected as the Boc derivative, and
2. the carboxyl group of leucine is protected as the methyl ester.
3. The two protected amino acids are coupled using DCC.
4. The Boc protecting group is removed by acid treatment.
5. The methyl ester is removed by basic hydrolysis.
Overall Steps in Peptide Synthesis

1. The amino group of alanine is protected as the Boc derivative, and
2. the carboxyl group of leucine is protected as the methyl ester.

3. The two protected amino acids are coupled using DCC.

4. The Boc protecting group is removed by acid treatment.

5. The methyl ester is removed by basic hydrolysis.
Automated Peptide Synthesis: The Merrifield Solid-Phase Technique

- Peptides are connected to beads of polystyrene, reacted, cycled and cleaved at the end.

1. A Boc-protected amino acid is covalently linked to the polystyrene polymer by formation of an ester bond (SN2 reaction).

2. The polymer-bonded amino acid is washed free of excess reagent and then treated with trifluoroacetic acid to remove the Boc group.

3. A second Boc-protected amino acid is coupled to the first by reaction with DCC. Excess reagents are removed by washing them from the insoluble polymer.

4. The cycle of deprotection, coupling, and washing is repeated as many times as desired to add amino acid units to the growing chain.

5. After the desired peptide has been made, treatment with anhydrous HF removes the final Boc group and cleaves the ester bond to the polymer, yielding the free peptide.
Protein Structure

• The **primary** structure of a protein is simply the amino acid sequence.

• The **secondary** structure of a protein describes how segments of the peptide backbone orient into a regular pattern.

• The **tertiary** structure describes how the entire protein molecule coils into an overall three-dimensional shape.

• The **quaternary** structure describes how different protein molecules come together to yield large aggregate structures.
α-Helix

• α-Helix stabilized by H-bonds between amide N–H groups and C=O groups four residues away
**β-Pleated Sheet**

- β-pleated sheet secondary structure is exhibited by polypeptide chains lined up in a parallel arrangement, and held together by hydrogen bonds between chains.
Denaturation of Proteins

- The tertiary structure of a globular protein is the result of many intramolecular attractions that can be disrupted by a change of the environment, causing the protein to become denatured.
- Solubility is drastically decreased as in heating egg white, where the albumins unfold and coagulate.
- Enzymes also lose all catalytic activity when denatured.
Enzymes and Coenzymes

- An **enzyme** is a protein that acts as a catalyst for a biological reaction.
- Most enzymes are specific for substrates while enzymes involved in digestion, such as papain attack many substrates.
Types of Enzymes by Function

- Enzymes are usually grouped according to the kind of reaction they catalyze, not by their structures

<table>
<thead>
<tr>
<th>Class</th>
<th>Some subclasses</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oxidoreductases</td>
<td>Dehydrogenases, Oxidases, Reductases</td>
<td>Introduction of double bond, Oxidation, Reduction</td>
</tr>
<tr>
<td>Transferases</td>
<td>Kinas, Transaminases</td>
<td>Transfer of phosphate group, Transfer of amino group</td>
</tr>
<tr>
<td>Hydrolases</td>
<td>Lipases, Nuclease, Protease</td>
<td>Hydrolysis of ester, Hydrolysis of phosphate, Hydrolysis of amide</td>
</tr>
<tr>
<td>Lyases</td>
<td>Deacarboxylases, Dehydrase</td>
<td>Loss of CO₂, Loss of H₂O</td>
</tr>
<tr>
<td>Isomerases</td>
<td>Epimerases</td>
<td>Isomerization of chirality center</td>
</tr>
<tr>
<td>Ligases</td>
<td>Carboxylases, Synthetase</td>
<td>Addition of CO₂, Formation of new bond</td>
</tr>
</tbody>
</table>
How Do Enzymes Work? Citrate Synthase

- Citrate synthase catalyzes a mixed Claisen condensation of acetyl CoA and oxaloacetate to give citrate
- Normally Claisen condensations require a strong base in an alcohol solvent but citrate synthetase operates in neutral solution

\[
\begin{align*}
\text{Oxaloacetate} & \quad + \quad \text{Acetyl CoA} & \quad \xrightarrow{\text{Citrate synthase}} & \quad \text{Citrate} \\
& \quad + \quad \text{HSCoA}
\end{align*}
\]
The Structure of Citrate Synthase

- Determined by X-ray crystallography
- Enzyme is very large compared to substrates, creating a complete environment for the reaction
Mechanism of Citrate Synthetase

• A cleft with functional groups binds oxaloacetate
• Another cleft opens for acetyl CoA with H 274 and D 375, which have carboxylate that abstract a proton from acetyl CoA
• The enolate (stabilized by a cation) adds to the carbonyl group of oxaloacetate
• The thiol ester in citryl CoA is hydrolyzed