Chapter 25: Amino Acids, Peptides and Proteins
[Sections: 25.1 – 25.8]

1. alpha (α) amino acids

- amino acids contain an amino group at the α-position relative to the carboxylic acid group
- there are 20 naturally-occurring amino acids that vary by the nature of the R group
- 19 of the amino acids are chiral because of the stereogenic α carbon with the exception of glycine (R=H)
- the R groups can be roughly categorized into 4 types: nonpolar, polar, acidic and basic
- humans can synthesize 10 of the amino acids from scratch, the other 10 (essential amino acids) must be derived from diet

2. Properties of α-amino acids

- pH > IP, the CO2- group remains deprotonated, and the NH₃⁺ group becomes deprotonated so that the amino acid is overall negatively charged
- pH < IP, the CO2- group becomes protonated, and the NH₃⁺ group remains protonated so that the amino acid is overall positively charged

A. What form predominates for lysine (IP = 9.74) in a solution of pH = 6?

B. An amino acid is predominantly negatively charged in a solution of pH = 8.2. What must be true about its IP?
The 20 Common Naturally-Occurring α-Amino Acids found in Proteins

*by the name denotes essential amino acids

**amino acids with non-polar side chains**

- **Glycine**  
  Gly or G  
  IP = 5.97  
- **Alanine**  
  Ala or A  
  IP = 6.01  
- **Valine***  
  Val or V  
  IP = 5.96  
- **Leucine***  
  Leu or L  
  IP = 5.98  
- **Isoleucine***  
  Ile or I  
  IP = 6.02  
- **Methionine***  
  Met or M  
  IP = 5.74  
- **Proline**  
  Pro or P  
  IP = 6.30  
- **Phenylalanine***  
  Phe or F  
  IP = 5.48  
- **Tryptophan***  
  Trp or W  
  IP = 5.89

**amino acids with polar side chains**

- **Asparagine**  
  Asn or N  
  IP = 5.41  
- **Glutamine**  
  Gln or Q  
  IP = 5.65  
- **Serine**  
  Ser or S  
  IP = 5.68  
- **Threonine***  
  Thr or T  
  IP = 5.60  
- **Tyrosine**  
  Tyr or Y  
  IP = 5.66  
- **Cysteine**  
  Cys or C  
  IP = 5.07

**amino acids with polar and acidic side chains**

- **Aspartic Acid**  
  Asp or D  
  IP = 2.77  
- **Glutamic Acid**  
  Glu or E  
  IP = 3.22

**amino acids with polar and basic side chains**

- **Arginine**  
  Arg or R  
  IP = 10.76  
- **Histidine***  
  His or H  
  IP = 7.59  
- **Lysine***  
  Lys or K  
  IP = 9.74

*by the name denotes essential amino acids
3. Analysis of amino acids: electrophoresis

- All 20 amino acids have a unique pI
- Electrophoresis exploits this difference and the resulting differences in behavior in response to an electric field to separate amino acid mixtures for analysis
- In cases where pI's are particularly close (e.g., glycine [MW = 75] pI = 5.97, leucine [MW = 131] pI = 5.98) differences in molecular weight also have an impact on rate/extent of movement

### Problems: 1, 2
5. Making peptides

CO₂H group from one α-amino acid

NH₂ group from another α-amino acid

• amide bond formed: peptide bond
• considered to be a "dipeptide"
• most often written with amino acid with the "free" NH₂ group at the left (N-terminal residue) and AA with free CO₂H group at the right (C-terminal residue)

Problem 1: need to convert OH of carboxylic acid group into a leaving group

Problem 2: need to limit reaction to one of the CO₂H groups and one of the NH₂ groups

To make a dipeptide from Alanine and Glycine:

* in order to ensure that only one CO₂H group and one NH₂ group react, the other groups must be protected from reaction
• using these methods and judicious protection/deprotection, dipeptides, tripeptides, tetra, penta, etc. (i.e., polypeptides) may be constructed sequentially

• proteins are polypeptides with ~50 AA residues. Proteins on average have 300 AA residues but can incorporate as many as 30,000

• the entire process has been mechanized via the Merrifield synthesis method that makes use of polymer supports
Determining the primary structure of a peptide

A. Short-chain polypeptides

- The primary structure of a peptide or protein is the sequence of amino acids (from N-terminal residue to C-terminal residue) that make up the peptide chain.

Leucine enkephalin
found in the brain; interacts with the same receptor as morphine and helps to control pain

Try–Gly–Gly–Phe–Leu

- Complete simultaneous cleavage of all of the peptide bonds is possible.
- All sequencing information is lost.


- Sequential removal of one AA at a time taking advantage of the free NH$_2$ group allows for identification of the N-terminal residue specifically.
- Is successful for determining sequence of ~ 50 AA's.
- Similar selective C-terminal residue analysis is also possible.

B. Long-chain polypeptides

- The primary structure of long chain proteins can be accomplished via partial hydrolysis of the polypeptide into shorter chains (<50 AAs in length) that can be sequenced as above.
- The individual short chains then need to be stitched together in a logical manner to provide the full sequence.

1 Tyr, 2 Cys, 1 Phe, 1 Pro
1 Gln, 1 Gly, 1 Arg, 1 Asn

- Polypeptides greater in length than ~ 50 AAs = proteins

Problems: 5,6
7. Secondary Structures of Proteins

- Protein secondary structure: general three-dimensional forms of local segments of proteins (e.g., alpha helices and beta sheets)

- right-handed
- \(~3.6\) AA per turn
- NH of each AA is hydrogen bound to the C=O 4 units away
- R groups point outward
- two or more protein chains line up side by side
- hydrogen bonding between NH and C=O of neighboring strand
- alkyl groups are generally positioned above and below the sheet

![Diagram of alpha helix and beta sheet](image)

8. Tertiary Structure of Proteins

- the protein's overall geometric shape
- non-regular but not random
- most stable arrangement for that sequence of AA residues
- hydrogen bonding and S-S [disulfide bonds between cysteine residues] play the major role in structural stability
- generally, the structure of enzymes have polar groups directed towards the outside of the structure, and nonpolar groups directed towards the interior which allows for water solubility
- change in solvent, pH, or temperature can alter the shape of the protein (unfolding), which is called "denaturization" and is generally irreversible

- the tertiary shape of the protein determines its behavior and specificity by creating "pockets" or "active sites" within the structure that recognize specific types of compound
Example: human cholinesterases in complex with tacrine

- part of a study to find drugs to aid in the battle against Alzheimer’s disease
- tacrine was one of the first drugs to be found beneficial in the treatment of Alzheimer’s disease, although it has been discontinued since 2013 due to concerns over safety
- human cholinesterase (PDB ID 4BDS) is a protein (enzyme) with 529 AA residues