Amino acid & Protein -1
BCH 3000

PRINCIPLES OF BIOCHEMISTRY

(Semester 2 - 2014/15)
Amino acids and proteins

- Draw a general amino acid and identify the two functional groups common to all.
- Classify each amino acid according to the chemical nature of its R group.
- Define the meaning of an essential amino acid.
- Draw the reaction that joins two amino acids to form a peptide bond.
- Describe and differentiate primary, secondary, tertiary, and quaternary protein structures.
- Describe and differentiate co-enzymes and prosthetic groups.
- List and discuss four forces that stabilize globular protein structure.
Proteins

Reverse transcription

DNA

RNA

Translation

protein

Transcription

Genotype

Genome

(similar in all cells)

Phenotype

Proteome

(unique to all cells)
Amino Acid Structure
Amino Acids

- There are twenty different amino acids that are common to all life.
- The number and arrangement of these twenty amino acids result in an infinite variety of proteins.
The general formula of an amino acid is

\[ \text{H}_2\text{N}-\text{C}^\text{R} \text{COOH} \]

- **amino group**
- **carboxyl group**
- **\( \alpha \)-carbon atom**
- **side-chain group**

R is commonly one of 20 different side chains. At pH 7 both the amino and carboxyl groups are ionized.
Simplest amino acid

Glycine
Glycine

amino group


carboxyl group

\( \text{H}_2\text{N} - \text{C} - \text{COOH} \)

\( \text{H}_3\text{N}^+ - \text{C} - \text{COO}^- \)

\( \alpha\)-carbon

side chain (R)

nonionized form

ionized form

pH 7
Alanine

Nonionized form (A)

Ionized form (B)

α-carbon
side chain (R)

amino group

carboxyl group

H₂\(\text{C} - \text{COO}^-\)

\(\text{H}_3\text{N}\)\(\text{C} - \text{COO}^-\)

©1998 JAFLAND
# Amino Acids

<table>
<thead>
<tr>
<th>Glycine</th>
<th>Tryptophan</th>
<th>Asparagine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine</td>
<td>Proline</td>
<td>Glutamine</td>
</tr>
<tr>
<td>Valine</td>
<td>Serine</td>
<td>Lysine</td>
</tr>
<tr>
<td>Leucine</td>
<td>Threonine</td>
<td>Arginine</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>Cysteine</td>
<td>Histidine</td>
</tr>
<tr>
<td>Methionine</td>
<td>Aspartic acid</td>
<td>Tyrosine</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>Phenylalanine</td>
<td></td>
</tr>
</tbody>
</table>
Amino acid abbreviations

- Ala = alanine
- Cys = cysteine
- Asp = aspartic acid
- Glu = glutamic acid
- Phe = phenylalanine
- Gly = glycine
- His = histidine
- Ile = isoleucine
- Lys = lysine
- Leu = leucine
- Met = methionine
- Asn = asparagine
- Pro = proline
- Glu = glutamine
- Arg = arginine
- Ser = serine
- Thr = threonine
- Val = valine
- Trp = tryptophan
- Tyr = tyrosine
## Common Amino Acids in Proteins

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Abbreviation</th>
<th>Structure</th>
<th>Amino Acid</th>
<th>Abbreviation</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycine</td>
<td>Gly</td>
<td>H—CH—COOH</td>
<td>*Isoleucine</td>
<td>Ile</td>
<td>CH₃—CH₂—CH—CH—COOH</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NH₂</td>
<td></td>
<td></td>
<td>CH₂</td>
</tr>
<tr>
<td>Alanine</td>
<td>Ala</td>
<td>CH₃—CH—COOH</td>
<td>Proline</td>
<td>Pro</td>
<td>H₂C—CH₂</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NH₂</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*Valine</td>
<td>Val</td>
<td>CH₃—CH—CH—COOH</td>
<td>*Phenyl-alanine</td>
<td>Phe</td>
<td>CH₂—CH—COOH</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CH₃</td>
<td></td>
<td></td>
<td>NH₂</td>
</tr>
<tr>
<td>*Leucine</td>
<td>Leu</td>
<td>CH₃—CH—CH₂—CH—COOH</td>
<td>*Methionine</td>
<td>Met</td>
<td>CH₃—S—CH₂CH₂—CH—COOH</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CH₃</td>
<td></td>
<td></td>
<td>NH₂</td>
</tr>
</tbody>
</table>
|            |              |           |            | *Tryptophan  | Trp      | \[\text{Indole}\]
|            |              |           |            |              |         |

Nonpolar R groups
## Common Amino Acids in Proteins

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Chemical Structure</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serine</td>
<td>Ser</td>
<td>HO—CH₂—CH—COOH</td>
</tr>
<tr>
<td>Asparagine</td>
<td>Asn</td>
<td>H₂N—C—CH₂—CH—COOH</td>
</tr>
<tr>
<td>*Thrreonine</td>
<td>Thr</td>
<td>CH₃—CH—CH—COOH</td>
</tr>
<tr>
<td>Glutamine</td>
<td>Gln</td>
<td>H₂N—C—CH₂CH₂—CH—COOH</td>
</tr>
<tr>
<td>Cysteine</td>
<td>Cys</td>
<td>HS—CH₂—CH—COOH</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>Tyr</td>
<td></td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>Glu</td>
<td>HO—C—CH₂CH₂—CH—COOH</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>Asp</td>
<td>HO—C—CH₂—CH—COOH</td>
</tr>
<tr>
<td>*Lysine</td>
<td>Lys</td>
<td>H₂N—CH₂CH₂CH₂CH₂—CH—COOH</td>
</tr>
<tr>
<td>‡Arginine</td>
<td>Arg</td>
<td>H₂N—C—NH—CH₂CH₂CH₂—CH—COOH</td>
</tr>
<tr>
<td>Histidine</td>
<td>His</td>
<td></td>
</tr>
</tbody>
</table>
Amino acids are classified by their R group

- Acidic amino acids
- Basic amino acids
- Nonpolar amino acids
- Polar amino acids
Acidic amino acids

- Aspartic acid (Asp, or D)
- Glutamic acid (Glu, or E)
Basic amino acids

Lysine (Lys, or K)

This group is very basic because its positive charge is stabilized by resonance.

Arginine (Arg, or R)

These nitrogens have a relatively weak affinity for H+ and are only partly protonated at neutral pH.

Histidine (His, or H)
Nonpolar amino acids

- Alanine (Ala, or A)
- Valine (Val, or V)
- Leucine (Leu, or L)
- Isoleucine (Ile, or I)
- Phenylalanine (Phe, or F)
- Proline (Pro, or P)
- Methionine (Met, or M)
- Tryptophan (Trp, or W)
- Glycine (Gly, or G)
- Cysteine (Cys, or C)
Amino Acid R-groups

<table>
<thead>
<tr>
<th>Non-Polar</th>
<th>Hydrophobic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tryptophan, Phenylalanine, Isoleucine, Tyrosine, Leucine, Valine, Methionine</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ambivalent</th>
<th>Charged</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycine, Threonine, Alanine</td>
<td>Arginine (+), Glutamic acid (-), Aspartic Acid (-), Lysine (+), Histidine (+)</td>
</tr>
</tbody>
</table>

### Polar

#### Uncharged
- Cysteine
- Proline
- Serine
- Glutamine
- Asparagine

#### Charged
- Arginine (+)
- Glutamic acid (-)
- Aspartic Acid (-)
- Lysine (+)
- Histidine (+)
# Hydrophobic Indexes

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Index</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arginine</td>
<td>-11.2</td>
<td>Arg [R]</td>
</tr>
<tr>
<td>Glutamic Acid</td>
<td>-9.9</td>
<td>Glu [E]</td>
</tr>
<tr>
<td>Aspartic Acid</td>
<td>-7.4</td>
<td>Asp [D]</td>
</tr>
<tr>
<td>Lysine</td>
<td>-4.2</td>
<td>Lys [K]</td>
</tr>
<tr>
<td>Histidine</td>
<td>-3.3</td>
<td>His [H]</td>
</tr>
<tr>
<td>Cysteine</td>
<td>-2.8</td>
<td>Cys [C]</td>
</tr>
<tr>
<td>Proline</td>
<td>-0.5</td>
<td>Pro [P]</td>
</tr>
<tr>
<td>Serine</td>
<td>-0.3</td>
<td>Ser [S]</td>
</tr>
<tr>
<td>Glutamine</td>
<td>-0.3</td>
<td>Gln [Q]</td>
</tr>
<tr>
<td>Asparagine</td>
<td>-0.2</td>
<td>Asn [N]</td>
</tr>
<tr>
<td>Glycine</td>
<td>0</td>
<td>Gly [G]</td>
</tr>
<tr>
<td>Threonine</td>
<td>0.4</td>
<td>Thr [T]</td>
</tr>
<tr>
<td>Alanine</td>
<td>0.5</td>
<td>Ala [A]</td>
</tr>
<tr>
<td>Methionine</td>
<td>1.3</td>
<td>Met [M]</td>
</tr>
<tr>
<td>Valine</td>
<td>1.5</td>
<td>Val [V]</td>
</tr>
<tr>
<td>Leucine</td>
<td>1.8</td>
<td>Leu [L]</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>2.3</td>
<td>Tyr [Y]</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>2.5</td>
<td>Ile [I]</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>2.5</td>
<td>Phe [F]</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>3.4</td>
<td>Trp [W]</td>
</tr>
</tbody>
</table>
Essential amino acids

- Definition - Those amino acids that cannot be synthesized in the body in sufficient quantities for anabolic needs.

- In humans,
  - Isoleucine
  - Leucine
  - Valine
  - Tryptophan
  - Methionine
  - Lysine
  - Phenylalanine
  - Threonine
  - Histidine
Zwitterion Form of Amino Acids

“zwitterion” = hybrid ion form that exists at pH = 7
Recall from Acid-Base Equilibria: Henderson-Hasselbalch Equation

\[ pH = pK_a + \log \frac{[X^-]}{[HX]} \]

or

\[ pH = pK_a + \log \frac{[\text{base}]}{[\text{acid}]} \]
\[ H_{2}A^{+} \xrightarrow{\text{pK}_{a1}} H^{+} \xrightarrow{\text{H}^{+}} HA^{0} \xrightarrow{\text{pK}_{a2}} A^{-} \]
Acid-Base Properties

- pK of a chemical group is the pH at which the group is half ionized.
- $pK_1 = \alpha$ - carboxylic acid ($pK \approx 2.2$)
  - Above pH 3.5 entirely in carboxylate form
- $pK_2 = \alpha$ - amino groups, ($pK \approx 9.4$)
  - Below pH 8.0 in ammonium ion form
- $pK_R = \text{side groups with acid-base properties}$

\[
pK_2 \quad H_3N^+ \quad \text{C} \quad \text{COO}^- \quad pK_1
\]
Acid-Base Properties

Example: Glycine

- Low pH: both acid-base groups fully protonated
  - Cationic form $\text{H}_3\text{NCH}_2\text{COOH}^+$
- High pH: loses $\text{H}^+$ in a step-wise fashion
  - Anionic form $\text{H}_2\text{NCH}_2\text{COO}^-$
pK's determined from Henderson-Hasselbalch equation

Correspond to midpoints of each leg of the titration curve

- $pK_1 = pH 2.35$:
  - $+\text{H}_3\text{NCH}_2\text{COOH}$ cation = $+\text{H}_3\text{NCH}_2\text{COO}^-$ zwitterion

- $pK_2 = pH 9.78$:
  - $\text{H}_2\text{NCH}_2\text{COO}^-$ anion = $+\text{H}_3\text{NCH}_2\text{COO}^-$ zwitterion
Titration curve of alanine. 

\[ \text{pK}_1 = 2.3 \]

\[ \frac{[\text{NH}_3\text{CHRCOOH}]}{[\text{NH}_2\text{CHRCOO}^-]} = 1 \]

\[ \text{pH} = 6.0 \rightarrow [\text{NH}_3\text{CHRCOO}^-] \]

\[ \frac{[\text{NH}_3\text{CHRCOO}^-]}{[\text{NH}_2\text{CHRCOO}^-]} = 1 \]

\[ \text{pK}_2 = 9.7 \]
Titration curves of glutamic acid, lysine, and histidine.
<table>
<thead>
<tr>
<th>Amino acid</th>
<th>PK₁ (α-COOH)</th>
<th>PK₂ (α-NH₂⁺)</th>
<th>PKᵣ (R Group)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine</td>
<td>2.35</td>
<td>9.87</td>
<td>-</td>
</tr>
<tr>
<td>Arginine</td>
<td>1.82</td>
<td>8.99</td>
<td>12.48</td>
</tr>
<tr>
<td>Asparagine</td>
<td>2.10</td>
<td>8.84</td>
<td>-</td>
</tr>
<tr>
<td>Aspartic Acid</td>
<td>1.99</td>
<td>9.90</td>
<td>3.90</td>
</tr>
<tr>
<td>Cysteine</td>
<td>1.92</td>
<td>10.78</td>
<td>8.33</td>
</tr>
<tr>
<td>Glutamic Acid</td>
<td>2.10</td>
<td>9.47</td>
<td>4.07</td>
</tr>
<tr>
<td>Glutamine</td>
<td>2.17</td>
<td>9.13</td>
<td>-</td>
</tr>
<tr>
<td>Glycine</td>
<td>2.35</td>
<td>9.78</td>
<td>-</td>
</tr>
<tr>
<td>Histidine</td>
<td>1.8</td>
<td>9.33</td>
<td>6.04</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>2.32</td>
<td>9.76</td>
<td>-</td>
</tr>
<tr>
<td>Leucine</td>
<td>2.33</td>
<td>9.74</td>
<td>-</td>
</tr>
<tr>
<td>Lysine</td>
<td>2.16</td>
<td>9.18</td>
<td>10.79</td>
</tr>
<tr>
<td>Methionine</td>
<td>2.13</td>
<td>9.28</td>
<td>-</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>2.16</td>
<td>9.18</td>
<td>-</td>
</tr>
<tr>
<td>Proline</td>
<td>1.95</td>
<td>10.65</td>
<td>-</td>
</tr>
<tr>
<td>Serine</td>
<td>2.19</td>
<td>9.21</td>
<td>~13</td>
</tr>
<tr>
<td>Threonine</td>
<td>2.09</td>
<td>9.10</td>
<td>~13</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>2.43</td>
<td>9.44</td>
<td>-</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>2.20</td>
<td>9.11</td>
<td>10.13</td>
</tr>
</tbody>
</table>
- **Physiological pH range**
  - Both carboxyl & amino groups ionized
- **Amphoteric**
  - Act as both an acid & base
  - Ampholytes (*amphoteric electrolytes*)
- **Zwitterions or Dipolar Ions**
  - Molecules with charged groups of opposite polarity
  - Physical properties characteristic of ionic compounds
- **Isoelectric Point:**
  - pH at which a molecule carries no net electric charge
  \[ pI = \frac{1}{2}(pK_i + pK_j) \]
  - For amino acids with only 2 pKs (2 ionizable groups) this equals \( pK_1 \) and \( pK_2 \).
  - For aspartic & glutamic acids, \( pK_i \) & \( pK_j \) are \( pK_1 \) and \( pK_R \).
  - For arginine, histidine & lysine, \( pK_i \) & \( pK_j \) are \( pK_R \) & \( pK_2 \).
Aas with ionizable R groups have a third pKa that varies with the R group.

### Table 4.1  
R groups also affect the pKa of the acid and amino groups.

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>α-COOH pKₐ</th>
<th>α-NH₃⁺ pKₐ</th>
<th>R group pKₐ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine</td>
<td>2.4</td>
<td>9.7</td>
<td>12.5</td>
</tr>
<tr>
<td>Arginine</td>
<td>2.2</td>
<td>9.0</td>
<td>12.5</td>
</tr>
<tr>
<td>Asparagine</td>
<td>2.0</td>
<td>8.8</td>
<td>12.5</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>2.1</td>
<td>9.8</td>
<td>12.5</td>
</tr>
<tr>
<td>Cystine</td>
<td>1.7</td>
<td>10.8</td>
<td>12.5</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>2.2</td>
<td>9.7</td>
<td>12.5</td>
</tr>
<tr>
<td>Glutamine</td>
<td>2.2</td>
<td>9.1</td>
<td>12.5</td>
</tr>
<tr>
<td>Glycine</td>
<td>2.3</td>
<td>9.6</td>
<td>12.5</td>
</tr>
<tr>
<td>Histidine</td>
<td>1.8</td>
<td>9.2</td>
<td>12.5</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>2.4</td>
<td>9.7</td>
<td>12.5</td>
</tr>
<tr>
<td>Leucine</td>
<td>2.4</td>
<td>9.6</td>
<td>12.5</td>
</tr>
<tr>
<td>Lysine</td>
<td>2.2</td>
<td>9.0</td>
<td>12.5</td>
</tr>
<tr>
<td>Methionine</td>
<td>2.3</td>
<td>9.2</td>
<td>12.5</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>1.9</td>
<td>9.1</td>
<td>12.5</td>
</tr>
<tr>
<td>Proline</td>
<td>2.1</td>
<td>10.6</td>
<td>12.5</td>
</tr>
<tr>
<td>Serine</td>
<td>2.2</td>
<td>9.2</td>
<td>~13</td>
</tr>
<tr>
<td>Threonine</td>
<td>2.6</td>
<td>10.4</td>
<td>~13</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>2.4</td>
<td>9.4</td>
<td>~13</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>2.2</td>
<td>9.1</td>
<td>12.5</td>
</tr>
<tr>
<td>Valine</td>
<td>2.3</td>
<td>9.6</td>
<td>12.5</td>
</tr>
</tbody>
</table>
The titration profiles of these AaS will have **three** pKa’s.
Amino acid & Protein -2
Recall: Amino Acids are Chiral*

*except glycine - R group is another H
A zwitterion is formed when a **proton** (a hydrogen nucleus) moves from the **carboxylic acid** group to the **amine** group.

The zwitterion ion forms when an amino acid is dissolved in water – and this is how they are usually found in nature.
Alanine is an amino acid. Its molecular formula is $C_3H_7O_2N$.

Make a molecular model of alanine and use it to:

1. investigate the shape of the molecule and explain why it can exist in two forms.
2. show how its zwitterion is formed.

Draw the structural formula for alanine. Click to see if you got it right.

Look at the picture of the zwitterion.

1. Where is the $-\text{NH}_3^+$ group?
2. Where is the $-\text{COO}^-$ group?

Click to see if you are right.
The shapes of some amino acid molecules

Click for a useful website that allows you to compare the shapes of amino acids
Formation of the peptide bond

Two amino acid molecules; the nature of the R group (R₁ and R₂) determines the amino acid

The molecules must be orientated so that the carboxylic acid group of one can react with the amine group of the other

The peptide bond forms with the elimination of a water molecule; it is another example of a condensation reaction
Hydrolysis of the peptide bond

The peptide bond holds two amino acid ‘residues’ together. It is a flat, rigid group.

A water molecule reacts with this group.

The two amino acids form or, if the peptide bond is somewhere in a long peptide chain, two smaller peptide molecules are formed.
Bonding within a peptide chain (intramolecular) and between one chain and another (intermolecular)

Hydrogen bonds
These form in all proteins. The hydrogen atom of the peptide link is attracted to the oxygen of another peptide link.

Covalent bonds
In a very small number of proteins, sulfur-sulfur covalent bonds (also called cystine bonds or disulfide bridges) are present.

Ionic bonds
If some of the amino acids in the proteins have carboxylic acid or amine side groups, an ionic bond can form.
Amino acid & Protein -3
## Types of Proteins

<table>
<thead>
<tr>
<th>Type</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structural</td>
<td>tendons, cartilage, hair, nails</td>
</tr>
<tr>
<td>Contractile</td>
<td>muscles</td>
</tr>
<tr>
<td>Transport</td>
<td>hemoglobin</td>
</tr>
<tr>
<td>Storage</td>
<td>milk</td>
</tr>
<tr>
<td>Hormonal</td>
<td>insulin, growth hormone</td>
</tr>
<tr>
<td>Enzyme</td>
<td>catalyzes reactions in cells</td>
</tr>
<tr>
<td>Protection</td>
<td>immune response</td>
</tr>
</tbody>
</table>
Amino Acids

• Building blocks of proteins
• Carboxylic acid group
• Amino group
• Side group R gives unique characteristics

\[
\begin{align*}
R & \quad \text{side chain} \\
\text{H}_2\text{H} & \quad \text{I} \\
\text{I} & \quad \text{C} \quad \text{—COOH} \\
\text{I} & \quad \text{H}
\end{align*}
\]
Examples of Amino Acids

Glycine

Alanine
Types of Amino Acids

Nonpolar

R = H, CH₃, alkyl groups, aromatic

Polar

R = –CH₂OH, –CH₂SH, –CH₂C–NH₂,
   (polar groups with –O-, -SH, -N-)

Polar/Acidic

R = –CH₂COOH, or -COOH

Polar/Basic

R = –CH₂CH₂NH₂
L-Form Amino Acid Structure

- **Carboxylic group**
- **Amino group**
- **R group**

H = Glycine
CH$_3$ = Alanine

Juang RH (2004) BCbasics
Mirror Images of Amino Acid

Mirror image

Same chemical properties
Stereo isomers

Juang RH (2004) BC basics
Amino Acid Subway Map

Central line

Basic

Arg [R]
Lys [K]
His [H]

Non-polar

Gly [G]
Ser [S]
Thr [T]

Polar

Ala [A]
Cys [C]
Met [M]

Hydroxy

South line

Trp [W]
Tyr [Y]
Phe [F]

Aromatic

Central line

Val [V]
Ile [I]
Leu [L]

Aliphatic

Circular line

Gln [Q]
Asn [N]
Asp [D]
Glu [E]

Amide

Nan-Kan line

Basic

Arg [R]
Lys [K]
His [H]

Non-polar

Gly [G]
Ser [S]
Thr [T]

Polar

Ala [A]
Cys [C]
Met [M]

Hydroxy

South line

Trp [W]
Tyr [Y]
Phe [F]

Aromatic

Central line

Val [V]
Ile [I]
Leu [L]

Aliphatic

Circular line

Gln [Q]
Asn [N]
Asp [D]
Glu [E]

Amide
**Classification of Amino Acids by Polarity**

<table>
<thead>
<tr>
<th>POLAR</th>
<th>Acidic</th>
<th>Neutral</th>
<th>Basic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp</td>
<td></td>
<td>Asn</td>
<td>Arg</td>
</tr>
<tr>
<td>Tyr</td>
<td>Cys</td>
<td>Ser</td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>Gln</td>
<td>Thr</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NON-POLAR</th>
<th>Ala</th>
<th>Ile</th>
<th>Leu</th>
<th>Met</th>
<th>Phe</th>
<th>Pro</th>
<th>Trp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Val</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Polar or non-polar, it is the bases of the amino acid properties.

Juang RH (2003) Biochemistry
Formation of Peptide Bonds by Dehydration

Amino acids are connected head to tail

1. NH₂ → COOH → NH₂
2. NH₂ → COOH

Carbodiimide

Dehydration

-H₂O

Juang RH (2004) BCbasics 56
Peptide Bond Is Rigid and Planar

Juang RH (2004) BCbasics
Protein Structure -4
Proteins

DNA → RNA → protein

Reverse transcription → Translation

Transcription

Genotype → Phenotype
Genome → Proteome

(similar in all cells) → (unique to all cells)
Proteins

- Proteins are large complex molecules composed of long chains of amino acids called polypeptide chains.
- Proteins are polymers, amino acid is the monomer.
- Carbon, hydrogen, oxygen, nitrogen, and sulfur are the elements found in proteins.
Protein Functions

- Enzymes
- Structural molecules
- Regulatory molecules
- Transport molecules
- Motor proteins

- Storage molecules
- Defense proteins
- Signaling proteins
- Receptor proteins
- Many other functions
Proteins

- Long polymers of amino acids joined by peptide bonds
- In the complete protein structure, the amino acid chain is twisted and folded into a specific three-dimensional shape
Peptide bonds

- Joins two amino acids into a dipeptide
- Bond forms between carboxyl group of one amino acid and amine group of the second amino acid
- Peptide bond forms by a condensation reaction losing a molecule of water with each bond
Peptide bond

Amino acids are commonly joined together by an amide linkage, called a peptide bond.

Peptide bond: The four atoms in each gray box form a rigid planar unit. There is no rotation around the C–N bond.
Peptide bond formation

Aspartate

Alanine

condensation

H₂O
Peptide bond formation

Aspartyl alanine, a dipeptide

Peptide bond

Primary Structure
**Dipeptide**

**Peptide bond resonance**

Asp<sub>α</sub>Val<sub>α</sub> is an example of a dipeptide.
α-helixes

Intra-chain
H-bonds

Secondary Structure
β-strands

Inter-chain
H-bonds

Secondary Structure
Tertiary structure

Hb monomer (or myoglobin)

Quaternary structure

Hb $\alpha_2\beta_2$ tetramer
### Protein Structure

<table>
<thead>
<tr>
<th><strong>Primary</strong> structure is the amino acid sequence.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Secondary</strong> structure is how the amino acids in sequence fold up locally. Examples are $\alpha$-helixes and $\beta$-strands and loops.</td>
</tr>
<tr>
<td><strong>Tertiary</strong> structure is the 3-dimensional folding of the secondary structural elements and connecting loops in space.</td>
</tr>
<tr>
<td><strong>Quaternary</strong> structure is the association of multiple subunits, each with a tertiary structure and each a unique gene product.</td>
</tr>
</tbody>
</table>
Electrostatic interactions involve the interaction of (+) and (-) charged side groups.

Hydrogen bonds involve sharing of a hydrogen atom between two electronegative atoms (e.g., O, N).

Van der Waal’s forces are weak forces based on optimal overlap of adjacent electronic orbitals. Can be repulsive.

Hydrophobic interactions are, by far, the most powerful force stabilizing protein structure. Basis of force is entropy gain realized by burying hydrophobic residues.
Tripeptide

The sequence of this tripeptide is histidine-cysteine-valine.
Hierarchy of Protein Structure

- 20 different amino acids: many combinations

Primary Structure
The order of amino acids: Protein sequence

Secondary Structure
Conformation varies depending on sequence

Tertiary Structure
Overall structure of the chain in full 3D
## Four Levels of Protein Structure

When a protein is dissected one observes four different levels or complexities of structure.

<table>
<thead>
<tr>
<th>Level</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1°</td>
<td>primary structure</td>
</tr>
<tr>
<td>2°</td>
<td>secondary structure</td>
</tr>
<tr>
<td>3°</td>
<td>tertiary structure</td>
</tr>
<tr>
<td>4°</td>
<td>quaternary structure</td>
</tr>
</tbody>
</table>
Give the name and structure of at least 2 examples of each of the following:

a. heterocyclic amino acid
b. aromatic amino acid
c. neutral amino acid
d. acidic amino acid
e. basic amino acid
f. sulfur containing amino acid

Important in determining the overall shape of the protein

R-groups of the amino acids in the polypeptide chain are responsible for the overall shape of the protein

N-terminus always on left
Secondary Protein Structure

The secondary structure of a protein is defined by the steric relationships between the amino acids that are close to each other in the primary amino acid sequence.

3 important examples of 2° structure

- α - helix
- β - pleated sheet
- β - turn (hairpin turn)
Secondary Structure

Local structure of consecutive amino acids

Common regular secondary structures

- α Helix
- β Sheet
- β turn

Amino acids have a greater propensity to form some secondary structures versus others

- Chemical properties
- Spatial constraints
Secondary Structure- α Helix

- α Helix
  - H-bond
  - 1.5 Å rise
  - 100-degree rotation
  - 5 Å
Three representations of an alpha helix
Beta pleated sheet

- Amino acid chain folds back on itself
- Also stabilized by hydrogen bonding between amino acids in parallel chains
Beta pleated sheet
Secondary Structure - β Sheet

- Oxygen
- Carbonyl C
- H Bond
- Nitrogen
- Carbon α
- Hydrogen
- R Group
Polypeptide with both configurations

α helix

β sheet

Domain
Tertiary Protein Structure

- Alpha helix or beta sheet is folded into a specific 3-dimensional shape
- Shape is stabilized by various interactions among the R-groups of the polypeptide chain
The tertiary structure of two proteases
Cytochrome b$_{592}$

An NAD binding domain

Varshah domain of IgG
Tertiary Structure

- Final conformation is lowest energy level
- Polypeptide chain will spontaneously fold into this conformation
- If the protein is denatured (unfolded by various chemical treatments), it will refold spontaneously into its original conformation
- Protein folding in a cell assisted by molecular chaperones
Quarternary protein structure

- Found in proteins formed from more than one polypeptide chain
- Each polypeptide chain is separately folded into its tertiary structure
- Separate polypeptide chains held together by various interactions between the chains
Subunits assemble to form quaternary structures.
Monomer subunits can assemble to form very high order structures.
Hemoglobin
What Causes $2^\circ$, $3^\circ$, $4^\circ$ Structure?

- intermolecular forces
  - peptide hydrogen bonds
  - side-chain hydrogen bonds
  - salt bridges
  - London dispersion forces

- metal ion coordination

- disulfide bridges (covalent interaction)
Forces in Proteins

- Metal ion coordination
- Hydrophobic interactions
- Disulfide bond
- Electrostatic attraction
- Side chain hydrogen bonding
Some Ways to Denature Proteins

- heat
- pH changes
- chemical reagents
  - urea
  - DMSO
  - HMPA
  - 2-mercaptoethanol
Give the name and structure of at least 2 examples of each of the following:

a. heterocyclic amino acid
b. aromatic amino acid
c. neutral amino acid
d. acidic amino acid
e. basic amino acid
f. sulfur containing amino acid
Show me a chemical structure of a tripeptide with 1 nonpolar, 1 polar, and 1 charge amino acids? Ensure that the charged amino acid is at the C-terminal and the polar amino acid is at the N-terminal.
1. How many peptide bonds are there?
2. What are amino acids?
3. How many α-carbon can you find?
4. How many chiral centers?
5. In a solution of these oligopeptides, what sort of interactions can develop?
Protein Structure - 5
Hierarchy of Protein Structure

• 20 different amino acids: many combinations

Primary Structure
The order of amino acids: Protein sequence

Secondary Structure
Conformation varies depending on sequence

Tertiary Structure
Overall structure of the chain in full 3D
Protein Sequences

A protein of $n$ residues $20^n$ possible sequences!

100 residue protein has $100^{20}$ possibilities $1.3 \times 10^{130}$!

The latest estimates indicate on 40,000 sequences in the human genome $\rightarrow$ **THERE MUST BE RULES!**

- 20 different amino acids: many combinations
Protein Sequences

A protein of n residues $20^n$ possible sequences!

100 residue protein has $100^{20}$ possibilities $1.3 \times 10^{130}$!

The latest estimates indicate on 40,000 sequences in the human genome $\rightarrow$ THERE MUST BE RULES!

- 20 different amino acids: many combinations

Amino terminus

Residue number

Carboxyl terminus

N

1

2

3

4

C

$\binom{N}{C}$
Protein Sequences

*Length is generally 100-1000 residues*

Minimum length for performing a function ~40

Molecular machines not perfect- errors

Function requires specific amino acid properties

- Not all amino acids are equally useful
- Abundant: Leu, Ala, Gly, Ser, Val, Glu
- Rare: Trp, Cys, Met, His

Post-translational modifications

- Addition of co-factors- metals, hemes
- Chemical modification- phosphate, glycos.
Protein Sequences

The pattern of amino acid side chains determines the secondary (and tertiary!!) structure

*Pattern is more important than exact sequence*

Reporting/Comparing Protein Sequences

<table>
<thead>
<tr>
<th></th>
<th>h-CaM</th>
<th>A</th>
<th>T</th>
<th>V</th>
<th>R</th>
<th>L</th>
<th>L</th>
<th>E</th>
<th>W</th>
<th>E</th>
<th>D</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>b-CaM</td>
<td>A</td>
<td>T</td>
<td>V</td>
<td>R</td>
<td>L</td>
<td>L</td>
<td>E</td>
<td>Y</td>
<td>K</td>
<td>D</td>
<td>L</td>
<td></td>
</tr>
</tbody>
</table>

5 conservative

10 non-conservative
Secondary Structure

Local structure of consecutive amino acids

Common regular secondary structures

- $\alpha$ Helix
- $\beta$ Sheet
- $\beta$ turn

Amino acids have a greater propensity to form some secondary structures versus others

- Chemical properties
- Spatial constraints
The Peptide Bond

- Peptide plane is flat
  $\omega$ angle $\sim 180^\circ$

- Partial double-bond:

\[
\begin{align*}
-C - \dot{\text{N}} - & \quad \longleftrightarrow \quad -C = \text{N} - \\
\text{H} & \quad \text{H} \\
\text{O} & \quad \text{O}^-
\end{align*}
\]

Resonance structures
- $\omega$ angle is fixed, $\phi$ and $\psi$ angles vary
- Many $\phi/\psi$ combinations cause atoms to collide
Side chains collision also limit $\phi/\psi$ combinations

Backbone restricted $\rightarrow$ Secondary structure limited
Specific Secondary Structures

Three acceptable backbone $\phi/\psi$ combinations

1. Right-hand helix: $\alpha$-helix (-40°, -60°)
2. Extended: antiparallel $\beta$-sheet (140°, -140°)
3. Left-hand helix (*rare*): $\alpha$-helix (45°, 45°)

- Glycine: special because it has no side chain!

Side chains positioned to minimize collisions $\rightarrow$

- amino acids prefer specific secondary structures

**Hydrogen bonds** between backbone atoms provide unique stability to secondary structures
Secondary Structure - β Sheet

- Oxygen
- Carbonyl C
- H Bond
- Nitrogen
- Carbon α
- Hydrogen
- R Group
Secondary Structure - β Turn

- Reverses direction of the chain
Titration of Amino Acids
- 6
Titration curves

Of amino acids and weak acids (acetic acid)
Titration

• Titration curves are produced by monitoring the pH of given volume of a sample solution after successive addition of acid or alkali

• The curves are usually plots of pH against the volume of titrant added or more correctly against the number of equivalents added per mole of the sample
Titration of acetic acid

• At the starting point the acid form predominates (CH₃COOH).
• As strong base is added (e.g. NaOH), the acid is converted to its conjugate base.
• At the mid point of the titration, where pH=pK, the concentrations of the acid and the conjugate base are equal.
• At the end point (equivalence point), the conjugate base predominates, and the total amount of OH added is equivalent to the amount of acid that was present in the starting point.
Titration

Determination of pKa values:

pKa values can be obtained from the titration data by the following methods:

1. The pH at the point of inflection is the pKa value and this may be read directly

2. By definition the pKa value is equal to the pH at which the acid is half titrated. The pKa can therefore be obtained from the knowledge of the end point of the titration.
Titration of amino acids

• Titration of glycine
• Titration of arginine
Titration

• When an amino acid is dissolved in water it exists predominantly in the isoelectric form.
• Upon titration with acid, it acts as a base, and upon titration with base, it acts as an acid (a compound that can act as either an acid or a base is known as an amphoteric compound).
• $^{+}\text{H}_3\text{N-CH}_2\text{-COO}^- + \text{HCl} \rightarrow ^{+}\text{H}_3\text{N-CH}_2\text{-COOH} + \text{Cl}^-$  
  (base)  (acid)  (1)

$^{+}\text{H}_3\text{N-CH}_2\text{-COO}^- + \text{NaOH} \rightarrow \text{H}_2\text{N-CH}_2\text{-COO}^- + \text{Na}^+ + \text{H}_2\text{O}$  
  (acid)  (base)  (2)

In this experiment, the amino acid represents either the $\text{A}^-$ or the HA form in the Henderson-Hasselbalch equation, depending on the titration.
Acid–base properties

- All of the amino acids have an acidic group (COOH) and a basic group (NH2) attached to the α carbon.
- Two of the amino acids have acidic side chains: aspartate and glutamate.
- Three of the amino acids have basic side chains: arginine, histidine, and lysine.
• All amino acids contain ionizable groups that act as weak acids or bases, giving off or taking on protons when the pH is altered.

These ionizations follow the Henderson-Hasselbalch equation:
\[ \text{pH} = \text{pK}_a + \log \frac{[\text{unprotonated form (base)\,}]}{[\text{protonated form (acid)\,}]}. \]
• When the conc of the unprotonated form equals that of the unprotonated form, the ratio of their concentrations equals 1, and log 1=0.

• Hence, pKa can be defined as the pH at which the concentrations of the protonated and unprotonated forms of a particular ionizable species are equal.

• The pKa also equals the pH at which the ionizable group is at its best buffering capacity; that is the pH at which the solution resists changes in pH most effectively.
• Consider applying the Henderson-Hasselbalch equation to the titration of glycine with acid and base.

• Glycine has two ionizable groups: a carboxyl group and an amino group, with pKa values of 2.4 and 9.6 respectively.

• In water at pH 6, glycine exists as a dipolar ion, or zwitterion, in which the carboxyl group is unprotonated (-COO\(^{-}\)) and the amino group is protonated to give the substituted ammonium ion (-NH\(_3\)^{+}).
• Addition of acid to the solution lowers the pH rapidly at first and then more slowly as the buffering action of the carboxyl is exerted.
• At pH 2.4 the pKa is reached, one-half the acid has been consumed, and the carboxyl group is half ionized and is most effective as a buffer.
• Titration of the amino group with base follows a similar curve into the alkaline region.
• The intersection between the titration of the carboxyl group and the titration of the amino group describes in this case the point at which glycine has no net charge, and is called the isoelectric point (pI).
The isoelectric point (pI)

- The isoelectric point, pI, is the pH of an aqueous solution of an amino acid at which the molecules have no net charge. In other words, the positively charged groups are exactly balanced by the negatively charged groups.

- For simple amino acids such as alanine, the pI is an average of the pKₐ's of the carboxyl (2.34) and ammonium (9.69) groups. Thus, the pI for alanine is calculated to be: \( \frac{2.34 + 9.69}{2} = 6.02 \).

- If additional acidic or basic groups are present as side-chain functions, the pI is the average of the pKₐ's of the two most similar acids.
Cont.. (pI)

• In the case of aspartic acid, the similar acids are the alpha-carboxyl function ($pK_a = 2.1$) and the side-chain carboxyl function ($pK_a = 3.9$), so $pI = (2.1 + 3.9)/2 = 3.0$.

• For arginine, the similar acids are the guanidinium species on the side-chain ($pK_a = 12.5$) and the alpha-ammonium function ($pK_a = 9.0$), so the calculated $pI = (12.5 + 9.0)/2 = 10.75$. 
Glycine

\[ \text{NH}_3 \quad \text{CH}_2 \quad \text{COOH} \quad \underset{pK_1}{\rightleftharpoons} \quad \text{NH}_3 \quad \text{CH}_2 \quad \text{COO}^- \quad \underset{pK_2}{\rightleftharpoons} \quad \text{NH}_2 \quad \text{CH}_2 \quad \text{COO}^- \]

\[ pK_1 = 2.34 \]

\[ pK_2 = 9.60 \]

\[ \text{pI} = 5.97 \]
• Most amino acids contain carboxyl and amino groups having pKa values similar to those of glycine.
• In addition to these groups, many amino acids contain other ionizable groups, which introduce other “steps” or pKa values into their titration curves.
Titration curves

• The pK is the pH at the midpoint of the buffering region (where the pH changes only slightly upon addition of either acid or base).

• The pK is the pH corresponding to the inflection point in the titration curve.

• The end point of a titration curve represents the observed end of the titration.

• The isoelectric point (isoelectric pH; pI) is the pH at which the amino acid has a net zero charge. For a simple diprotic amino acid, the pI falls halfway between the two pK values. For acidic amino acids, the pI is given by $\frac{1}{2}(pK_1 + pK_2)$ and for basic amino acids it’s given by $\frac{1}{2}(pK_2 + pK_3)$.
Arginine

\[ \text{pK}_a^1 = 2.1 \]
\[ \text{pK}_a^2 = 9.0 \]
\[ \text{pK}_a^3 = 12.5 \]

\[ \text{pI} = 10.8 \]

\( \text{pH} \) vs. Equivalents of \( \text{OH}^- \)

Chemical structures:

- \( \text{H}_2\text{N} - \text{H} \)
- \( \text{H}_2\text{N} - \text{H} \)
- \( \text{H}_2\text{N} - \text{H} \)
- \( \text{CO}_2\text{H} \)
- \( \text{H}_3\text{N} - \text{H} \)
- \( \text{H}_2\text{N} - \text{NH}_2 \)
- \( \text{H}_2\text{N} - \text{NH} - \text{CH}(\text{CH}_2)_3 \)

Chemical properties:

- pH > 13
- pH ≈ 10.8
- pH < 2
The diagram illustrates the ionization states of lysine (Lys) under different pH conditions. The pH scale is shown on the y-axis, ranging from 0 to 14, and the equivalents of OH⁻ added are shown on the x-axis. The graph shows three pK values: pK₁, pK₂, and pK₃, corresponding to the ionization of the amino, carboxyl, and imidazole groups of lysine, respectively. The isoelectric point is indicated by the curve's steep change in pH with the addition of OH⁻ equivalents.