Chapter 5: Amino Acids, Peptides & Proteins

- Amino acids share common functional groups
  - Amino, Carboxyl & H bonded to C<sub>α</sub>
- Distinct chemistry of αα’s result of side chains
  - αα’s categorized on basis of R group (side chain)
    - Nonpolar, aromatic, polar, (+)-charged, (-)-charged
- αα’s can act as both weak acids & weak bases
  - Some R groups can also ionize (HA<sup>-</sup> + H<sup>+</sup> A<sup>-</sup>)

- Proteins & polypeptides are polymers of αα’s
  - Peptides < 100 residues, Proteins > 100 residues
- Proteins are studied using variety of methods
  - Solubility, size, shape, charge, binding patterns
- Protein structure is defined at 4 levels
  - Primary (1<sup>st</sup>) Secondary (2<sup>nd</sup>) Tertiary (3<sup>rd</sup>) Quarternary (4<sup>th</sup>)
- Proteins with similar functions often have similar ααsequences

Amino Acids - the alphabet of protein structure

![Amino acid structure](image)

Amino acids share common structural features

3-letter and 1-letter amino acid abbreviations

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>3-letter</th>
<th>1-letter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine</td>
<td>Ala</td>
<td>A</td>
</tr>
<tr>
<td>Leucine</td>
<td>Leu</td>
<td>L</td>
</tr>
<tr>
<td>Proline</td>
<td>Pro</td>
<td>P</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>Trp</td>
<td>W</td>
</tr>
</tbody>
</table>

The alpha carbon is a chiral center

![Alpha carbon structure](image)
### Stereisomerism in amino acids

<table>
<thead>
<tr>
<th>αCHO</th>
<th>CHO</th>
</tr>
</thead>
<tbody>
<tr>
<td>HO⁻²C⁻H</td>
<td>H⁻C⁻OH</td>
</tr>
<tr>
<td>³CH₂OH</td>
<td>CH₂OH</td>
</tr>
<tr>
<td>L-Glyceraldehyde</td>
<td>D-Glyceraldehyde</td>
</tr>
<tr>
<td>COO⁻</td>
<td>COO⁻</td>
</tr>
<tr>
<td>H₂N⁻C⁻H</td>
<td>H⁻C⁻NH₃</td>
</tr>
<tr>
<td>CH₃</td>
<td>CH₃</td>
</tr>
</tbody>
</table>

L-Alanine  D-Alanine

### Amino Acids are classified according to their R group

<table>
<thead>
<tr>
<th>Nonpolar, Aliphatic R groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>COO⁻</td>
</tr>
<tr>
<td>H₂N⁻C⁻H</td>
</tr>
<tr>
<td>Glycine</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Aromatic R groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>COO⁻</td>
</tr>
<tr>
<td>H₂N⁻C⁻H</td>
</tr>
<tr>
<td>Phenylalanine</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Polar, Uncharged R groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>COO⁻</td>
</tr>
<tr>
<td>H₂N⁻C⁻H</td>
</tr>
<tr>
<td>Serine</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Positively Charged (Basic) R groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>COO⁻</td>
</tr>
<tr>
<td>H₂N⁻C⁻H</td>
</tr>
<tr>
<td>Lysine</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Negatively Charged (Acidic) R groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>COO⁻</td>
</tr>
<tr>
<td>H₂N⁻C⁻H</td>
</tr>
<tr>
<td>Aspartate</td>
</tr>
</tbody>
</table>

Fig 5.5
Nonstandard amino acid residues

Amino acids can act as acids and bases

Nonionic
Zwitterionic
form
form

Zwitterions - dipolar ions

Amino Acids have characteristic titration curves

Effects of chemical environment on $pK_a$

$\alpha\alpha$’s have at least 2 regions of buffering power

Isoelectric point calculations (pI)

(1) determine the number of ionizable functional groups (-COOH, -NH$_3^+$, -OH, -SH, etc.)

(2) Draw the series of proton losses from low to high pH

(3) Determine which species has a net charge of 0

(4) Average the pK’s which bracket that ionization
**Histidine - Biological Proton donor/Acceptor**

Only αα with R group pK$_n$ near physiological pH (7.4)

**Peptides and Proteins**

αα’s are covalently joined through amide linkages

**Peptide & Protein Terminology**

- Oligopeptide - few αα’s
- Polypeptide - > 20 αα’s
- Peptides, proteins have amino terminal (N-terminal) and carboxyl-terminal (C-terminal) ends
- Peptides & proteins also have isolectric points (pI’s)

**Biologically Active Peptides**

- Aspartame (Nutrasweet) 2 αα’s
- Oxytocin (uterine contractions) 9 αα’s
- Bradykinin (inflammation inhibitor) 9 αα’s
- Insulin (sugar uptake) 30/21 αα’s

Compare to proteins:

- Cytochrome C (energy metabolism) 104 αα’s
- Titin (muscle protein) 27000 αα’s
Proteins can have single or multiple polypeptide chains

- Multiple subunits = Multimeric proteins

- Subunits can be same or different
  - Ex: actin (identical subunits)
  - Ex: Hemoglobin (different subunits)

- Dimers, trimers, tetramers, etc. refer to number of subunits in multimeric proteins

- Subunits held together by noncovalent and covalent (disulfide bonds) linkages

Polypeptides have characteristic αα composition

- Hydrolysis yields characteristic proportion of αα’s

- The standard 20 αα’s almost never occur in equal proportions

- Proteins with different functions will differ significantly in their respective proportions

- Crude yet powerful method of protein identification

---

Disulfide bonds

```
\[ \text{Cysteine} \quad \begin{array}{c}
\text{COO}^- \\
\text{HLN}^– \\
\text{CH}_2
\end{array} \quad \begin{array}{c}
\text{COO}^- \\
\text{HLN}^– \\
\text{CH}_2
\end{array} \quad \text{Cystine} \]
```

*Fig. 5.7*

---

Chemical modification of proteins

<table>
<thead>
<tr>
<th>Conjugated Proteins</th>
<th>Class</th>
<th>Prosthetic group(s)</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lipoproteins</td>
<td>Lipids</td>
<td></td>
<td>α-Lipoprotein of blood</td>
</tr>
<tr>
<td>Cytocarboxylic</td>
<td>Carboxylates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phosphoproteins</td>
<td>Phosphate groups</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hemoproteins</td>
<td>Heme (iron porphyrins)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flavoproteins</td>
<td>Flavin nucleotides</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metalloproteins</td>
<td>Inter</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metalloproteins</td>
<td>Zinc</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Calcium</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Molybdenum</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Copper</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 5.4*

---

Levels of structure in Proteins

![Levels of structure in Proteins](image)

1° 2° 3° 4°

---

Working with proteins

Proteins must be purified before they can be characterized

- Knowledge of structure and function

Proteins are purified based on their physical and chemical properties

- Size (fractionation)
- Solubility (function of pH, temperature, salt)
- Charge (binding to oppositely charged compounds)
- Binding properties (ligands)
Chromatography
Takes advantage of differences in:
- Charge
- Size
- Binding affinity

Mobile (liquid) and Stationary (matrix) phases

Unseparated proteins can be quantified by measuring catalytic activity

Activity = total units of enzyme in solution
Specific Activity = number of units of enzyme per milligram of protein

Proteins can be purified several thousand fold via combination of chromatography steps

<table>
<thead>
<tr>
<th>Table 5-5</th>
<th>A Purification Table for a Hypothetical Enzyme*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedure</td>
<td>Fraction volume (ml)</td>
</tr>
<tr>
<td>1. Crude cellular extract</td>
<td>1,400</td>
</tr>
<tr>
<td>2. Precipitate with ammonium sulfate</td>
<td>280</td>
</tr>
<tr>
<td>3. Ion-exchange chromatography</td>
<td>90</td>
</tr>
<tr>
<td>4. Size-exclusion chromatography</td>
<td>80</td>
</tr>
<tr>
<td>5. Affinity chromatography</td>
<td>6</td>
</tr>
</tbody>
</table>

*All data represent the status of the sample after the designated procedure has been carried out. Activity and specific activity are defined on page 137.

Electrophoresis - Isoelectric point (pI)

Purified proteins
- Size (electrophoresis)
- Charge (electrophoresis or chromatography)
- Catalytic activity (enzyme assays)
- Binding properties (ligands)

Detailed Biochemical analysis (Structure/Function)
- 1° - amino acid sequence
- 2° - Circular dichroism
- 3° - X-ray crystallography, NMR

Covalent Structure of Proteins

Electrophoresis: Separation & Characterization

Molecular weight

- Myosin 200,000
- maltodextrin 150,000
- Glycogen phosphorylase 97,000
- Bovine serum albumin 68,000
- Ovalbumin 45,000
- Carbonic anhydrase 21,000
- Ribozyme trypsin inhibitor 21,000
- Lysozyme 14,400
First degree provides important Biochemical Information

- Insights into 3D structure
  - α-helix, β-sheet, etc.
- Cellular location of protein
  - Cytosol, plasma membrane, nucleus, etc.
- Evolutionary relationships
- Genetic disease

Determining the 1°

1. Purification of protein
2. Break disulfide (-S-S-) bonds
3. Cleave protein into smaller peptide fragments
   - Enzymatic digestion
4. Chemical modification & hydrolysis
5. Identification of individual amino acids
   - Isoelectric point, etc.

First degree provides information on structure & function

- Cellular location or chemical modification
  - N-terminal sequences - signals for export to nucleus, etc.
  - Ser, Thr, Tyr context - phosphorylation signals
  - Asn, Ser, Thr context - glycosylation signals
- Comparison to sequences w/known structure
  - Functional similarity
  - Evolutionary relationship

First degree sequence information (contd.)

- Structural domains
  - Catalytic sites: KTGGL (glucose pocket)
  - LxxxxLxxxxxL: “leucine zipper”
  - NNRKN (Basic (+) residues): DNA binding domain
  - Etc.
- Genetic disease
  - Sickle cell anemia, Cystic Fibrosis, MD
  - Comparison to healthy individuals
  - Single amino acid changes in many cases

Protein homology & Polymorphism

- Homologous proteins share a significant amount of sequence identity (>25%)
- Evolutionarily related
- Usually perform the same function in different species
- Most proteins are polymorphic - exhibit variation in ααα sequence within species

Positions which vary in ααα sequence are called variable residues

Gradation in variation:
- Conservative (hydrophobic for hydrophobic ααα)
- Nonconservative (Polar for hydrophobic, etc.)
### Biomedical & related benefits

- Evolutionary relationships
  - Anthropology
- Forensic Science
  - Identification of individuals, populations
- Proteomics
  - Examination of the expression of all proteins in a cell
  - Comparison of healthy/diseased states - what proteins are (not) expressed?
  - Provide targets for drug development

### Chapter 5 - Summary

- 20 Standard αα’s found in proteins
  - α-COOH group
  - α-NH₂ group
  - Distinctive side chain (R group)
- α-carbon (central carbon) is asymmetric in all αα’s (except glycine)
  - 2 stereoisomeric forms (D- and L-)
  - Only L-form in proteins

### Summary (contd.)

- αα’s are often characterized by their isoelectric point (pI) - pH where they have no net charge
- αα’s are covalently joined through peptide bonds
  - Amino (N-) and Carboxy (C-) terminal ends
- Proteins are often conjugated to other molecules
  - Metal ions
  - Lipids
  - Carbohydrates
  - Etc.

### Summary (contd.)

- 4 levels of protein structure
  - Primary, secondary, tertiary & quaternary
- Protein structural & functional analysis
  - Solubility (Precipitation with salts)
  - Chromatography (Size, Charge, Binding affinity)
  - Electrophoresis (Size and charge)
- 1° provides important Biochemical Information
  - 3D structure, active sites, targeting signals, etc.
- Protein homology and polymorphism