I. PROTEIN COMPOSITION AND STRUCTURE: A REVIEW OF THE BASICS

Proteins occupy a central position in the structure and function of all living organisms. Some proteins serve as structural components while others function in communication, defense, and cell regulation. The enzyme proteins act as biological catalysts which control the pace and nature of essentially all biochemical events. Indeed, although DNA serves as the genetic blueprint of a cell, none of the life processes would be possible without the proteins.

Amino Acids - Building Blocks of Proteins

The fundamental unit of proteins is the amino acid. The common amino acids have the general structure shown in Figure 1. Each amino acid has an amino group (NH₂) and a carboxylic acid group (COOH) attached to a central carbon atom called the alpha carbon. Also attached to the alpha carbon are a hydrogen atom and an R group or side chain.

![Figure 1. General Structure of Alpha-Amino Acids.](image)

The C stands for a carbon atom; C* is the alpha carbon; H is hydrogen; N is nitrogen, O is oxygen, -NH₂ is an amino group and -COOH is the carboxylic acid group. R is a general term for any one of several different side chains that determine the nature of different amino acids.

There are 20 amino acids commonly found in proteins and these differ from each other in the nature of the R-groups attached to the alpha carbon. A convenient classification of amino acids depends on the number of acidic and basic groups that are present. Thus the neutral amino acids contain one amino and one carboxyl group. The acidic amino acids have an excess of acidic carboxyl over amino groups. The basic amino acids possess an excess of basic amino groups. Table 1 lists the major amino acids found in proteins.
Table 1. Amino Acids Found in Proteins

<table>
<thead>
<tr>
<th>Neutral Amino Acids</th>
<th>Acidic Amino Acids</th>
<th>Basic Amino Acids</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycine</td>
<td>Alanine</td>
<td>Aspartic Acid</td>
</tr>
<tr>
<td>Valine</td>
<td>Leucine</td>
<td>Glutamic Acid</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>Serine</td>
<td></td>
</tr>
<tr>
<td>Threonine</td>
<td>Cysteine</td>
<td></td>
</tr>
<tr>
<td>Methionine</td>
<td>Phenylalanine</td>
<td></td>
</tr>
<tr>
<td>Tyrosine</td>
<td>Tryptophan</td>
<td></td>
</tr>
<tr>
<td>Proline</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The Peptide Bond and the Primary Structure of Proteins

Proteins are composed of amino acids linked into chains by peptide bonds as shown in Figure 2. Two amino acids joined by a single peptide bond form a dipeptide; three amino acids form a tripeptide, and a large number of amino acids joined together constitute a polypeptide. A protein is a polypeptide chain that contains more than 50-100 amino acids. The monomer units in the chain are known as amino acid residues. The average protein contains about 350 amino acid residues although proteins with as many as 1000 residues and those with as few as 100 are not uncommon.

The sequence of amino acids along a polypeptide chain is referred to as the primary structure of the protein. The primary structure of the protein myoglobin is given in Figure 3. This protein serves to bind and store oxygen in muscle. The primary structure of over 500 different proteins is now known.

Figure 3. The Primary Structure of Whale Myoglobin.

This protein consists of 153 amino acid residues. Each residue in the protein is given a 3-letter abbreviation (thus: lysine = Lys; valine = Val). Proteins are always written with the free amino or N-terminus toward the left.

Three-Dimensional Protein Structure

In the cell, the polypeptide chain is folded into a highly ordered shape or conformation. Most proteins are globular in shape and these proteins are usually soluble in water or in aqueous media containing salts. This group includes the enzymes, antibodies, and a variety of other proteins. Less frequently, proteins are long and fibrous and most of these elongated molecules are insoluble in water and serve a role in the maintenance of cell structure.

The three-dimensional structure of a protein is due to the type and sequence of its constituent amino acids. Since the amino acid sequence of each protein is unique, it follows that different proteins assume different shapes. Thus, there is a remarkable diversity of three-dimensional protein forms. The conformation of a protein is usually of critical importance in the protein's function. For example, a protein can be unfolded into a polypeptide chain that has lost its original shape. In general, proteins such as enzymes are rendered nonfunctional upon unfolding because functional activity is dependent on the proteins native shape. This process is called denaturation. Most proteins can be denatured by heating, by certain detergents, and by extremes of pH. The ionic detergent, sodium dodecyl sulfate (SDS), is often used to denature proteins. The denaturing treatment can frequently be reversed, for example by removing the detergent or by neutralizing the pH. During this renaturing process, the polypeptide chain spontaneously refolds into its original conformation and the protein regains its biological activity. A similar folding process occurs in the cell for when a polypeptide is constructed on the ribosomes, it folds into a biologically active conformation. Thus, the three-dimensional folding of a protein and its biological properties are directed by the sequence of amino acid residues along the polypeptide chain.

Biochemists have identified three structural levels that define the three-dimensional shape of a protein. These levels of organization are secondary structure, tertiary structure, and quaternary structure. Figure 4 shows examples of these levels of organization. The major force involved in the formation and maintenance of these structures are various types of weak, noncovalent bonds that are formed between the amino acid residues and between the amino acid residues and water. Although a noncovalent bond typically has less than 1/20 the strength of a covalent bond, a large number of noncovalent bonds participate in the folding of a single protein into its native
Secondary Structure

The spatial arrangement of the protein backbone that is generated from folding of the polypeptide chain is called the secondary structure of the protein. The secondary structures of proteins are stabilized by hydrogen bonds in which a hydrogen serves as a bridge between oxygen and nitrogen atoms (-C=O HN-). A common secondary structure is the \( \alpha \)-helix which consists of a single polypeptide chain coiled into a rigid cylinder. In the \( \alpha \)-helix, each peptide bond along the polypeptide is itself hydrogen bonded to other peptide bonds. Many enzymes contain small regions of \( \alpha \)-helices, while long sections of the \( \alpha \)-helix arc often found in proteins involved in cell structure. Another type of secondary structure of proteins is the \( \beta \)-sheet, which is a central organizing feature of enzymes, antibodies, and most other proteins that perform nonstructural functions. Here, a single polypeptide chain folds back and forth upon itself to produce a rather rigid sheet. Hydrogen bonds between neighboring polypeptide chains are a major stabilizing force for the \( \beta \)-sheet conformation.

Tertiary Structure

The tertiary structure of a protein describes the detailed features of the three dimensional
conformation of the polypeptide chain. It is brought about by the interactions between the amino acid side chains that cause the folding and bending of α-helix and β-sheet segments of the protein. One very important interaction at this level of organization involves the hydrophobic and hydrophilic side chains of the amino acid residues. Hydrophobic amino acids, such as phenylalanine and leucine, show limited solubility in water. Thus, these hydrophobic residues in a protein tend to cluster on the inside of the protein in order to avoid contact with the aqueous environment. Hydrophilic amino acids such as glutamic acid and lysine are readily soluble in water, and thus these amino acids arrange themselves on the surface of the protein molecule, where they can interact with water and with other hydrophilic side chains. The consequence of these interactions is that a polypeptide chain typically folds spontaneously into a stable, usually globular structure, with the hydrophobic side chains packed into the central core of the protein and the hydrophilic side chains forming the irregular, external surface.

**Quaternary Structure**

Some proteins contain more than one polypeptide chain. For example, each molecule of human hemoglobin consists of four polypeptide chains which are held together by a variety of noncovalent bonds. The arrangement of the polypeptide in such proteins is called the quaternary structure.

**II. Theoretical Aspects of Electrophoresis**

**General Description**

Electrophoresis is the movement of charged molecule under the influence of an electric field. Because amino acids and proteins are charged molecules, they migrate in an electric field at appropriate pH values. In the most common form of electrophoresis, the sample is applied to a stabilizing medium that serves as a matrix for the buffer in which the sample molecules travel. The agarose gel is frequently used in the electrophoretic separation of native (nondenatured) proteins since low percentage gels (e.g. 1% agarose) form a sponge-like network which serves as a medium for the buffer, but has pores large enough to allow even the largest proteins to pass unimpeded.

A diagram of the essential components of an agarose electrophoretic system is shown in Figure 6. The agarose gel, containing preformed sample wells, is submerged in buffer that is contained within the electrophoretic chamber. Samples to be separated are then loaded into the sample wells. Current from the power supply travels to the negative electrode (cathode), supplying electrons to the conductive buffer solution, gel, and positive electrode (anode), thus completing the circuit.
Separation of Amino Acids and Proteins by Electrophoresis

All amino acids contain at least one amino and one carboxyl group. In acid solutions, the amino groups are positively charged while the carboxyls are not ionized (Figure 5). Therefore, in strong acid solutions, amino acids are positively charged and migrate in an electric field to the negative electrode. In basic or alkaline solutions, the carboxyls are negatively charged while amino groups are not ionized. It follows then that in strong alkaline solutions, amino acids are negatively charged, and migrate to the positive electrode during electrophoresis.

Figure 5. Amino Acids in Acid and Alkaline Solutions.
Thus there must be an intermediate pH at which each amino acid bears no net charge and does not migrate in an electric field. The pH at which an amino acid or protein does not migrate in an electric field is called the isoelectric point. Most neutral amino acids have isoelectric points around pH 6.0. The isoelectric points of aspartic acid and glutamic acid, however, are close to pH 3. Therefore, at pH 6, these acidic amino acids carry a negative charge and migrate, to the positive electrode during electrophoresis. The isoelectric points of the basic amino acids, lysine and arginine, are pH 9.7 and 10.8, respectively. These amino acids carry a positive charge at pH 6, and, hence migrate to the negative electrode. These differences in charge permit the electrophoretic separation of acidic, neutral, and basic amino acids at pH 6, as illustrated in Figure 7.

The positively and negatively charged side chains of proteins cause them to migrate like amino acids in an electric field. The electrochemical character of a protein is dependent primarily on the numerous positively charged ammonium groups (-NH$_3^+$) of lysine and arginine and the negatively charged carboxyl groups (-COO$^-$) of aspartic acid and glutamic acid. The isoelectric points of most proteins is in the range of pH 5 to 7.
Electrophoresis of proteins is usually performed at a pH above the isoelectric point of most proteins. At pH 8.6 most proteins are negatively charged and when applied to sample wells at the negative electrode end of the gel, they travel towards the positive electrode. The rate of migration of a protein species in an electric field depends upon its net charge; the higher the charge the faster the protein will travel. For example, serum albumin, which has an isoelectric point of 4.8, will carry a strong negative charge in a buffer of pH 8.6 as compared to myoglobin, which has an isoelectric point of 7.2. Therefore, at pH 8.6 albumin will migrate toward the positive electrode at a much faster rate than myoglobin. These considerations form the basis for the electrophoretic separation of proteins according to net charge.

III. PRACTICAL ASPECTS OF ELECTROPHORESIS

Electrophoresis Equipment

A horizontal electrophoresis unit is composed of an acrylic cell with central platform, platinum electrodes, four removable get casting trays, four sample well forming combs and a safety lid with power cords. The four gels are made in the casting trays and then placed on the central platform of the electrophoresis cell. Each gel contains 8 separate sample wells.

The power supply produces a constant voltage output of 85 or 170 volts. Voltage selection is controlled by the switch located in the center of the front panel. The ammeter, also located on the front panel, permits the current to be monitored during an electrophoretic run. The unit can reach a maximum of 500 mA.
Electrophoresis Chemicals

Agarose Gel

A number of different types of stabilizing supports have been used in the electrophoretic separation of proteins. These include filter paper, cellulose acetate, and gels composed of either starch, polyacrylamide, agar or agarose. The agarose gel is an ideal solid support for the separation of proteins on the basis of charge, and agarose gel electrophoresis is used extensively in both biochemical and clinical laboratories. Agarose is a natural polysaccharide of galactose and 3,6-anhydrogalactose derived from agar, which in turn is obtained from certain marine red algae. Agarose gels are made by dissolving the dry powder in boiling buffer, pouring the gels into casting trays, and allowing them to set by cooling at room temperature. The gel forms a porous sponge-like network which serves to hold the buffer and proteins during the electrophoretic separation. Agarose gels are frequently run in the submarine mode where the gel is completely immersed in buffer. This feature reduces heat development in the gel, which could otherwise lead to protein band distortion.

Electrophoresis Buffer

The migration of a protein through an agarose gel is dependent primarily on the net electric charge on the protein, the intensity of the electric field, and the pH and ionic strength of the buffer. The pH of the electrophoresis buffer is critical because it determines the net charge on the protein molecules in the sample. The Tris-Glycine buffer used in these experiments is pH 8.6, which is higher than the isoelectric points of most proteins. Therefore, most proteins carry a net negative charge at pH 8.6, and migrate toward the positive electrode during the electrophoretic separation. The ionic strength of the buffer solution is also important in electrophoresis. High ionic strength buffers permit fast migration and can promote the sharpening of protein zones. However, high ionic strength buffers may also cause high heat production which can lead to band distortion. The moderate ionic strength buffer used in these experiments permits optimal resolution of protein bands in minimal time.

Staining and Destaining

Most protein are not colored, and therefore it is necessary to visualize them in some way in order to determine their position in the agarose gel after electrophoresis. The most commonly used stain for the detection of proteins is Coomassie blue, and this stain has been incorporated into the staining solution that you will use. The staining solution also contains acetic acid which serves to precipitate and immobilize the proteins in the structure of the gel matrix after electrophoresis. The acid serves to fix the proteins in the gel so that the protein bands do not become blurred by diffusion. After the proteins in the gel have been stained, the unbound dye must be rinsed from the gel by a process known as destaining. A dilute solution of acetic acid and methanol is often used for the destaining of the agarose gels.
Sample Buffer

The protein samples in these experiments are loaded into the wells of the agarose gel as 10-20% glycerol solutions. The viscous glycerol ensures that the samples will layer smoothly at the bottom of the sample wells. The sample buffer also contains the tracking dye, bromophenol blue. As will be described below, this dye enables the investigator to follow the progress of an electrophoretic run.

IV.  PROCEDURES FOR THE PREPARATION, ELECTROPHORESIS AND STAINING OF AGAROSE GELS

Pouring the Agarose Gels

1. Place the casting tray on a level work surface and place a precleaned glass slide into the gel support deck.

2. Seal both ends of the gel support deck with tape. The tape must be firmly pressed against the edges of the deck to ensure a tight seal.

3. Dispense 15 ml of electrophoresis buffer into a 25 ml glass test tube and add 0.18 grams of agarose. The agarose can be weighed out directly on an appropriate balance. Gently swirl the glass tube until the agarose forms a suspension.

4. Place the test tube into a boiling water bath and heat the solution until the agarose is completely melted. At this time, the agarose solution should be absolutely clear.

5. Pour the melted agarose directly from the test tube onto the casting deck. Insert the comb into the casting tray slots and push down gently on the top of the comb until resistance is encountered. The teeth of the comb will come to rest in the melted agarose about 0.2 mm above the surface of the glass plate.

6. After the gel has cooled for at least 15 minutes, remove the tape strips and carefully lift the comb straight up and away from the casting tray. The gel is now ready for sample application. Gels can also be stored for up to one week before use. For gel storage, the comb is left in place and the tray containing the gel and comb is wrapped in plastic wrap and placed in the refrigerator.

Sample Application

Carefully direct the filling end of the micropipette into the top of the sample well and slowly eject 15 μl of the sample into the well. Duplicate samples should not be placed in adjacent lanes.

Electrophoresis
1. Transfer the four casting trays with gels to the central platform of the electrophoresis cell and position them such that the sample wells are closest to the black (negative) electrode. Upon electrophoresis, most proteins will then migrate from the negative (black) towards the positive (red) electrode.

2. Slowly fill the electrophoresis chamber with electrophoresis buffer until the four gels are covered with a 1/4 cm layer of buffer. Approximately 2.5 liters of buffer are required.

3. Place the electrophoresis cell lid in position.

4. With the power supply off, connect the cables from the cell to the power supply, red to red (positive) and black to black (negative).

5. Push the rocker switches on the power supply to "on" and "170V". The voltage will now remain constant at 170 volts during the run.

6. Unless otherwise indicated, electrophorese until the bromophenol blue in the sample solution has migrated to within 1/4 cm of the positive electrode end of the gel. At 170 V, this takes approximately 50 minutes.

7. At the termination of the electrophoretic separation, shutoff the power supply, disconnect the cables and remove the gel casting trays containing the gels.

**Staining and Destaining**

1. Carefully slide the agarose gel out of the casting deck and place it in a staining dish. Do not put the glass slide in the staining dish.

2. Cover the gel with staining solution, making certain that the agarose does not stick to the tray and all gel surfaces are exposed to the stain.

3. After allowing 1 hour to 1 day for staining, decant and discard the stain, rinse the gel and dish with water, and add about 100 ml of destaining solution.

4. Change the destaining solution after at least 1 day.

5. When the background stain has been reduced sufficiently, place the staining dish over a light source such as a desk lamp or light box and note the position of the stained protein bands.

6. The gels can be stored in a small plastic bag with a few mls of destaining solution. Alternatively, place the stained gel on a dry glass slide and smooth with a gloved index finger to remove air bubbles between the gel and slide. Allow the gel to dry onto the slide for 3-4 days. Cover the dry gel film and glass slide with plastic wrap.