



Consortium for Educational Communication

Module on
Analysis of proteins

By
UZMA POTTOO

Lab.Assistant
Department of Food Science and
Technology
University of Kashmir



INTRODUCTION

Proteins are polymers of amino acids. Twenty different types of amino acids occur naturally in proteins. Proteins differ from each other according to the type, number and sequence of amino acids that make up the polypeptide backbone. As a result, they have different molecular structures, nutritional attributes and physiochemical properties. Proteins are important constituents of foods for a number of different reasons. They are a major source of energy, and contain essential amino-acids, such as lysine, tryptophan, methionine, leucine, isoleucine and valine, which are essential to human health, but which the body cannot synthesize. Proteins also form the major structural components of many natural foods, often determining their overall texture, e.g., tenderness of meat or fish products. Isolated proteins are often used in foods as ingredients because of their unique functional properties, i.e., their ability to provide desirable appearance, texture or stability. Typically, proteins are used as gelling agents, emulsifiers, foaming agents and thickeners. Many food proteins are enzymes which are capable of enhancing the rate of certain biochemical reactions. These reactions can have either a favorable or detrimental effect on the overall properties of foods. Food analysts are interested in knowing the total concentration, type, molecular structure and functional properties of the proteins in foods.

STRUCTURE OF PROTEINS

Primary Structure:

There are 20 different standard L- α -amino acids used by cells for protein construction. Amino acids, as their name indicates, contain both a basic amino group and an acidic carboxyl group. This difunctionality allows the individual amino acids to join together in long chains by forming peptide bonds amide bonds between the $-NH_2$ of one amino acid and the $-COOH$ of another. Sequences with fewer than 50 amino acids are generally referred to as peptides, while the terms protein or polypeptide are used for longer sequences. A protein can be made up of one or more polypeptide molecules. The end of the peptide or protein sequence with a free carboxyl group is called the carboxy-terminus or C-terminus. The terms amino-



terminus or N-terminus describe the end of the sequence with a free α -amino group. The amino acids differ in structure by the substituent on their side chains. These side chains confer different chemical, physical and structural properties to the final peptide or protein. Each amino acid has both a one-letter and three-letter abbreviation. These abbreviations are commonly used to simplify the written sequence of a peptide or protein. Depending on the side-chain substituent, an amino acid can be classified as being acidic, basic or neutral. Although 20 amino acids are required for synthesis of various proteins found in humans, we can synthesize only 10. The remaining 10 are called essential amino acids and must be obtained in the diet. The amino acid sequence of a protein is encoded in DNA. Proteins are synthesized by a series of steps called transcription (the use of a DNA strand to make a complementary messenger RNA strand - mRNA) and translation (the mRNA sequence is used as a template to guide the synthesis of the chain of amino acids which make up the protein). Often, post-translational modifications, such as glycosylation or phosphorylation, occur which are necessary for the biological function of the protein. While the amino acid sequence makes up the primary structure of the protein, the chemical/biological properties of the protein are very much dependent on the three-dimensional or tertiary structure.

Secondary Structure:

Stretches or strands of proteins or peptides have distinct characteristic local structural conformations or secondary structure, dependent on hydrogen bonding. The two main types of secondary structure are the α -helix and the β -sheet. The α -helix is a right-handed coiled strand. The side-chain substituents of the amino acid groups in an α -helix extend to the outside. Hydrogen bonds form between the oxygen of the C=O of each peptide bond in the strand and the hydrogen of the N-H group of the peptide bond four amino acids below it in the helix. The hydrogen bonds make this structure especially stable. The sidechain substituents of the amino acids fit in beside the N-H groups. The hydrogen bonding in a β -sheet is between strands (inter-strand) rather than within strands (intra-strand). The sheet conformation consists of pairs of strands lying side-by-side. The carbonyl oxygens in one strand hydrogen bond with the amino hydrogens of the adjacent strand. The two strands can be either parallel or anti-parallel depending on whether the



strand directions (N-terminus to C-terminus) are the same or opposite. The anti-parallel β -sheet is more stable due to the more well-aligned hydrogen bonds.

Tertiary Structure:

The overall three-dimensional shape of an entire protein molecule is the tertiary structure. The protein molecule will bend and twist in such a way as to achieve maximum stability or lowest energy state. Although the three-dimensional shape of a protein may seem irregular and random, it is fashioned by many stabilizing forces due to bonding interactions between the side-chain groups of the amino acids. Under physiological conditions, the hydrophobic side-chains of neutral, non-polar amino acids such as phenylalanine or isoleucine tend to be buried on the interior of the protein molecule thereby shielding them from the aqueous medium. The alkyl groups of alanine, valine, leucine and isoleucine often form hydrophobic interactions between one-another, while aromatic groups such as those of phenylalanine and tryosine often stack together. Acidic or basic amino posed on the surface of the protein as they are hydrophilic. The formation of disulfde bridges by oxidation of the sulfhydryl groups on cysteine is an important aspect of the stabilization of protein tertiary structure, allowing different parts of the protein chain to be held together covalently. Additionally, hydrogen bonds may form between different side-chain groups. As with disulfde bridges, these hydrogen bonds can bring together two parts of a chain that are some distance away in terms of sequence.

Quaternary Structure:

Many proteins are made up of multiple polypeptide chains, often referred to as protein subunits. These subunits may be the same (as in a homodimer) or different (as in a heterodimer). The quaternary structure refers to how these protein subunits interact with each other and arrange themselves to form a larger aggregate protein complex. The final shape of the protein complex is once again stabilized by various interactions, including hydrogen-bonding, disulfde-bridges and salt bridges. The four levels of protein structure are shown in Figure1. (In downloads)



GENERAL METHODS OF PROTEIN ANALYSIS

1) Chemical Methods:

Determination of Overall Protein Concentration:

1.1 Kjeldahl Method

The Kjeldahl method was developed in 1883 by a brewer called Johann Kjeldahl. A food product is digested with a strong acid so that it releases nitrogen which can be determined by a suitable titration technique. The amount of protein present is then calculated from the nitrogen concentration of the food. The same basic approach is still used today, although a number of improvements have been made to speed up the process and to obtain more accurate measurements. It is usually considered to be the standard method of determining protein concentration. Because the Kjeldahl method does not measure the protein content directly a conversion factor (F) is needed to convert the measured nitrogen concentration to a protein concentration. A conversion factor of 6.25 (equivalent to 0.16 g nitrogen per gram of protein) is used for many applications, however, this is only an average value, and each protein has a different conversion factor depending on its amino-acid composition. The Kjeldahl method can conveniently be divided into three steps: digestion, neutralization and titration. (In downloads)

Principle

•Digestion:

The food sample to be analyzed is weighed into a digestion flask and then digested by heating it in the presence of sulfuric acid (an oxidizing agent which digests the food), anhydrous sodium sulfate (to speed up the reaction by raising the boiling point) and a catalyst, such as copper, selenium, titanium, or mercury (to speed up the reaction). Digestion converts any nitrogen in the food (other than that which is in the form of nitrates or nitrites) into ammonia and other organic matter to CO₂ and H₂O. Ammonia gas is not liberated in an acid solution because the ammonia is in the form of the ammonium ion (NH₄⁺) which binds to the sulfate ion (SO₄²⁻) and thus remains in solution:





●Neutralisation:

After the digestion has been completed the digestion flask is connected to a receiving flask by a tube. The solution in the digestion flask is then made alkaline by addition of sodium hydroxide, which converts the ammonium sulfate into ammonia gas:



The ammonia gas that is formed is liberated from the solution and moves out of the digestion flask and into the receiving flask - which contains an excess of boric acid. The low pH of the solution in the receiving flask converts the ammonia gas into the ammonium ion, and simultaneously converts the boric acid to the borate ion:



●Titration:

The nitrogen content is then estimated by titration of the ammonium borate formed with standard sulfuric or hydrochloric acid, using a suitable indicator to determine the end-point of the reaction.



The concentration of hydrogen ions (in moles) required to reach the end-point is equivalent to the concentration of nitrogen that was in the original food (Equation 3). The following equation can be used to determine the nitrogen concentration of a sample that weighs m grams using a xM HCl acid solution for the titration:



Where v_s and v_b are the titration volumes of the sample and blank, and 14g is the molecular weight of nitrogen N. A blank sample is usually run at the same time as the material being analyzed to take into account any residual nitrogen which may be in the reagents used to carry out the analysis. Once the nitrogen content has been determined it is converted to a protein content using the appropriate conversion factor: %Protein = F X % N.



Advantages and Disadvantages:

●**Advantages.** The Kjeldahl method is widely used internationally and is still the standard method for comparison against all other methods. Its universality, high precision and good reproducibility have made it the major method for the estimation of protein in foods.

●**Disadvantages.** It does not give a measure of the true protein, since all nitrogen in foods is not in the form of protein. Different proteins need different correction factors because they have different amino acid sequences. The use of concentrated sulfuric acid at high temperatures poses a considerable hazard, as does the use of some of the possible catalysts. The technique is time consuming to carry-out.

1.2 The Combustion Method

The combustion method involves burning a sample in an oxygen-rich atmosphere at high temperatures and analyzing the resulting gases. This process has three stages.

●**Combustion:** Once the sample is weighed and purged of any atmospheric gases, it is heated in a high-temperature furnace and rapidly combusted in the presence of pure oxygen at about 1,000°C. Cupric oxide may be used to complete the oxidation.

●**Reduction and adsorption:** The combustion products mainly carbon dioxide, water, nitrogen dioxide, and nitrogen gas are collected and allowed to equilibrate. An aliquot of the gas mixture is passed over hot copper to remove any oxygen and catalytically convert nitrogen dioxide to nitrogen. The sample is then passed through a trap that removes carbon dioxide and water.

●**Quantitation:** The total nitrogen is measured by thermal conductivity.

1.3 Enhanced Dumas method

Recently, an automated instrumental technique has been developed which is capable of rapidly measuring the protein



concentration of food samples. This technique is based on a method first described by a scientist called Dumas over a century and a half ago. It is beginning to compete with the Kjeldahl method as the standard method of analysis for proteins for some foodstuffs due to its rapidness.

Principle:

A sample of known mass is combusted in a high temperature (about 900 °C) chamber in the presence of oxygen. This leads to the release of CO₂, H₂O and N₂. The CO₂ and H₂O are removed by passing the gasses over special columns that absorb them. The nitrogen content is then measured by passing the remaining gasses through a column that has a thermal conductivity detector at the end. The column helps separate the nitrogen from any residual CO₂ and H₂O that may have remained in the gas stream. The instrument is calibrated by analyzing a material that is pure and has a known nitrogen concentration, such as EDTA (= 9.59%N). Thus the signal from the thermal conductivity detector can be converted into a nitrogen content.

Advantages and Disadvantages:

●**Advantages:** It is much faster than the Kjeldahl method (under 4 minutes per measurement, compared to 1-2 hours for Kjeldahl). It doesn't need toxic chemicals or catalysts. Many samples can be measured automatically. It is easy to use.

●**Disadvantages:** High initial cost. It does not give a measure of the true protein, since all nitrogen in foods is not in the form of protein. Different proteins need different correction factors because they have different amino acid sequences. The small sample size makes it difficult to obtain a representative sample.

2) Methods using UV-visible spectroscopy

Spectroscopic Methods

A number of methods have been devised to measure protein concentration, which are based on UV-visible spectroscopy. These methods use either the natural ability of proteins to absorb (or scatter) light in the UV-visible region of the electromagnetic spectrum, or they chemically or physically modify proteins to make them absorb (or scatter) light in this region. The basic principle behind each of these tests is similar. First of all a



calibration curve of absorbance (or turbidity) versus protein concentration is prepared using a series of protein solutions of known concentration. The absorbance (or turbidity) of the solution being analyzed is then measured at the same wavelength, and its protein concentration determined from the calibration curve. The main difference between the tests are the chemical groups which are responsible for the absorption or scattering of radiation, e.g., peptide bonds, aromatic side-groups, basic groups and aggregated proteins. (In downloads)

A number of the most commonly used UV-visible methods for determining the protein content of foods are highlighted below:

2.1 Biuret Method:

A violet-purplish color is produced when cupric ions (Cu^{2+}) interact with peptide bonds under alkaline conditions. The biuret reagent(Dissolve 3gm of copper sulphate and 9 gm of sodium potassium tartarate in 500ml of 0.2mol/litre Sodium hydroxide,add 5 g of potassium iodide and make upto 1 litre with 0.2 sodium hydroxide) It is mixed with a protein solution and then allowed to stand for 15-30 minutes before the absorbance is read at 540 nm. The major advantage of this technique is that there is no interference from materials that adsorb at lower wavelengths, and the technique is less sensitive to protein type because it utilizes absorption involving peptide bonds that are common to all proteins, rather than specific side groups. However, it has a relatively low sensitivity compared to other UV-visible methods.

2.2 Lowry Method:

The Lowry method combines the biuret reagent with another reagent (the Folin-Ciocalteu phenol reagent) which reacts with tyrosine and tryptophan residues in proteins. This gives a bluish color which can be read somewhere between 500 - 750 nm depending on the sensitivity required. There is a small peak around 500 nm that can be used to determine high protein concentrations and a large peak around 750 nm that can be used to determine low protein concentrations. This method is more sensitive to low concentrations of proteins than the biuret method.



2.3.DYE binding methods:

A known excess of a negatively charged (anionic) dye is added to a protein solution whose pH is adjusted so that the proteins are positively charged (i.e. < the isoelectric point). The proteins form an insoluble complex with the dye because of the electrostatic attraction between the molecules, but the unbound dye remains soluble. The anionic dye binds to cationic groups of the basic amino acid residues (histidine, arginine and lysine) and to free amino terminal groups. The amount of unbound dye remaining in solution after the insoluble protein-dye complex has been removed (e.g., by centrifugation) is determined by measuring its absorbance. The amount of protein present in the original solution is proportional to the amount of dye that bound to it:

$\text{dye}_{\text{bound}} = \text{dye}_{\text{initial}} - \text{dye}_{\text{free}}$.

2.4Turbimetric method:

Protein molecules which are normally soluble in solution can be made to precipitate by the addition of certain chemicals, e.g., trichloroacetic acid. Protein precipitation causes the solution to become turbid. Thus the concentration of protein can be determined by measuring the degree of turbidity.

Advantages and Disadvantages:

●Advantages: UV-visible techniques are fairly rapid and simple to carry out, and are sensitive to low concentrations of proteins.

●Disadvantages: For most UV-visible techniques it is necessary to use dilute and transparent solutions, which contain no contaminating substances which absorb or scatter light at the same wavelength as the protein being analyzed. The need for transparent solutions means that most foods must undergo significant amounts of sample preparation before they can be analyzed, e.g., homogenization, solvent extraction, centrifugation, filtration, which can be time consuming and laborious. In addition, it is sometimes difficult to quantitatively extract proteins from certain types of foods, especially after they have been processed so that the proteins become aggregated or covalently bound with other substances. In addition the absorbance depends on the type of protein analyzed (different



proteins have different amino acid sequences).

PROTEIN SEPARATION METHODS

1) Methods Based On Different Solubility Characteristics

Proteins can be separated by exploiting differences in their solubility in aqueous solutions. The solubility of a protein molecule is determined by its amino acid sequence because this determines its size, shape, hydrophobicity and electrical charge. Proteins can be selectively precipitated or solubilized by altering the pH, ionic strength, dielectric constant or temperature of a solution. These separation techniques are the most simple to use when large quantities of sample are involved, because they are relatively quick, inexpensive and are not particularly influenced by other food components. They are often used as the first step in any separation procedure because the majority of the contaminating materials can be easily removed.

Salting out:

Proteins are precipitated from aqueous solutions when the salt concentration exceeds a critical level, which is known as salting-out, because all the water is "bound" to the salts, and is therefore not available to hydrate the proteins. Ammonium sulfate $[(\text{NH}_4)_2\text{SO}_4]$ is commonly used because it has a high water-solubility, although other neutral salts may also be used, e.g., NaCl or KCl. Generally a two-step procedure is used to maximize the separation efficiency. In the first step, the salt is added at a concentration just below that necessary to precipitate out the protein of interest. The solution is then centrifuged to remove any proteins that are less soluble than the protein of interest. The salt concentration is then increased to a point just above that required to cause precipitation of the protein. This precipitates out the protein of interest (which can be separated by centrifugation), but leaves more soluble proteins in solution. The main problem with this method is that large concentrations of salt contaminate the solution, which must be removed before the protein can be resolubilized, e.g., by dialysis or ultrafiltration.

Isoelectric Precipitation:



The isoelectric point (pI) of a protein is the pH where the net charge on the protein is zero. Proteins tend to aggregate and precipitate at their pI because there is no electrostatic repulsion keeping them apart. Proteins have different isoelectric points because of their different amino acid sequences (i.e., relative numbers of anionic and cationic groups), and thus they can be separated by adjusting the pH of a solution. When the pH is adjusted to the pI of a particular protein it precipitates leaving the other proteins in solution.

2)Separation due to Different Adsorption Characteristics:

Adsorption chromatography involves the separation of compounds by selective adsorption-desorption at a solid matrix that is contained within a column through which the mixture passes. Separation is based on the different affinities of different proteins for the solid matrix. Affinity and ion-exchange chromatography are the two major types of adsorption chromatography commonly used for the separation of proteins. Separation can be carried out using either an open column or high-pressure liquid chromatography.

Ion Exchange Chromatography: (In downloads)

Ion exchange chromatography relies on the reversible adsorption-desorption of ions in solution to a charged solid matrix or polymer network. This technique is the most commonly used chromatographic technique for protein separation. A positively charged matrix is called an anion-exchanger because it binds negatively charged ions (anions). A negatively charged matrix is called a cation-exchanger because it binds positively charged ions (cations). The buffer conditions (pH and ionic strength) are adjusted to favor maximum binding of the protein of interest to the ion-exchange column. Contaminating proteins bind less strongly and therefore pass more rapidly through the column. The protein of interest is then eluted using another buffer solution which favors its desorption from the column (e.g., different pH or ionic strength).

Affinity Chromatography: (In downloads)



Affinity chromatography uses a stationary phase contained in a column that consists of a ligand covalently bound to a solid support. The ligand is a molecule that has a highly specific and unique reversible affinity for a particular protein. The sample to be analyzed is passed through the column and the protein of interest binds to the ligand, whereas the contaminating proteins pass directly through. The protein of interest is then eluted using a buffer solution which favors its desorption from the column. This technique is the most efficient means of separating an individual protein from a mixture of proteins, but it is the most expensive, because of the need to have columns with specific ligands bound to them.

Both ion-exchange and affinity chromatography are commonly used to separate proteins and amino-acids in the laboratory. They are used less commonly for commercial separations because they are not suitable for rapidly separating large volumes and are relatively expensive.

3) Separation Due to size Differences:

Proteins can also be separated according to their size. Typically, the molecular weights of proteins vary from about 10,000 to 1,000,000 daltons. In practice, separation depends on the Stokes radius of a protein, rather than directly on its molecular weight. The Stokes radius is the average radius that a protein has in solution, and depends on its three dimensional molecular structure. For proteins with the same molecular weight the Stokes radius increases in the following order: compact globular protein < flexible random-coil < rod-like protein.

●Dialysis:

Dialysis is used to separate molecules in solution by use of selectively semipermeable membranes that permit the passage of molecules smaller than a certain size through, but prevent the passing of larger molecules. A protein solution is placed in dialysis tubing which is sealed and placed into a large volume of water or buffer which is slowly stirred. Low molecular weight solutes flow through the bag, but the large molecular weight protein molecules remain in the bag. Dialysis is a relatively slow method, taking up to 12 hours to be completed. It is therefore



most frequently used in the laboratory. Dialysis is often used to remove salt from protein solutions after they have been separated by salting-out, and to change buffers.

● **Ultrafiltration:**

A solution of protein is placed in a cell containing a semipermeable membrane, and pressure is applied. Smaller molecules pass through the membrane, whereas the larger molecules remain in the solution. The separation principle of this technique is therefore similar to dialysis, but because pressure is applied separation is much quicker. Semipermeable membranes with cutoff points between about 500 to 300,000 are available. That portion of the solution which is retained by the cell (large molecules) is called the retentate, whilst that part which passes through the membrane (small molecules) forms part of the ultrafiltrate. Ultrafiltration can be used to concentrate a protein solution, remove salts, exchange buffers or fractionate proteins on the basis of their size. Ultrafiltration units are used in the laboratory and on a commercial scale.

4) Separation by Electrophoresis

Electrophoresis relies on differences in the migration of charged molecules in a solution when an electrical field is applied across it. It can be used to separate proteins on the basis of their size, shape or charge. (In downloads)

Non-denaturing electrophoresis:

In non-denaturing electrophoresis, a buffered solution of native proteins is poured onto a porous gel (usually polyacrylamide, starch or agarose) and a voltage is applied across the gel. The proteins move through the gel in a direction that depends on the sign of their charge, and at a rate that depends on the magnitude of the charge, and the friction to their movement:

Proteins may be positively or negatively charged in solution depending on their isoelectric points (pI) and the pH of the solution. A protein is negatively charged if the pH is above the pI, and positively charged if the pH is below the pI. The magnitude of the charge and applied voltage will determine how far proteins



migrate in a certain time. The higher the voltage or the greater the charge on the protein the further it will move. The friction of a molecule is a measure of its resistance to movement through the gel and is largely determined by the relationship between the effective size of the molecule, and the size of the pores in the gel. The smaller the size of the molecule, or the larger the size of the pores in the gel, the lower the resistance and therefore the faster a molecule moves through the gel. Gels with different porosities can be purchased from chemical suppliers, or made up in the laboratory. Smaller pores sizes are obtained by using a higher concentration of cross-linking reagent to form the gel. Gels may be contained between two parallel plates, or in cylindrical tubes. In non-denaturing electrophoresis the native proteins are separated based on a combination of their charge, size and shape.

Denaturing Electrophoresis:

In denaturing electrophoresis proteins are separated primarily based on their molecular weight. Proteins are denatured prior to analysis by mixing them with mercaptoethanol, which breaks down disulfide bonds, and sodium dodecyl sulfate (SDS), which is an anionic surfactant that hydrophobically binds to protein molecules and causes them to unfold because of the repulsion between negatively charged surfactant head-groups. Each protein molecule binds approximately the same amount of SDS per unit length. Hence, the charge per unit length and the molecular conformation is approximately similar for all proteins. As proteins travel through a gel network they are primarily separated on the basis of their molecular weight because their movement depends on the size of the protein molecule relative to the size of the pores in the gel smaller proteins moving more rapidly through the matrix than larger molecules. This type of electrophoresis is commonly called sodium dodecyl sulfate -polyacrylamide gel electrophoresis, or SDS-PAGE.

To determine how far proteins have moved a tracking dye is added to the protein solution, e.g., bromophenol blue. This dye is a small charged molecule that migrates ahead of the proteins. After the electrophoresis is completed the proteins are made visible by treating the gel with a protein dye such as Coomassie Brilliant Blue or silver stain. The relative mobility(R_m) of each protein band is calculated:

Electrophoresis is often used to determine the protein



composition of food products. The protein is extracted from the food into solution, which is then separated using electrophoresis. SDS-PAGE is used to determine the molecular weight of a protein by measuring R_m , and then comparing it with a calibration curve produced using proteins of known molecular weight: a plot of \log (molecular weight) against relative mobility is usually linear. Denaturing electrophoresis is more useful for determining molecular weights than non-denaturing electrophoresis, because the friction to movement does not depend on the shape or original charge of the protein molecules.

Isoelectric Focussing Electrophoresis:

This technique is a modification of electrophoresis, in which proteins are separated by charge on a gel matrix which has a pH gradient across it. Proteins migrate to the location where the pH equals their isoelectric point and then stop moving because they are no longer charged. This method has the highest resolution of all techniques used to separate proteins. Gels are available that cover a narrow pH range (2-3 units) or a broad pH range (3-10 units) and one should therefore select a gel which is most suitable for the proteins being separated.

Two Dimensional Electrophoresis:

Isoelectric focussing and SDS-PAGE can be used together to improve resolution of complex protein mixtures. Proteins are separated in one direction on the basis of charge using isoelectric focussing, and then in a perpendicular direction on the basis of size using SDS-PAGE (Indownloads)