Introduction to Proteins

Proteins are polymers of amino acids. Each amino acid contains a central carbon, a hydrogen, a carboxyl group, an amino group, and a variable R group. The R group specifies which class of amino acids it belongs to: electrically charged hydrophilic side chains, polar but uncharged side chains, nonpolar hydrophobic side chains, and special cases.

Proteins have a variety of function in cells. Major functions include acting as enzymes, receptors, transport molecules, regulatory proteins for gene expression, and so on. Enzymes are biological catalysts that speed up a chemical reaction without being permanently altered. They have "active sites" where the substrate/reactant binds, and they can be either activated or inhibited (competitive and/or noncompetitive inhibitors).

Protein Molecule

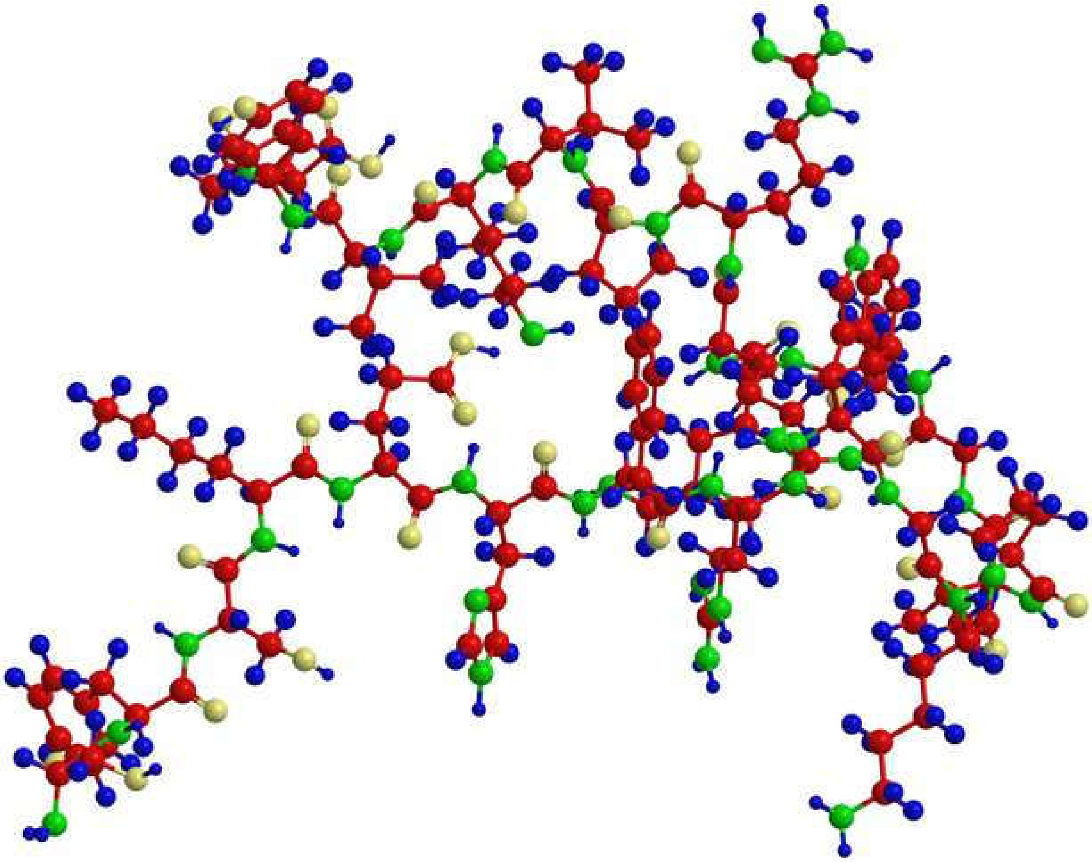
A protein molecule is very large compared with molecules of sugar or salt and consists of many amino acids joined together to form long chains, much as beads are arranged on a string. There are about 20 different amino acids that occur naturally in proteins.

Proteins of similar function have similar amino acid composition and sequence. Although it is not yet possible to explain all of the functions of a protein from its amino acid sequence, established correlations between structure and function can be attributed to the properties of the amino acids that compose proteins.

Plants can synthesize all of the amino acids; animals cannot, even though all of them are essential for life. Plants can grow in a medium containing inorganic nutrients that provide nitrogen, potassium, and other substances essential for growth.

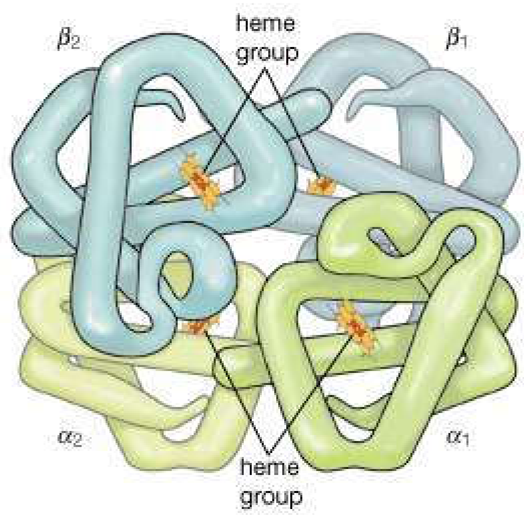
They utilize the carbon dioxide in the air during the process of photosynthesis to form organic compounds such as carbohydrates. Animals, however, must obtain organic nutrients from outside sources.

Because the protein content of most plants is low, very large amounts of plant material are required by animals, such as ruminants (e.g., cows), that eat only plant material to meet their amino acid requirements. Nonruminant animals, including humans, obtain proteins principally from animals and their products—e.g., meat, milk, and eggs. The seeds of legumes are increasingly being used to prepare inexpensive protein-rich food (see human nutrition).



The protein content of animal organs is usually much higher than that of the blood plasma. Muscles, for example, contain about 30 percent protein, the liver 20 to 30 percent, and red blood cells 30 percent. Higher percentages of protein are found in hair, bones, and other organs and tissues with a low water content. The quantity of free amino acids and peptides in animals is much smaller than the amount of protein; protein molecules are produced in cells by the stepwise alignment of amino acids and are released into the body fluids only after synthesis is complete.

The high protein content of some organs does not mean that the importance of proteins is related to their amount in an organism or tissue; on the contrary, some of the most important proteins, such as enzymes and hormones, occur in extremely small amounts. The importance of proteins is related principally to their function. All enzymes identified thus far are proteins. Enzymes, which are the catalysts of all metabolic reactions, enable an organism to build up the chemical substances necessary for life—proteins, nucleic acids, carbohydrates, and lipids—to convert them into other substances, and to degrade them. Life without enzymes is not possible. There are several **protein** hormones with important regulatory functions. In all vertebrates, the respiratory protein hemoglobin acts as oxygen carrier in the blood, transporting oxygen from the lung to body organs and tissues. A large group of structural proteins maintains and protects the structure of the animal body.



Almost everything that occurs in the cell involves one or more proteins. Proteins provide structure, catalyze cellular reactions, and carry out a myriad of other tasks. Their central place in the cell is reflected in the fact that genetic information is ultimately expressed as protein. For each protein there is a segment of DNA (a gene; see Chapters 12 and 23) that encodes information specifying its sequence of amino acids. There are thousands of different kinds of proteins in a typical cell, each encoded by a gene and each performing a specific function. Proteins are among the most abundant biological macromolecules and are also extremely versatile in their functions.

Proteins Have Many Different Biological Functions

Enzymes

The most varied and most highly specialized proteins are those with catalytic activity-the enzymes. Virtually all the chemical reactions of organic biomolecules in cells are catalyzed by enzymes. Many thousands of different enzymes, each capable of catalyzing a different kind of chemical reaction, have been discovered in different organisms (Fig. 6-la).

Transport Proteins

Transport proteins in blood plasma bind and carry specific molecules or ions from one organ to another. Hemoglobin of erythrocytes (Fig. 6-lb) binds oxygen as the blood passes through the lungs, carries it to the peripheral tissues, and there releases it to participate in the energy-yielding oxidation of nutrients. The blood plasma contains lipoproteins, which carry lipids from the liver to other organs. Other kinds of transport proteins are present in the plasma membranes and intracellular membranes of all organisms; these are adapted to bind glucose, amino acids, or other substances and transport them across the membrane.

Nutrient and Storage Proteins

The seeds of many plants store nutrient proteins required for the growth of the germinating seedling. Particularly well-studied examples are the seed proteins of wheat, corn, and rice. Ovalbumin, the major protein of egg white, and casein, the major protein of milk, are other examples of nutrient protein. The ferritin found in some bacteria and in plant and animal tissues stores iron.

Contractile or Motile Proteins

Some proteins endow cells and organisms with the ability to contract, to change shape, or to move about. Actin and myosin function in the contractile system of skeletal muscle and also in many nonmuscle cells. Tubulin is the protein from which microtubules are built. Microtubules act in concert with the protein dynein in flagella and cilia to propel cells.

Structural Proteins

Many proteins serve as supporting filaments, cables, or sheets, to give biological structures strength or protection. The major component of tendons and cartilage is the fibrous protein collagen, which has very high tensile strength. Leather is almost pure collagen. Ligaments contain elastin, a structural protein capable of stretching in two dimensions. Hair, fingernails, and feathers consist largely of the tough, insoluble protein keratin. The major component of silk fibers and spider webs is fibroin (Fig. 6-le). The wing hinges of some insects are made of resilin, which has nearly perfect elastic properties.

Defense Proteins

Many proteins defend organisms against invasion by other species or protect them from injury. The immunoglobulins or antibodies, specialized proteins made by the lymphocytes of vertebrates, can recognize and precipitate or neutralize invading bacteria, viruses, or foreign proteins from another species. Fibrinogen and thrombin are blood-clotting proteins that prevent loss of blood when the vascular system is injured. Snake venoms, bacterial toxins, and toxic plant proteins, such as ricin, also appear to have defensive functions (Fig. 6-lf). Some of these, including fibrinogen, thrombin, and some venoms, are also enzymes.

Regulatory Proteins

Some proteins help regulate cellular or physiological activity. Among them are many hormones. Examples include insulin, which regulates sugar metabolism, and the growth hormone of the pituitary. The cellular response to many hormonal signals is often mediated by a class of GTP-binding proteins called G proteins (GTP is closely related to ATP, with guanine replacing the adenine portion of the molecule; see Figs. 1-12 and 3-16b.) Other regulatory proteins bind to DNA and regulate the biosynthesis of enzymes and RNA molecules involved in cell division in both prokaryotes and eukaryotes (Fig. 6-ig).

Other Proteins

There are numerous other proteins whose functions are rather exotic and not easily classified. Monellin, a protein of an African plant, has an intensely sweet taste. It is being studied as a nonfattening, nontoxic food sweetener for human use. The blood plasma of some Antarctic fish contains antifreeze proteins, which protect their blood from freezing.

Proteins Are Very Large Molecules

How long are the polypeptide chains in proteins? Table 6-1 shows that human cytochrome c has 104 amino acid residues linked in a single chain; bovine chymotrypsinogen has 245 amino acid residues. Probably near the upper limit of size is the protein apolipoprotein B, a cholesterol-transport protein with 4,536 amino acid residues in a single polypeptide chain of molecular weight 513,000. Most naturally occurring polypeptides contain less than 2,000 amino acid residues.

Table 6–1 Molecular data on some proteins

	Molecular weight	Number of residues	Number of polypeptide chains
Insulin (bovine)	5,733	51	2
Cytochrome *c* (human)	13,000	104	1
Ribonuclease A (bovine pancreas)	13,700	124	1
Lysozyme (egg white)	13,930	129	1
Myoglobin (equine heart)	16,890	153	1
Chymotrypsin (bovine pancreas)	21,600	241	3
Chymotrypsinogen (bovine)	22,000	245	1
Hemoglobin (human)	64,500	574	4
Serum albumin (human)	68,500	~550	1
Hexokinase (yeast)	102,000	~800	2
Immunoglobulin G (human)	145,000	~1,320	4
RNA polymerase (*E. coli*)	450,000	~4,100	5
Apolipoprotein B (human)	513,000	4,536	1
Glutamate dehydrogenase (bovine liver)	1,000,000	~8,300	~40

Some proteins consist of a single polypeptide chain, but others, called multisubunit proteins, have two or more (Table 6-1). The individual polypeptide chains in a multisubunit protein may be identical or different. If at least some are identical, the protein is sometimes called an oligomeric protein and the subunits themselves are referred to as protomers. The enzyme ribonuclease has one polypeptide chain. Hemoglobin has four: two identical α chains and two identical β chains, all four held together by noncovalent interactions.

The molecular weights of proteins, which can be determined by various physicochemical methods, may range from little more than 10,000 for small proteins such as cytochrome c (104 residues), to more than 106 for proteins with very long polypeptide chains or those with several subunits. The molecular weights of some typical proteins are given in Table 6-1. No simple generalizations can be made about the molecular weights of proteins in relation to their function.

One can calculate the approximate number of amino acid residues in a simple protein containing no other chemical group by dividing its molecular weight by 110. Although the average molecular weight of the 20 standard amino acids is about 138, the smaller amino acids predominate in most proteins; when weighted for the proportions in which the various amino acids occur in proteins (see Table 5-1), the average molecular weight is nearer to 128. Because a molecule of water (Mr 18) is removed to create each peptide bond, the average molecular weight of an amino acid residue in a protein is about 128 - 18 = 110. Table 6-1 shows the number of amino acid residues in several proteins.

Proteins Have Characteristic Amino Acid Compositions

As is true for simple peptides, hydrolysis of proteins with acid or base yields a mixture of free α-amino acids. When completely hydrolyzed, each type of protein yields a characteristic proportion or mixture of the different amino acids. Table 6-2 shows the composition of the amino acid mixtures obtained on complete hydrolysis of human cytochrome c and of bovine chymotrypsinogen, the inactive precursor of the digestive enzyme chymotrypsin. These two proteins, with very different functions, also differ significantly in the relative numbers of each kind of amino acid they contain. The 20 amino acids almost never occur in equal amounts in proteins. Some amino acids may occur only once per molecule or not at all in a given type of protein; others may occur in large numbers.

Table 6–3 Conjugated proteins

Class	Prosthetic group	Example
Lipoproteins	Lipids	β_1-Lipoprotein of blood
Glycoproteins	Carbohydrates	Immunoglobulin G
Phosphoproteins	Phosphate groups	Casein of milk
Hemoproteins	Heme (iron porphyrin)	Hemoglobin
Flavoproteins	Flavin nucleotides	Succinate dehydrogenase
Metalloproteins	Iron	Ferritin
	Zinc	Alcohol dehydrogenase
	Calcium	Calmodulin
	Molybdenum	Dinitrogenase
	Copper	Plastocyanin

The aggregate biochemical picture of protein structure and function is derived from the study of many individual proteins. To study a protein in any detail it must be separated from all other proteins in a cell, and techniques must be available to determine its properties. The necessary methods come from protein chemistry, a discipline as old as biochemistry itself and one that retains a central position in biochemical research. Modern techniques are providing ever newer experimental insights into the critical relationship between the structure of a protein and its function.

Table 6–2 Amino acid composition of two proteins

Amino acid	Number of residues per molecule of protein: Human cytochrome c	Bovine chymotrypsinogen
Ala	6	22
Arg	2	4
Asn	5	15
Asp	3	8
Cys	2	10
Gln	2	10
Glu	8	5
Gly	13	23
His	3	2
Ile	8	10
Leu	6	19
Lys	18	14
Met	3	2
Phe	3	6
Pro	4	9
Ser	2	28
Thr	7	23
Trp	1	8
Tyr	5	4
Val	3	23
Total	104	245

Proteins Can Be Separated and Purified.

Methods for separating proteins take advantage of properties such as charge, size, and solubility, which vary from one protein to the next. Because many proteins bind to other biomolecules, proteins can also be separated on the basis of their binding properties. The source of a protein is generally tissue or microbial cells. The cells must be broken open and the protein released into a solution called a crude extract. If necessary, differential centrifugation can be used to prepare subcellular fractions or to isolate organelles (see Fig. 2-24). Once the extract or organelle preparation is ready, a variety of methods are available for separation of proteins. Ion-exchange chromatography (see Fig. 5-12) can be used to separate proteins with different charges in much the same way that it separates amino acids. Other chromatographic methods take advantage of differences in size, binding affinity, and solubility (Fig. 6-2). Non chromatographic methods include the selective precipitation of proteins with salt, acid, or high temperatures.

The approach to the purification of a "new" protein, one not previously isolated, is guided both by established precedents and common sense. In most cases, several different methods must be used sequentially to completely purify a protein. The choice of method is somewhat empirical, and many protocols may be tried before the most effective is determined. Trial and error can often be minimized by using purification procedures developed for similar proteins as a guide. Published purification protocols are available for many thousands of proteins. Common sense dictates that inexpensive procedures be used first, when the total volume and number of contaminants is greatest. Chromatographic methods are often impractical at early stages because the amount of chromatographic medium needed increases with sample size. As each purification step is completed, the sample size generally becomes smaller (Table 6-4) and more sophisticated (and expensive) chromatographic procedures can be applied.

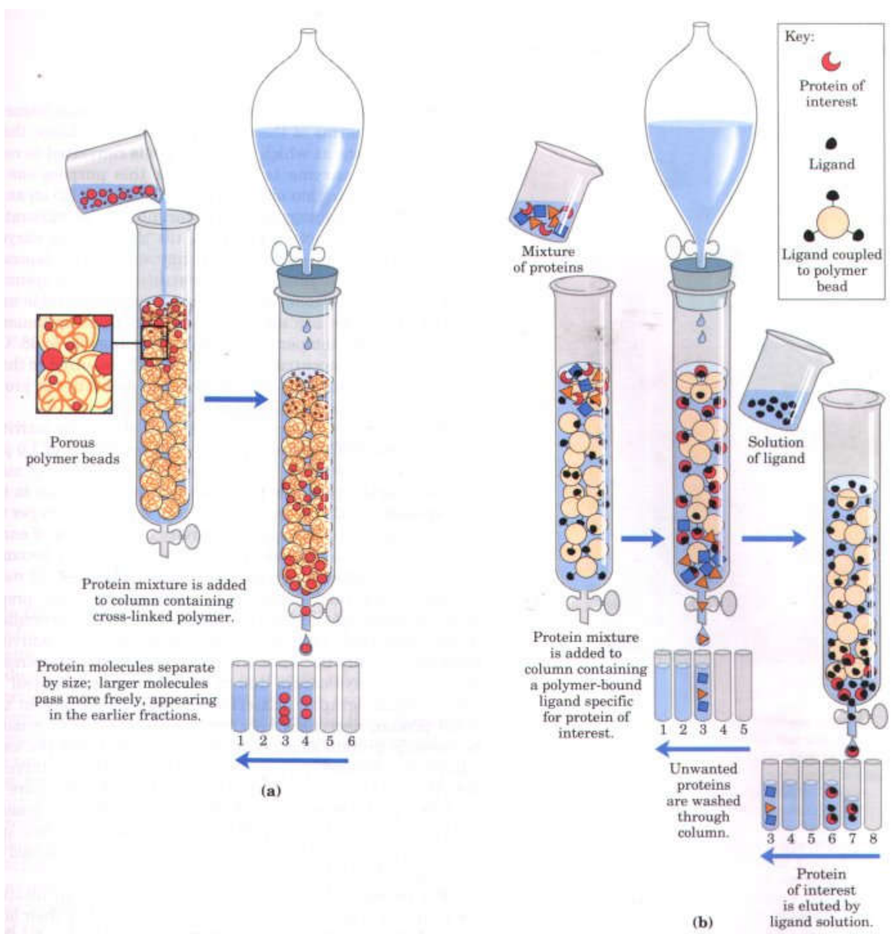


Figure 6-2 two types of chromatographic methods used in protein purification. (a) Size-exclusion chromatography; also called gel filtration. This method separates proteins according to size. The column contains a cross-linked polymer with pores of selected size. Larger proteins migrate faster than smaller ones, because they are too large to enter the pores in the beads and hence take a more direct route through *the column. The smaller proteins enter the pores and are slowed by the more labyrinthian path they take through the column. (b) Affinity chromatography separates proteins by their binding specificities. The proteins retained on the column are those that bind specifically to a ligand cross-linked to the beads. (In biochemistry, the term "ligand" is used to refer to a group or molecule that is bound.) After nonspecific proteins are washed through the column, the bound protein of particular interest is eluted by a solution containing free ligand.*

Table 6–4 A purification table for a hypothetical enzyme*

Procedure or step	Fraction volume (ml)	Total protein (mg)	Activity (units)	Specific activity (units/mg)
1. Crude cellular extract	1,400	10,000	100,000	10
2. Precipitation	280	3,000	96,000	32
3. Ion-exchange chromatography	90	400	80,000	200
4. Size-exclusion chromatography	80	100	60,000	600
5. Affinity chromatography	6	3	45,000	15,000

* All data represent the status of the sample *after* the procedure indicated in the first column has been carried out.