Contents

CONTENTS

C H A P T E R

10

Proteins Classification

THE HORSE, LIKE ALL ANIMALS, IS POWERED BY THE MOLECULAR MOTOR PROTEIN, MYOSIN

A portion of myosin moves dramatically (as shown above) in response to ATP binding, hydrolysis, and release, propelling myosin along an action filament. This molecular movement is translated into movement of the entire animal, excitingly depicted in da Vinci's rearing horse. (Left) Leonardo da vinct's "Study of a rearing horse" for the battle of Anghiari (C. 1504) from the Royal Collection ® Her Royal Majesty Queen Elizabeth II.)

The proteins are all remarkably similar in structure insofar as they contain amino acids. But as little is known so far regarding their structure, a classification based totally on this criterion is not possible. However, a few systems based on one or the other cirterion are given below.

CLASSIFICATION BASED ON THE SOURCE OF PROTEIN MOLECULE

Since long, the proteins have been traditionally divided into two well-defined groups: animal proteins and plant proteins. **Animal proteins** are the proteins derived from animal sources such as eggs, milk, meat and fish. They are usually called *higher-quality proteins* because they contain (and hence supply) adequate amounts of all the essential amino acids. On the other hand, **plant proteins** are called *lower-quality proteins* since they have a low content (limiting amount) of one or more of the essential amino acids. The *four most common limiting amino acids are methionine, lysine, threonine and tryptophan* (Table 10–1). Although plant proteins have limiting amounts of some (but not all) amino acids, it should not be construed that they are poor protein sources.

Table 10–1. Limiting amino acids in some plant proteins

Food	*Amino acid(s)*
Cereal grains and millets	Lysine, Threonine
Rice and soybeans	Methionine
Legumes (peas and beans)	Methionine, Tryptophan
Groundnuts	Methionine, Lysine, Threonine
Sunflower seeds	Lysine
Green leafy vegetables	Methionine

CLASSIFICATION BASED ON THE SHAPE OF PROTEIN MOLECULE

On the basis of the shape of protein molecule, the proteins have been grouped under two categories : globular and fibrous.

1. Globular or Corpuscular Proteins.

These have an axial ratio (length : width) of less than 10 (usually not over 3 or 4) and, henceforth, possess a relatively spherical or ovoid shape. These are usually soluble in water or in aqueous media containing acids, bases, salts or alcohol, and diffuse readily. *As a class, globular proteins are more complex in conformation than fibrous proteins, have a far greater variety of biological functions and are dynamic rather than static in their activities.* Tertiary and quaternary structures are usually associated with this class of proteins. Nearly all enzymes are globular proteins, as are protein hormones, blood transport proteins, antibodies and nutrient storage proteins.

A simple functional classification of globular proteins is not possible because of 2 reasons :

(*a*) Firstly, these proteins perform a variety of different functions.

(*b*) Secondly, many widely-differing globular proteins perform almost similar functions.

However, Conn and Stumpf (1976) have classified globular proteins as follows:

Globular Proteins
- Cytochrome C
- Blood proteins
  - Serum albumin
  - Glycoproteins
  - Antibodies (= Immunoglobulins)
  - Hemoglobin
  - Hormones
- Enzymes
- Nutrient proteins

2. Fibrous or Fibrillar Proteins.

These have axial ratios greater than 10 and, henceforth, resemble long ribbons or fibres in shape. *These are mainly of animal origin and are insoluble in all common solvents* such as water, dilute acids, alkalies and salts and also in organic solvents. Most fibrous proteins serve in a structural or protective role.

The fibrous proteins are extremely strong and possess two important properties which are characteristic of the elastomers. These are:

(*a*) They can *stretch* and later recoil to their original length.

(*b*) They have a tendency to *creep*, *i.e.*, if stretched for a long time, their basic length increases and equals the stretched length but, if the tension on the two ends of the fibril is relaxed, they creep to their shorter and shorter length. A large scar, for example, creeps to a smaller size if there is no tension on the scar. On the contrary, if the scar is in a region of high tension, the scar becomes larger and larger as happens in the skin of a person gradually becoming obese.

It is a heterogeneous group and includes the proteins of connective tissues, bones, blood vessels, skin, hair, nails, horns, hoofs, wool and silk. The important examples are:

I. **Collagens.** These are of mesenchymal origin and form the major proteins of white connective tissues (tendons*, cartilage) and of bone. *More than half the total protein in mammalian body is collagen*; acted upon by boiling in water, dilute acids or alkalies to produce the soluble gelatins; *unique in containing high contents (12%) of hydroxyproline*; poor in sulfur since cysteine and cystine are lacking.

Tendons connect muscle to bone, whereas ***ligaments*** attach one bone to another. Tendons are nonelastic but ligaments can be stretched, because they contain, in addition to collagen fibres, the protein elastin.

II. **Elastins.** Also of mesenchymal origin; form the major constituents of yellow elastic tissues (ligaments, blood vessels); differ from collagens in not being converted to soluble gelatins.

III. **Keratins.** These are of ectodermal origin; form the major constituents of epithelial tissues (skin, hair, feathers, horns, hoofs, nails); *usually contain large amounts of sulfur in the form of cystine*– human hair has about 14% cystine.

IV. **Fibroin.** It is the principal constituent of the fibres of silk; composed mainly of glycine, alanine and serine units.

CLASSIFICATION BASED ON COMPOSITION AND SOLUBILITY

This is nowadays the most accepted system of classification and is based on the proposals made by the committees of *British Physiological Society* (1907) and the *American Physiological Society* (1908). The system divides the proteins into 3 major groups, based on their composition *viz*., simple, conjugated and derived.

A. Simple Proteins or Holoproteins.

These are of globular type except for scleroproteins which are fibrous in nature. This group includes proteins containing only amino acids, as structural components. On decomposition with acids, these liberate the constituent amino acids.

These are further classified mainly on their solubility basis as follows:

1. Protamines and histones. These are basic proteins and occur almost entirely in animals, mainly in sperm cells; possess simplest structure and lowest molecular weight (approximately 5,000); soluble in water; unlike most other proteins, not coagulated by heat; strongly basic in character owing to high content of basic amino acids (lysine, arginine); form salts with mineral acids and nucleic proteins. Protamines are virtually devoid of sulfur and aromatic amino acids. Histones are somewhat weaker bases and are, therefore, insoluble in NH_4OH solution, whereas the protamines are soluble.

e.g., protamines—*clupeine* from herring sperm, *salmine* from *salmon* sperm, *sturine* from sturgeon and *cyprinine* from carp,

histones—*nucleohistones* of nuclei; *globin* of hemoglobin.

2. Albumins. These are widely distributed in nature but more abundant in seeds; soluble in water and dilute solutions of acids, bases and salts; precipitated with a saturated solution of an acid salt like $(NH_4)_2.SO_4$ or a neutral salt like $Na_2.SO_4$; coagulated by heat.

e.g., leucosine in cereals, *legumeline* in legumes, *ovalbumin* from white of egg, *serum albumin* from blood plasma, *myosin* of muscles and *lactalbumin* of milk whey.

3. Globulins. These are of two types— pseudoglobulins and euglobulins*. Euglobulins are more widely distributed in nature than the pseudoglobulins; either soluble

Pseudoglobulins and euglobulins differentiated on the basis of solubility behaviour. A pseudoglobulin is soluble at very low ionic strength, whereas euglobulin remains sparingly soluble until the ionic strength is raised. Euglobulins remain soluble at the isoelectric point (pI), whereas pseudoglobulins do not.

(pseudoglobulins) or insoluble (euglobulins) in water; precipitated with half saturated solution of $(NH_4)_2.SO_4$; coagulated by heat.

e.g., pseudoglobulins— *pseudoglobulin* of milk whey.

euglobulins— *serum globulin* from blood plasma, *ovoglobulin* from eggwhite; *myosinogen* from muscle; globulins of various plant seeds like hemp (*edestin*), soybeans (*glycinine*), peas (*legumine*), peach (*amandine*), oranges (*pomeline*); also potato (*tuberin*).

4. **Glutelins.** These have been isolated only from plant seeds; insoluble in water, dilute salt solutions and alcohol solutions but soluble in dilute acids and alkalies; coagulated by heat.

e.g., glutenin from wheat, *glutelin* from corn, *oryzenin* from rice, etc.

5. **Prolamines.** These have also been isolated only from plant seeds; insoluble in water and dilute salt solutions but soluble in dilute acids and alkalies and also in 60 – 80% alcohol solutions; not coagulated by heat

e.g., *gliadin* from wheat, *zein* from corn, *hordein* from oat, etc.

Some biochemists like Karlson (1968) are of the viewpoint that glutelins and prolamines should not be granted the status of exclusive classes since they are small groups of vegetable proteins occurring in grain kernels.

6. **Scleroproteins or Albuminoids.** These occur almost entirely in animals and are, therefore, commonly known as the '*animal skeleton proteins*'; insoluble in water, dilute solution of acids, bases and salts and also in 60–80% alcohol solutions; not attacked by enzymes.

e.g., collagen of bones, *elastin* in ligaments, *keratin* in hair and horry tissues and *fibroin* of silk.

B. Conjugated or Complex Proteins or Heteroproteins.

These are also of globular type except for the pigment in chicken feathers which is probably of fibrous nature. These are the proteins linked with a separable nonprotein portion called *prosthetic group*. The prosthetic group may be either a metal or a compound. On decomposition with acids, these liberate the constituent amino acids as well as the prosthetic group.

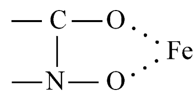
Their further classification is based on the nature of the prosthetic group present. The various divisions are metalloproteins, chromoproteins, glycoproteins, phosphoproteins, lipoproteins and nucleoproteins. (Instead of metalloproteins, chromoproteins etc., the terms metalloproteids, chromoproteids etc., are sometimes used.)

1. **Metalloproteins.** These are the proteins linked with various metals. These may be of stable nature or may be more or less labile. Based on their reactivity with metal ions, the metalloproteins may be classified into 3 groups:

I. *Metals strongly bound by proteins*. Some heavy metals (Hg, Ag, Cu, Zn) become strongly binded with proteins like *collagen, albumin, casein* etc., through the —SH radicals of the side chains. Some other proteins have strong binding affinities for Fe (*siderophilin**) and Cu (*ceruloplasmin***). In these cases, the following pattern of binding may be present:

Siderophilin, also called as **transferrin**, is an important metalloprotein and constitutes about 30% of the total plasma protein. It has a molecular weight of about 90,000 and is capable of binding 2 atoms of iron per mole. It facilitates iron transport.

Ceruloplasmin is an important blue copper-binding protein in the blood of humans and other vertebrates. This protein contains about 90% of copper in serum. It has a molecular weight of about 150,000 and contains 8 atoms of copper per mole. Ceruloplasmin is only one of the many sialoglycoproteins whose removal from the bloodstream is triggered by the loss of sialic acid units. It probably functions by reversibly releasing and binding copper at various sites in the body, whereby regulating copper absorption. In its deficiency, the Wilson's disease develops in man which is characterized by hepatolenticular degeneration.



II. *Metals bound weakly by proteins.* Ca belongs to this category. Here the binding takes place with the help of radicals possessing the electron charge.

III. *Metals which do not couple with proteins.* Na and K belong to this group. These form compounds with nucleic acids where apparently electrostatic bonds are present.

2. **Chromoproteins.** These are proteins coupled with a coloured pigment. Such pigments have also been found among the enzymes like catalase, peroxidase and flavoenzymes. Similarly, chlorophyll is present in leaf cells in the form of a protein, the chloroplastin. The chloroplastin dissolves in water as a colloid and is readily denatured.

e.g., myoglobin, hemoglobin, hemocyanin, hemoerythrin, cytochromes, flavoproteins, catalase, etc.

3. **Glycoproteins and Mucoproteins.** These are the proteins containing carbohydrate as prosthetic group. Glycoproteins contain small amounts of carbohydrates (less than 4%), whereas mucoproteins contain comparatively higher amounts (more than 4%).

e.g., glycoproteins— *egg albumin*, elastase certain *serum globulins* and also certain *serum albumins*.

mucoproteins— *ovomucoid* from egg-white, *mucin* from saliva and *Dioscorea* tubers, *osseomucoid* from bone and *tendomucoid* from tendon.

Elastase, a secreted glycoprotein, showing linked carbohydrates on its surface. Elastase is a protease found in serum. Note that the oligosaccharide chains have substantial size even for this protein, which has a relatively low level of glycosylation.

4. **Phosphoproteins.** These are proteins linked with phosphoric acid; mainly acidic.

e.g., casein from milk and *ovovitellin* from egg yolk.

5. **Lipoproteins.** Proteins forming complexes with lipids (cephalin, lecithin, cholesterol) are called lipoproteins; soluble in water but insoluble in organic solvents.

e.g., lipovitellin and *lipovitellenin* from egg yolk; lipoproteins of blood.

The lipoproteins are in reality the temporary intermediates in the process of transfer of lipids from the site of absorption to the site of utilization. The classification of lipoproteins is frequently based on an operational definition, *i.e.*, the migration of the fraction in a density gradient separation. On this basis, the lipoproteins have been classified into following 4 categories (Table 10–2):

(*a*) Very high density lipoproteins (VHDLs). These have densities greater than 1.21.

(*b*) High density lipoproteins (HDLs). These possess density range of 1.063 to 1.21.

(*c*) Low density lipoproteins (LDLs). Their densities range between 1.05 and 1.063.

(*d*) Very low density lipoproteins (VLDLs). Their density range is from 0.93 to 1.05.

Table 10–2 also lists the compositions of the fractions floated at the respective densities, along with proposed functions.

Contents

Table 10–2. Classification of lipoproteins based on density

Name	*Density range*	*Composition by weight (in blood plasma)*	*Half-life*	*Function*
Very high density lipoprotein	> 1.21	62% protein 28% lipid	> 5 days	Phospholipid transport (contains 8–15% of serum phospholipids)
High density lipoprotein	1.063 – 1.21	50% protein 50% lipid	3 – 5 days	Cholesterol transport Lipid transport
Low density lipoprotein	1.05 – 1.063	22% protein 12% triglyceride 20% phospholipids 46% cholesterol	3 – 5 days	Lipid transport (congenital absence accompanied by intestinal malabsorption and distorted shape of erythrocyte)
Very low density lipoprotein	0.93 – 1.05	Largely protein, triglyceride and phospholipid; composition variable	3 – 4 hours	Triglyceride transport (related to chylomicron units)

(Adapted from Mallette MF, Clagett CO, Phillips AT and McCarl RL, 1979)

6. **Nucleoproteins.** These are compounds containing nucleic acid and protein, esp., protamines and histones. These are usually the salt-like compounds of proteins since the two components have opposite charges and are bound to each other by electrostatic forces. They are present in nuclear substances as well as in the cytoplasm. These may be considered as the sites for the synthesis of proteins and enzymes.

e.g., nucleoproteins from yeast and thymus and also viruses which may be regarded as large molecules of nucleoproteins; *nucleohistones* from nuclei-rich material like glandular tissues; *nuclein*.

C. Derived Proteins.

These are derivatives of proteins resulting from the action of heat, enzymes or chemical reagents. This group also includes the artificially-produced polypeptides.

I. *Primary derived proteins.* These are derivatives of proteins in which the size of protein molecule is not altered materially.

1. **Proteans.** Insoluble in water; appear as first product produced by the action of acids, enzymes or water on proteins.

e.g., edestan derived from edestin and *myosan* derived from myosin.

2. **Metaproteins or Infraproteins.** Insoluble in water but soluble in dilute acids or alkalies; produced by further action of acid or alkali on proteins at about 30–60°C.

e.g., acid and *alkali metaproteins.*

3. **Coagulated Proteins.** Insoluble in water; produced by the action of heat or alcohol on proteins.

e.g., *coagulated eggwhite.*

II. *Secondary derived proteins.* These are derivatives of proteins in which the hydrolysis has certainly occurred. The molecules are, as a rule, smaller than the original proteins.

Contents

1. **Proteoses.** Soluble in water; coagulable by heat; produced when hydrolysis proceeds beyond the level of metaproteins; *primary proteoses* are salted out by half saturation with $(NH_4)_2.SO_4$ and precipitated by HNO_3 and picric acid; *secondary proteoses* are salted out only by complete saturation with $(NH_4)_2.SO_4$ but are not precipitated by HNO_3 or picric acid.

e.g., albumose from albumin; *globulose* from globulin.

2. **Peptones.** Soluble in water; noncoagulable by heat; produced by the action of dilute acids or enzymes when hydrolysis proceeds beyond proteoses; neither salted out by $(NH_4)_2SO_4$ nor precipitated by HNO_3 or picric acid.

3. **Polypeptides.** These are combinations of two or more amino acid units. In fact, *the proteins are essentially long chain polypeptides*.

Drawbacks. Although widely accepted, the system outlined above has certain discrepancies:

1. The classification is arbitrary.
2. The criterion of solubility is not well demarcated as some globulins (pseudoglobulins) are also soluble in water.
3. Protamines and histones should have been kept under derived proteins.
4. The group metaproteins is an artificial assemblage.

The two (second and third) systems of classification, described above, may be merged into one, as shown on page 197.

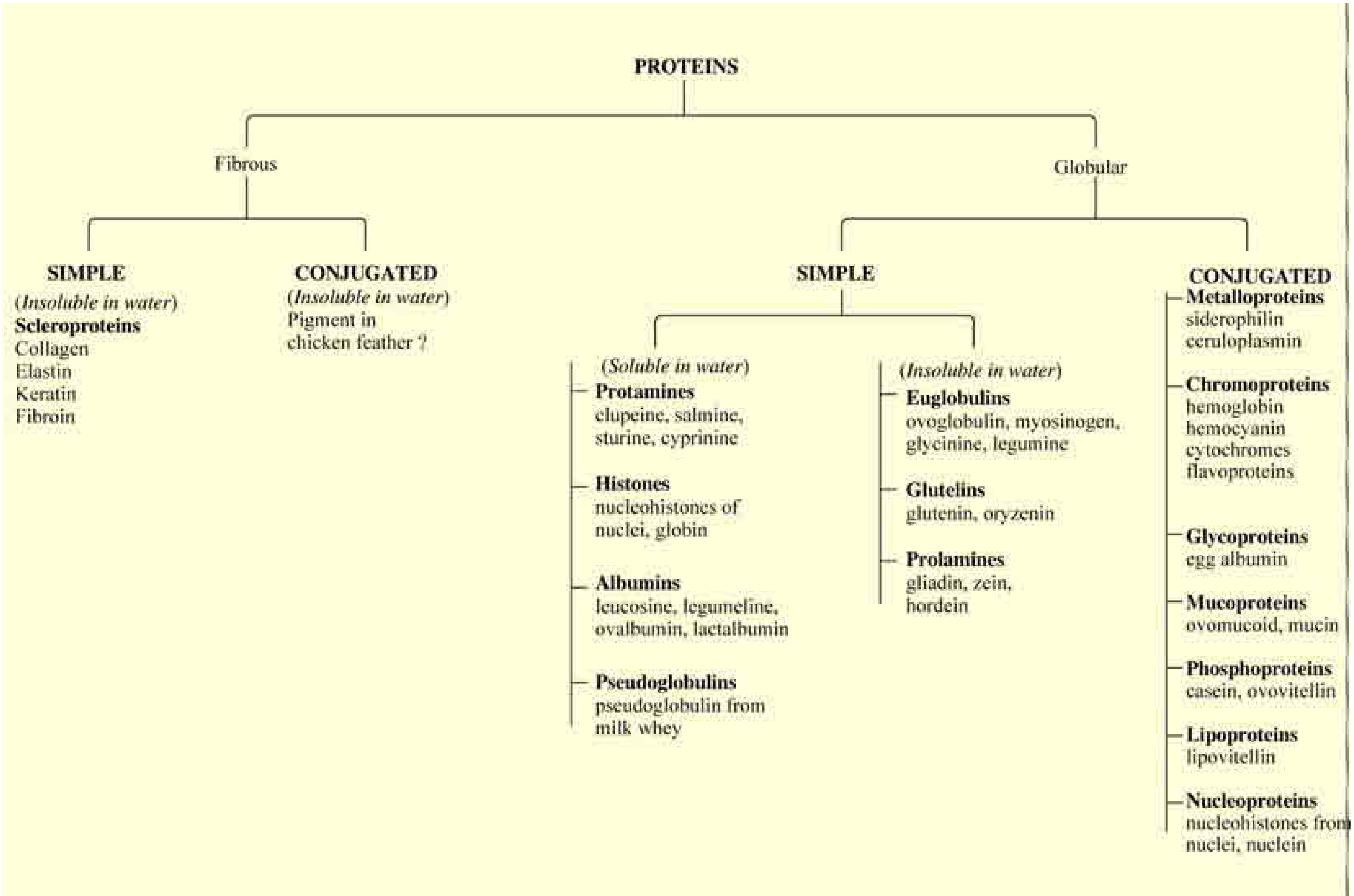
CLASSIFICATION BASED ON BIOLOGICAL FUNCTION

Depending upon their physical and chemical structure and location inside the cell, different proteins perform various functions. As such diverse proteins may be grouped under following categories, based on the metabolic functions they perform (Table 10–3):

Table 10–3. Classification of proteins on the basis of their biological functions

Class of protein	*Function*	*Examples*
Enzymic proteins	Biological catalysts	Urease, Amylase, Catalase, Cytochrome C, Alcohol dehydrogenase.
Structural proteins	Strengthening or protecting biological structures	Collagen, Elastin, Keratin, Fibroin
Transport or carrier proteins	Transport of ions or molecules in the body	Myoglobin, Hemoglobin, Ceruloplasmin, Lipoproteins
Nutrient and storage proteins	Provide nutrition to growing embryos and store ions	Ovalbumin, Casein, Ferritin
Contractile or motile proteins	Function in the contractile system	Actin, Myosin, Tubulin
Defense proteins	Defend against other organisms	Antibodies, Fibrinogen, Thrombin
Regulatory proteins	Regulate cellular or metabolic activities	Insulin, G proteins, Growth hormone
Toxic proteins	Hydrolyze (or degrade) enzymes	Snake venom, Ricin.

1. **Enzymic proteins.** The most varied and most highly specialized proteins are those with catalytic activity– the enzymes. Virtually, all the chemical reactions of organic biomolecules are catalyzed by the enzymes. *Nearly all enzymes are globular proteins.* Chemically, some enzymes are simple proteins,


PROTEINS
Fibrous
Globular
SIMPLE
(Insoluble in water)
Scleroproteins
Collagen
Elastin
Keratin
Fibroin
CONJUGATED
(Insoluble in water)
Pigment in
chicken feather ?
SIMPLE
(Soluble in water)
Protamines
clupeine, salmine,
sturine, cyprinine
Histones
nucleohistones of
nuclei, globin
Albumins
leucosine, legumeline,
ovalbumin, lactalbumin
Pseudoglobulins
pseudoglobulin from
milk whey
(Insoluble in water)
Euglobulins
ovoglobulin, myosinogen,
glycinine, legumine
Glutelins
glutenin, oryzenin
Prolamines
gliadin, zein,
hordein
CONJUGATED
Metalloproteins
siderophilin
ceruloplasmin
Chromoproteins
hemoglobin
hemocyanin
cytochromes
flavoproteins
Glycoproteins
egg albumin
Mucoproteins
ovomucoid, mucin
Phosphoproteins
casein, ovovitellin
Lipoproteins
lipovitellin
Nucleoproteins
nucleohistones from
nuclei, nuclein

containing only amino acid residues; others are complex proteins, containing a major protein part (apoenzyme) and a small nonprotein part (prosthetic group) associated with the protein unit. Enzymes catalyze a variety of reactions. *Urease, amylase, catalase, cytochrome C, alcohol dehydrogenase* are some of the examples of enzymic proteins.

2. **Structural proteins.** The structural proteins are usually inert to biochemical reactions. They maintain the native form and position of the organs. The cell wall and primary fibrous constituents of the cell have structural proteins. *Collagen*, which has very high tensile strength, is the most abundant protein of animals. It is found in connective tissue such as tendons, cartilage, matrix of bones and cornea of the eye. *Leather is almost pure collagen*. Ligaments contain *elastin*, a structural protein capable of stretching in two dimensions. α *keratin* constitutes almost the entire dry weight of hair, wool, feathers, nails, claws, quills, scales, horns, hooves, tortoise shell and much of the outer layer of skin. The major component of silk fibres and spider webs is *fibroin*. The wing hinges of some insects are made of *resilin*, which has nearly perfect elastic properties.

3. **Transport or carrier proteins.** Certain proteins, specially in the animals, are involved in the transport of many essential biological factors to various parts of the organisms. *Hemoglobin* of erythrocytes binds O_2 as the blood passes through the lungs, carries it to the peripheral tissues, and there releases it to participate in the 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. *Ceruloplasmin* transports copper in blood.

4. **Nutrient and storage proteins.** *Ovalbumin* is the major protein of eggwhite. The milk protein, *casein* stores amino acids. The seeds of many plants store nutrient proteins, required for the growth of the germinating seedlings. *Ferritin*, found in some bacteria and in plant and animal tissues, stores iron.

5. **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.

6. **Defense proteins.** Many proteins defend organism against invasion by other species or protect them from injury. The *antibodies* (or immunoglobulins), the specialized proteins made by the lymphocytes of vertebrates, can precipitate or neutralize invading bacteria, viruses or foreign proteins from another species. *Fibrinogen* and *thrombin*, although enzymic, are blood-clotting proteins that prevent loss of blood when the vascular system is injured.

7. **Regulatory proteins.** Some proteins help regulate cellular or physiological activity. Among them are many hormones, such as *insulin* which regulates sugar metabolism, and *growth hormone* which is required for bone growth. The cellular response to many hormonal signal is often mediated by a class of GTP-binding proteins called *G proteins*. Other regulatory proteins bind to DNA and regulate the biosynthesis of enzymes and RNA molecules involved in cell division.

8. **Toxic proteins.** Some proteins act as toxic substances, such as *snake venom*, *bacterial toxins* and toxic plant proteins like *ricin*. These toxic proteins also have defensive functions.

EGG PROTEINS

The eggs and milk are consumed as food by an appreciable portion of the global population. Hence, the egg and milk proteins deserve a special mention. Eggs contain 2 types of proteins : the egg white protein and egg yolk protein. The various components present in the egg white protein are listed in Table 10–4.

Contents

Table 10–4. Protein composition of egg white

Protein	*Percentage*
Total protein	10-11% (on wet basis) 82.8% (on dry basis)
Ovalbumin	70% of total proteins
Conalbumin	9%
Ovomucoid	13%
Globulins	
Lysozyme (G_1)	2.6%
(G_2) (G_3)	7%
Mucin	2%
Avidin	0.06%

(*Adapted from Fevold HL, 1951*)

When egg yolk is diluted with water, proteins precipitate. When the yolk is heated, the proteins undergo heat denaturation and precipitate. Egg yolk contains 2 phosphoproteins, lipovitellin and lipovitellenin, which differ from each other in the following characteristics :

	Lipovitellin	***Lipovitellenin***
Lipid concentration :	17-18%	36-41% (mainly lecithins)
Phosphorus concentration :	1%	0.29%
Protein present :	Vitellin	Vitellenin

Egg yolk also contains water-soluble protein that does not precipitate on dilution of the yolk. This fraction is called ***livetin***. The enzymatic activity of yolk lies in this fraction.

Whole egg is an excellent food because it is a very rich source not only of protein and lipid but also of most of the vitamins (except vitamin C) and most of the required minerals (except calcium).

MILK PROTEINS

Milk, a source of nourishment for all mammals, is composed, in part, of a variety of proteins.

Milk (Fig. 10-2) contains about 0.6–0.7% protein which is not precipitated on acidification to pH 4.7. This represents about 20% of the protein contained in skim milk. These ***whey proteins*** are separated into 2 fractions : ***lactalbumin*** and ***lactoglobulin***.

The name ***casein*** is assigned to the fraction precipitated by acidifying milk to a pH of 4.7. It is present in cow's milk (3-3.5%) and human milk (0.3-0.6%). Casein may be further purified by redissolving and precipitating again. It is of 3 types : α, β and γ. These differ one another in their molecular weights, their rate of migration in an electric field and their phosphorus content.

REFERENCES

See list following Chapter 11.