DEPARTMENT OF IMB B.D. COLLEGE PATAN – 1

B.Sc PART I

PROTEIN by Deepak Kumar Ratan

Proteins Their Biological Functions and Structure

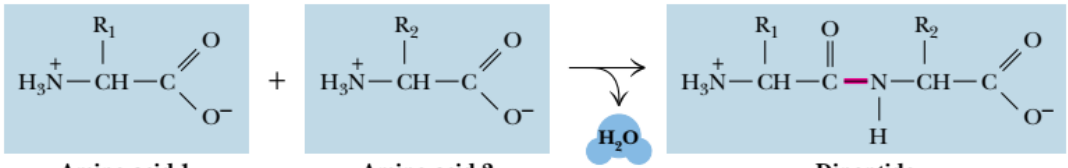
Proteins are a diverse and abundant class of biomolecules, constituting more than 50% of the dry weight of cells. This diversity and abundance reflect the central role of proteins in virtually all aspects of cell structure and function. An extraordinary diversity of cellular activity is possible only because of the versatility inherent in proteins, each of which is specifically tailored to its biological role. The pattern by which each is tailored resides within the genetic information of cells, encoded in a specific sequence of nucleotide bases in DNA. Each such segment of encoded information defines a gene, and expression of the gene leads to synthesis of the specific protein encoded by it, endowing the cell with the functions unique to that particular protein. Proteins are the agents of biological function; they are also the expressions of genetic information.

STRUCTURE OF PROTEIN

1. PRIMARY STRUCTURE
2. SECONDARY STRUCTURE
3. TERTIARY STRUCTURE
4. QUATENARY STRUCTURE

1.PRIMARY STRUCTURE:

Primary proteins are unbranched polymers of amino acids linked head to tail, from carboxyl group to amino group, through formation of covalent **peptide bonds,** a type of amide linkage.

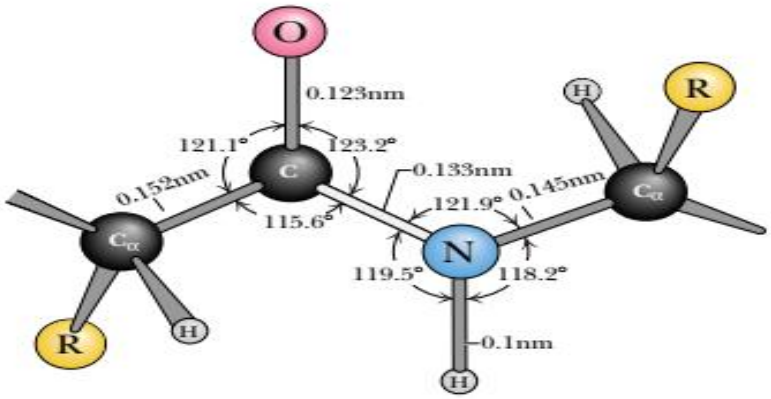


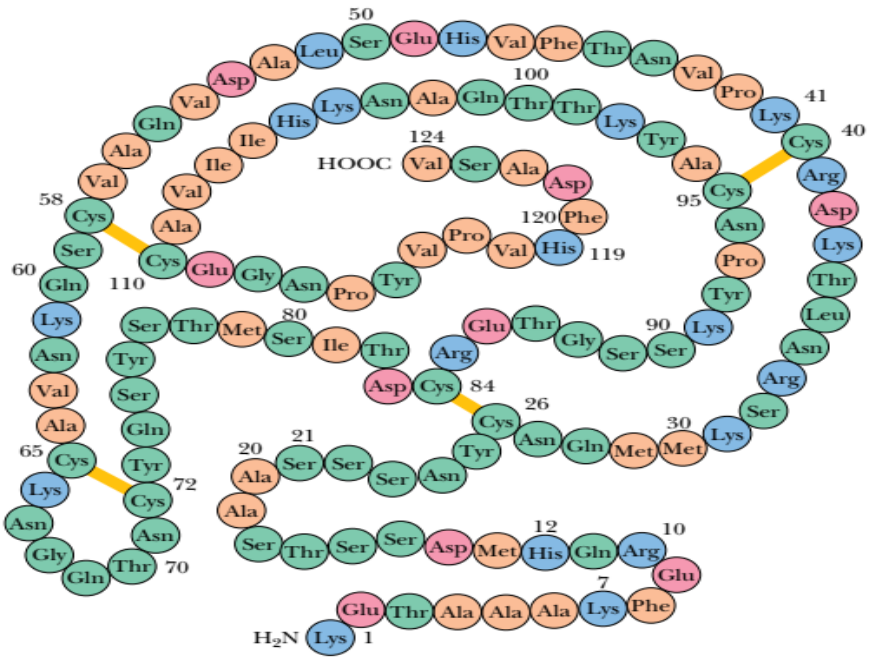
Peptide formation is the creation of an amide bond between the carboxyl group of one amino acid and the amino group of another amino acid. R1 and R2 represent the R groups of two different amino acids.

Peptide bond formation results in the release of H2O. The peptide “backbone” of a protein consists of the repeated sequence – N- Cα- C-, where the N represents the amide nitrogen, the C_α is the α-carbon atom of an amino acid in the polymer chain, and the final C is the carbonyl carbon of the amino acid, which in turn is linked to the amide N of the next amino acid. The peptide bond is shown in its usual trans conformation of carbonyl O and amide H. The C_α atoms are the -carbons of two adjacent amino acids joined in peptide linkage. The

dimensions and angles are the average values observed by crystallographic analysis of amino acids and small peptides. The peptide bond is the light grey bond between C and N. Peptide bond resonance has several important consequences. First, it restricts free rotation around the peptide bond and leaves the peptide backbone with only two degrees of freedom per amino acid group: rotation around the N-C_α bond and rotation around the

Cα-C bond. Second, the six atoms composing the peptide bond group tend to be coplanar, forming the so-called **amide plane** of the polypeptide backbone (Figure 5.4). Third, the C-N bond length is 0.133 nm, which is shorter than normal C-N bond lengths (for example, the C=N bond of 0.145 nm) but longer than typical CPN bonds (0.125 nm). The peptide bond is estimated to have 40% double-bond character.

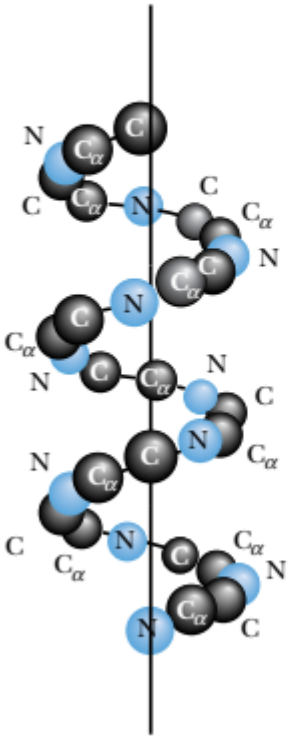




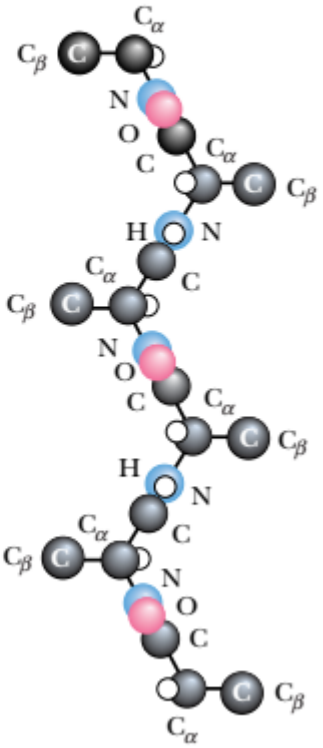
Example of primary structure of protein: Bovine pancreatic ribonuclease A contains 124 amino acid residues. And Insulin Consists of 51 Amino acid.

2. Secondary Structure: Structures resulting from these interactions constitute **secondary structure** for proteins. When a number of hydrogen bonds form between portions of the peptide chain in this manner, two basic types of structures can result α-helices and β-pleated sheets. Linus Pauling, Robert Corey, and their colleagues at the California Institute of Technology summarized a large volume of crystallographic data in a set of dimensions for polypeptide chains and describing the secondary structure of protein.

α-Helix
Only the N — C_α — C backbone is represented. The vertical line is the helix axis.



β-Strand
The N — C_α — C_O backbone as well as the C_β of R groups are represented here. Note that the amide planes are perpendicular to the page.



Forces Influencing Protein Structure

Several different kinds of noncovalent interactions are of vital importance in protein structure. Hydrogen bonds, hydrophobic interactions, electrostatic bonds, and van der Waals forces are all noncovalent in nature, yet are extremely important influences on protein conformations. The stabilization free energies afforded by each of these interactions may be highly dependent on the local environment within the protein, but certain generalizations can still be made.

Hydrogen Bonds

Hydrogen bonds are generally made wherever possible within a given protein structure. In most protein structures that have been examined to date, component atoms of the peptide backbone tend to form hydrogen bonds with one another. Furthermore, side chains capable of forming H bonds are usually located on the protein surface and form such bonds primarily with the water solvent.

Hydrophobic Interactions

Hydrophobic "bonds," or, more accurately, *interactions,* form because nonpolar side chains of amino acids and other nonpolar solutes prefer to cluster in a nonpolar environment rather than to intercalate in a polar solvent such as water. The forming of hydrophobic bonds minimizes the interaction of nonpolar residues with water and is therefore highly favorable. Such clustering is entropically driven. The side chains of the amino acids in the interior or core of the protein structure are almost exclusively hydrophobic. Polar amino acids are almost never found in the interior of a protein, but the protein surface may consist of both polar and nonpolar residues.

Electrostatic Interactions

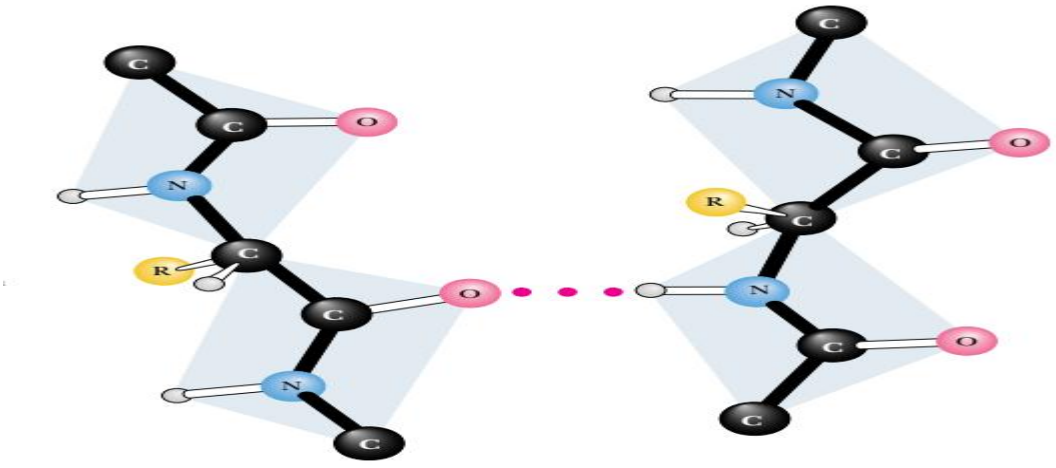
Ionic interactions arise either as electrostatic attractions between opposite charges or repulsions between like charges. Amino acid side chains can carry positive charges, as in the case of lysine, arginine, and histidine, or negative charges, as in aspartate and glutamate. In addition, the NH2-terminal and COOH-terminal residues of a protein or peptide chain usually exist in ionized states and carry positive or negative charges, respectively.

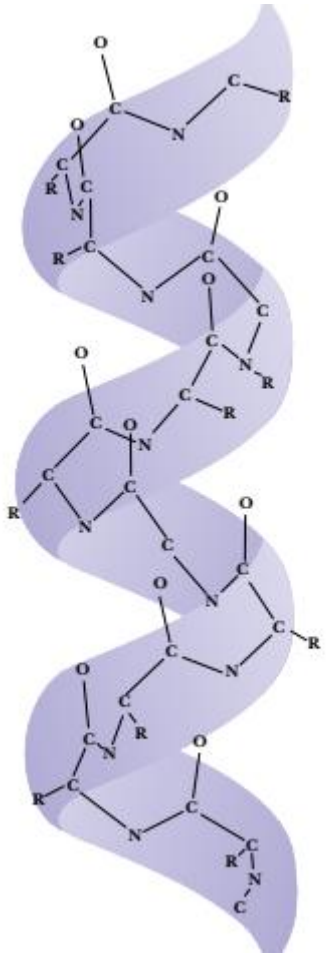
Van der Waals Interactions

Both attractive forces and repulsive forces are included in van der Waals interactions. The attractive forces are due primarily to instantaneous dipole-induced dipole interactions that arise because of fluctuations in the electron charge distributions of adjacent non bonded atoms. Individual van der Waals interactions are weak ones (with stabilization energies of 4.0 to 1.2 kJ/mol), but many such interactions occur in a typical protein, and, by sheer force of numbers, they can represent a significant contribution to the stability of a protein.

The Alpha-Helix STRUCTURE:

The carbonyl oxygen and amide hydrogen of the peptide bond could participate in H bonds either with water molecules in the solvent or with other H-bonding groups in the peptide chain. In nearly all proteins, the carbonyl oxygens and the amide protons of many peptide bonds participate in H bonds that link one peptide group.



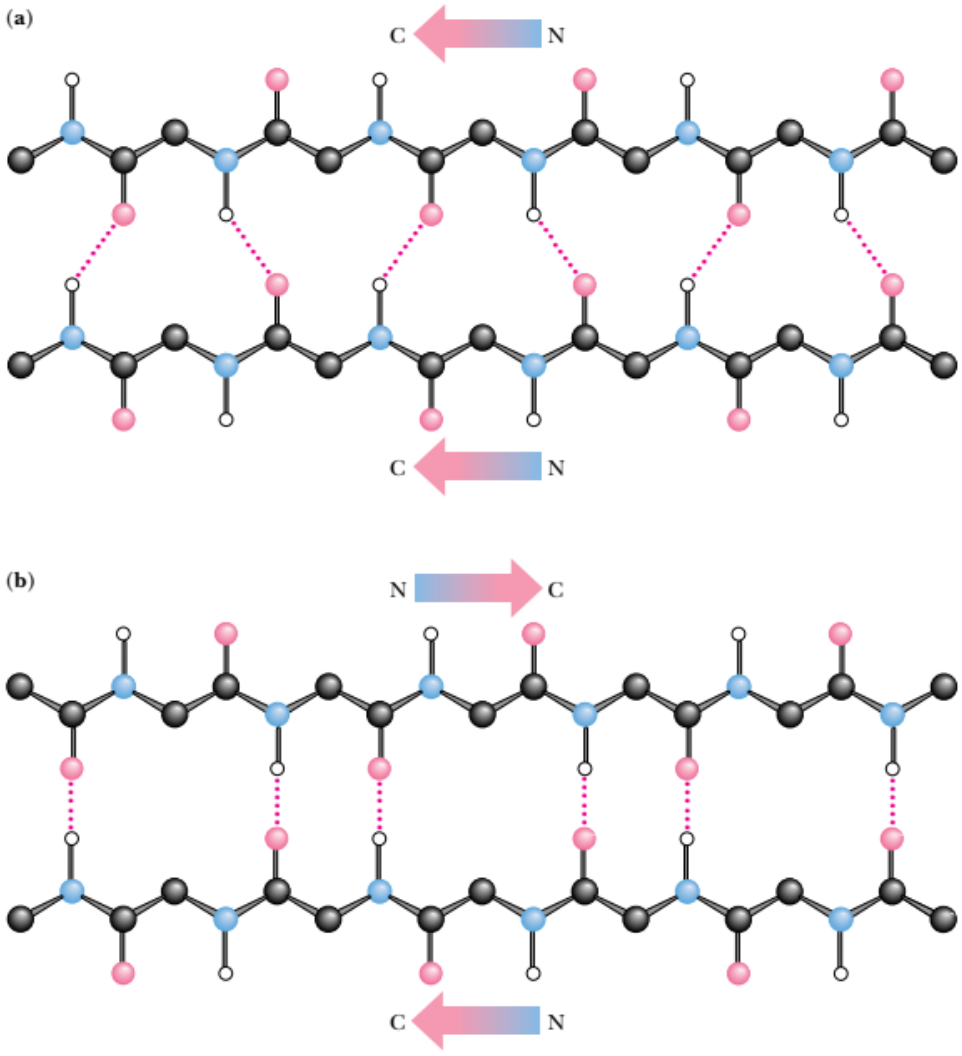


- One turn of the helix represents 3.6 amino acid residues.
- Each amino acid residue extends **1.5 Å (0.15 nm)** along the helix axis.
- With **3.6 residues per turn,** this amounts to 3.6 1.5 Å or **5.4 Å (0.54 nm)** of travel along the helix axis per turn.
- The helix is about 6 Å in diameter.
- The side chains, extending outward from the core structure of the helix, are removed from steric interference with the polypeptide backbone.
- As can be seen from above Figure each peptide carbonyl is hydrogen bonded to the peptide N- H group four residues farther up the chain. Note that all of the H bonds lie parallel to the helix axis and that all of the carbonyl groups are pointing in one direction along the helix axis while the N-H groups are pointing in the opposite direction.
- Recall that the entire path of the peptide backbone can be known if the ϕ and Ψ twist angles are specified for each residue. The α –helix is formed if the values of Ψ are approximately 60° and the values of are in the range of 45 to 50°.

The Beta-Pleated Sheet

- Another type of structure commonly observed in proteins also forms because of local, cooperative formation of hydrogen bonds.
- That is the Beta pleated sheet, or β-structure, often called the β **-pleated sheet.** This structure was also first postulated by Pauling and Corey in 1951 and has now been observed in many natural proteins.

- A -pleated sheet can be visualized by laying thin, pleated strips of paper side by side to make a “pleated sheet” of paper.

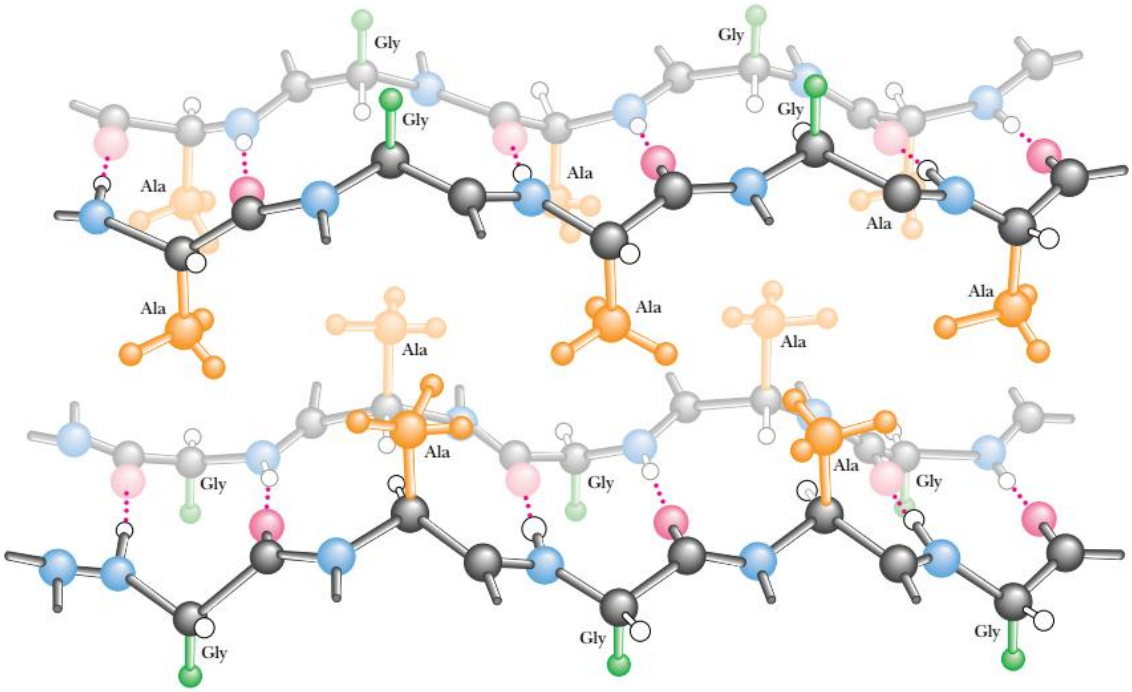


The arrangement of hydrogen bonds in (a) parallel and (b) antiparallel β-pleated sheets.

Antiparallel pleated sheets are the fundamental structure found in silk, with the polypeptide chains forming the sheets running parallel to the silk fibers. The silk fibers thus formed have properties consistent with those of the β-sheets that form them. They are quite flexible but cannot be stretched or extended to any appreciable degree. Antiparallel structures are also observed in many other proteins, including immunoglobulin G, superoxide dismutase from bovine erythrocytes, and concanavalin A. Many proteins, including carbonic anhydrase, egg lysozyme, and glyceraldehyde phosphate dehydrogenase, possess both -helices and -pleated sheet structures within a single polypeptide chain.

3. Tertiary Structure:

The folding of a single polypeptide chain in three-dimensional space is referred to as its **tertiary structure.** All of the information needed to fold the protein into its native tertiary structure is contained within the primary structure of the peptide chain itself. It soon became apparent that the proteins knew how they were supposed to fold into tertiary structure. Fibrous proteins, Globular proteins, and membrane proteins. Fibrous proteins contain polypeptide chains organized approximately parallel along a single axis, producing long fibers or large sheets. Such proteins tend to be mechanically strong and resistant to solubilization in water and dilute salt solutions. Fibrous proteins often play a structural role in nature.



The **fibroin** proteins found in silk fibers represent another type of fibrous protein. These are composed of stacked antiparallel -sheets, as shown in above figure the polypeptide sequence of silk proteins, there are large stretches in which every other residue is a glycine. The β-keratins found in bird feathers are also made up of stacked -sheets.

Collagen: A Triple Helix

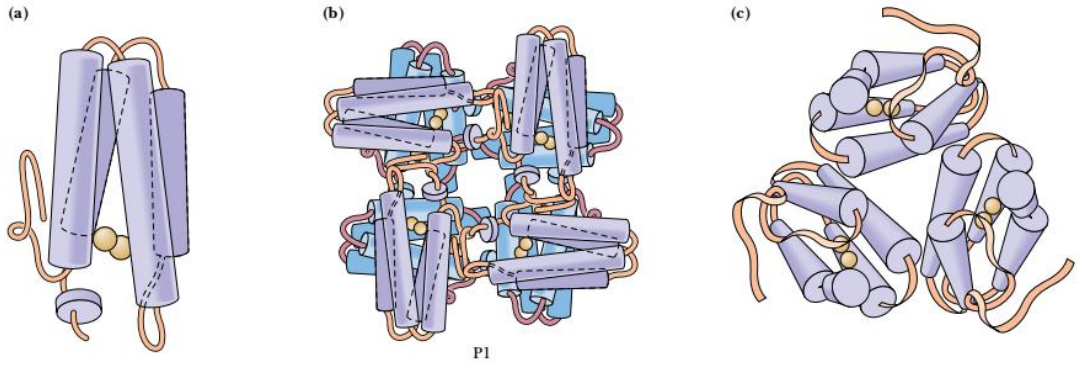
Collagen is a rigid, inextensible fibrous protein that is a principal constituent of connective tissue in animals, including tendons, cartilage, bones, teeth, skin, and blood vessels. The basic structural unit of collagen is **tropocollagen,** which has a molecular weight of 285,000 and consists of three intertwined polypeptide chains, each about 1000 amino acids in length. Tropocollagen molecules are about 300 nm long and only about 1.4 nm in diameter.

Collagen-Related Diseases

Collagen provides an ideal case study of the molecular basis of physiology and disease. For example, the nature and extent of collagen cross-linking depends on the age and function of the tissue. Collagen from young animals is predominantly uncross linked and can be extracted in soluble form, whereas collagen from older animals is highly cross-linked and thus insoluble. The loss of flexibility of joints with aging is probably due in part to increased cross-linking of collagen. Several serious and debilitating diseases involving collagen abnormalities are known. Lathyrism occurs in animals due to the regular consumption of seeds of Lathyrus odoratus, the sweet pea, and involves weakening and abnormalities in blood vessels, joints, and bones. A number of rare genetic diseases involve collagen abnormalities, including Marfan's syndrome and the Ehlers–Danlos syn dromes, which result in hyperextensible joints and skin.

4.Quaternary Structure

Many proteins exist in nature as **oligomers,** complexes composed of (often symmetric) noncovalent assemblies of two or more monomer subunits. The way in which separate folded monomeric protein subunits associate to form the oligomeric protein constitutes the **quaternary structure** of that protein.



The oligomeric states of hemerythrin in various marine worms. (a) The hemerythrin in *Thermiste zostericola* crystallized as a monomer; (b) the octameric hemerythrin crystallized from *Phascolopsis gouldii;* (c) the trimeric hemerythrin crystallized from *Siphonosoma* collected in mangrove swamps in Fiji.

Biological Functions of Proteins and Some Representative Examples

Functional Class	Examples
Enzymes	Ribonuclease Trypsin Phosphofructokinase Alcohol dehydrogenase Catalase "Malic" enzyme
Regulatory proteins	Insulin Somatotropin Thyrotropin *lac* repressor NF1 (nuclear factor 1) Catabolite activator protein (CAP) AP1
Transport proteins	Hemoglobin Serum albumin Glucose transporter
Storage proteins	Ovalbumin Casein Zein Phaseolin Ferritin
Contractile and motile proteins	Actin Myosin Tubulin Dynein Kinesin
Structural proteins	α-Keratin Collagen Elastin Fibroin Proteoglycans
Scaffold proteins	Grb 2 crk shc stat IRS-1
Protective and exploitive proteins	Immunoglobulins Thrombin Fibrinogen Antifreeze proteins Snake and bee venom proteins Diphtheria toxin Ricin
Exotic proteins	Monellin Resilin Glue proteins

Regulatory Proteins

A number of proteins do not perform any obvious chemical transformation but nevertheless can regulate the ability of other proteins to carry out their physiological functions. Such proteins are referred to as **regulatory proteins.** A well-known example is insulin, the hormone regulating glucose metabolism in animals.

Transport Proteins

A third class of proteins is the **transport proteins.** These proteins function to transport specific substances from one place to another. One type of transport is exemplified by the transport of oxygen from the lungs to the tissues by haemoglobin

Storage Proteins

Proteins whose biological function is to provide a reservoir of an essential nutrient are called **storage proteins.** Because proteins are amino acid polymers and because nitrogen is commonly a limiting nutrient for growth, organisms have exploited proteins as a means to provide sufficient nitrogen in times of need. For example, ovalbumin, the protein of egg white.

Contractile and Motile Proteins

Certain proteins endow cells with unique capabilities for movement. Cell division, muscle contraction, and cell motility represent some of the ways in which cells execute motion. The **contractile** and **motile proteins** underlying these motions share a common property: they are filamentous or polymerize to form filaments. Examples include actin and myosin.

Structural Proteins

An apparently passive but very important role of proteins is their function in creating and maintaining biological structures. **Structural proteins** provide strength and protection to cells and tissues. Monomeric units of structural proteins typically polymerize to generate long fibers (as in hair) or protective sheets of fibrous arrays, as in cowhide (leather). - Keratins are insoluble fibrous proteins making up hair, horns, and fingernails. Collagen, another insoluble fibrous protein, is found in bone, connective tissue, tendons, cartilage, and hide, where it forms inelastic fibrils of great strength.

Exotic Proteins : Some proteins display rather exotic functions that do not quite fit the previous classifications. Monellin, a protein found in an African plant, has a very sweet taste and is being considered as an artificial sweetener for human consumption.

GLYCOPROTEINS. Glycoproteins are proteins that contain carbohydrate.

Proteins destined for an extracellular location are characteristically glycoproteins. For example, fibronectin and proteoglycans are important components of the extracellular matrix that surrounds the cells of most tissues in animals. Immunoglobulin G molecules are the principal antibody species found circulating free in the blood plasma. Many membrane proteins are glycosylated on their extracellular segments.

LIPOPROTEINS. Blood plasma lipoproteins are prominent examples of the class of proteins conjugated with lipid. The plasma lipoproteins function primarily in the transport of lipids to sites of active membrane synthesis. Serum levels of low density lipoproteins (LDLs) are often used as a clinical index of susceptibility to vascular disease.

NUCLEOPROTEINS. Nucleoprotein conjugates have many roles in the storage and transmission of genetic information. Ribosomes are the sites of protein synthesis. Virus particles and even chromosomes are protein–nucleic acid complexes.

PHOSPHOPROTEINS. These proteins have phosphate groups esterified to the hydroxyls of serine, threonine, or tyrosine residues. Casein, the major protein of milk, contains many phosphates and serves to bring essential phosphorus to the growing infant.

METALLOPROTEINS. Metalloproteins are either metal storage forms, as in the case of ferritin, or enzymes in which the metal atom participates in a catalytically important manner. We encounter many examples throughout this book of the vital metabolic functions served by metalloenzymes.

HEMOPROTEINS. These proteins are actually a subclass of metalloproteins because their prosthetic group is **heme,** the name given to iron protoporphyrin IX. Because heme-containing proteins enjoy so many prominent biological functions, they are considered a class by themselves.

FLAVOPROTEINS. Flavin is an essential substance for the activity of a number of important oxidoreductases. We discuss the chemistry of flavin and its derivatives, FMN and FAD, in the chapter on electron transport and oxidative phosphorylation