Molecular Biology
Lecture 3

- **EUKARYOTIC CHROMOSOME STRUCTURE**

The mitotic The familiar picture of a chromosome (*Fig. 1*) is actually that of its most highly **chromosome** condensed state at **mitosis**. As the daughter chromosomes are pulled apart by the **mitotic spindle** at cell division. The structure in *Fig. 1* actually illustrates two identical **sister chromatids**, the products of replication of a single chromosome, joined at their **centromeres**. The tips of the chromosomes are the **telomeres**, which are also the ends of the DNA molecule. The chromosomal loops fan out from a central scaffold or nuclear matrix region consisting of protein. The loops consist of chromatin in the 30 nm fiber form.

Nuclear matrix

The nuclear matrix is a scaffold of insoluble protein fibers which acts as an organizational framework for nuclear processes, including DNA replication. Huge **replication factories** containing all the enzymes and DNA associated with the replication forks of all replicons within a cluster are immobilized on the matrix, and the DNA moves through these sites as it replicates.

The centromere

The **centromere** is the constricted region where the two sister chromatids are joined in the metaphase chromosome. This is the site of assembly of the **kinetochore**, a protein complex which attaches to the **microtubules** of the mitotic spindle. The microtubules act to separate the chromatids at **anaphase**.

Telomeres

Telomeres are specialized DNA sequences that form the ends of the linear DNA molecules of the eukaryotic chromosomes. A telomere consists of up to hundreds of copies of a short repeated sequence (5′-TTAGGG-3′ in humans), which is synthesized by the enzyme **telomerase** in a mechanism independent of normal DNA replication. The telomeric DNA forms a special secondary structure, the function of which is to protect the ends of the chromosome proper from degradation.

Heterochromatin

Heterochromatin comprises a portion of the chromatin which is highly compacted. It can be visualized under the microscope as dense regions at the periphery of the nucleus, and probably consists of closely packed regions of 30 nm fiber. It has been shown more recently that heterochromatin is transcriptionally inactive.

Euchromatin

The rest of the chromatin, which is not visible as heterochromatin, is known as **euchromatin**, and is the region where all transcription takes place. Euchromatin is not homogeneous, however, consisting of chromosomal loops compacted in 30 nm fibers, and regions where genes are actively being transcribed or are destined to be transcribed in that cell type, where the 30 nm fiber has been dissociated to the ‘beads on a string’ structure.

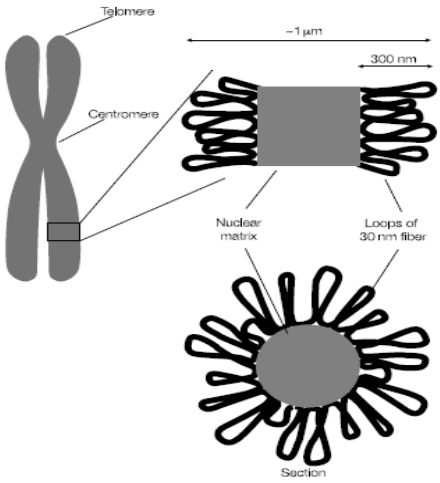
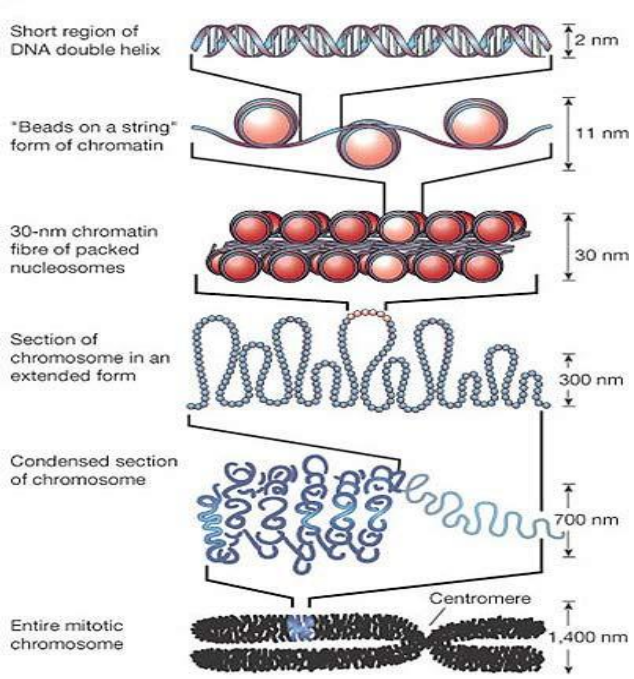


Fig. 1. The structure of the mitotic chromosome.



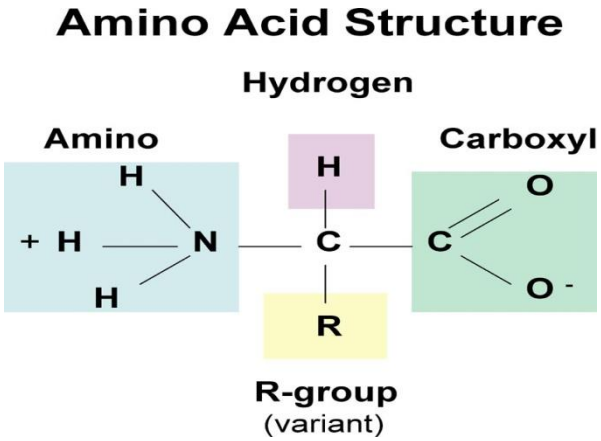
- **PROTEINS**

Proteins are polymers of amino acids linked together by peptide bonds.

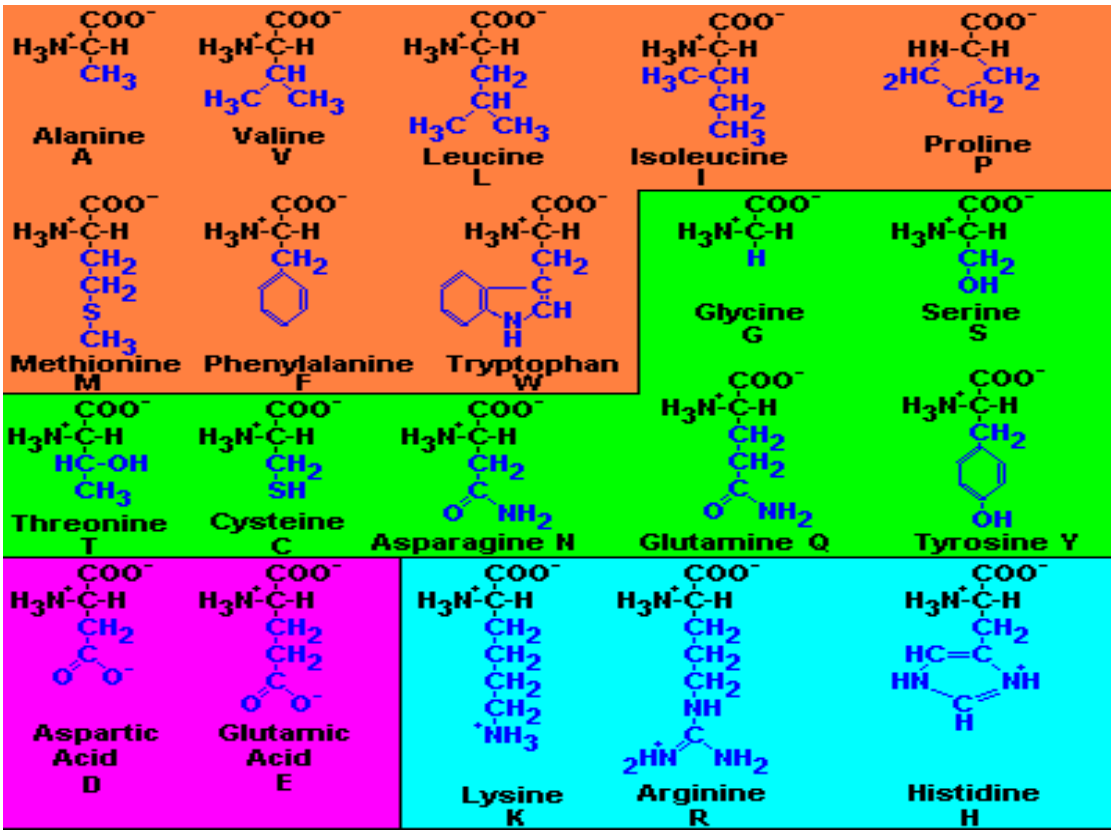
Amino acids

Structure

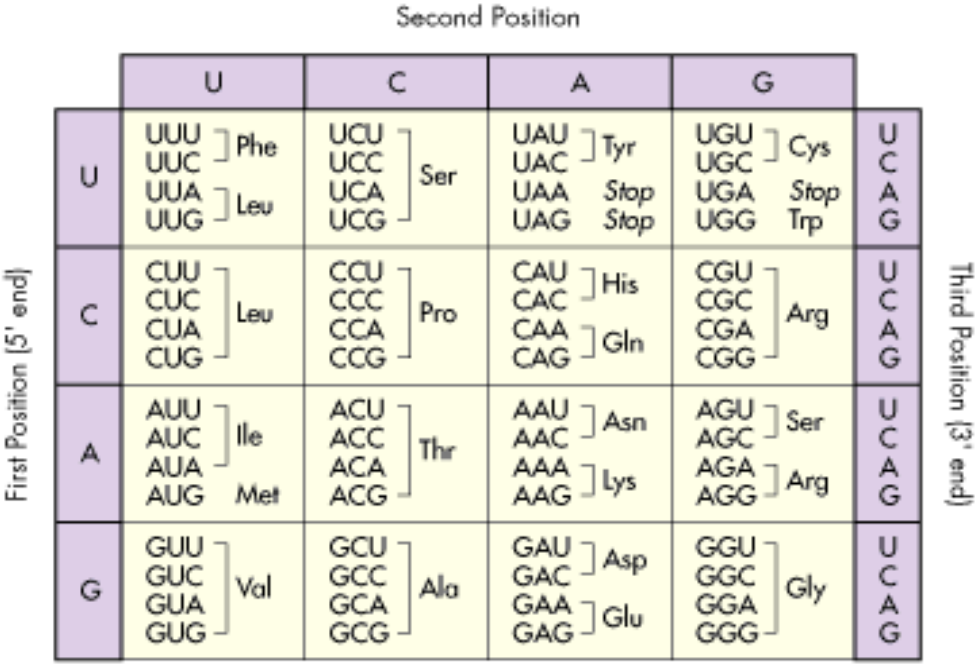
Proteins are polymers of L-amino acids. Apart from proline, all of the 20 amino acids found in proteins have a common structure in which a carbon atom (the α-carbon) is linked to a carboxyl group, a primary amino group, a proton and a **side chain** (R) which is different in each amino acid (*Fig. 1*).



Except in glycine, the α -carbon atom is asymmetric – it has four chemically different groups attached. Thus, amino acids can exist as pairs of optically active stereoisomers (D- and L-). However, only the L-isomers are found in proteins. Amino acids are dipolar ions (**zwitterions**) in aqueous solution and behave as both acids and bases (they are **amphoteric**). The side chains differ in size, shape, charge and chemical reactivity, and are responsible for the differences in the properties of different proteins (*Fig. 2*).



Amino Acid	3 Letter	1 Letter	Side chain polarity	Side chain charge
Alanine	Ala	A	nonpolar	neutral
Arginine	Arg	R	polar	positive
Asparagine	Asn	N	polar	neutral
Aspartic acid	Asp	D	polar	negative
Cysteine	Cys	C	nonpolar	neutral
Glutamic acid	Glu	E	polar	negative
Glutamine	Gln	Q	polar	neutral
Glycine	Gly	G	nonpolar	neutral
Histidine	His	H	polar	positive
Isoleucine	Ile	I	nonpolar	neutral
Leucine	Leu	L	nonpolar	neutral
Lysine	Lys	K	polar	positive
Methionine	Met	M	nonpolar	neutral
Phenylalanine	Phe	F	nonpolar	neutral
Proline	Pro	P	nonpolar	neutral
Serine	Ser	S	polar	neutral
Threonine	Thr	T	polar	neutral
Tryptophan	Trp	W	nonpolar	neutral
Tyrosine	Tyr	Y	polar	neutral
Valine	Val	V	nonpolar	neutral

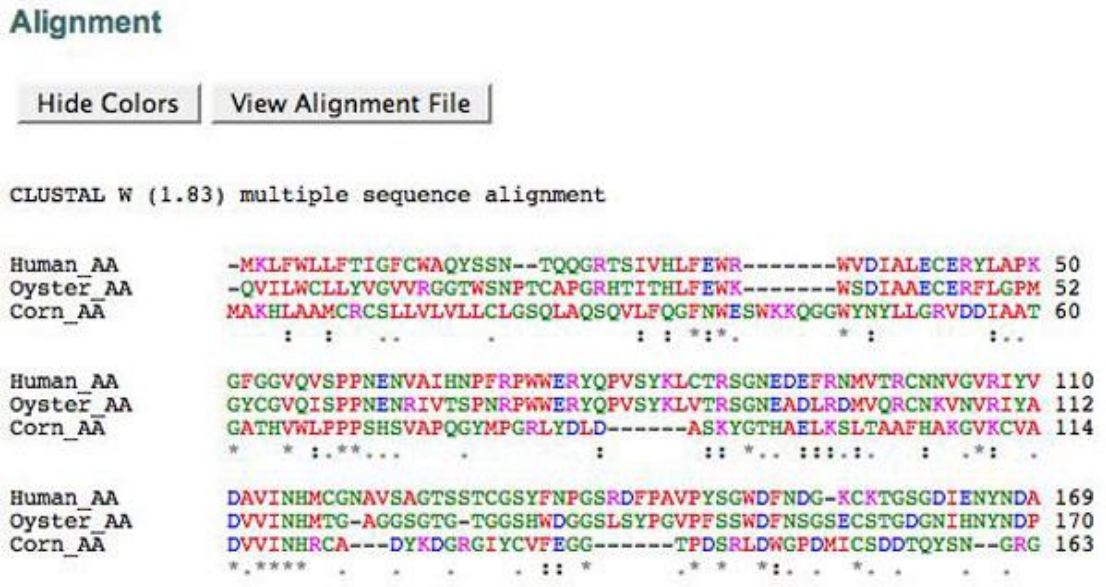


Sizes and shapes

Two broad classes of protein may be distinguished. **Globular proteins** are folded compactly and behave in solution more or less as spherical particles; most enzymes are globular in nature. **Fibrous proteins** have very high axial ratios (length/width) and are often important structural proteins, for example silk fibroin and keratin in hair and wool. Molecular masses can range from a few thousand Daltons (Da) (e.g. the hormone insulin with 51 amino acids and a molecular mass of 5734 Da) to at least 5 million Daltons in the case of the enzyme complex pyruvate dehydrogenase. Some proteins contain bound nonprotein material, either in the form of small **prosthetic groups**, which may act as co-factors in enzyme reactions, or as large associations (e.g. the lipids in **lipoproteins** or the carbohydrate in **glycoproteins**.

Primary structure

The α-carboxyl group of one amino acid is covalently linked to the α-amino group of the next amino acid by an amide bond, commonly known as a **peptide bond** when in proteins. When two amino acid **residues** are linked in this way the product is a **dipeptide**. Many amino acids linked by peptide bonds form a **polypeptide**. The repeating sequence of α-carbon atoms and peptide bonds provides the **backbone** of the polypeptide while the different amino acid **side chains** confer functionality on the protein. The amino acid at one end of a polypeptide has an unattached α-amino group while the one at the other end has a free α-carboxyl group. Hence, polypeptides are directional, with an **N terminus** and a **C terminus**. Sometimes the N terminus is **blocked** with, for example, an acetyl group. The sequence of amino acids from the N to the C terminus is the **primary structure** of the polypeptide. Typical sizes for single polypeptide chains are within the range 100–1500 amino acids, though longer and shorter ones exist.



Secondary Structure

Polypeptides can fold into a number of regular structures. The right-handed α-helix has 3.6 amino acids per turn and is stabilized by hydrogen bonds between peptide N–H and C=O groups three residues apart. Parallel and antiparallel β-pleated sheets are stabilized by hydrogen bonds between different portions of the polypeptide chain.

Tertiary structure

The different sections of secondary structure and connecting regions fold into a well-defined tertiary structure, with hydrophilic amino acids mostly on the surface and hydrophobic ones in the interior. The structure is stabilized by noncovalent interactions and sometimes disulfide bonds. Denaturation leads to loss of secondary and tertiary structure.

What do proteins do ?

- **Enzymes.** all enzymes are proteins. These can enhance the rate of biochemical reactions

by several orders of magnitude. Binding of the **substrate** involves various noncovalent interactions with specific amino acid side chains, including van der Waals forces, hydrogen bonds, salt bridges and

hydrophobic forces. Specificity of binding can be extremely high, with only a single substrate binding (e.g. glucose oxidase binds only glucose), or it can be group-specific (e.g. hexokinase binds a variety of hexose sugars). Side chains can also be directly involved in catalysis, for example by acting as nucleophiles, or proton donors or abstractors.

● **Signaling**. Receptor proteins in cell membranes can bind **ligands** (e.g. hormones) from the extracellular medium and, initiate reactions within the cell in response to that ligand. Ligand binding is similar to substrate binding but the ligand usually remains unchanged. Some hormones are themselves small proteins, such as insulin and growth hormone.

● **Transport and storage**.

Hemoglobin transports oxygen in the red blood cells while **transferrin** transports iron to the liver. Once in the liver, iron is stored bound to the protein **ferritin**. Dietary fats are carried in the blood by **lipoproteins**. Many other molecules and ions are transported and stored ina protein-bound form. This can enhance solubility and reduce reactivity until they are required.

● **Structure and movement**. **Collagen** is the major protein in skin, bone and connective tissue, while hair is made mainly from **keratin**. There are also many structural proteins within the cell, for example in the **cytoskeleton**. The major muscle proteins **actin** and **myosin** form sliding filaments which are the basis of muscle contraction.

● **Nutrition**. **Casein** and **ovalbumin** are the major proteins of milk and eggs, respectively, and are used to provide the amino acids for growth of developing offspring. Seed proteins also provide nutrition for germinating plant embryos.

● **Immunity**. **Antibodies**, which recognize and bind to bacteria, viruses and other foreign material (the **antigen**) are proteins.

● **Regulation**. **Transcription factors** bind to and modulate the function of DNA. Many other proteins modify the functions of other molecules by binding to them.

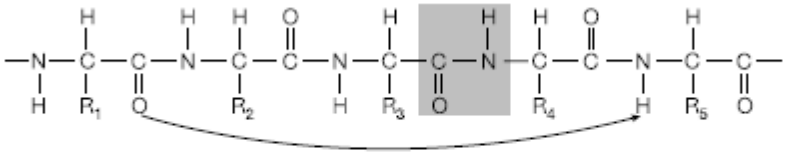


Fig. 1. Section of a polypeptide chain. The peptide bond is boxed. In the α-helix, the CO group of amino acid residue n is hydrogen-bonded to the NH group of residue n + 4 (arrowed).

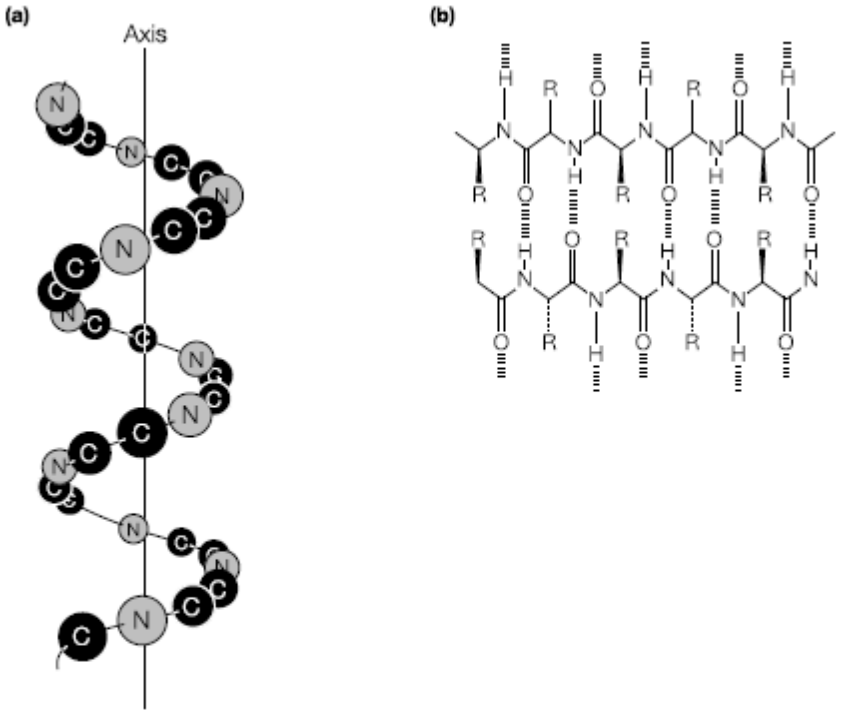


Fig. 2. (a) α-Helix secondary structure. Only the α-carbon and peptide bond carbon and nitrogen atoms of the polypeptide backbone are shown for clarity. (b) Section of a β-sheet secondary structure.

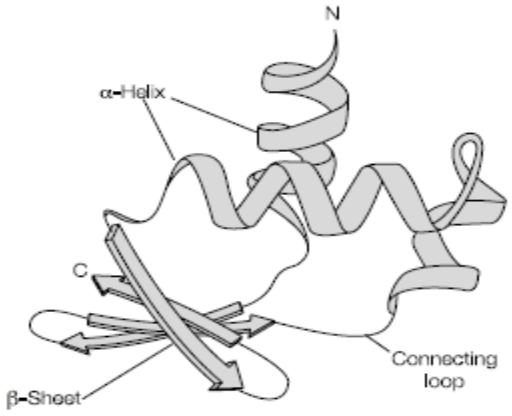


Fig. 3. Schematic diagram of a section of protein tertiary structure.