

# CHEMISTRY OF THE CELL

*Ho Huynh Thuy Duong*

*University of Science*

According to conventional definitions, a living organism is composed of one or more cells, can grow, reproduce and response to stimuli, and has some other characteristics.

These characteristics are mainly based on the **cell** – the basic unit of life

☞ ...

☞ Robert Hooke (1635 – 1703) termed “*cell*” (*cella*, a Latin word means “small container”) what he saw through a microscope when he observed a slice of cork.

☞ René Dutrochet (1776-1847) declared in 1824 that “*The cell is the fundamental element in the structure of living bodies, forming both animals and plants through juxtaposition*”.

☞ Theodore Schwann (1810-1882) created the term “*Cell theory*” and declared that plants consisted of cells. A similar declaration was made by Matthias Schleiden (1804-1881) on the units forming animals.

☞ Rudolf Virchow (1821-1902) stated that “*Every cell comes from a cell*”

***WHAT IS A CELL MADE OF ?***

# **CHEMISTRY OF THE CELL**

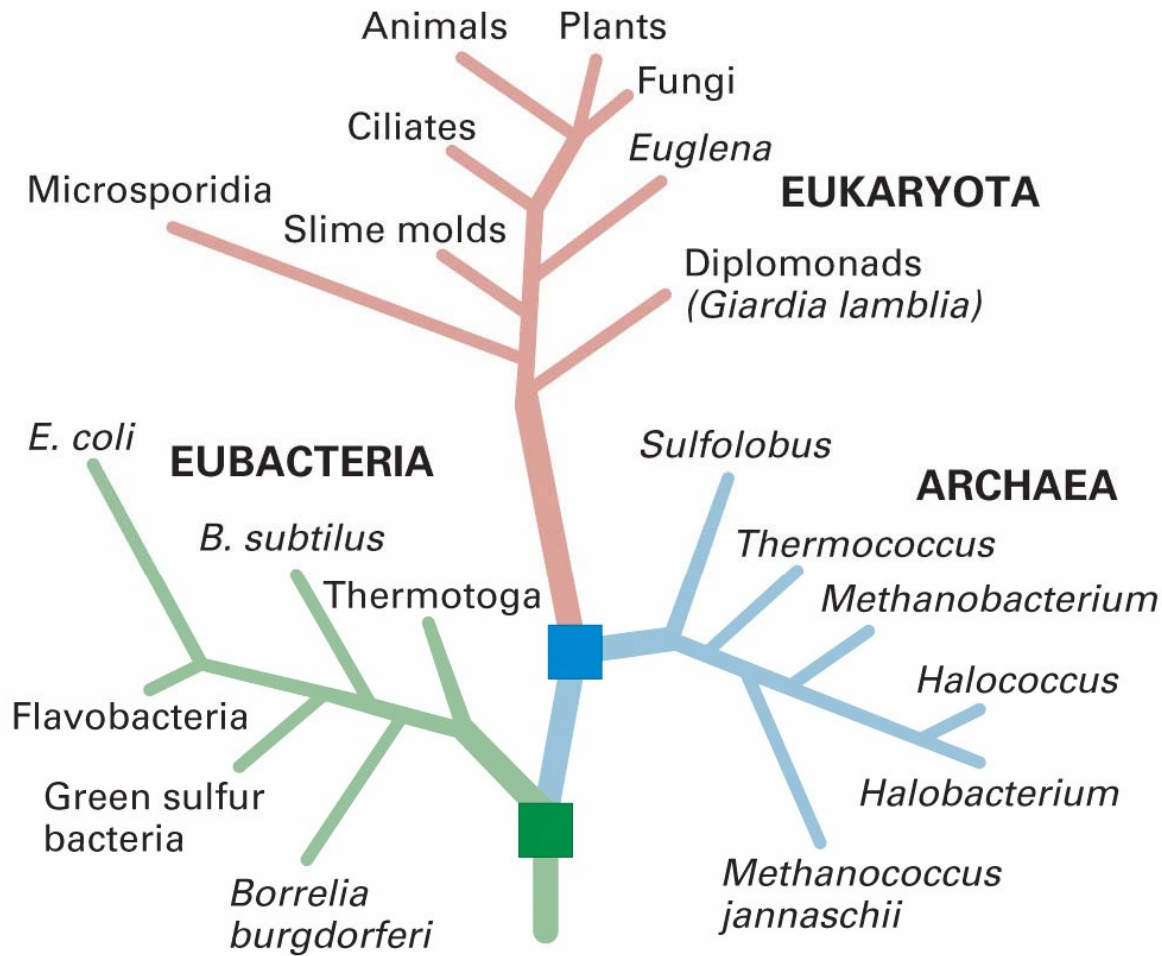
**∞ PROKARYOTIC AND EUKARYOTIC CELLS**

**∞ BIOLOGICAL MACROMOLECULES : POLYSACCHARIDES, LIPIDS, PROTEINS AND NUCLEIC ACIDS**

**∞ WEAK CHEMICAL INTERACTIONS DETERMINE MACROMOLECULE STRUCTURE AND FUNCTIONS**

**∞ ORGANIZATION OF PROKARYOTIC AND EUKARYOTIC GENOME**

# THE TREE OF LIFE



■ Presumed common progenitor of all extant organisms

■ Presumed common progenitor of archaeobacteria and eukaryotes

# PROKARYOTIC AND EUKARYOTIC CELL

∞ The Phylogenetic Tree of Life has three main branches : **Bacteria, Archaea and Eukaryota.**

**Bacteria and Archaea (Archaeobacteria) are classified as Prokaryotes. Archaea are similar to Bacteria for metabolism and energy production but are more related to Eukaryotes for DNA replication, transcription and translation.**

∞ Some characteristics of prokaryotic cell and eukaryotic cell are listed below :

## PROKARYOTA

### Archaea – Bacteria

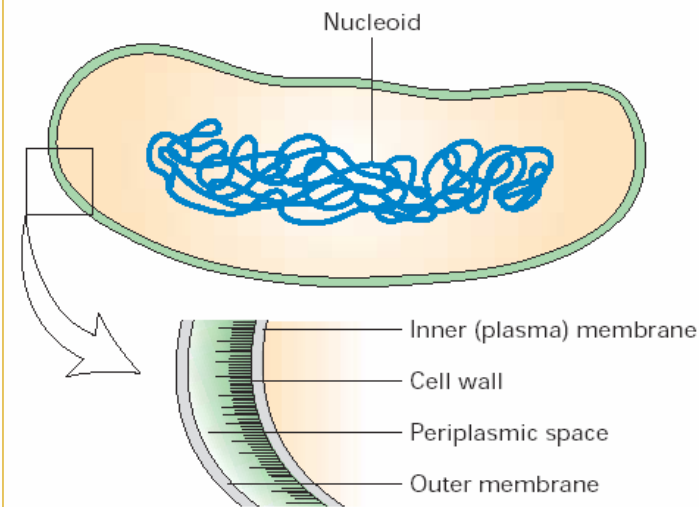
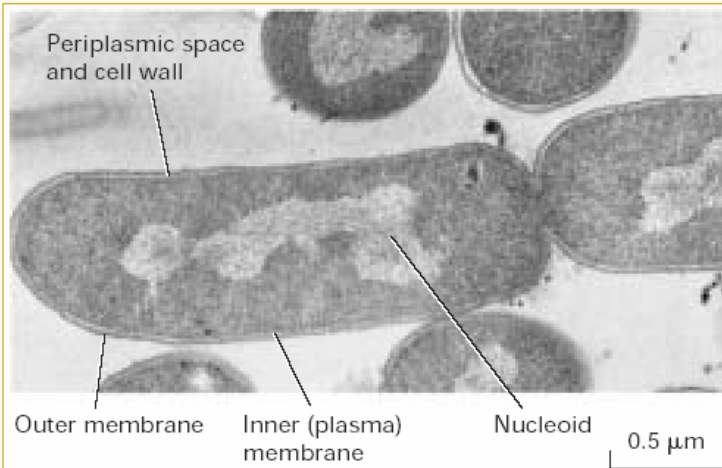
- **No nuclear compartment**
- **Unicellular organisms**
- **Small cell**
- **Small and compact genome**
- **Lack of cellular organelles**
- **Varied ecological niches and varied biochemical capabilities, especially for Archaea (Archaeobacteria)**

## EUKARYOTA

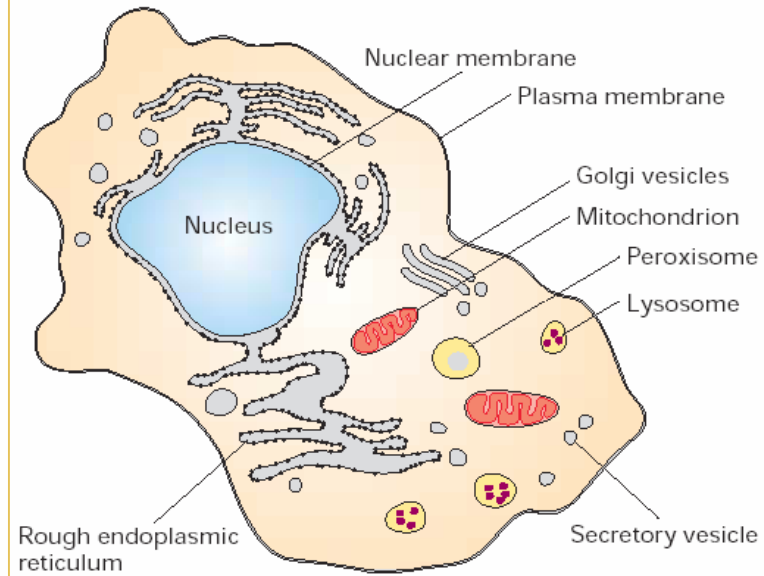
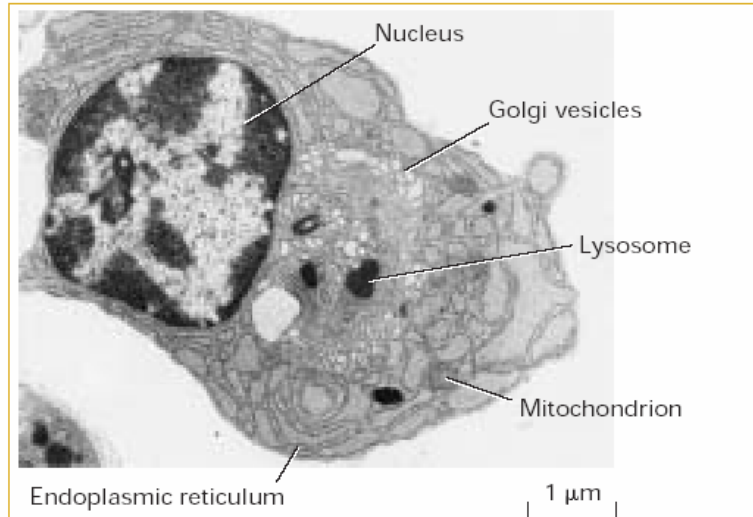
- **Nucleus**
- **Usually multicellular organisms**
- **Larger cell**
- **Big genome containing important amount of noncoding DNA**
- **Cellular organelles (mitochondria, Golgi apparatus, reticulum endoplasmic, chloroplasts, ...)**

# PROKARYOTIC & EUKARYOTIC CELLS

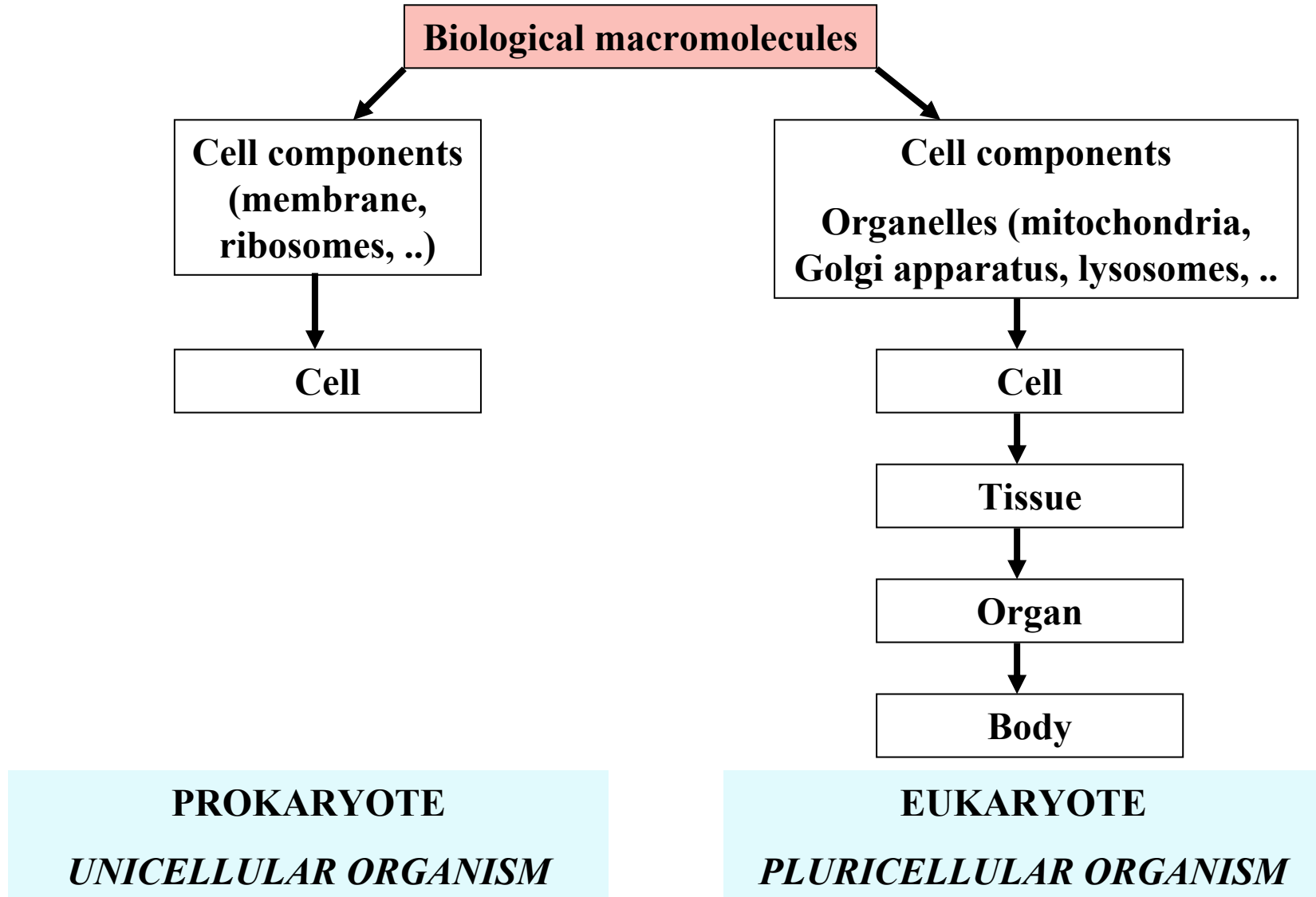
(a) Prokaryotic cell



(b) Eukaryotic cell



# ORGANIZATION OF LIVING ORGANISMS



# BIOLOGICAL MACROMOLECULES

☞ The main biological macromolecules include **Proteins, Nucleic Acids, Lipids and Polysaccharides**

☞ Biological macromolecules are polymers constituted of chemically similar monomers linked to each other by covalent (high-energy) bonds :

□-□-□-□ **Proteins** are polymers of amino acids linked by peptide bonds --

○-○-○-○ **Nucleic acids** (DNA, RNA) are polymers of nucleotides linked by phosphodiester bonds --

◇-◇-◇-◇ **Polysaccharides** are polymers of simple sugars linked by glycosidic bonds -

△-△-△-△ **Lipids** are not strictly macromolecules, constituted only of a few fatty acids

☞ The structure and shape of biological macromolecules which determine their functions are mainly defined by weak (low-energy) chemical bonds.

☞ Weak chemical bonds could be established between different parts of the macromolecule (intra-molecular interaction) or between different macromolecules (inter-molecular interaction) to form large macromolecular assemblies, e.g ribosomes comprising rRNA and proteins, chromatin constituted of DNA and histones, ...



# POLYSACCHARIDES

Simple sugars forming polysaccharides can be linked by different kinds of glycosidic bond, e.g  $\alpha(1\rightarrow4)$  or  $\beta(1\rightarrow4)$  linkages are found in linear polymers such as cellulose whereas  $\alpha(1\rightarrow6)$  linkages form branched polysaccharides such as amylopectin which is one of the two components of starch.

The main functions of polysaccharides :

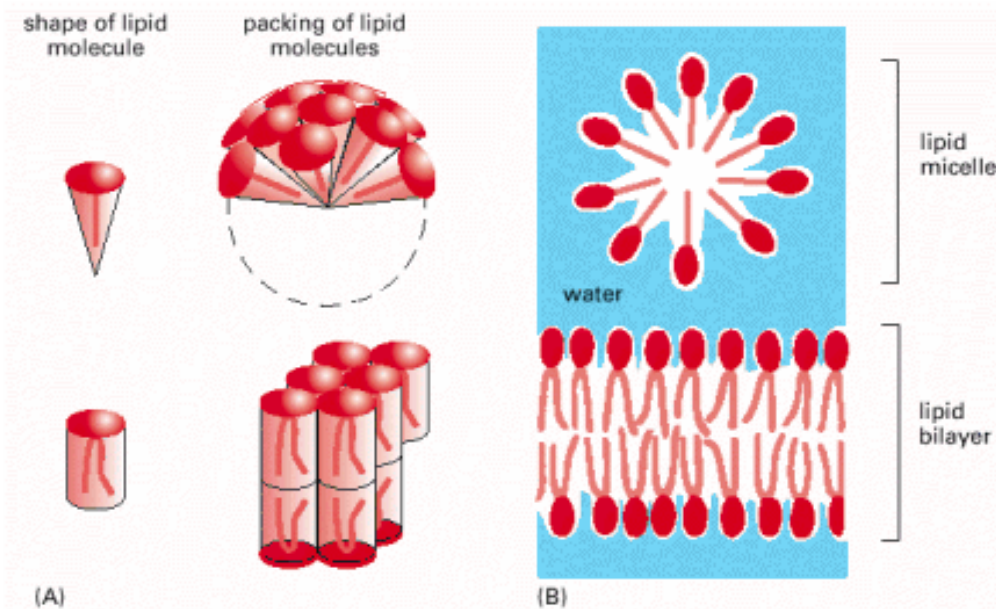
∞ **Food molecules** of the cell : Glucose is the principal food molecule of many cell types.



∞ **Cellular stock of energy**, e.g starch in plants and glycogen in animals

∞ Other functions of polysaccharides are : **structural components** of some cells e.g cellulose in plants or chitin in fungal cell wall and insects, components of complex macromolecules such as glycoproteins involving in **cell-cell recognition**, ..

# LIPIDS



*Packing arrangements of lipid molecules in an aqueous environment. (A) Wedge-shaped lipid molecules (above) form micelles, whereas cylinder-shaped phospholipid molecules (below) form bilayers. (B) A lipid micelle and a lipid bilayer seen in cross section.*

*Lipid molecules spontaneously form one or other of these structures in water, depending on their shape.*

*“Copyright 2002 from Molecular Biology of the Cell by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC.”*

The main functions of lipids :

☞ **Source of energy** for the cell ; lipids can produce twice more energy than sugar (weight per weight)

☞ **Construction of cell membrane and membrane of other cellular organelles** ; due to their hydrophobic nature, they can form stable bilayers which are the basis of all biological membranes.

April 2009

# THE 20 AMINO ACIDS CONSTITUTING PROTEINS

POLAR AMINO ACIDS				NONPOLAR AMINO ACIDS			
AMINO ACID			SIDE CHAIN	AMINO ACID			SIDE CHAIN
Aspartic acid	Asp	D	negative	Alanine	Ala	A	nonpolar
Glutamic acid	Glu	E	negative	Glycine	Gly	G	nonpolar
Arginine	Arg	R	positive	Valine	Val	V	nonpolar
Lysine	Lys	K	positive	Leucine	Leu	L	nonpolar
Histidine	His	H	positive	Isoleucine	Ile	I	nonpolar
Asparagine	Asn	N	uncharged polar	Proline	Pro	P	nonpolar
Glutamine	Gln	Q	uncharged polar	Phenylalanine	Phe	F	nonpolar
Serine	Ser	S	uncharged polar	Methionine	Met	M	nonpolar
Threonine	Thr	T	uncharged polar	Tryptophan	Trp	W	nonpolar
Tyrosine	Tyr	Y	uncharged polar	Cysteine	Cys	C	nonpolar

*“Copyright 2002 from Molecular Biology of the Cell by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC.”*

☞ Amino acids with charged (negative or positive) side chains (R) e.g.  $-\text{CH}_2\text{COO}^-$  (Asp),  $-(\text{CH}_2)_4\text{NH}_3^+$  (Lys), ... have role in catalytic activities of many enzymes and form salt bridges in proteins.

☞ Amino acids with uncharged polar side chains :  $-\text{CH}_2\text{OH}$  (Ser),  $-\text{CH}(\text{OH})\text{CH}_3$  (Thr), ... form hydrogen bonds with water and are usually found on the outside of protein molecules.

☞ Amino acids with nonpolar side chains :  $-\text{H}$  (Gly),  $-\text{CH}_3$  (Ala), ...are clustered on the inside of proteins.

# PROTEINS

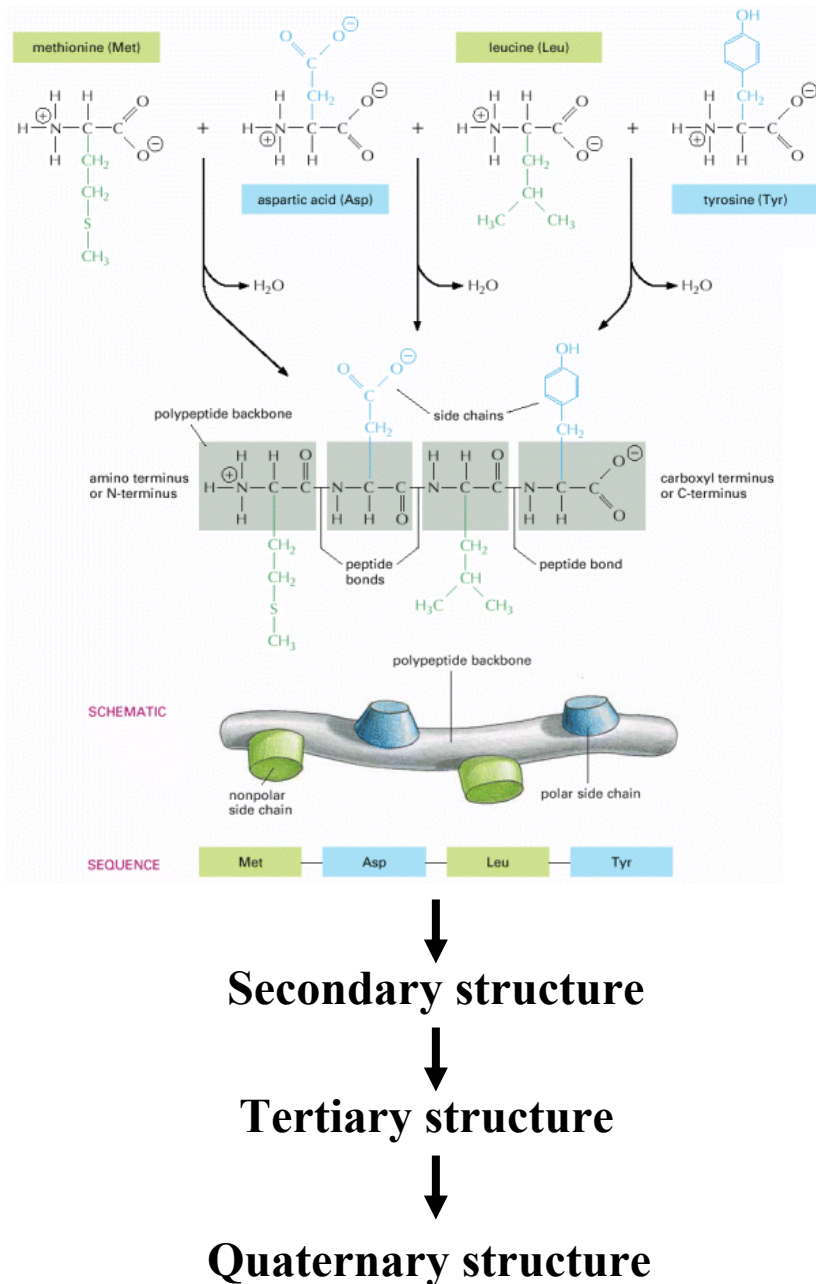
Proteins have 4 structure levels ranging from primary to quaternary structure

☞ Primary structure : Sequence of amino acids linked together by peptide bonds to form polypeptides ; typical sizes are within 100 – 1500 amino acids.

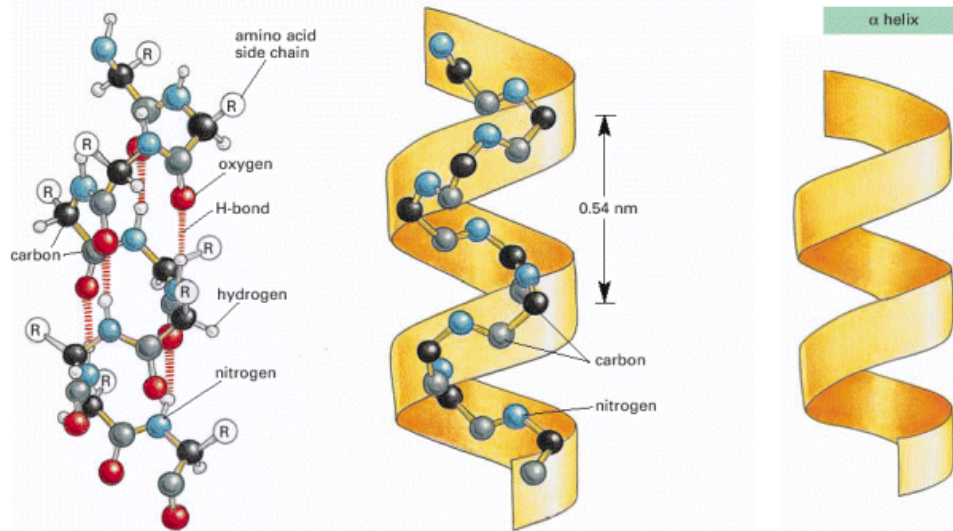
☞ Secondary structure : Regular structure due to folding of the primary sequence and maintained by hydrogen bonds ; the main secondary structures include  $\alpha$ -helix and  $\beta$ -sheet.

☞ Tertiary structure : three-dimensional structure due to the combination of different sections of  $\alpha$ -helix,  $\beta$ -sheet and connecting loops of the polypeptide chain. Tertiary structures are hold by various types of weak chemical bonds.

☞ Quaternary structure : constituted of some similar or different polypeptide chains, holded by noncovalent bonds and disulfide bonds between cysteins from different polypeptide chains.



# SECONDARY STRUCTURE OF PROTEINS : $\alpha$ -HELIX



*One of the two main secondary structures,  $\alpha$ -helix, is a right-handed helix with 3.6 amino acid residues per turn formed by polypeptide backbone and held by hydrogen bonding between the N-H group of an amino acid and another one three residues away*

*“Copyright 2002 from Molecular Biology of the Cell by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC.”*



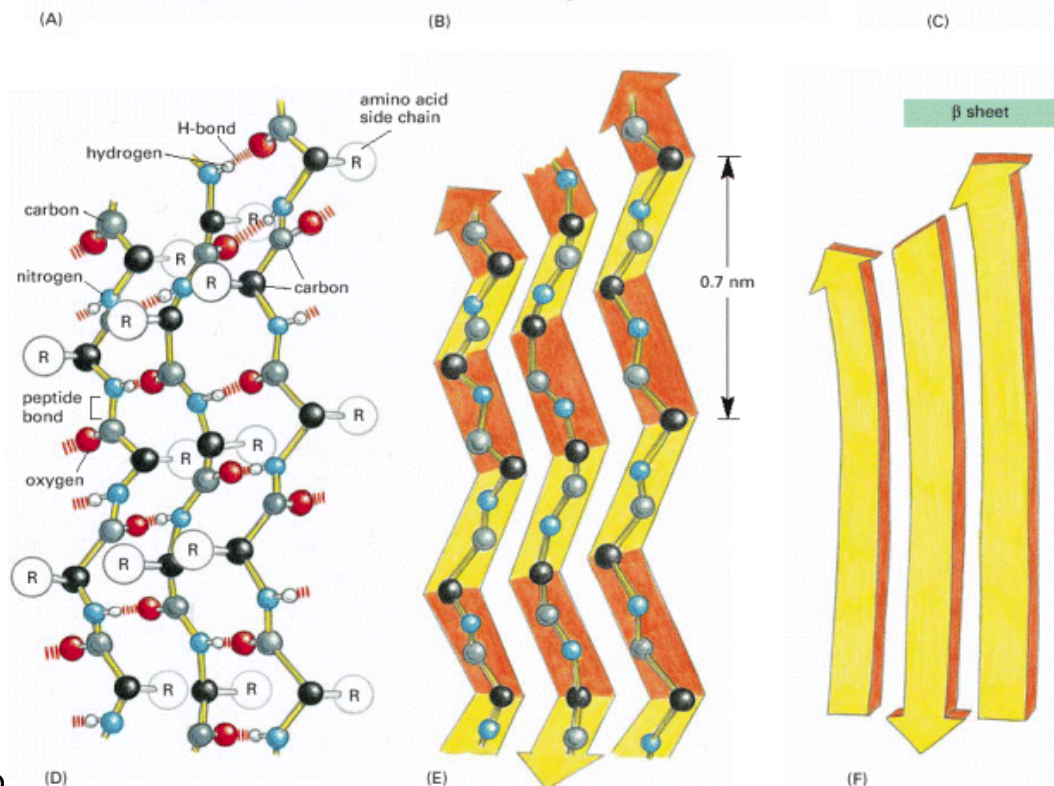
*Leucine zipper (green) of a DNA-binding protein, constituted of 2  $\alpha$ -helices, binds to a DNA molecule (red)*

*“Copyright 2002 from Molecular Biology of the Cell by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC.”*

# SECONDARY STRUCTURE OF PROTEINS : $\beta$ -SHEET

A  **$\beta$ -sheet** is formed by hydrogen bonding between the N-H and C=O groups of a section with complementary groups of another section. These two sections are parts of the same polypeptide chain.

The arrow  $\Rightarrow$  shows the amino to carboxyl direction of a section of the polypeptide chain.



If adjacent sections have the same direction, the  $\beta$ -sheet is **parallel**.

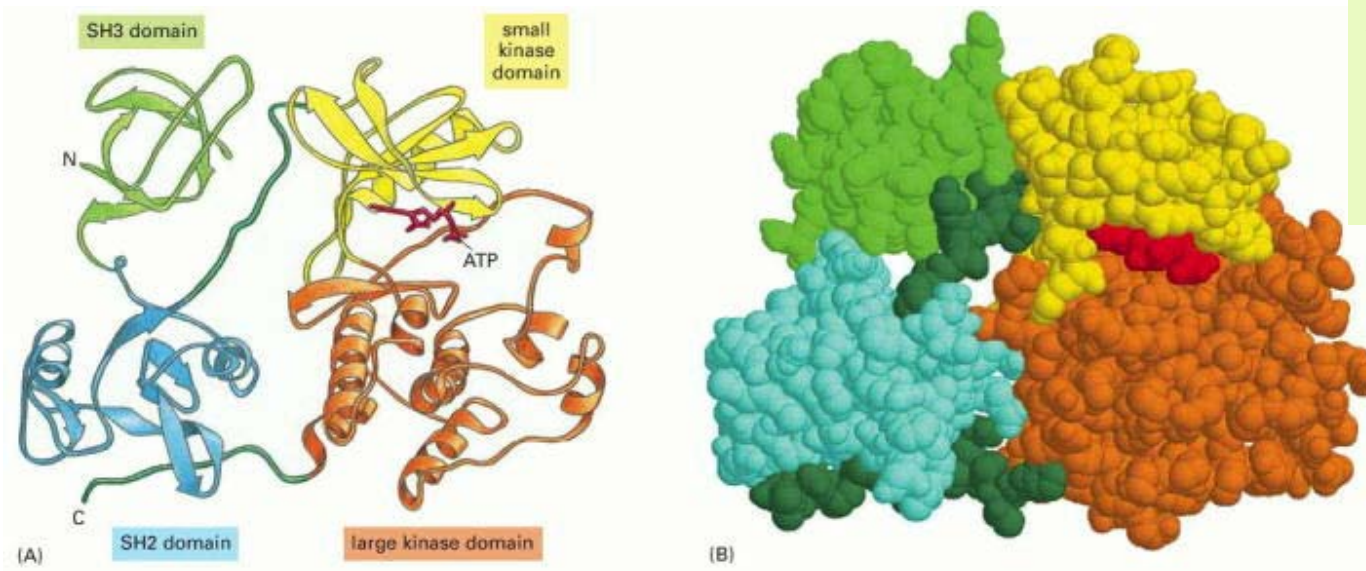
The  $\beta$ -sheet is **antiparallel** if adjacent sections run in opposite directions (*figure in the left*)

$\beta$ -sheets confer solidity to structural proteins, e.g silk proteins

Ap

“Copyright 2002 from Molecular Biology of the Cell by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC.”

# TERTIARY AND QUATERNARY STRUCTURE OF A PROTEIN



*Src protein tertiary structure presented as (A) ribbon model and (B) spacing-filling model*

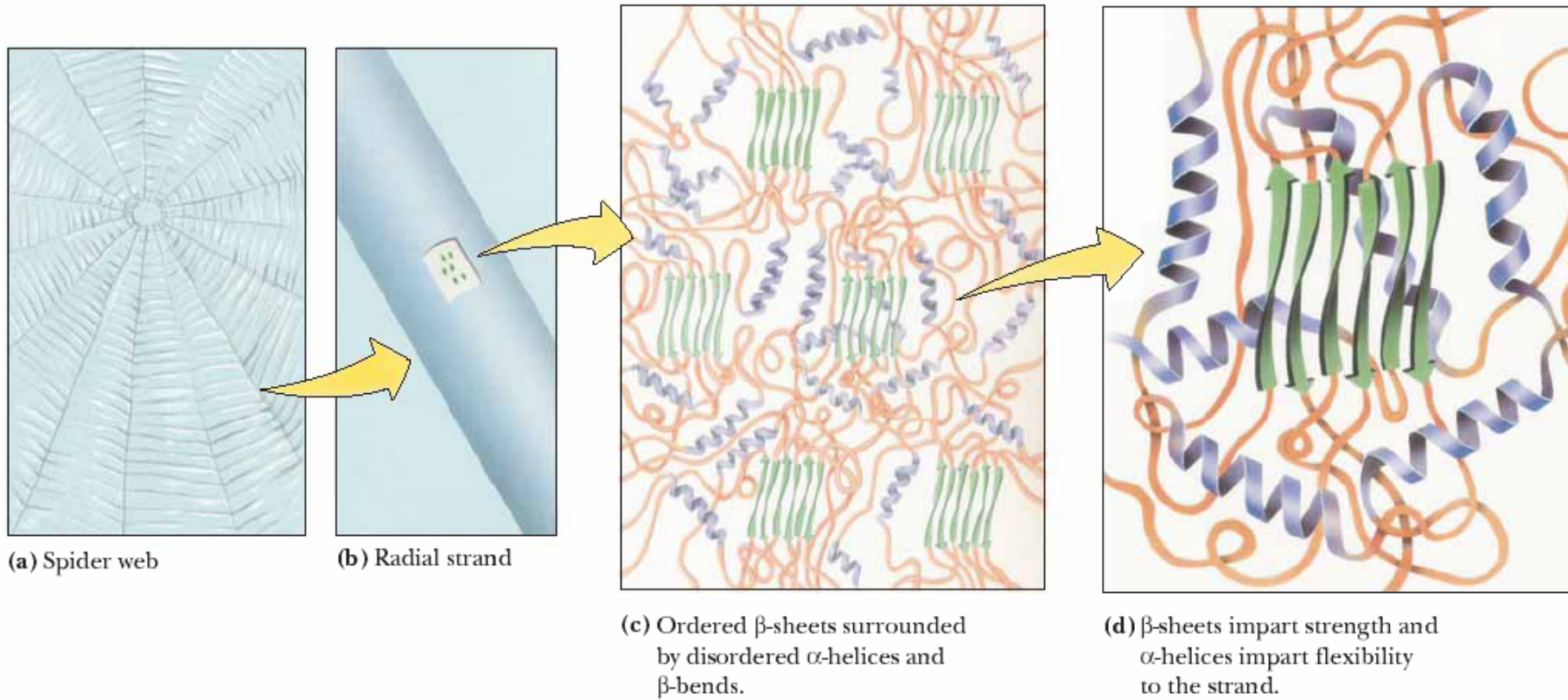
$\alpha$ -helices and  $\beta$ -sheets, linked together by turns or loops having no regular secondary structure, form the three-dimensional **tertiary structure**

*“Copyright 2002 from Molecular Biology of the Cell by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC.”*

Many proteins are composed of more than one polypeptide chain ; they have **quaternary structure**. **Quaternary structure** can (1) make very big proteins, or (2) create new proteins which have various activities by associating different polypeptide chains, each having a defined activity.

*Haemoglobin is composed of 2  $\alpha$ -globin and 2  $\beta$ -globin chains. Each chain is folded in tertiary structure. Haemoglobin, as a tetramer, is more effective in oxygen transport than an ~~as a~~ monomer form.*

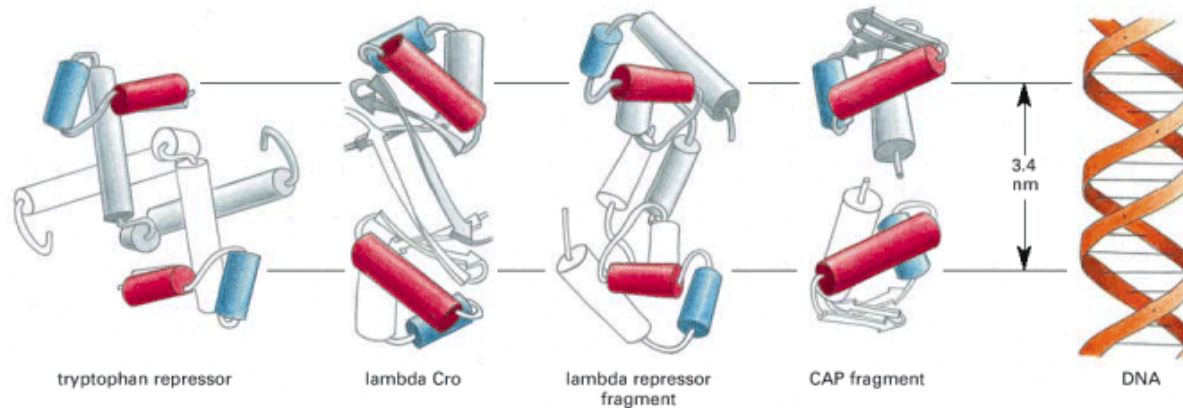
# SPIDER WEB STRUCTURE





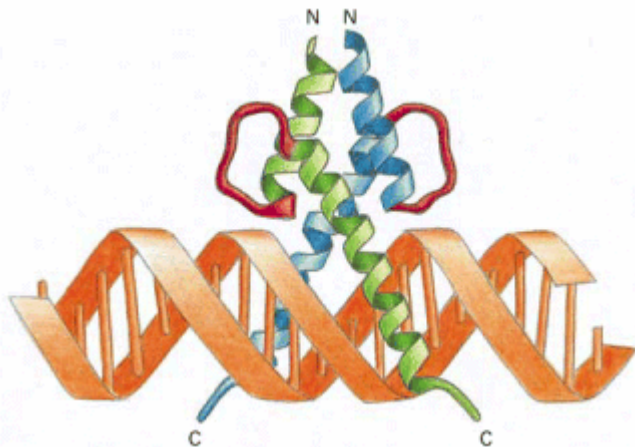
# PROTEINS ARE MODULAR

Although composed of limited number of **structural motifs**, proteins are very diversified. Protein diversity is due to various combinations of these basic structural motifs.



*“Copyright 2002 from Molecular Biology of the Cell by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC.”*

*The figure above shows some DNA –binding proteins having a **helix-turn-helix structural motif** with a recognition helix (red) linked to another helix (blue). This motif is related to DNA-binding activities.*

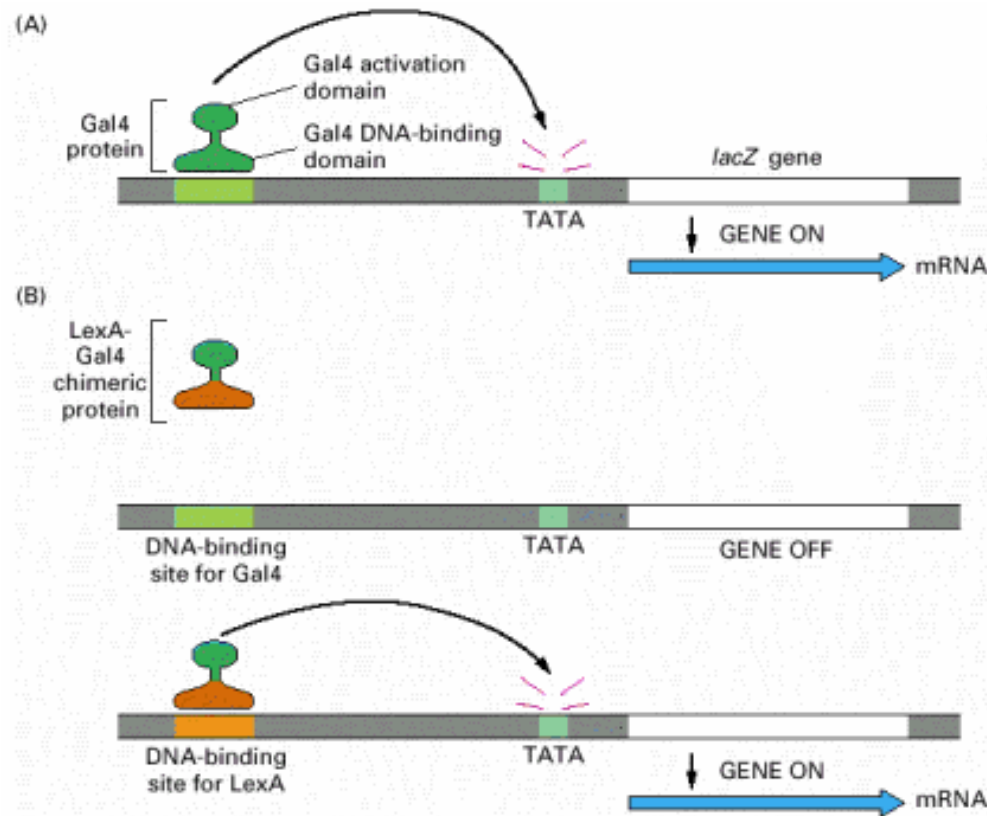


*Another example of structural motif, a **leucine zipper dimer** constituted of helix-loop-helix motifs usually found in DNA-binding proteins*

*“Copyright 2002 from Molecular Biology of the Cell by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC.”*

## PROTEINS ARE MODULAR (continued)

Proteins are **modular** - they contain different **domains** (modules) each having defined activity. Different proteins are created by different combinations of these domains.



*“Copyright 2002 from Molecular Biology of the Cell by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC.”*

***Gal 4 protein contains 2 domains : (1) DNA-binding domain and (2) Activation domain.***

***When replacing the Gal4 DNA-binding domain by LexA DNA-binding domain, the Gal4 protein will activate gene controlled by LexA promoter. In this example, a protein having new characteristic is created by combining an existing domain with a new domain***

# NUCLEIC ACIDS

Nucleic acids are polymers of **nucleotides**.

Each nucleotide comprises a heterocyclic base, a pentose sugar and a phosphate. Bases are of two categories : purines (adenine – A and guanine - G) and pyrimidines (cytosine – C, thymine – T, uracil –U). Bases exist in two tautomeric forms : amino (common) and imino (rare) for C, T, U ; keto (common) and enol (rare) for A, G.

Some **bases** can be **modified**. In DNA, modification is essentially methylation on the N-6 position of adenine, the 5-position and the 4-amino group of cytosine. In RNA, modifications are much more diverse (see “Protein synthesis”)

**Nucleoside** = Base + Sugar

**Nucleotide** = Base + Sugar + Phosphate

In DNA (deoxyribonucleic acid), there are 4 bases – A, G, C T and a deoxyribose ; whereas in RNA (ribonucleic acid), T is replaced by U and the sugar is a ribose.

Nucleic acids consist of nucleotides linked by **phosphodiester bonds**. Phosphodiester bond is formed by a phosphate linked to a pentose at 5' position and the next pentose at 3' position. Phosphodiester linkage confers a polarity to nucleic acid sequence. Conventionally, nucleic acid sequence is written in a 5'-3' direction, e.g 5' CAATAGCCATTAGCA 3'.

**Nucleic acids** are negatively charged in aqueous solution due to their phosphate. 19

# THE DNA DOUBLE HELIX

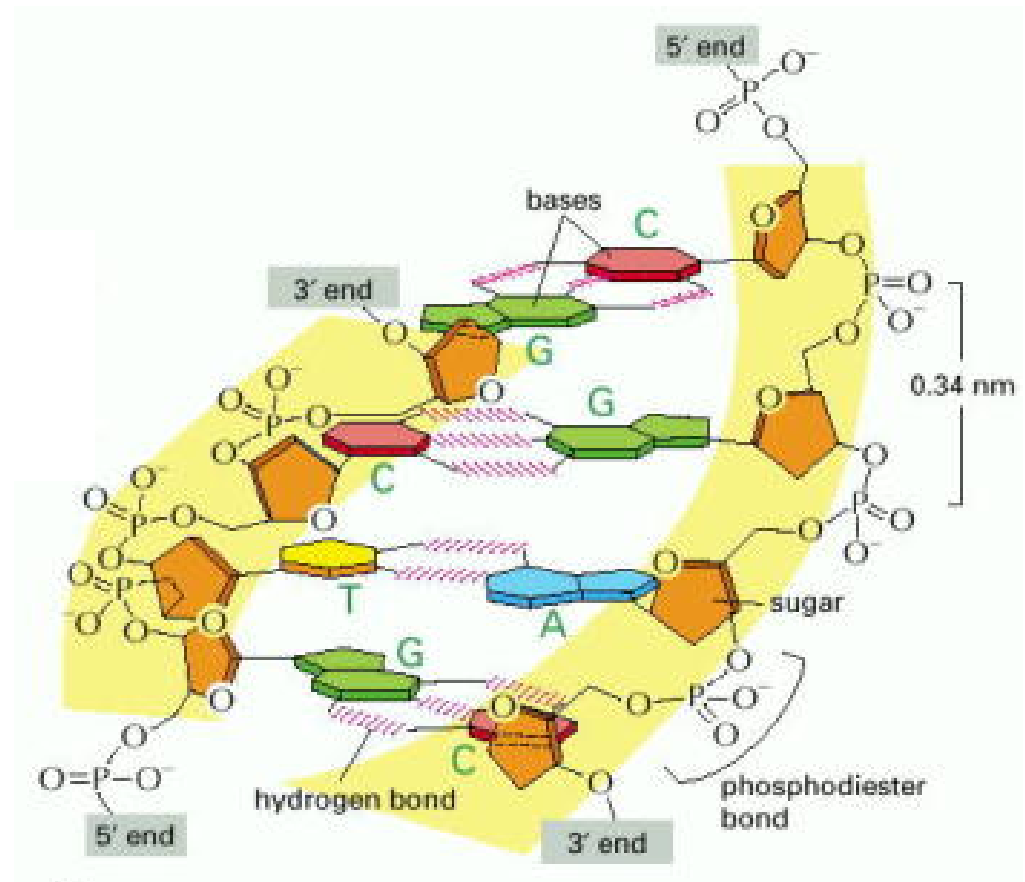
☞ The two strands of DNA have complementary sequence and antiparallel orientation.

☞ The two DNA strands are stabilized by two forces : H bondings and stacking interactions of bases which stack above each other in the double helix.

☞ The strand complementarity is due to complementarity of shape and hydrogen bonding (H bonding) between A and T, C and G. These pairings are called “**Watson-Crick pairings**”. There are 2 H bondings between A and T, and 3 H bondings between C and G.

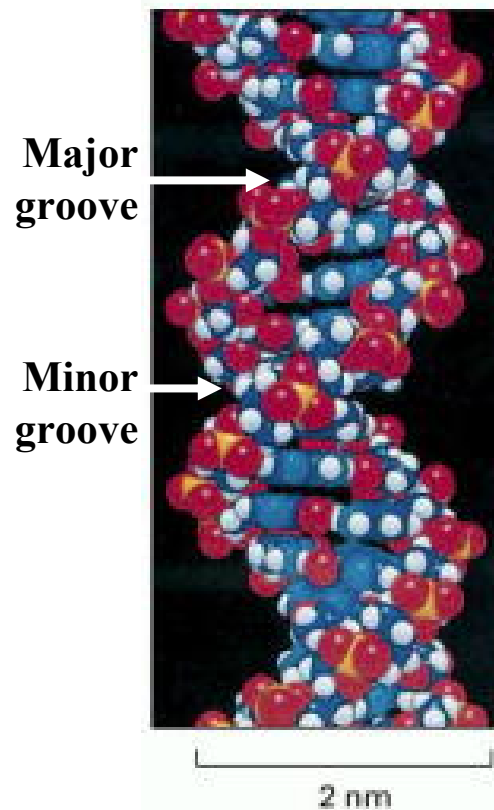
☞ The two strands of DNA, held by weak chemical bondings, can be separated from each other by heating or increasing pH → DNA is said to be **denatured**.

DNA denaturation is reversible → the two strands can reassociate.



“Copyright 2002 from *Molecular Biology of the Cell* by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC.”

## THE DNA DOUBLE HELIX (continued)



**B form**

∞ DNA is usually a right-handed helix and can exist in different conformations. B form is the major form, existing under normal physiological conditions, A form is present in some DNA:protein complexes. In high concentrations of positively charged ions, DNA can take Z form which is a left-handed helix and makes up a small proportion of DNA in the cell.

∞ Linear DNA molecules are found in eukaryotic chromosomes. Most of bacterial chromosomes, many viral DNA and plasmids (extrachromosomal DNA) are circular. Some DNA molecules are alternatively linear and circular, e.g  $\lambda$  phage DNA is linear in the virion particle, and is circular when injected into *E. coli* cell.

*“Copyright 2002 from Molecular Biology of the Cell by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC.”*

## THE DNA DOUBLE HELIX (continued)

☞ In circular double-stranded DNA, the two single strands are twisted around one another. To separate these two strands, one has to pass one strand through the other some many times. The number of times that a strand must be passed through the other for complete separation of the two strands is called the linking number ( $Lk$ ).

**Linking number** is the sum of twist and writhe. **Twist ( $Tw$ )** is the number of times a strand wraps around the other strand. **Writhe ( $Wr$ )** is the number of times the axis of the double helix crosses over itself in three dimensional space.

$$Lk = Tw + Wr$$

☞ When  $Lk < Lk^0$  and  $\Delta Lk < 0$ , DNA is negatively supercoiled

*$Lk^0$  : Linking number of closed circular double stranded DNA under physiological conditions*

☞ In the cell, DNA is mostly negatively supercoiled. Negative supercoiled DNA can be untwisted leading to partial unwinding of the double helix → the two strands can be easily separated ; this facilitates many processes including DNA replication and transcription.

# RNA

☞ RNA differs from DNA at three points :

1. Pentose sugar is a ribose instead of a deoxyribose
2. Uracil replaces Thymine
3. RNA is a single-stranded polynucleotide chain

☞ The single-stranded RNA is usually folded back on itself to form secondary and complex tertiary structures.

☞ The principal roles of RNA :

1. An intermediate to transfer genetic information from the genome to the protein synthesis machinery in the cytoplasm : **mRNA** (message RNA)
2. An adaptor between codon in the polynucleotide chain and amino acid in the polypeptide chain : **tRNA** (transfer RNA)
3. A structural component as well as an enzyme catalyzing the formation of peptide bonds : **rRNA** (ribosomal RNA).

***HOW ARE THE BIOLOGICAL MACROMOLECULES ASSOCIATED  
TOGETHER TO FORM HIGHER STRUCTURES ?***

**ROLE OF WEAK CHEMICAL BONDS IN  
LIVING SYSTEMS**



# WEAK CHEMICAL (NONCOVALENT) BONDS

**Chemical bonds** : forces which hold atoms together to form molecules

There are two types of chemical bonds : (1) **covalent bonds**, such as peptide bonds or phosphodiester bonds, respectively form the polypeptide and polynucleotide chains, and (2) **noncovalent (weak) bonds**.

Weak bondings differ from covalent bondings in many characteristics :

1. **Bond strength** : 1 – 7 kcal/mol for weak bonds compared to 90 kcal/mol for covalent bonds. Weak bonds are constantly made and broken in physiological conditions.
2. **Distance between two atoms engaged in a bonding** : atoms linked by covalent bonds are closer to each other than those connected by weak bonding. Typical noncovalent bonds have a length two- to three-times larger than that of covalent bonds.
3. **Maximum number of bonds a single atom can make is called its valence**, e.g Oxygen has a valence of 2 which means it can be engaged maximally in two bondings. Covalent bondings strictly depend on the valence of atoms participating in the linkage. Some of weak bondings, such as van der Waals bonding, are only limited by the surface getting in contact between the two participants.
4. **Bond angle formed by two covalent bonds from a single atom always have a defined value**, e.g the bond angle formed by the two bonds between an oxygen and two hydrogen atoms is 104.45 °C. Weak chemical bonds usually do not have fixed bond angle.

# ***WEAK BONDS ARE NOT STRONG AND CONSTANT BUT ARE CRUCIAL FOR LIVING SYSTEMS***

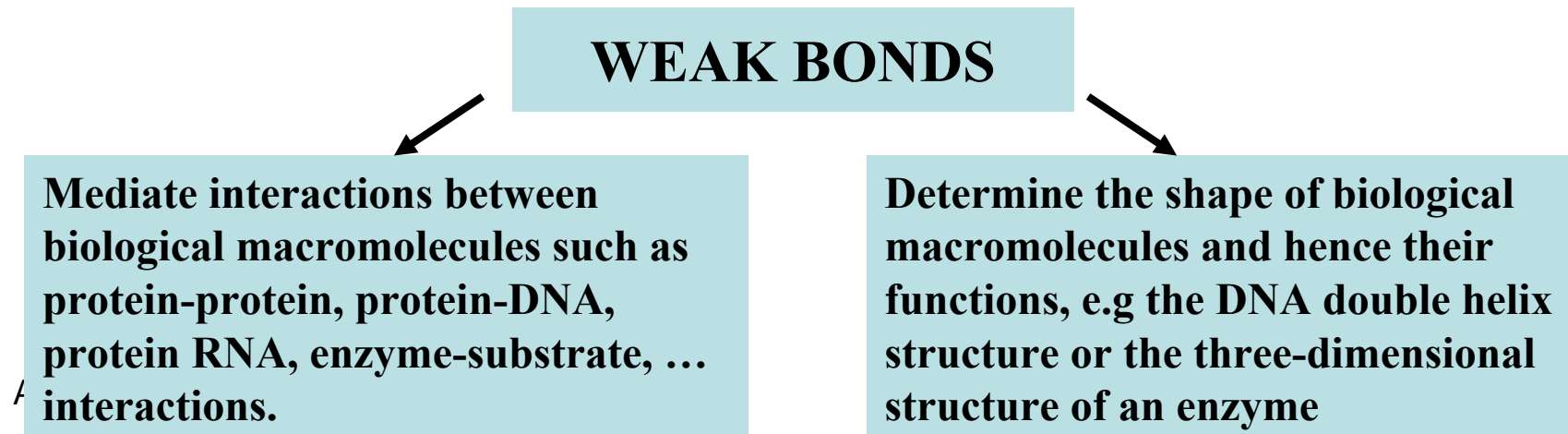
## ***WHY ?***

∞ Weak bonds are easily made and broken

→ They allow high diffusion rates which are essential for maintaining vital reactions of the cell. If all cell components are linked by covalent bonds, the cell would have the state of formaldehyde-fixed cell.

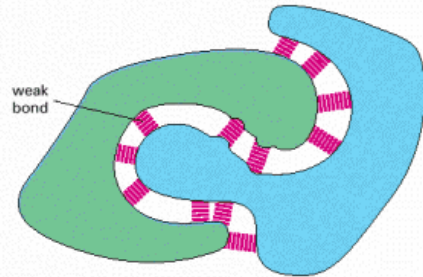
→ They participate in highly reversible reactions, e.g the denaturation and renaturation of double-stranded DNA ; or the interaction between enzymes and their substrates.

∞ Weak bonds are significant when interacting surfaces between components are structurally complementary such as lock-and-key model → weak bonds determine the specificity of interaction between components, e. g the antigen/antibody interaction.



### WEAK CHEMICAL BONDS

Organic molecules can interact with other molecules through short-range noncovalent forces.



Weak chemical bonds have less than 1/20 the strength of a strong covalent bond. They are strong enough to provide tight binding only when many of them are formed simultaneously.

### HYDROGEN BONDS

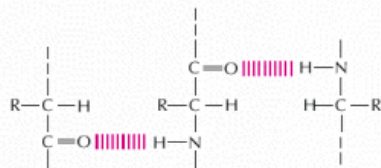
As already described for water (see Panel 2-2) **hydrogen bonds** form when a hydrogen atom is "sandwiched" between two electron-attracting atoms (usually oxygen or nitrogen).

Hydrogen bonds are strongest when the three atoms are in a straight line:

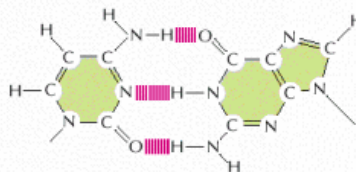


Examples in macromolecules:

Amino acids in polypeptide chains hydrogen-bonded together.

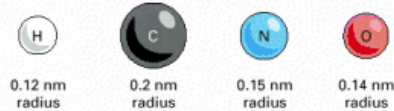


Two bases, G and C, hydrogen-bonded in DNA or RNA.



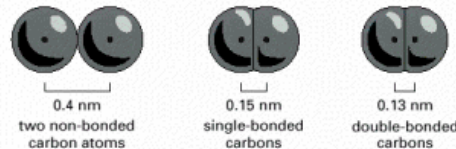
### VAN DER WAALS ATTRACTIONS

If two atoms are too close together they repel each other very strongly. For this reason, an atom can often be treated as a sphere with a fixed radius. The characteristic "size" for each atom is specified by a unique **van der Waals radius**. The contact distance between any two non-covalently bonded atoms is the sum of their van der Waals radii.



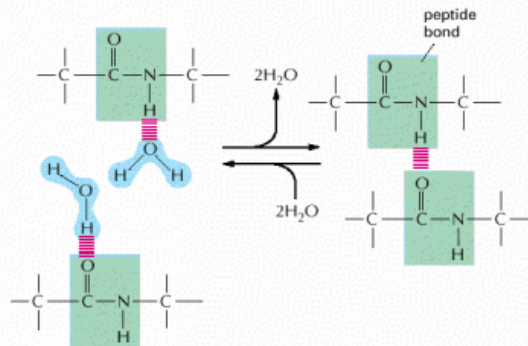
At very short distances any two atoms show a weak bonding interaction due to their fluctuating electrical charges. The two atoms will be attracted to each other in this way until the distance between their nuclei is approximately equal to the sum of their van der Waals radii. Although they are individually very weak, **van der Waals attractions** can become important when two macromolecular surfaces fit very close together, because many atoms are involved.

Note that when two atoms form a covalent bond, the centers of the two atoms (the two atomic nuclei) are much closer together than the sum of the two van der Waals radii. Thus,



### HYDROGEN BONDS IN WATER

Any molecules that can form hydrogen bonds to each other can alternatively form hydrogen bonds to water molecules. Because of this competition with water molecules, the hydrogen bonds formed between two molecules dissolved in water are relatively weak.



## TYPES OF WEAK BONDS – VAN DER WAALS BONDS

The main types of weak bonds in living organisms are : van der Waals bonds, ionic/hydrogen bonds and hydrophobic bonds.

∞ **Van der Waals bonding** is formed between two nonpolar atoms. It has a force of 1 kcal/mol. The distance between the two components is the sum of their van der Waals radii.

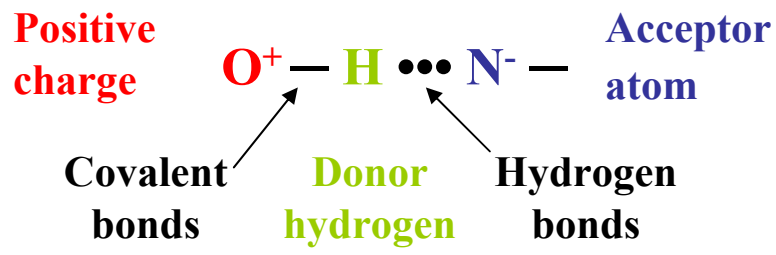
*When two nonpolar atoms come close to each other enough, they are simultaneously affected by attractive and repulsive forces. Their van der Waals radii result from the balance between these two forces.*

"Copyright 2002 from *Molecular Biology of the Cell* by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC."

# TYPES OF WEAK BONDS – *HYDROGEN BONDS*

∞ **Van der Waals bonds** are the weakest bonds and are effective only when the number of bonds established between the two participants are important which means when there is a complementarity of surface structures between them , e.g antigen-antibody interaction → van der Waals bonds play an important role in determining the specificity of some interactions.

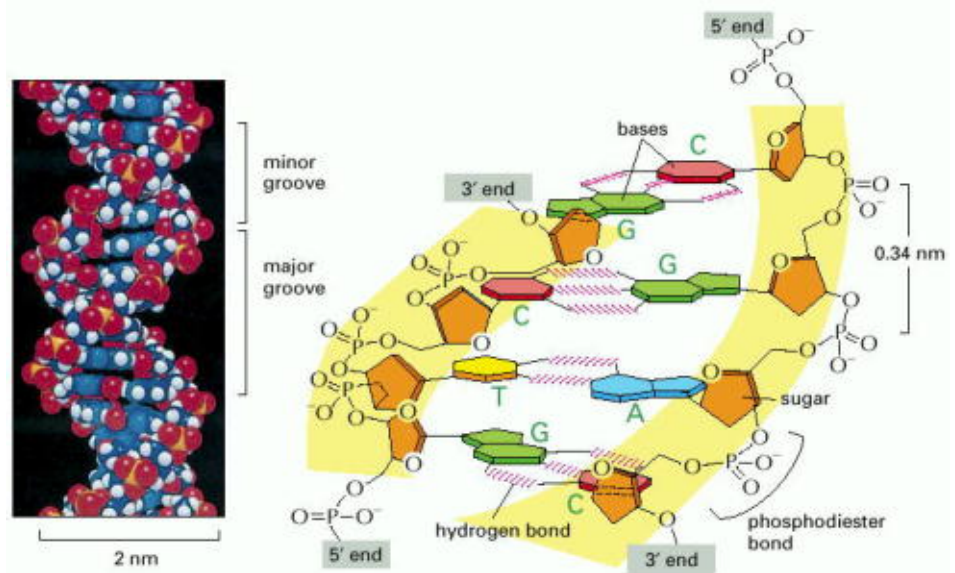
∞ **Hydrogen bonds (H bonds)** are formed between a donor hydrogen atom covalently linked to a positive charge and a negatively charged acceptor atom.



H bonds are directional bonds with the positive charge, donor hydrogen and the acceptor atom ranging on a straight line.

H bonds are stronger than van der Waals bond with strength range between 12 and 30 kcal/mol (in a nonaqueous environment).

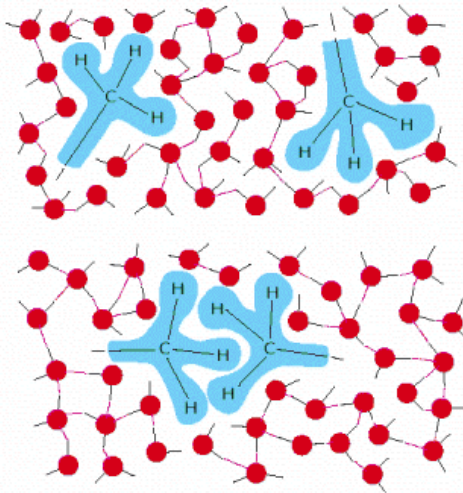
H bonds have important roles in structure formation of macromolecules as well as in the interactions between them.



“Copyright 2002 from *Molecular Biology of the Cell* by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC.”

*H bonds between A and T, C and G maintain the DNA double helix*

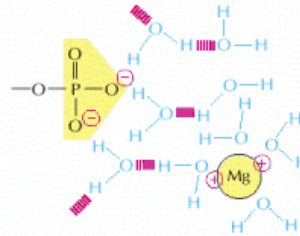
## HYDROPHOBIC FORCES



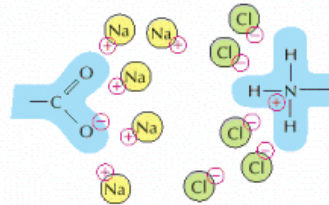
Water forces hydrophobic groups together, because doing so minimizes their disruptive effects on the hydrogen-bonded water network. Hydrophobic groups held together in this way are sometimes said to be held together by “hydrophobic bonds,” even though the attraction is actually caused by a repulsion from the water.

## IONIC BONDS IN AQUEOUS SOLUTIONS

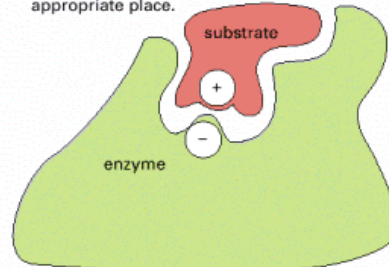
Charged groups are shielded by their interactions with water molecules. Ionic bonds are therefore quite weak in water.



Similarly, other ions in solution can cluster around charged groups and further weaken ionic bonds.



Despite being weakened by water and salt, ionic bonds are very important in biological systems; an enzyme that binds a positively charged substrate will often have a negatively charged amino acid side chain at the appropriate place.



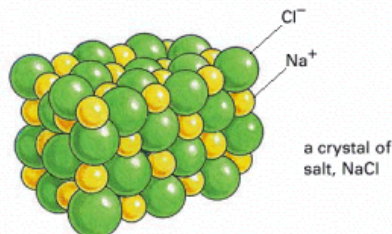
## IONIC BONDS

Ionic interactions occur either between fully charged groups (ionic bond) or between partially charged groups.



The force of attraction between the two charges,  $\delta^+$  and  $\delta^-$ , falls off rapidly as the distance between the charges increases.

In the absence of water, ionic forces are very strong. They are responsible for the strength of such minerals as marble and agate.



# TYPES OF WEAK BONDS - IONIC BONDS

⌘ **Ionic bonds** are formed between two oppositely charged chemical groups. Bond energy is about 5 kcal/mol.

Some ionic bonds with hydrogen donor are also hydrogen bonds, e.g



In aqueous environment, ionic bondings do not affect the shape of biological macromolecules because all charged groups of these molecules are entirely surrounded by a layer of water molecules and are engaged in hydrogen bonding with these molecules.

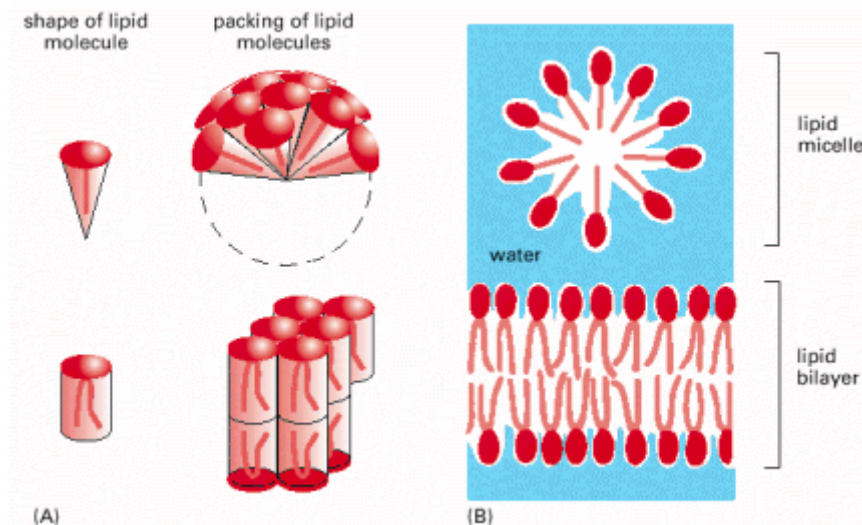
“Copyright 2002 from *Molecular Biology of the Cell* by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC.”

# TYPES OF WEAK BONDS – *HYDROPHOBIC INTERACTION*

☞ **Hydrophobic interactions** are tendency of nonpolar groups to arrange themselves to maximally limiting their contact with water rather than “real” bonds. The real bonding between nonpolar groups is van der Waals bonding.

Water molecules tend to exclude nonpolar molecules from water network and hence strengthen the binding of these nonpolar groups to each other. The binding energy is thus 2 to 3 kcal/mol greater in aqueous than in nonaqueous environment.

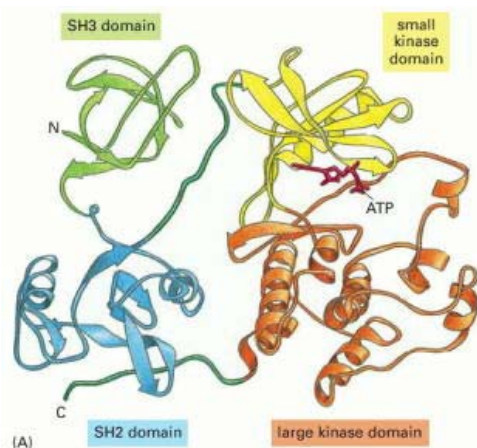
Hydrophobic interactions are crucial in the stabilization of biological macromolecules. The hydrophobic bondings between nonpolar chains of phospholipid molecules stabilize biological membranes. The folding of some proteins to avoid the contact of their hydrophobic segments with surrounding water molecules also stabilizes protein structure.



*Hydrophobic bonds between nonpolar chains of phospholipids molecules stabilize the lipid bilayer forming biological membranes.*

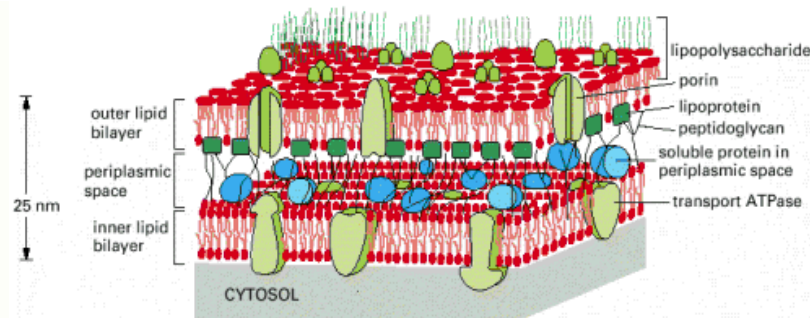
# FUNCTIONS OF WEAK BONDS

Determining the shape of macromolecules



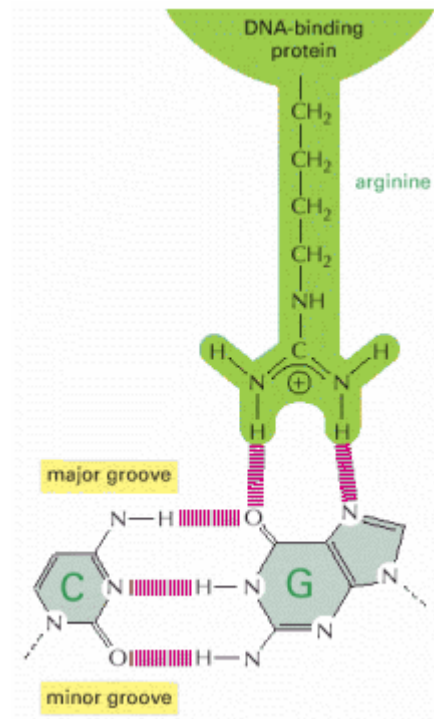
*The three-dimensional structure of a protein is determined by weak bonds between different segments of the molecule*

Forming large macromolecular assemblies



*Biological membranes are formed by the association of different types of membrane proteins, lipid bilayer and other components. These components are linked together by numerous weak bonds.*

Mediating interactions between macromolecules



*Protein-DNA interaction through H bonds between guanine and arginine*

April 2009

*“Copyright 2002 from Molecular Biology of the Cell by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC.”*

# **ORGANIZATION OF PROKARYOTIC AND EUKARYOTIC GENOMES**



# CHROMATIN, CHROMOSOME

☞ In the cell, the DNA-protein complex is called **chromatin**. Depending on the degree of compaction, chromatin can be classified as “**euchromatin**” or “**heterochromatin**”, named on their degree of staining with basic dyes - slightly (euchromatin) or darkly (heterochromatin).

Most part of chromatin is euchromatin. Euchromatin is much more less compacted than heterochromatin and is the genetically active part of the genome, mainly composed of actively transcribed genes. Constitutive heterochromatin is essentially located at the telomeres and centromeres, some is dispersed all over the chromosomes. This dispersed heterochromatin is mostly composed of repetitive sequences. Non-constitutive heterochromatin originates from temporally non transcribed euchromatin.

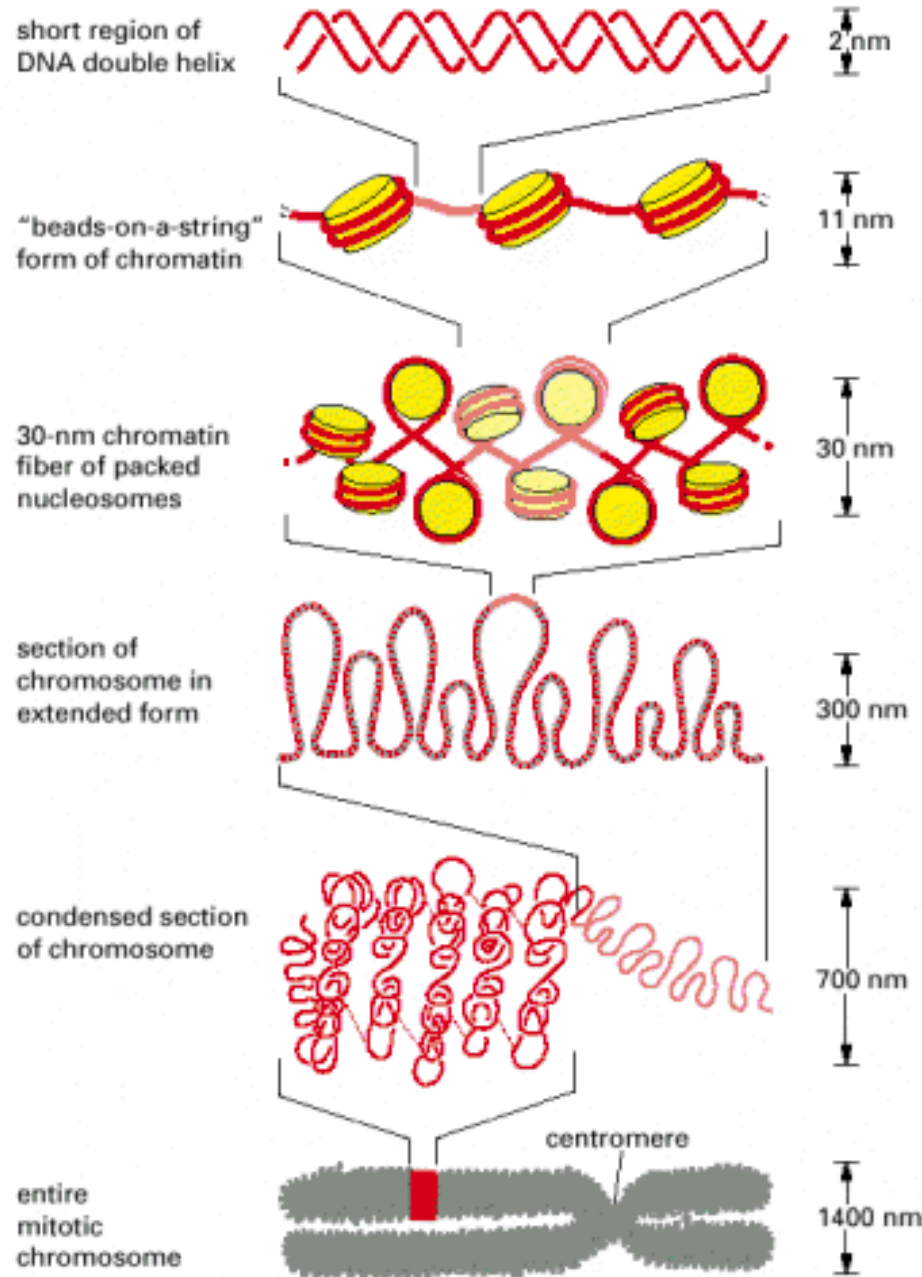
The genome of prokaryote usually exists as a single DNA molecule whereas eukaryotic genome is distributed into many DNA molecules ; each DNA molecule associated with its proteins is called a **chromosome**.

☞ The role of associated proteins is to **compact** DNA molecule in order to :

1. Make long DNA molecule fitted in a small space which is the cell (prokaryote), or the nucleus (eukaryote). *E. coli* chromosome measuring 1 mm must fit in a 1  $\mu\text{m}$ -length cell. In human, a haploid chromosome set is constituted of  $3 \cdot 10^9$  bp  $\rightarrow$  a diploid set has  $2 \times 3 \cdot 10^9$  bp, each bp has a thickness of  $3.4 \text{ \AA}$  ( $1 \text{ \AA} = 10^{-10} \text{ m}$ )  $\rightarrow$  in a somatic cell, the nucleus contains a 2m-length DNA !
2. Protect the long DNA molecule from damages.
3. Facilitate the replication and the recombination of DNA as a whole entity during cell division.
4. Facilitate gene expression through selectively decompacting individual regions in the genome.

☞ The majority of eukaryotic cells are **diploid**, which means that they contain 2 copies of each chromosome, each coming from one parent. **Haploid** cells (egg or spermatozoid) are those containing only one set of each chromosome.

# COMPACTING LEVEL OF GENOMIC DNA



NET RESULT: EACH DNA MOLECULE HAS BEEN PACKAGED INTO A MITOTIC CHROMOSOME THAT IS 10,000-FOLD SHORTER THAN ITS EXTENDED LENGTH

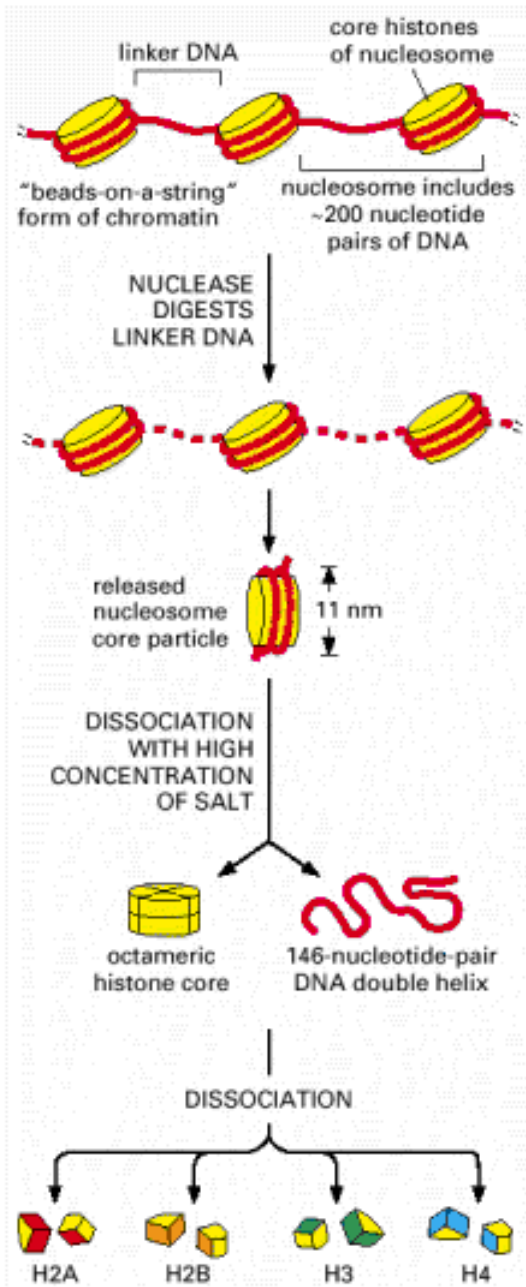
Prokaryotic DNA has a low degree of compacting. The most abundant DNA-associated proteins in prokaryotes are HU and H-NS proteins. These proteins play many roles : condensing DNA, regulating DNA replication and expression, ..

*E. coli* chromosome is constituted of many DNA domains (loops) with the end of each loop binding to a protein-membrane scaffold.

Eukaryotic DNA which is 100 to 1000 times longer than prokaryotic DNA is highly compacted. Compaction is achieved at different levels :

Nucleosome → 30-nm fiber → Higher ordered structures : chromosomal loop, condensed mitotic chromosome

# NUCLEOSOME : THE BUILDING UNIT OF EUKARYOTIC CHROMOSOME

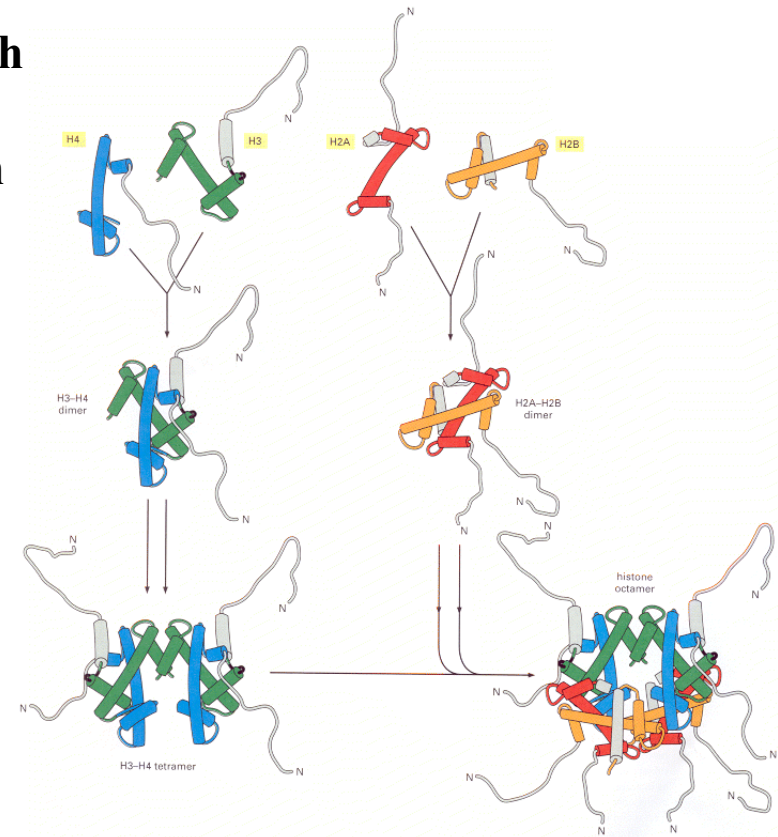


When DNA is digested with nucleases, **nucleosomes** are released from the chromatin

Each nucleosome core consists of 8 histones including 2 H2A, 2H2B, 2 H3, 2 H4.

The length of DNA fragment wrapping 1.65 times around a nucleosome core is 146 bp

Nucleosomes are linked by **linker DNA**.



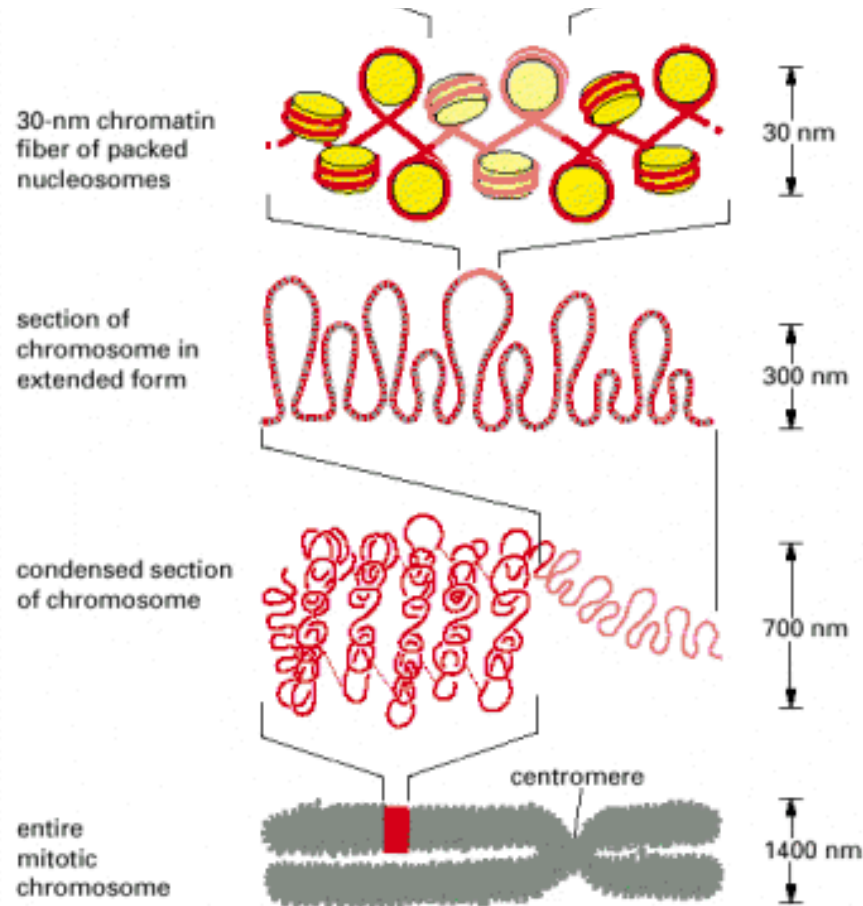
*“Copyright 2002 from Molecular Biology of the Cell by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC.”*

Each nucleosome core is formed by the association of 2 H3-H4 dimers → H3-H4 tetramer. The adding of 2 H2A-H2B dimers to the H3-H4 tetramer scaffold complete the assembly

*“Copyright 2002 from Molecular Biology of the Cell by Alberts et al. Reproduced by permission of Garland Science/Taylor & Francis LLC.”*

### ☞ The 30-nm fiber

The next step in DNA packaging is the binding of histone H1. H1 binds simultaneously to linker DNA and a site at the middle of the DNA fragment wrapping around nucleosome core → H1 binding tightens DNA wrapping around nucleosome and compacts DNA more.



NET RESULT: EACH DNA MOLECULE HAS BEEN PACKAGED INTO A MITOTIC CHROMOSOME THAT IS 10,000-FOLD SHORTER THAN ITS EXTENDED LENGTH

## HIGHER COMPACTING LEVEL OF EUKARYOTIC DNA

In *in vitro* experiments, increasing salt concentration and adding histone H1 lead to the formation of 30-nm fiber.

Two models explain this structure : solenoid and zigzag models. Solenoid is a superhelix constituted of six nucleosomes per turn. In the “zigzag” model (*Figure*), linker DNAs cross the central axis of the 30-nm fiber at each turn.

☞ Higher compacting level : **Loops** formed by the 30-nm fiber. These loops, up to 100 kb, are held at their base by a protein complex called nuclear matrix. Loops form an array of about 300 nm (*Figure*).

☞ The most condensed state of DNA is **mitotic chromosome** (*Figure*)

***WHY IS THE PROKARYOTIC GENOME LESS COMPACT THAN THE  
EUKARYOTIC GENOME ?***

# GENOME CONTENT

☞ Genome size is related to the complexity of the organism because a complex organism may need more genes participating in the formation of different parts of its organism.

Genome size : 0.58 Mb (1Mb =  $1.10^6$  nucleotides) – 6.7 Mb in prokaryotes ; 12 Mb – 12,000 Mb in eukaryotes.

Number of genes : about 500 – 6,000 genes in prokaryotes ; about 6,000 - 45,000 genes in eukaryotes.

Nevertheless, the correlation between genome size/number of genes with the structural complexity of an organism is not perfect ; e.g in human there are 27,000 genes whereas maize has > 45,000 genes

☞ The different ratio of the number of genes/ genome size between prokaryotes and eukaryotes is due to the fact that prokaryotic genome is nearly entirely composed of coding sequences (genes) whereas eukaryotic genomes contain a very small percentage of coding sequences (about 1%) with the remaining constituted of non-coding sequences.

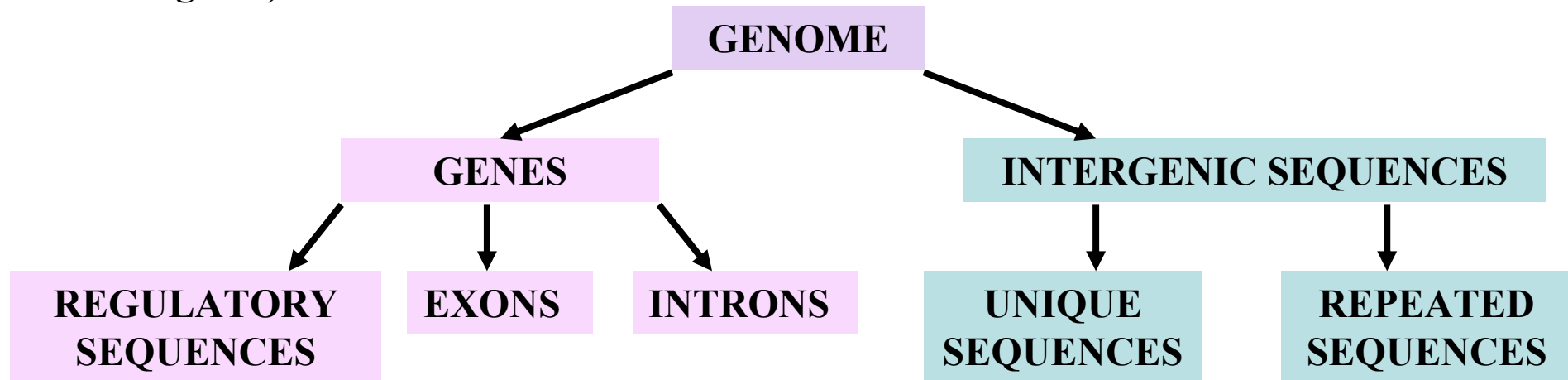
→ **Gene density** : number of genes / genome size (genes/Mb) ratio

Organisms with increasing structural complexity usually have decreasing gene density ;

e.g gene density in *E. coli* (prokaryote), *S. cerevisiae* (unicellular eukaryote) and human (eukaryote) is respectively 950, 480 and 9.3.

# EXAMPLE OF HUMAN GENOME STRUCTURE

The human genome is composed of **genes** and **intergenic sequences** (DNA sequences lying between genes) :



Except of a small percentage (1%) of coding sequences (exons), the human genome is mostly composed of **non-coding sequences** including :

☞ **Introns** : interspersed non-coding sequences found inside each gene which make up 95% of the gene ; the remaining 5% is made up of exons

☞ **Regulatory sequences** : sequences controlling gene expression (promoters, enhancers, silencers, ...)

☞ **Unique intergenic sequences** are composed of mutant genes, gene fragments, pseudogenes, ... which are non functional copies of gene

☞ **Repeated intergenic sequences** : microsatellite DNAs which are very short sequences (< 13bp) and dispersed repeats which consist of transposable elements

# SUMMARY

☞ Prokaryotic and eukaryotic cells are made of biological macromolecules which are formed by chemically similar monomers.

☞ Biological macromolecules include : polysaccharides, lipids, proteins and nucleic acids

Proteins are composed of 20 types of amino acids and have four structure level from primary structure to quaternary structure. Primary structure is a chain of amino acids linked together by peptide bonds, a covalent bond. Secondary, tertiary and quaternary structures are determined by weak bonds. Functions of proteins are determined by their three-dimensional structure.

Nucleic acids include DNA and RNA. They are chains formed by nucleotides linked together by a covalent bond, the phosphodiester bond. In the cell, DNA associates with specific proteins to form chromosomes. Eukaryotic chromosomes have different level of compacting. Eukaryotes have a cell density less important than that of prokaryotes due to the presence of different types of non-coding sequences.

☞ The shape as well as the function of biological macromolecules are determined by weak bonds. Weak bonds also participate in the formation of large macromolecule assemblies. Finally, weak bonds control the interactions between different macromolecules

Weak bonds include : van der Waals bonds, hydrogen bonds, ionic and hydrophobic bonds