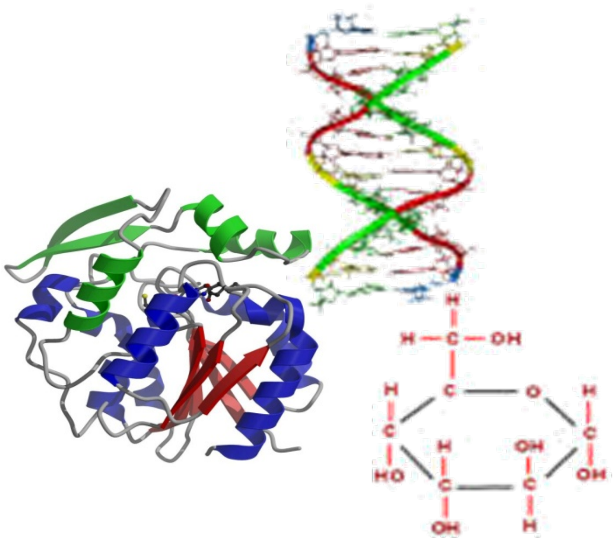
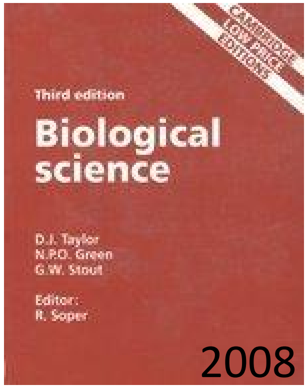
Kusuma School of Biological Sciences

SBL-100
Introductory Biology for Engineers
L-2

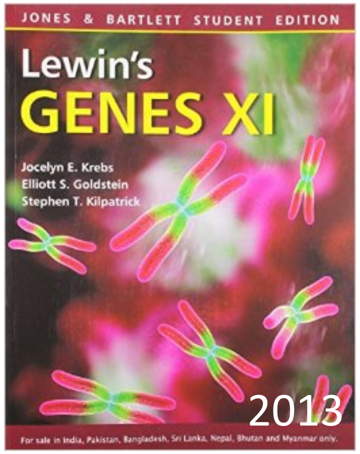
What's happening inside the cells? The Molecules of Life





An introductory biology text book

Disclaimer. **These slides are compiled from diverse textbooks and resources in public domain, exclusively for teaching purposes . The author does not claim copyrights.**



The chemical composition of any cell – be it animal or plant reveals a variety of inorganic & organic compounds essential to the establishment and maintenance of the living state.

Inorganic precursors (Carbon dioxide, water, ammonia, nitrate, nitrogen, etc.)

↓

Metabolites (pyruvate, succinate, acetyl-CoA etc.)

↓

Building blocks (Amino acids, Sugars, Nucleotides, Fatty acids, Glycerol..)

↓

Biomolecules (Proteins, Nucleic Acids, Polysaccharides, Lipids)

↓

Molecular Assemblies (Ribosomes, Cytoskeleton)

↓

Organelle (Nucleus, Mitochondria, Golgi, etc.)

↓

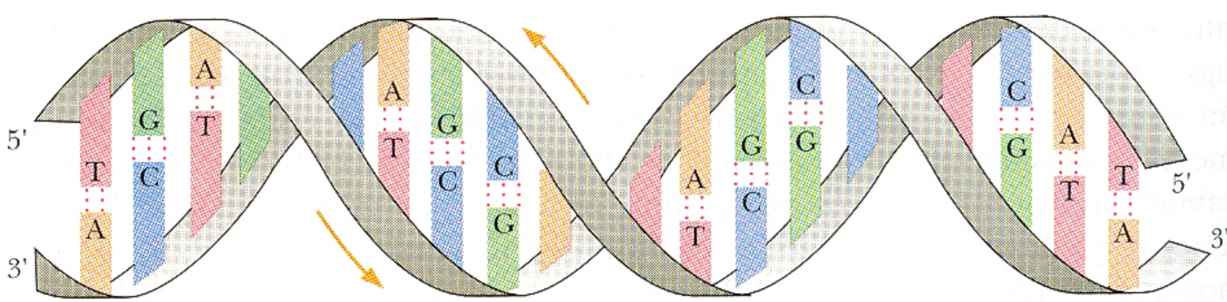
Cell → **Tissues** → **Organs** → **Living organisms**

Molecular organization in the cell is a hierarchy

Biomolecule / Function / Special Characteristics

DNA	Information storage	Genetic material
RNA	Messenger of information	
Protein	Expressed information	Cell growth, metabolic capability & morphology
Carbohydrates		Energy fuel
Lipids		Energy suppliers

Nucleic acids are polymers of nucleotides which in turn consist of a nitrogenous base, a pentose sugar and a phosphate. The sugar is a ribose in RNAs and deoxyribose in DNA. Unlike DNA, RNAs are normally single stranded. Certain forms and regions of RNA do assume complex shapes notably transfer RNA (tRNA), ribosomal RNA (rRNA). Another form of RNA is the messenger RNA (mRNA). All these RNAs are involved in protein synthesis.

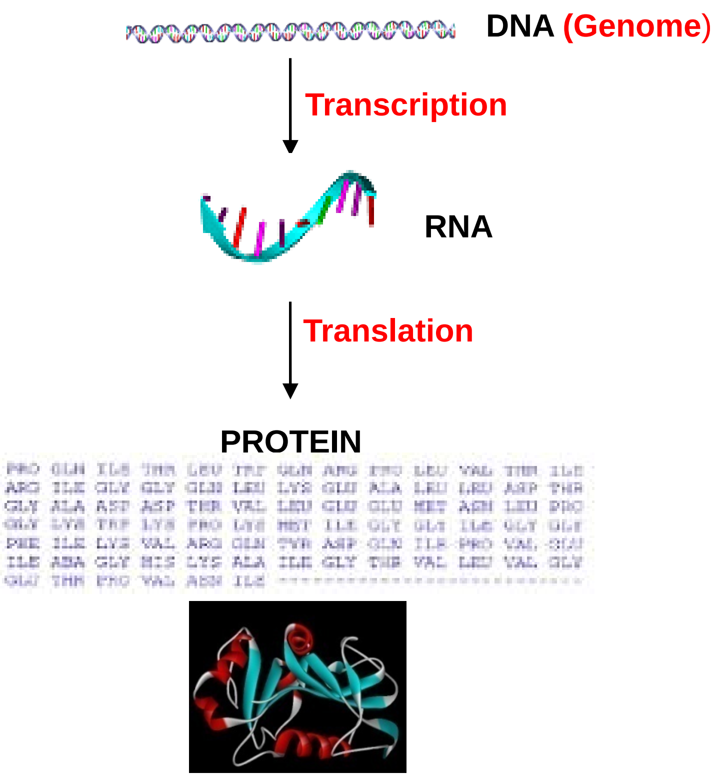


DNA double helix

***DNA, RNA sequences have a directionality**. They are read and written in the 5' → 3' direction unless specified otherwise. Protein sequences have N → C directionality. More about these where covalent connectivities are discussed.*

Kusuma School of Biological Sciences

Central Dogma of Life…



Central dogma of modern biology

Proteins (polymers made of amino acids) are classified as :

(i) **Structural proteins**. Collagen is a component of connective tissue, bone, tendons, cartilage. Alpha keratin is a component of skin, feathers, nails, hair, horn. Elastin is associated with elastic connective tissues such as ligaments and muco-proteins with mucous secretions, viral coat proteins to wrap nucleic acid of viruses.

(ii) **Enzymes**. These are proteins with catalytic ability. Trypsin catalyzes hydrolysis of protein, glutamine synthetase catalyzes synthesis of glutamine from glutamic acid and ammonia.

(iii) **Hormones**. Insulin and glucagon help to regulate glucose metabolism. ACTH stimulates growth and activity of adrenal cortex.

(iv) **Transport proteins**. Haemoglobin transports oxygen in vertebrate blood. Haemocyanin transports oxygen in some non-vertebrates. Myoglobin transports oxygen in muscles.

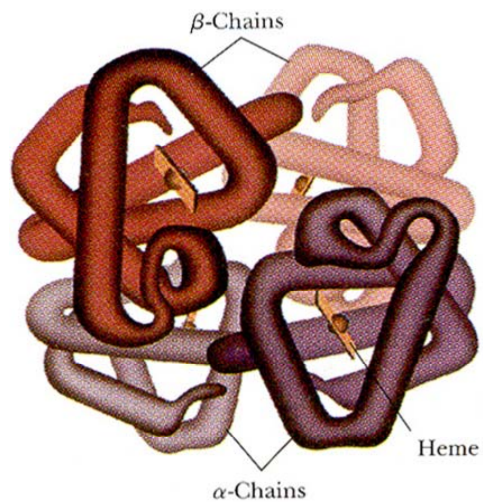
(v) **Protective proteins**. Antibodies form complexes with foreign proteins. Fibrinogen is a precursor of fibrin in blood clotting. Thrombin is involved in clotting mechanism.

(vi) **Contractile proteins**. Myosin is involved in moving filaments in myofibril of sarcomere and Actin in stationary filaments in myofibrilin sarcomere.

(vii) **Storage proteins**. Some examples are ovalabumin in egg white, casein in milk.

(viii) **Toxins**. Snake venom functions as an enzyme. Diphtheria toxin is made by the diphtheria bacteria.

VLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLSFPTTKTYFPHFDLSHGSAQVKGHGK……
……………………………………………………………….LHAHKLRVDPVNFKLLSHCLLVTLAAH
LPAEFTPAVHASLDKFLASVSTVLTSKYR



Hemoglobin

Sequence dictates structure &

Structure dictates function

Amino acid sequence to quarternary structure of a protein

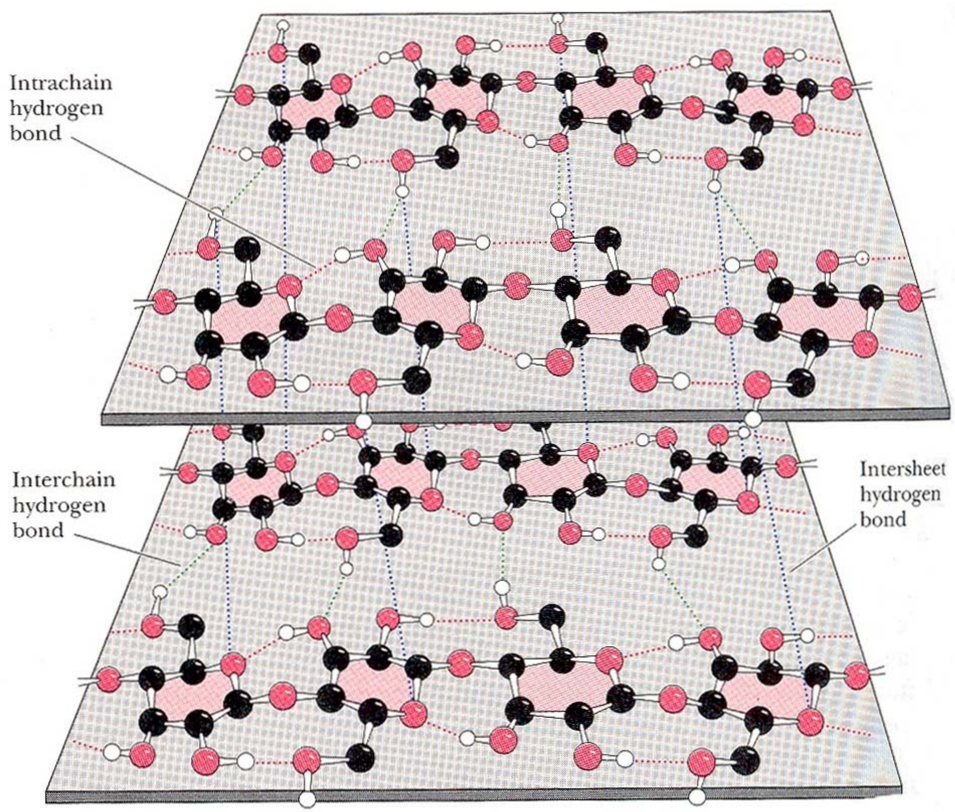
Carbohydrates. These are substances with general formula $C_x(H_2O)_y$ and hence their name. All carbohydrates are aldehydes (-CHO) or ketones (>C=O) and contain several hydroxyl (-OH) groups. Interactions and functions of carbohydrates are determined by these groups. Carbohydrates are divided into three classes, mono-, di- and polysaccharides.

(i) **Monosaccharides**. (a) Trioses $C_3H_6O_3$. Some examples are glyceraldehydes and dihydroxyacetone which are intermediates in respiration, photosynthesis. (b) Pentoses $C_5H_{10}O_5$. Some examples are ribose and ribulose which are involved in synthesis of nucleic acids. Ribose is a constituent of RNA, deoxyribose of DNA. Pentoses are involved in synthesis of some coenzymes eg. NAD, NADP, coenzyme A, FAD, FMN, synthesis of AMP, ADP and ATP, synthesis of polysaccharides. Ribulose biphosphate is the CO_2 acceptor in photosynthesis. (c) Hexoses $C_6H_{12}O_6$. Some examples are glucose, fructose, galactose and mannose. These act as sources of energy when oxidized in respiration. These are involved in synthesis of di, oligo and poly saccharides. Derivatives of monosaccharides include glycerol used in lipid synthesis. Vitamin C (ascorbic acid) is derived from a hexose. Deoxyribose is used in DNA synthesis. Galactoseamine is used in synthesis of cartilage.

(ii) **Disaccharides**. Most common disaccharides are maltose (glucose + glucose), lactose (glucose + galactose) and sucrose (glucose + fructose).
Maltose is formed by the action of amylases (enzymes) on starch during digestion. It is then converted to glucose. Lactose is found exclusively in milk. Sucrose or cane sugar is most abundant in plants. Sucrose is non-reducing and relatively inert metabolically.

(iii) **Polysaccharides** (eg. starch, glycogen, cellulose) are all polymers of glucose. They function chiefly as food and energy stores because of their size, their insolubility in water which means no osmotic or chemical influence in the cell, their compact shapes and their easy conversion into sugars by hydrolysis when required. Starch is a major fuel store in plants. Glycogen is the animal equivalent of starch. Glycogen is chiefly stored in the liver and muscles. About 50% of carbon found in plants is in cellulose and it is the most abundant organic molecule on Earth.

Carbohydrates



Structure of a polysaccharide (cellulose)

Lipids. Lipids are esters of fatty acid (containing a –COOH group and mostly 16 or 18 carbon atoms) and an alcohol (containing an -OH group) mostly glycerol.

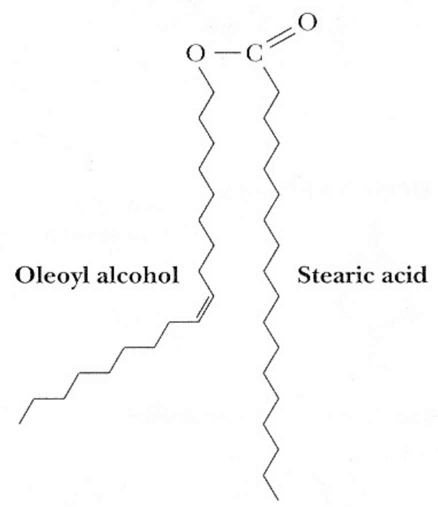
Fatty acids containing one or more double bonds (>C=C<) such as oleic acid are called unsaturated and those which lack a double bond are called saturated. Unsaturated fatty acids (olive oil) melt at lower temperatures than saturated ones (stearic acid, palmitic acid).

Glycerol has three OH groups all of which can react with a fatty acid molecule to form triglycerides which are the most common lipids. Triglycerides are insoluble in water and less dense than water and hence float.

A major function of lipids is to act as energy stores. Lipids have a higher calorific value than carbohydrates. Animals store extra fat when hibernating. Fat is also found below the dermis of the skin of mammals particularly in cold climates.

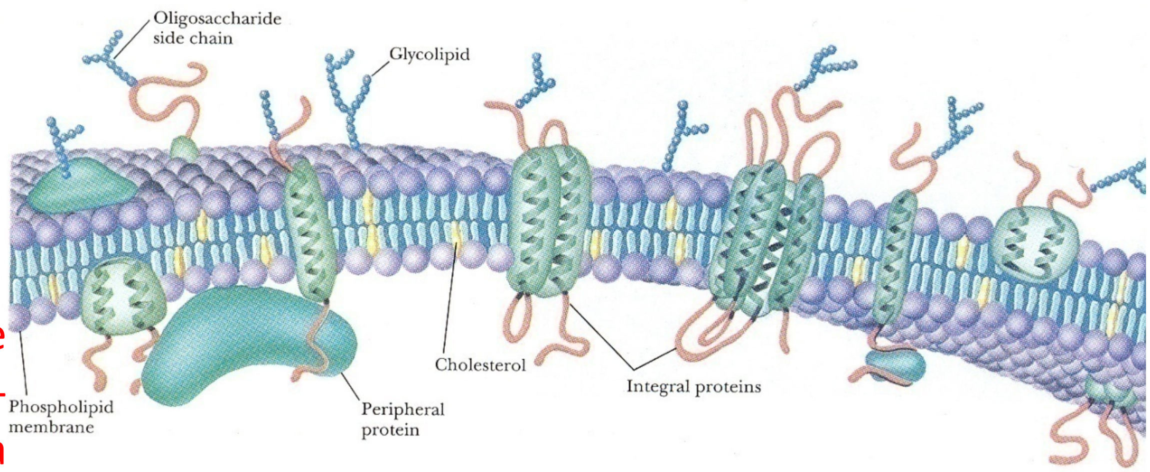
Plants store oils rather than fats. Seeds (coconut, castor bean, soyabean, sunflower, fruits, chloroplasts) are rich in oils.

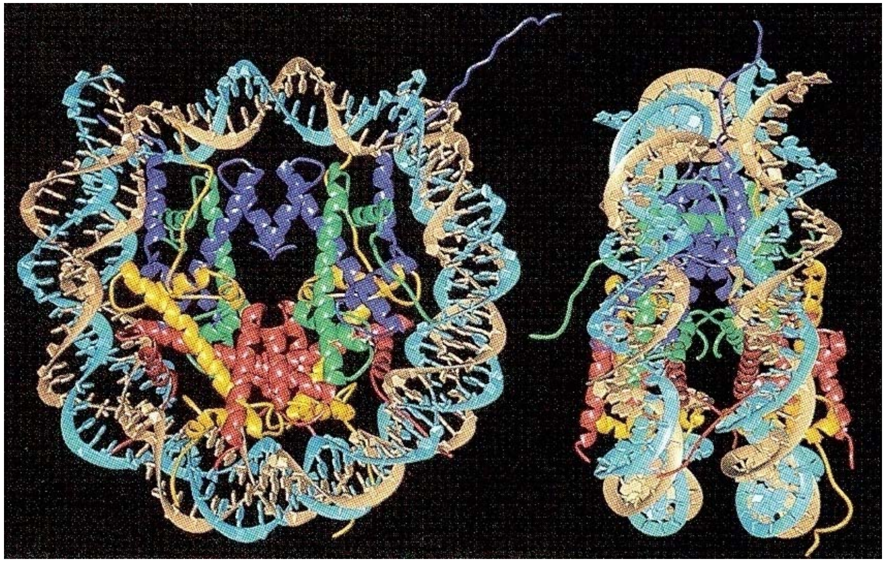
Solid triglycerides at 20^{o}C are classified as fats and liquids as oils. Waxes are esters of fatty acids with long chain alcohols.



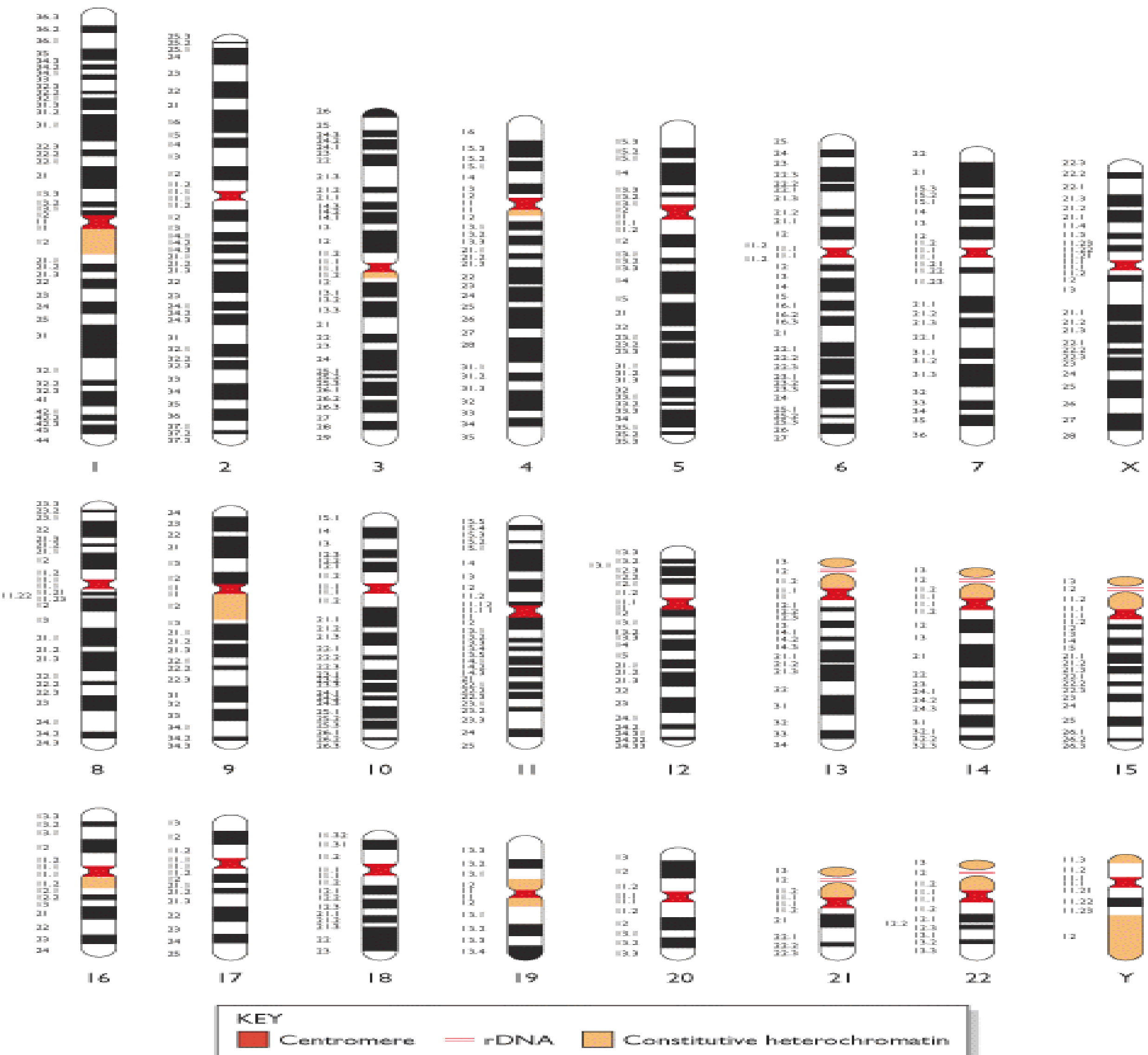
Structure of a lipid: Oleoyl alcohol esterified with stearic acid

Biomolecules interact with one another non-covalently in a majority of the cases or covalently in a few, to form macromolecular (supramolecular) complexes .





Representative supramolecular assemblies. (a) Membrane (b) Chromatin



The human karyogram. The chromosomes are shown with the G-banding pattern obtained after Giemsa staining. Chromosome numbers are given below and the band numbers to the left.

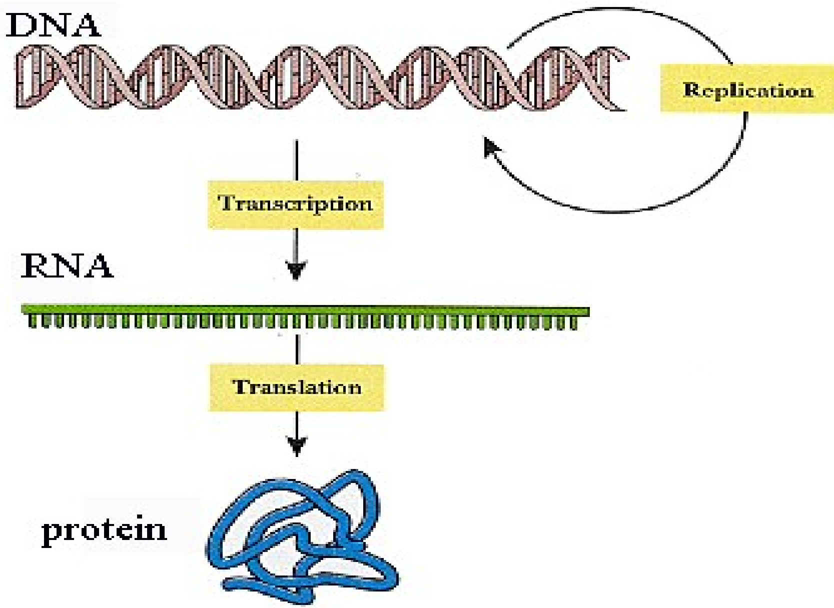
X: 196 MB (data size) (~ 5% of the genome);
Y: 25 MB

Organization of the human genome As a species, we all share in essence the same genome. At this collective level, the genome is the common heritage of humanity. In contrast, apart from identical twins, individuals exhibit significant variation. Each of the 30 trillion cells in the adult human body has its own copy or copies of the genome, the only exception being those of few types, such as red blood cells, that lack nuclei. The human genome is made up of two distinct components –

The nuclear genome comprises approximately 3 billion base pairs (bp) or 3000Mb (megabase pairs). The nuclear genome is divided into 24 linear DNA molecules, each contained in different chromosomes. These 24 chromosomes consist of 22 autosomes and the two sex chromosomes, X and Y. The vast majority of cells in human are diploid and so have two copies of each autosomes, plus two sex chromosomes, XX for females or XY for males -46 chromosomes in all. These cells are called somatic cells, in contrast to sex cells or gametes, which are haploid and have just 23 chromosomes.

The mitochondrial genome, a circular DNA molecule of 16.6 kb, many copies of which are located in the mitochondria. Unlike nuclear DNA, **mitochondrial DNA is only inherited from the maternal chromosomes**. This is because mitochondria are only found in the female gametes or egg cells, not in the male gamete or sperm. Large numbers of mitochondria are found in the tail of sperm, providing them with an engine that generates the energy needed. However, when the sperm enters the egg during fertilization, the tail falls off. Mitochondrial DNA also does not recombine, so there is no shuffling of genes as there is in nuclear genes.

What is happening inside the cells?



Essentials:

1. DNA is made of 4 bases:
   A, G, C, T.

Watson-Crick pairing states
A pairs with T(U) and T(U) with A.
G pairs with C and C with G.

2. Proteins are made of 20 Amino acids.

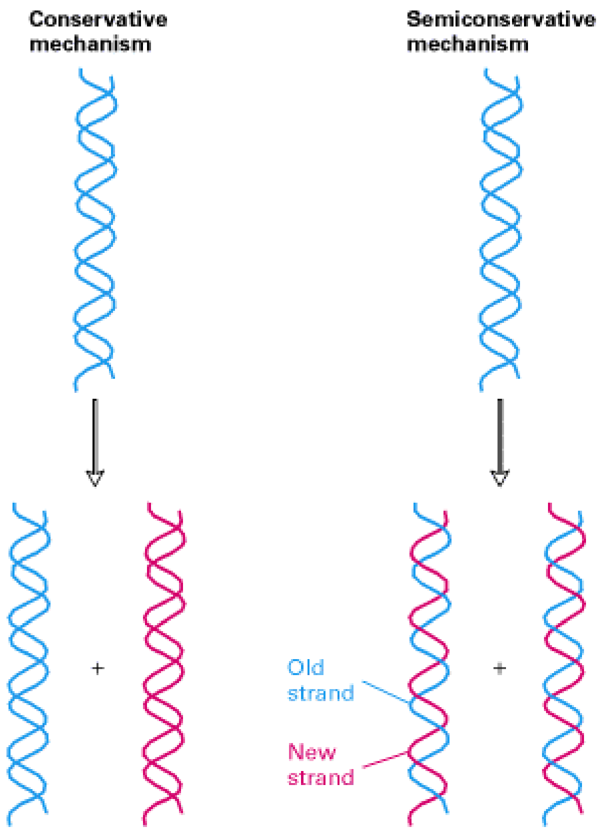
Genetic code maps the correspondence between bases and amino acids.

A depiction of gene expression (the central dogma), summarized as **DNA (gene) makes RNA & RNA makes proteins**, the two steps being called transcription and translation.

DNA carries genes which code for several types of RNAs such as mRNA, tRNA, rRNA, micro RNA etc.. Only mRNA gets converted into proteins.

RNA viruses pose an exception to central dogma in that RNA of virus gets converted to DNA within the host with the help of reverse transcriptase enzyme of the virus. The DNA of the virus then follows the central dogma using host cell machinery.

Replication



Error rate:
1 in 10 billion bases

Speed:
1000 bases /sec (Ecoli)
50 bases/sec
(Humans)

The Enzymes helping DNA Replication

DNA Helicases
DNA single-stranded binding proteins
DNA Gyrase

DNA Polymerase – (for elongation, proof reading & repair) **(Arthur Kornberg, Nobel in Physiology or Medicine, 1959)**

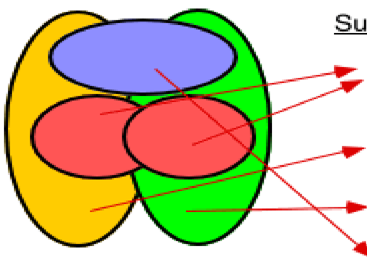
Primase
DNA Ligase

The first evidence supporting a semiconservative mechanism came from a classic experiment by Meselson & Stahl. Apparently all cellular DNA in both prokaryotic and eukaryotic cells is replicated by a semiconservative mechanism.

Kusuma School of Biological Sciences

Transcription

Prokaryotic RNA Polymerase: Holoenzyme Enzyme

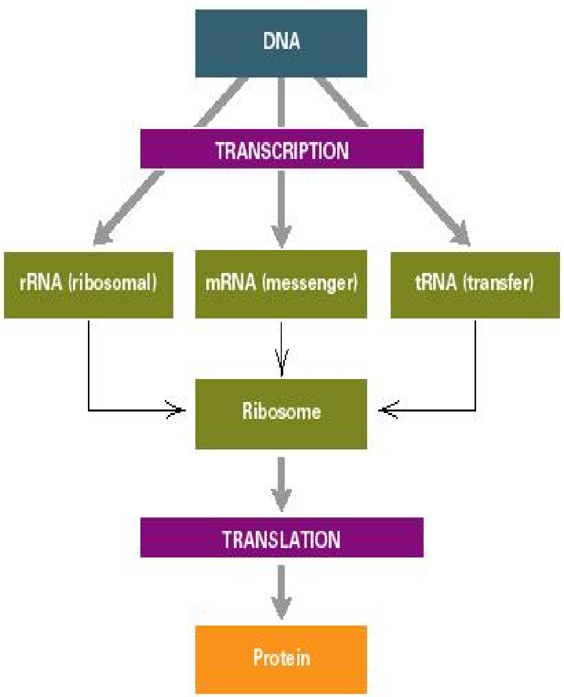


Subunit	Size	#/Molecule	Function
α	36.5 kD	2	chain initiation and interaction with regulatory proteins
β	151 kD	1	chain initiation and elongation
β'	155 kD	1	DNA binding
σ	70 kD	1	promoter recognition

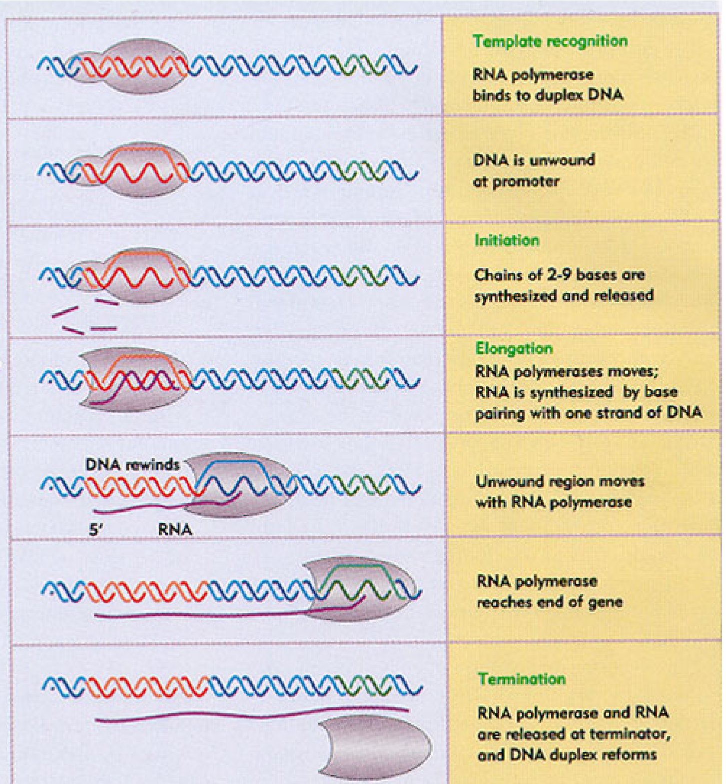
Roger Kornberg, Nobel in Chemistry, 2006. (son of Arthur Kornberg of the previous slide)

Error rate: 1 in 100,000 bases

Speed: ~60 bases/sec



Transcription has four stages, which involve different types of interaction between RNA polymerase and DNA. The enzyme binds to the promoter and melts DNA, remains stationary during initiation, moves along the template during elongation, and dissociates at termination. (The change in shape of RNA polymerase between initiation and elongation indicates the involvement of different forms of the enzyme, as described in the next section.)



Kusuma School of Biological Sciences

Translation

The Nobel Prize in Physiology or Medicine 1968

"for their interpretation of the genetic code and its function in protein synthesis"

Robert W Holley
Cornell U., USA
b. 1922 (USA)

Har Gobind khorana
U. Wisconsin, USA
b. 1922 (India)

Marshall W Nirenberg
NIH, USA
•b. 1927 (USA)

The Genetic Code

	Second position				
First position	**U**	**C**	**A**	**G**	**Third position**
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	**stop**	**stop**	A
	Leu	Ser	**stop**	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

How does it work?

DNA Sequence: atggccctgtggatgcgcctcctgcccctgctggcgctgctggccctctggggacctgac................

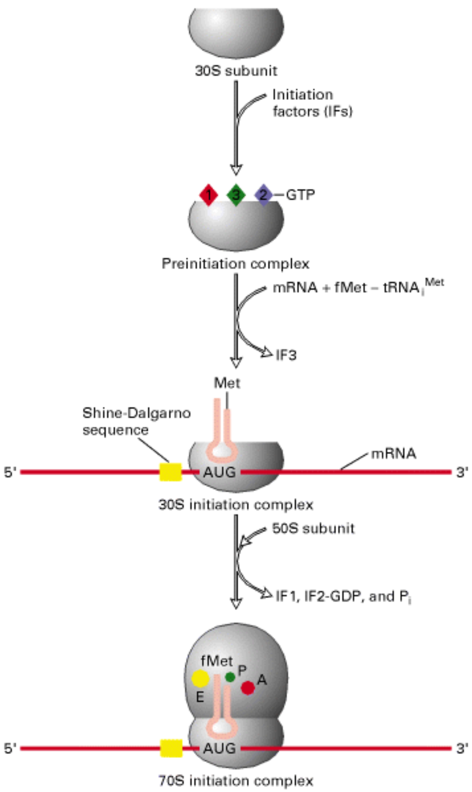
mRNA sequence: auggcccuguggaugcgccuccugccccugcuggcgcugcuggcccucuggggaccugac......

Protein sequence: M A L W M R L L P L L A L L A L W G P D........

Kusuma School of Biological Sciences

Translation

Molecular Engineering – par excellence !



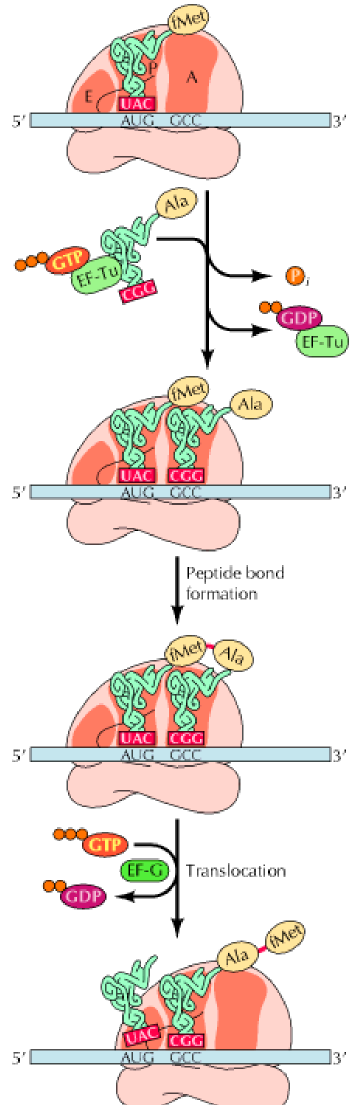
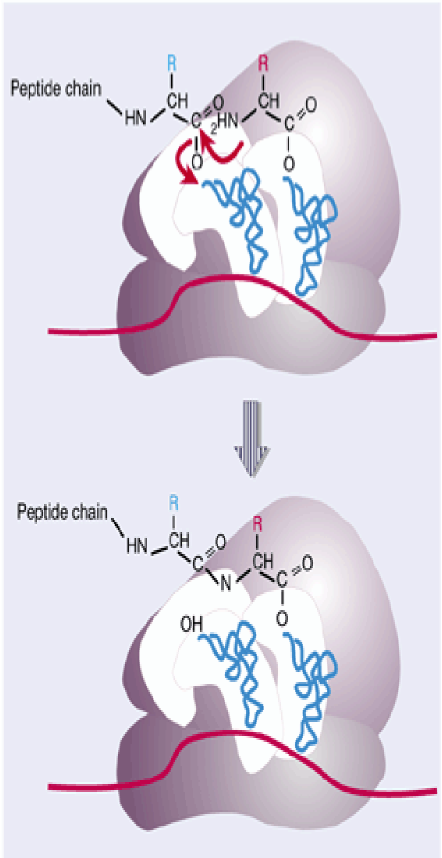
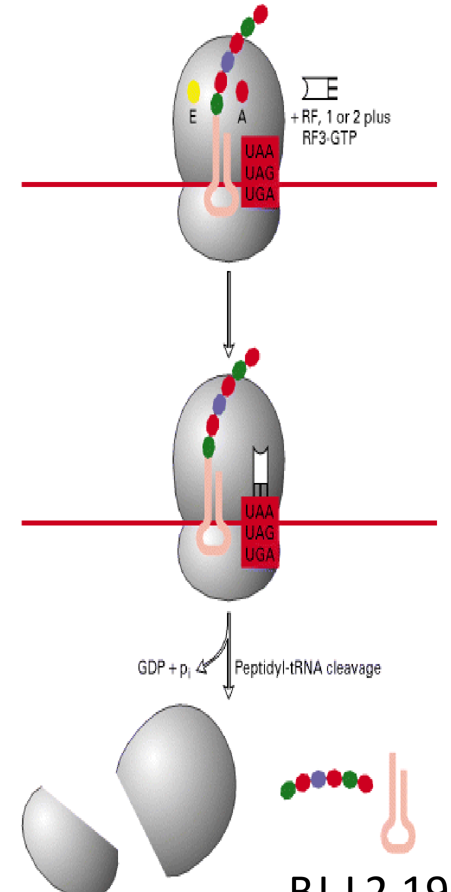


Figure Peptide bond formation takes place by reaction between the polypeptide of peptidyl-tRNA in the P site and the amino acid of aminoacyl-tRNA in the A site.



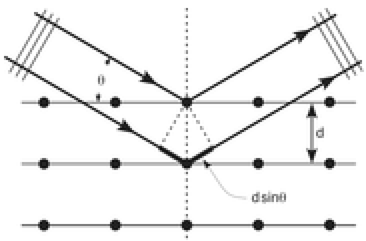
Error rate ~
1 in 5,000 amino acids

Speed ~ 20 AAs/sec



Kusuma School of Biological Sciences

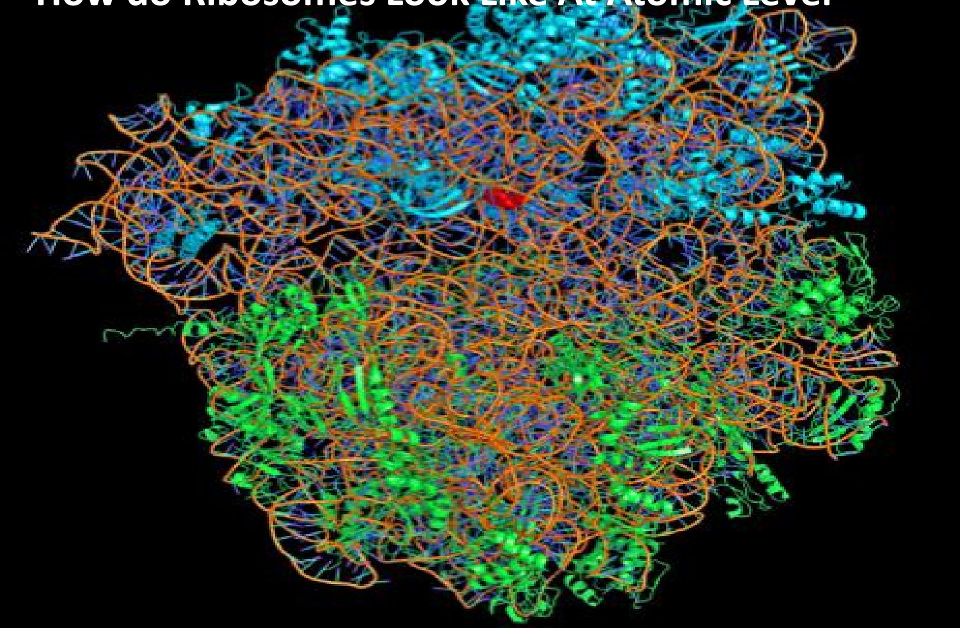
Seeing is believing: "one picture is worth a thousand words"



$$2d\sin\theta = n\lambda$$

X-ray crystallography is a method of determining the arrangement of atoms within a crystal, in which a beam of X-rays strikes a crystal and diffracts into many specific directions. From the angles and intensities of these diffracted beams, a crystallographer can produce a three-dimensional picture of the density of electrons within the crystal. From this electron density, the mean positions of the atoms in the crystal can be determined, as well as their chemical bonds etc..

How do Ribosomes Look Like At Atomic Level



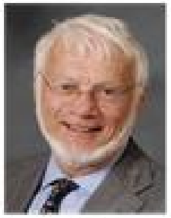
The Nobel Prize in Chemistry 2009

"for studies of the structure and function of the ribosome"

Venkatraman Ramakrishnan

MRC Laboratory of Molecular Biology Cambridge, United Kingdom

b. 1952 (India)



Thomas A. Steitz

Yale University New Haven, CT, USA

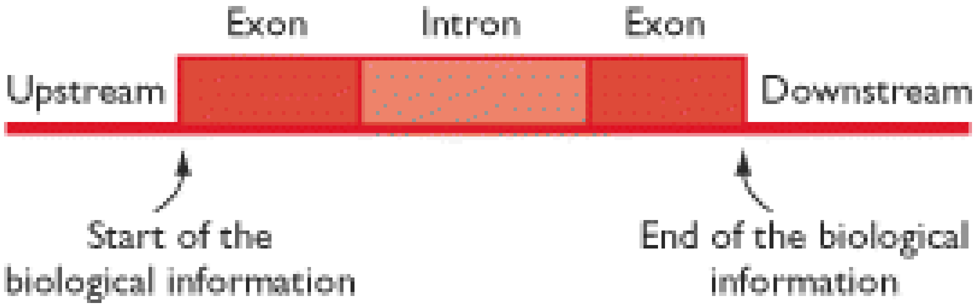
b. 1940 (USA)

Ada E. Yonath

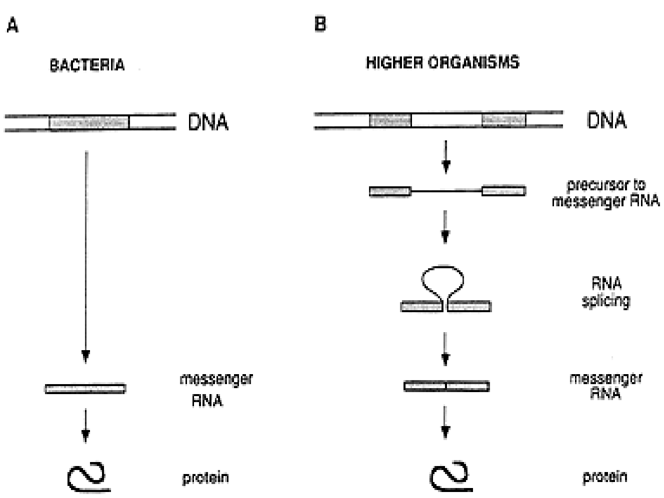
Weizmann Institute of Science Rehovot Israel

b. 1939 (Israel)

The Nobel Prize in Physiology or Medicine 1993 was awarded jointly to R. J. Roberts and P. A. Sharp *"for their discoveries of split genes"*



Sequence of DNA comprising an interrupted gene in eukaryotes.



"Gene structure and the flow of genetic information in bacteria (A) and higher organisms (B). In bacteria, the genetic information is stored as a continuous segment of DNA, and the messenger RNA can immediately direct the synthesis of the corresponding protein. In higher organisms, the gene is usually split, and the messenger RNA has to be processed by splicing before it can be translated into a protein."

This gives rise to the possibility of a single gene coding for multiple proteins via alternative splicing.

Kusuma School of Biological Sciences

Regulation of Gene Expression

Each of the three enzymes synthesized in response to lactose is encoded by a separate gene. The three structural genes (so named because they encode a product) are arranged in tandem on the bacterial chromosome. The parallel regulation of β-galactosidase, the permease, and the transacetylase suggested that the expression of genes encoding these enzymes is controlled by a common mechanism. Francois Jacob and Jacques Monod proposed the ***operon model*** *to account for this parallel regulation. The genetic elements of the model are a regulator gene, a regulatory DNA sequence called an operator site, and a set of structural genes.*

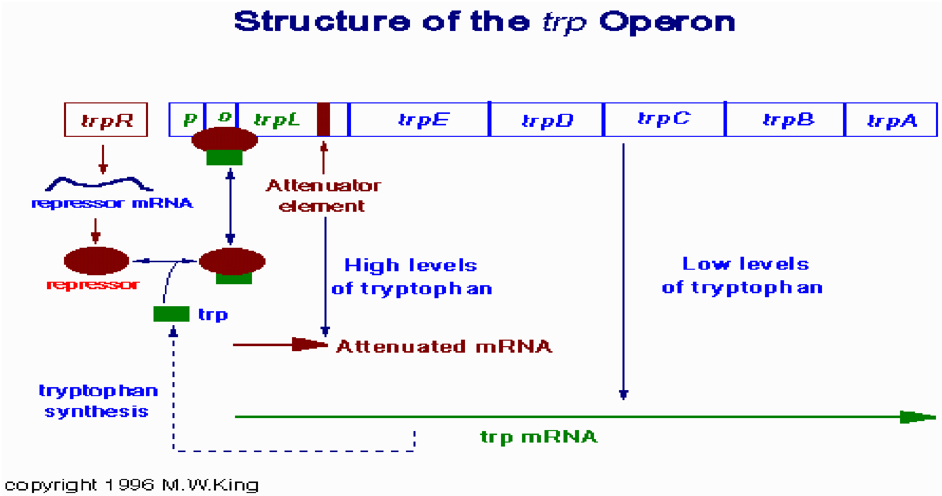
The Nobel Prize in Physiology or Medicine 1965 

"for their discoveries concerning genetic control of enzyme and virus synthesis"

Francois Jacob
Institut Pasteur, France
b. 1920 (France)

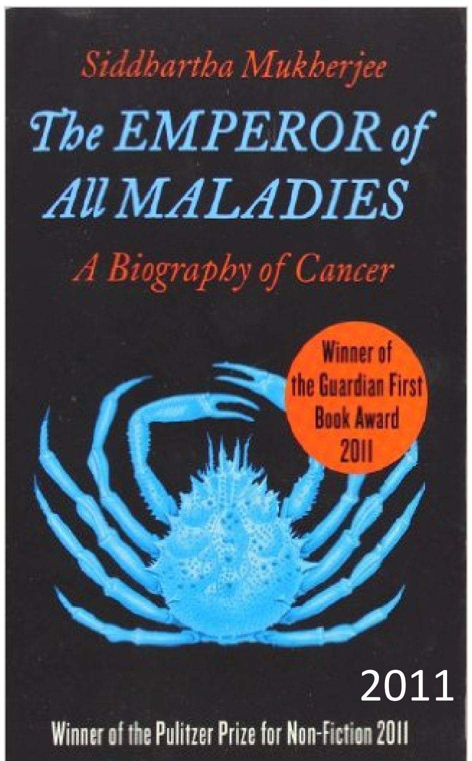
Andre Lwoff
Institut Pasteur, France
b. 1902 (France)

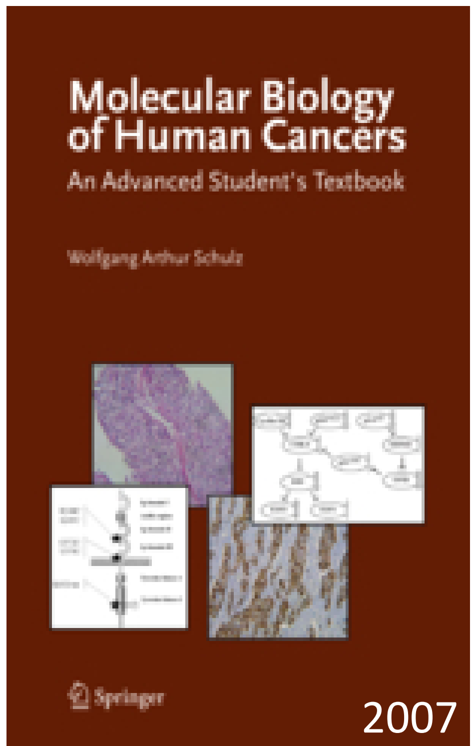
Jacques Monod
Institut Pasteur, France
b. 1910 (France)



When normal cell functioning goes out of control…

At least 200 types of cancers are known. They all share one or more of the following features: (a) Increased cell proliferation; (b) Insufficient apoptosis; (c) Altered cell and tissue differentiation; (d) Altered metabolism; (e) Genomic instability; (f) Immortalization; (g) Metastasis…





Several reasons for cancer:

P53 is a tumor suppressor protein and considered as the guardian of the genome. Mutations in P53 gene could cause cancer.

BRCA1 & BRCA2 are tumor supressor proteins and mutations in their genes could cause cancer.

Genetic, environmental , life style and other factors are implicated as causes…

Treatments:
Surgery, Irradiation, Chemotherapy
Not enough

DNA Repair

"Each day our DNA is damaged by UV radiation, free radicals and other carcinogenic substances, but even without such external attacks, a DNA molecule is inherently unstable. Thousands of spontaneous changes to a cell's genome occur on a daily basis. Furthermore, defects can also arise when DNA is copied during cell division, a process that occurs several million times every day in the human body. The reason our genetic material does not disintegrate into complete chemical chaos is that a host of molecular systems continuously monitor and repair DNA.....". (www.Nobelprize.org). Sleep Well !

The Nobel Prize in Chemistry 2015

"for mechanistic studies of DNA repair"

Tomas Lindahl
Francis Crick Institute, UK
b. 1938 (Sweden)

Paul Modrich
Howard Hughes Inst., USA
b. 1946 (USA)

Aziz Sancar
U of NC, USA
•b. 1946 (Turkey)

Assignment 4 (2020). How to cure cancer and diabetes? List all the genetic networks, biochemical pathways involved in causing these diseases/disorders and design a foolproof strategy for eliminating them.

Assignment 5 (2020). Design an alternative strategy for protein synthesis from DNA (gene) (without using known transcription and translation methods).

Assignment 6 (2020). How much of noncoding DNA in humans is really "junk DNA" or the incipient DNA of a superman or superwoman?