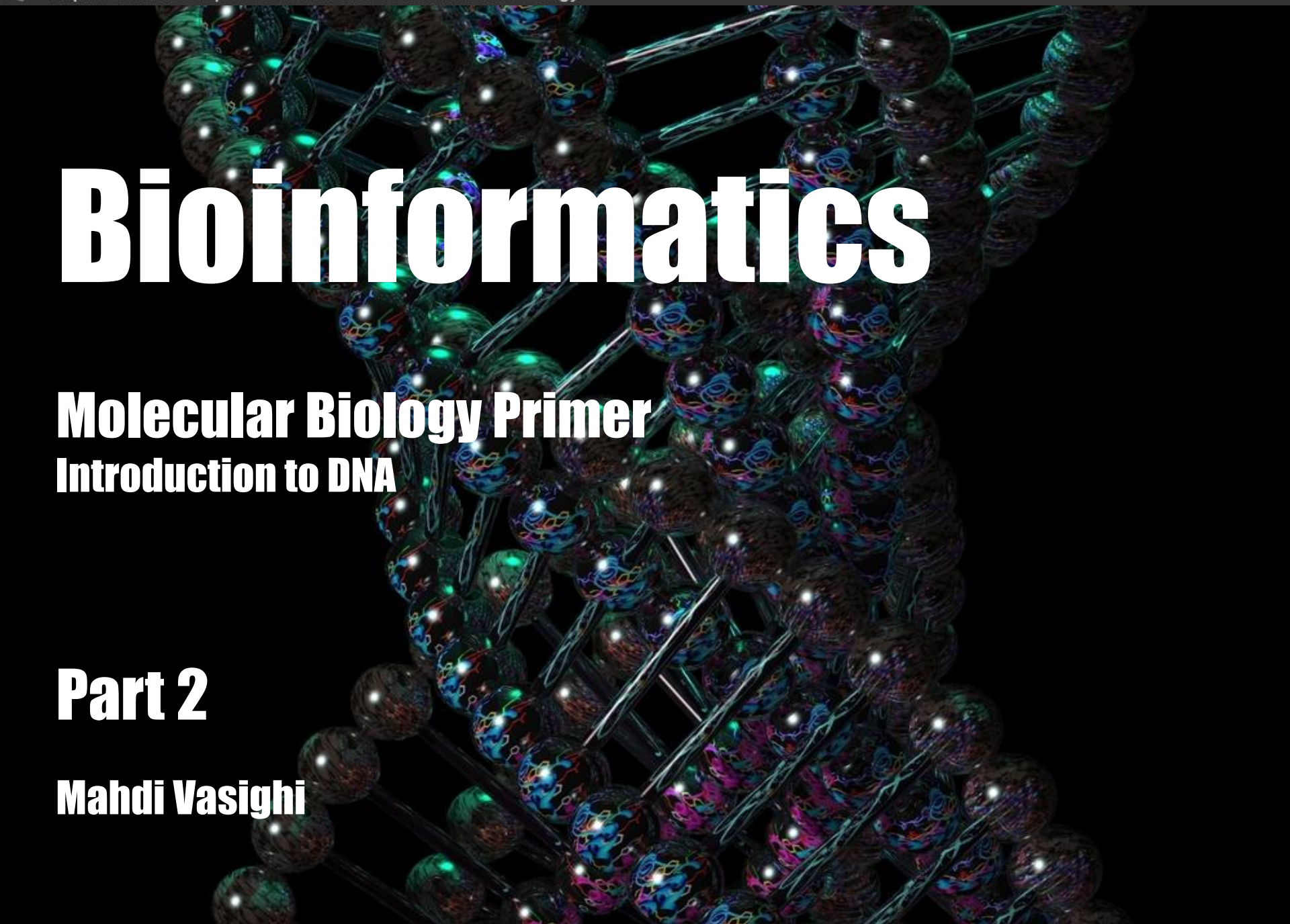


Bioinformatics

Molecular Biology Primer
Introduction to DNA

Part 2

Mahdi Vasighi

Molecular Biology: an Information Science

- the single most important scientific advance of the last century, it would likely be symbolized in the image of the double helix of DNA

- Its dual meaning is: “life propagates as pure information, which is encoded in physical molecules”

molecular biology is an information science as much as it is a physical science.

Molecular Biology: an Information Science

- Molecules are lifeless but molecules compose living things

- hydrogen (H), oxygen (O), carbon (C), and nitrogen (N) constitute >99% of human body

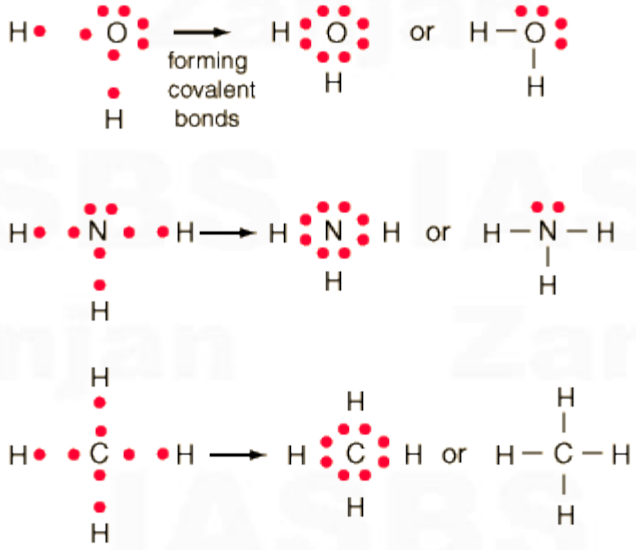
- Why are H, O, C, N so suitable to chemistry of life?
  - covalent bonds by electron-pair sharing!
  - the lightest elements of the periodic table capable of forming covalent bonds
- Two more elements appear in significant fractions: P (phosphorus) and S (sulfur)

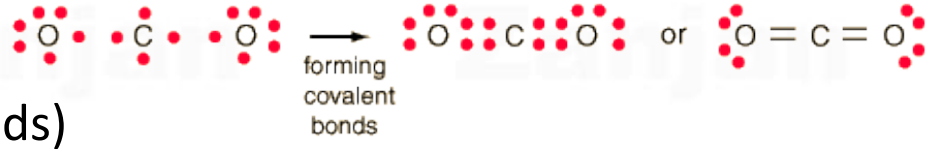
- Some other elements appear only in small traces: Cl (chlorine), Ca (calcium), Mg (magnesium), Cu (copper), Fe (iron), Mn (manganese), Zn(zinc), and Co (cobalt).

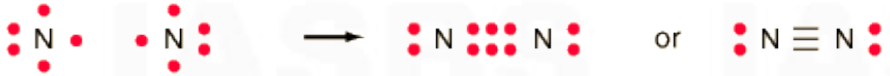
Molecular Biology: an Information Science

I	II	III	IV	V	VI	VII	0
H							He
Li	Be	B	C	N	O	F	Ne
Na	Mg	Al	Si	P	S	Cl	Ar
K	Ca	Ga	Ge	As	Se	Br	Kr
Rb	Sr	In	Sn	Sb	Te	I	Xe
Cs	Ba	Tl	Pb	Bi	Po	At	Rn

Metal Metalloid Nonmetal



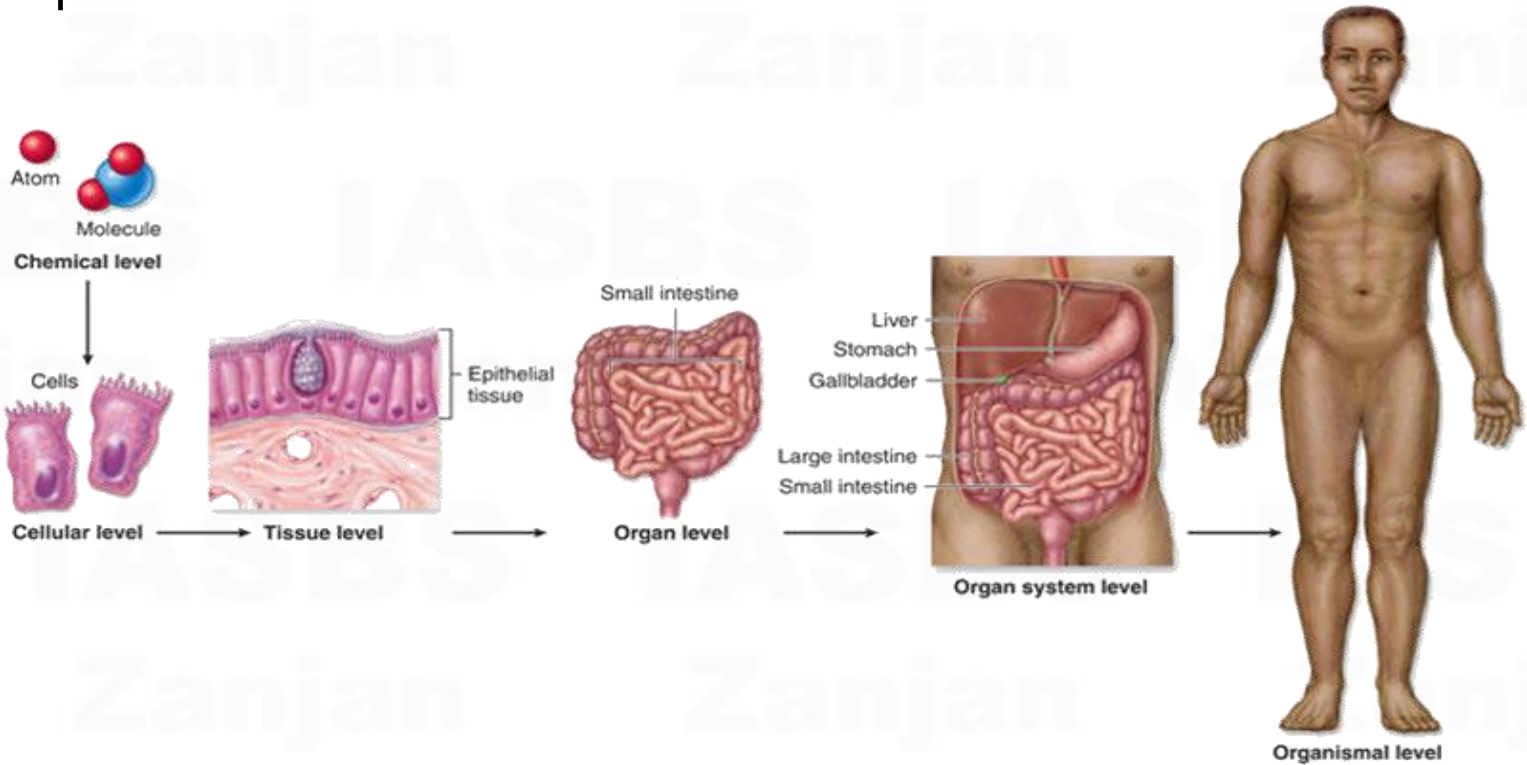




C can share 4 electrons (can form 4 bonds)
O has 2 unpaired electrons
N has 3 unpaired electrons
H has 1 unpaired electron

Basics of molecular Biology

- What makes living things distinct?
  - they can grow
  - they can replicate themselves
  - they can respond to stimuli
  - they can perform metabolism



What is a Cell?

Cells are the smallest structures capable of basic life processes, such as:

- growth
- metabolism (taking in nutrients and expelling waste)
- Stimulus response,
- Reproducing

Type of Cells

All living things fall into one of the two categories:

- **Prokaryotes**
- **Eukaryotes**

pro = means “prior to”
eu = means “true”
karyote = means “nucleus”

The distinction is based on whether or not a cell has a nucleus.

Prokaryotic cells do not have nuclei, while eukaryotic cells do. Also, eukaryotic cells have organelles.

Type of Cells

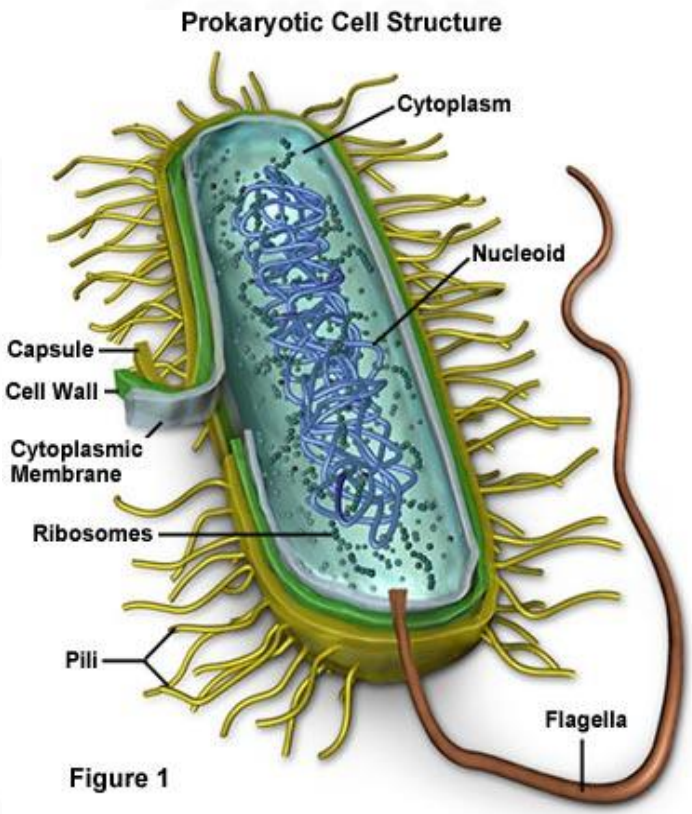
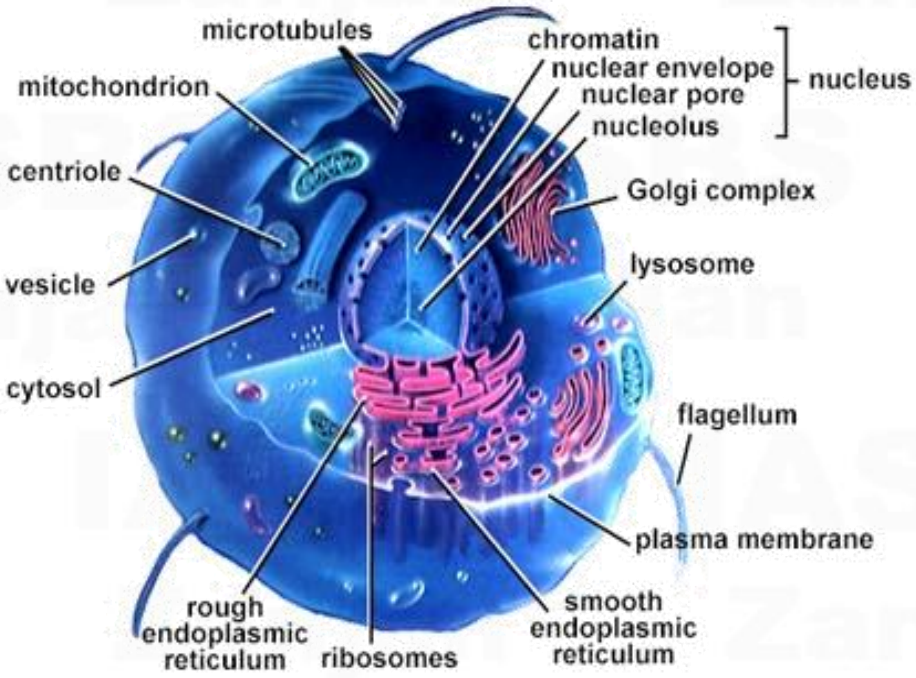


Figure 1

Prokaryotic Cell

- **Cell wall:** a rigid framework of polysaccharide cross-linked by short peptide chains;
- **Cell membrane:** bilayer; highly selective and controls the entry of most substances into the cell;
- **Nucleoid** (DNA): repository of the cell's genetic information; contains a single tightly coiled DNA
- **Ribosomes:** sites where proteins are synthesized; a bacterial cell has about 15,000 ribosomes
- **Storage granules:** granules where polymerized metabolites are stored (e.g. sugars)
- **cytosol** or cytoplasmic matrix: the site of intermediary metabolism

Type of Cells



Eukaryotic Cell

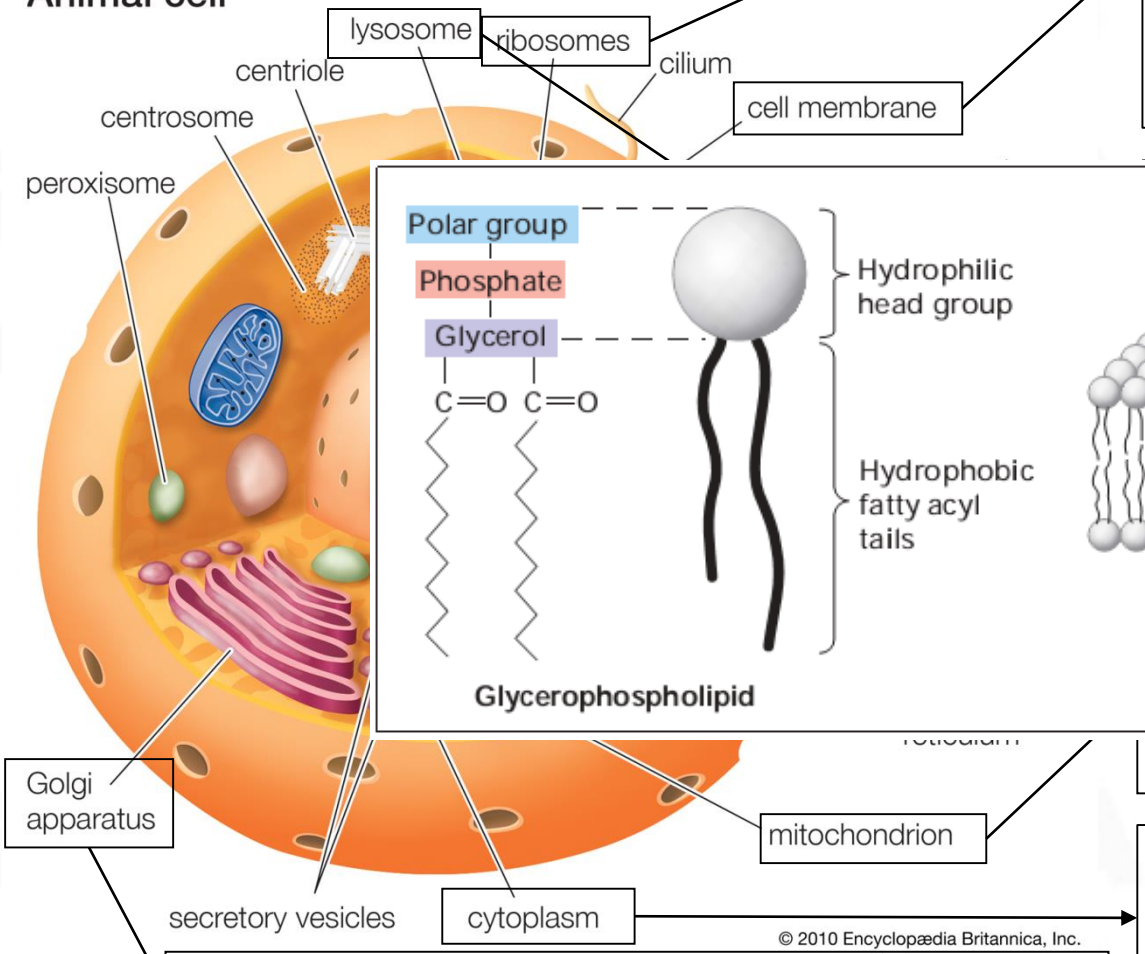
- much larger in size (1,000 to 10,000 times larger than prokaryotic cells)
- much more complex metabolic processes are organized into compartments, with each compartment dedicated to a particular function (enabled by a system of membranes)
- possess a nucleus, the repository of cell's genetic material which is distributed among a few or many chromosomes

Cell Organelles

Ribosome assembles proteins

Membranes are the boundaries between the cell and the outside world. All present day cells have a phospholipids cell membrane. Phospholipids are lipids (oils or fats) with a phosphate group attached.

Animal cell



that powers the cell.

Cytoplasm is the name for the gel-like collection of substances inside the cell.

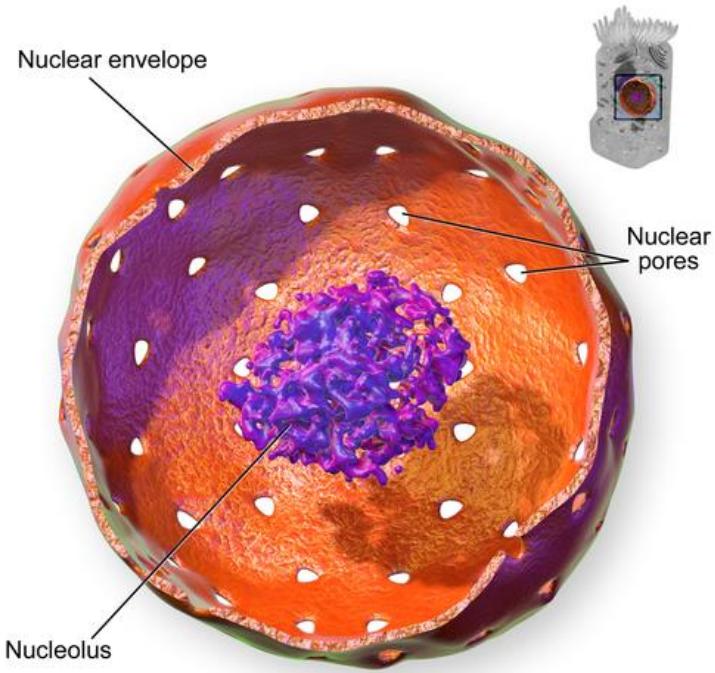
Golgi Apparatus is the delivery system for the cell

Nucleus

The nucleus of eukaryotic cells is composed primarily of protein and **deoxyribonucleic acid** or **DNA**.

- Spherical shape
- Denser than surrounding cytoplasm

The DNA is organized into linear units called **chromosomes**, also known as **chromatin** when the linear units are not obvious.

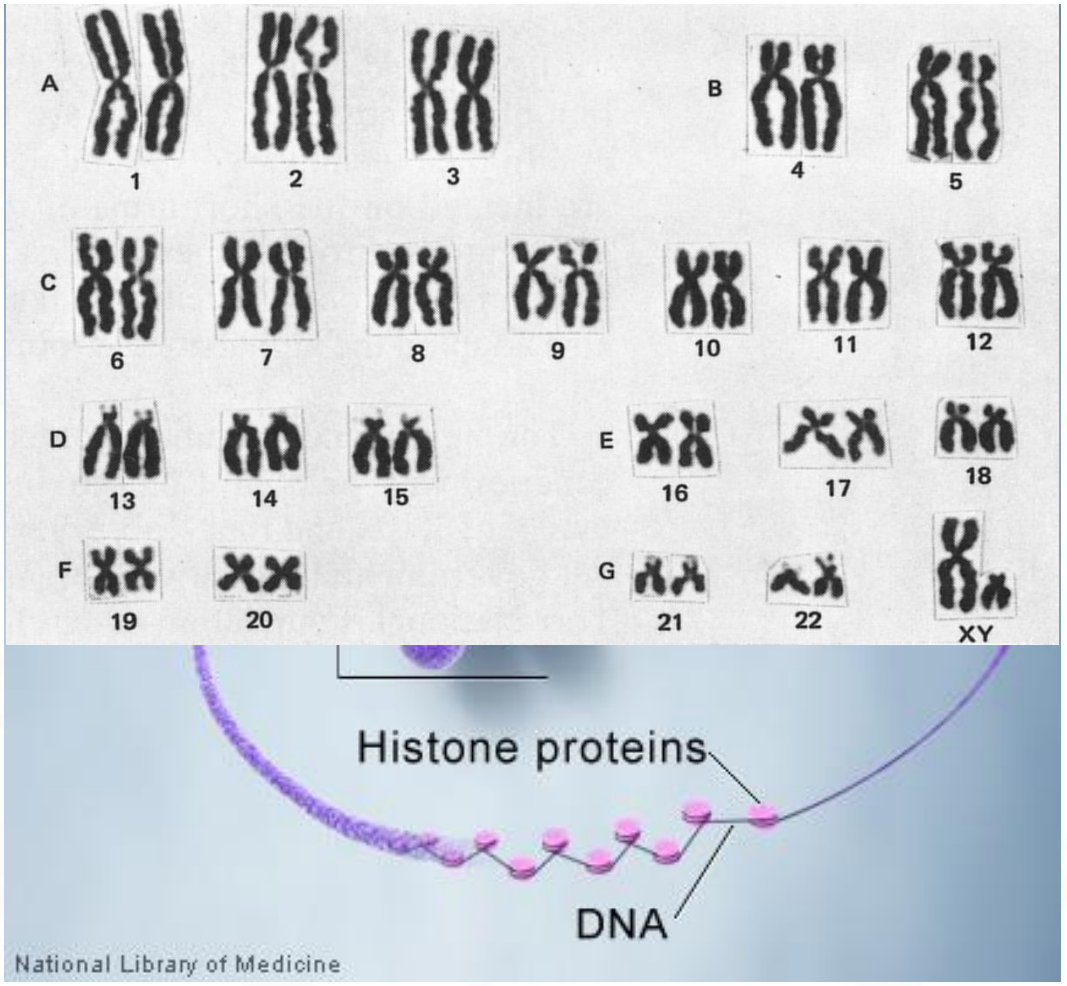


Nucleus

Chromosome

A chromosome is an organized structure of DNA and protein found in cells.

It is a single piece of coiled DNA containing many genes and regulatory elements.

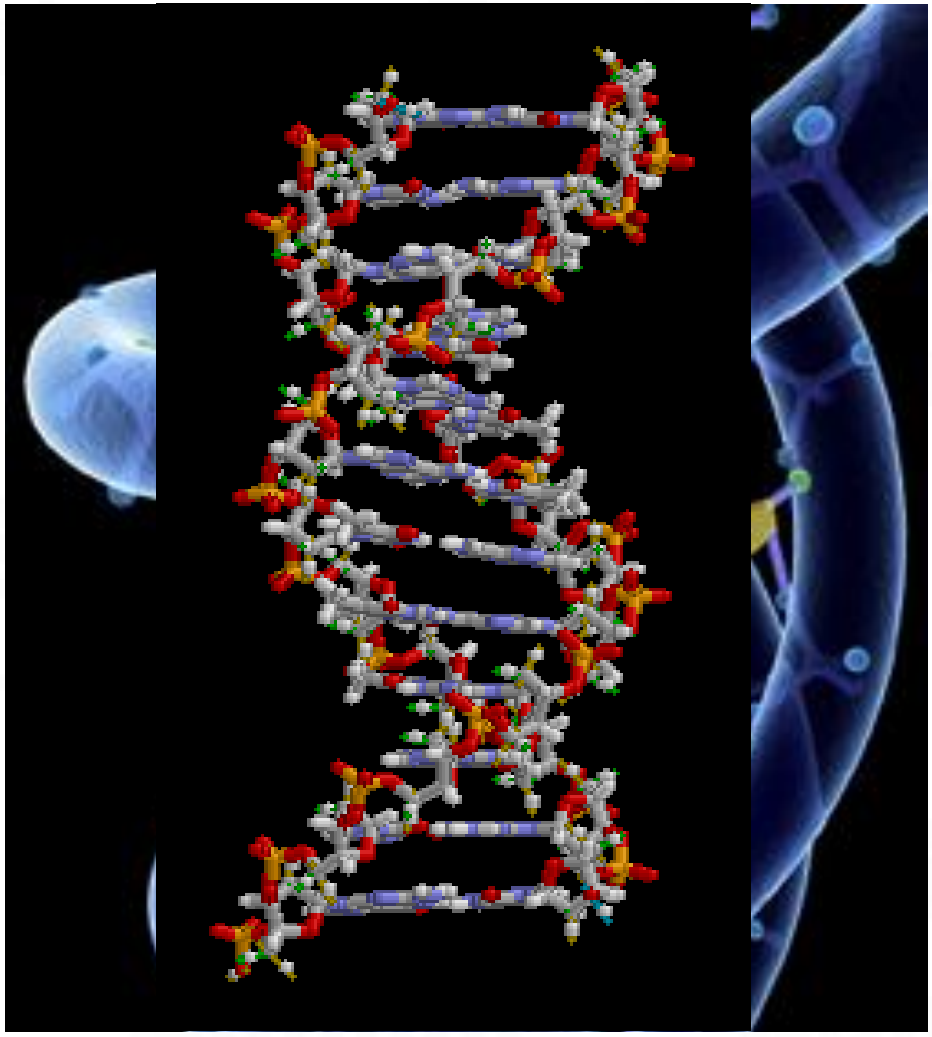


Deoxyribonucleic Acid : DNA

All the information directing every cell function is stored in large **DNA** molecules found in the nucleus.

The DNA segments that carry genetic information are called **genes**.

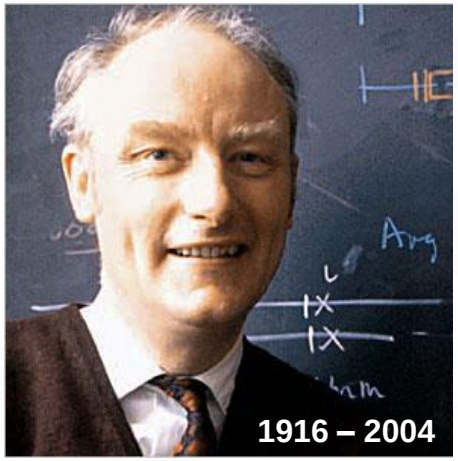
DNA consists of two long polymers of simple units called **nucleotides**.



Deoxyribonucleic Acid : DNA

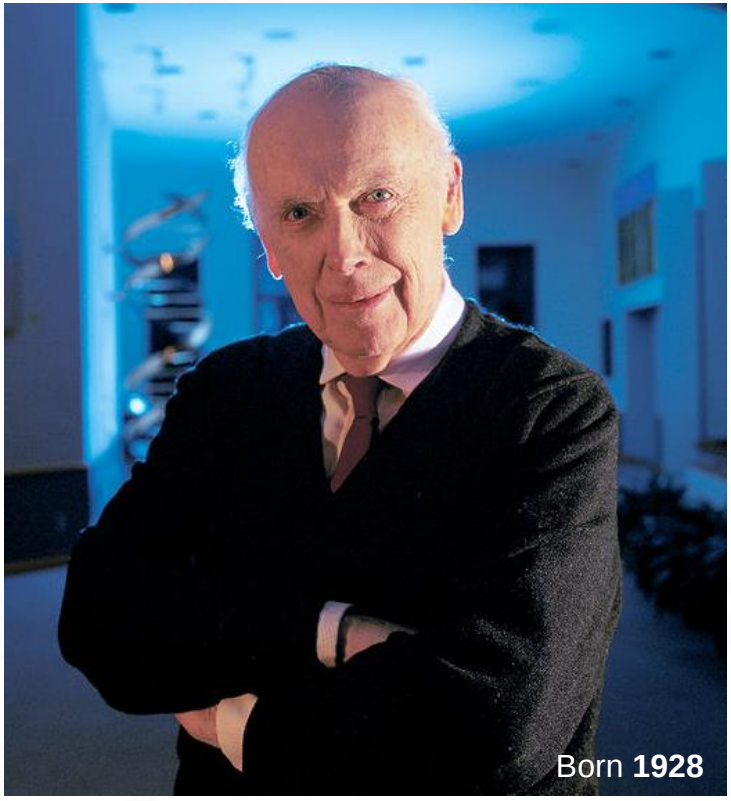

1920 –1958

The double helix model was supported by the work of **Rosalind Franklin** and **Maurice Wilkins**


1916 – 2004


1916 – 2004

The structure of DNA was discovered in 1953 by **Francis Crick** and **James Watson**


Born **1928**

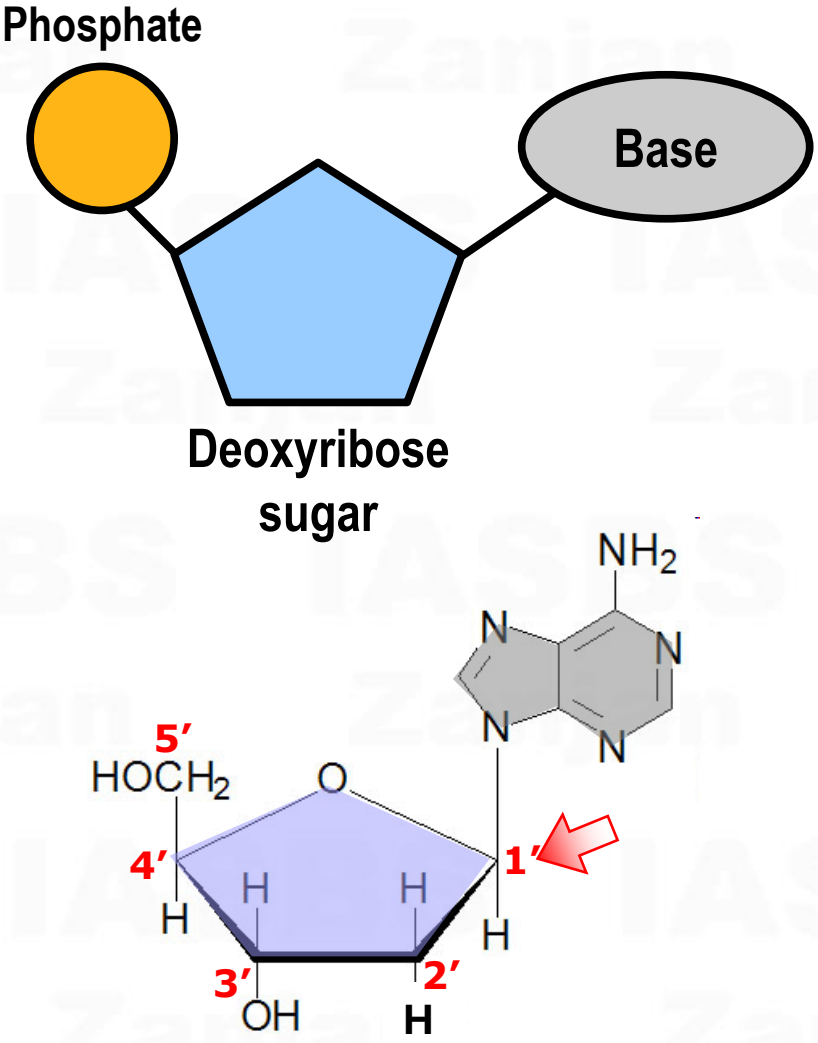
Deoxyribonucleic Acid : DNA

A nucleotide is composed of:

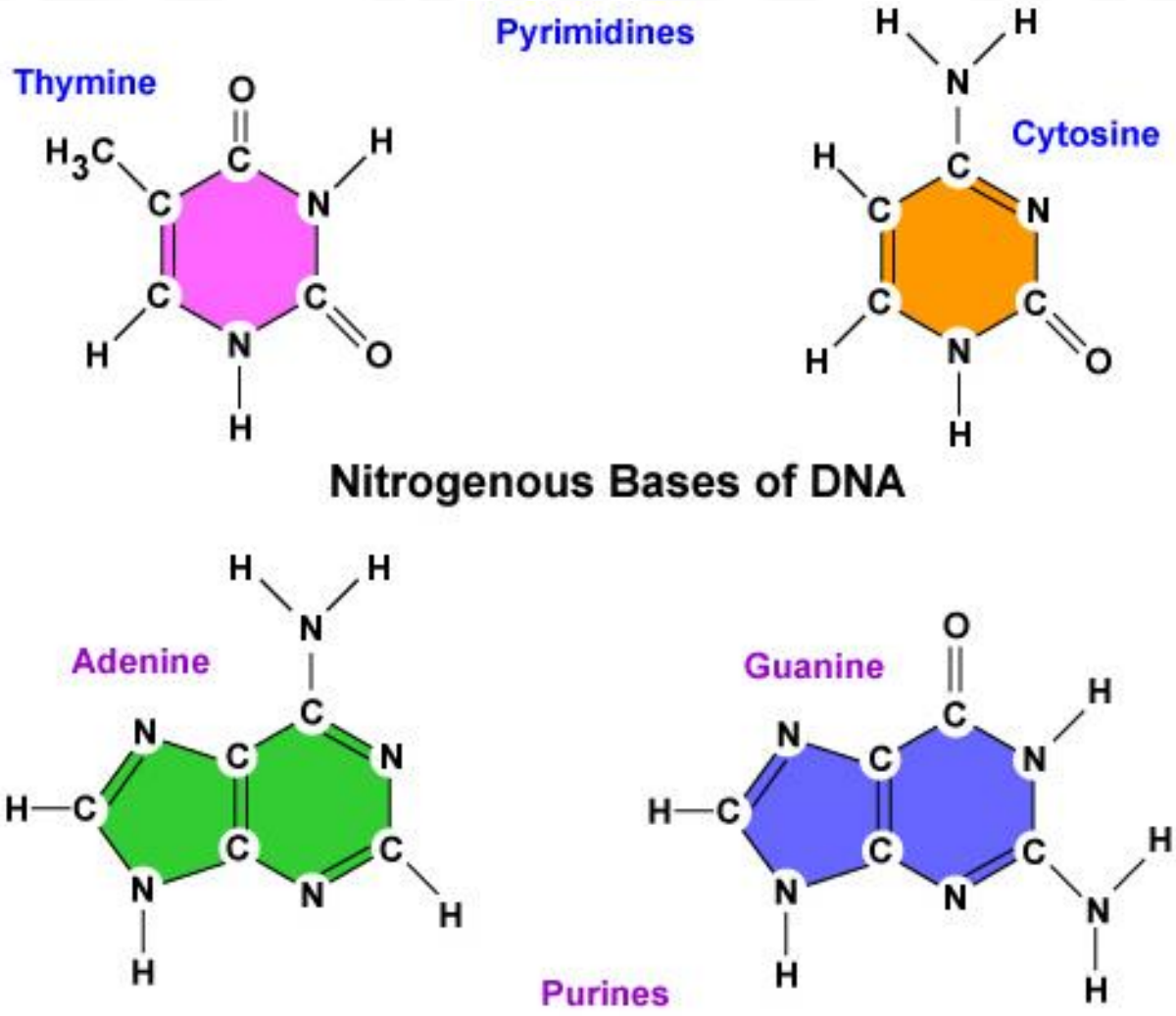
- a nitrogenous base
- a five-carbon sugar
- a phosphate groups

Sugar + Base = ***Nucleoside***

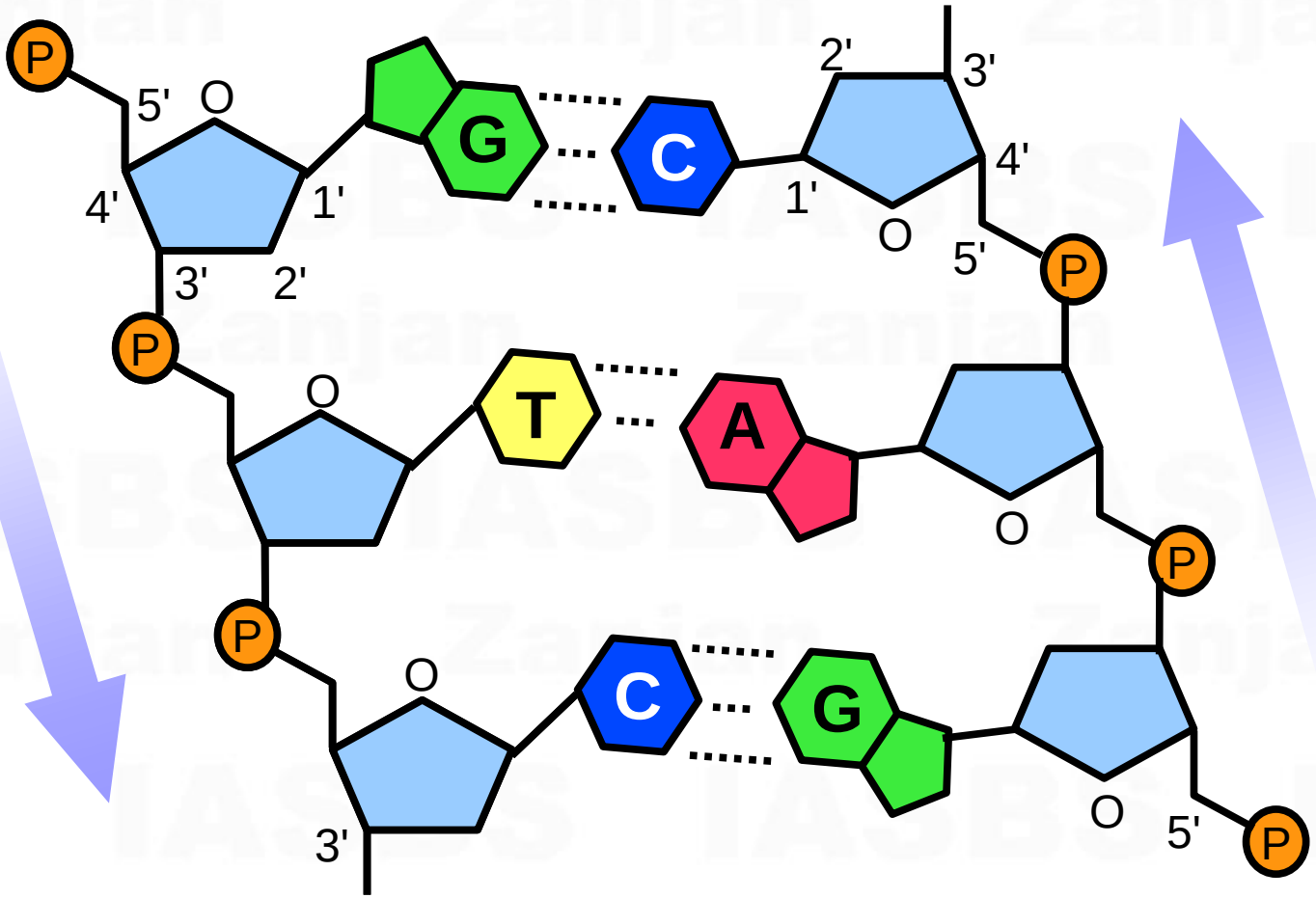
a phosphoric acid residue to the 5’ carbon of a nucleoside to make a ***Nucleotide***.



Nitrogenous Bases

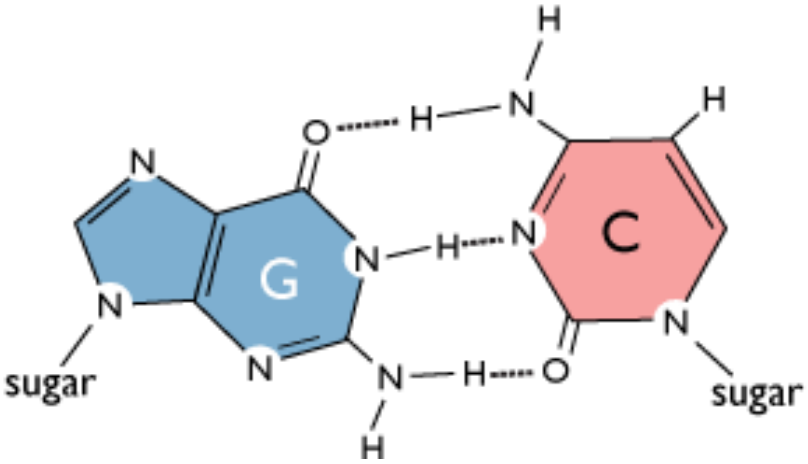


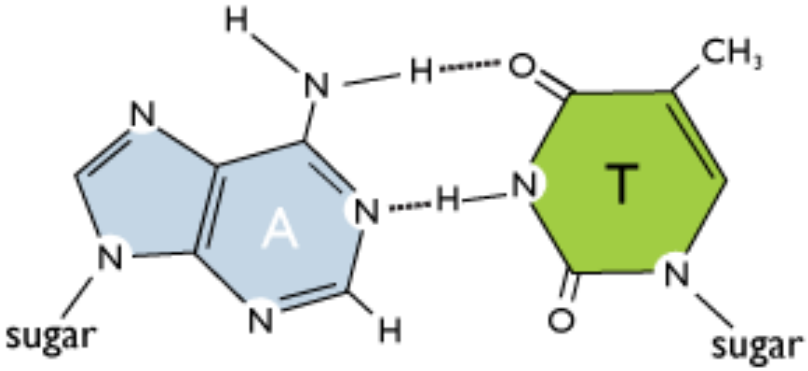
Nitrogenous Bases



The orientation of the two strands is antiparallel

Nitrogenous Bases





Guanine must pair with **Cytosine**

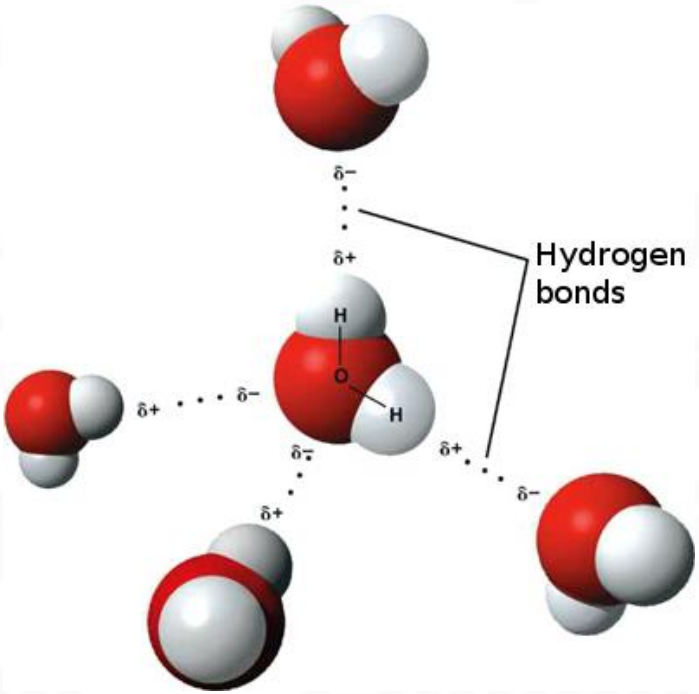
Adenine must pair with **Thymine**

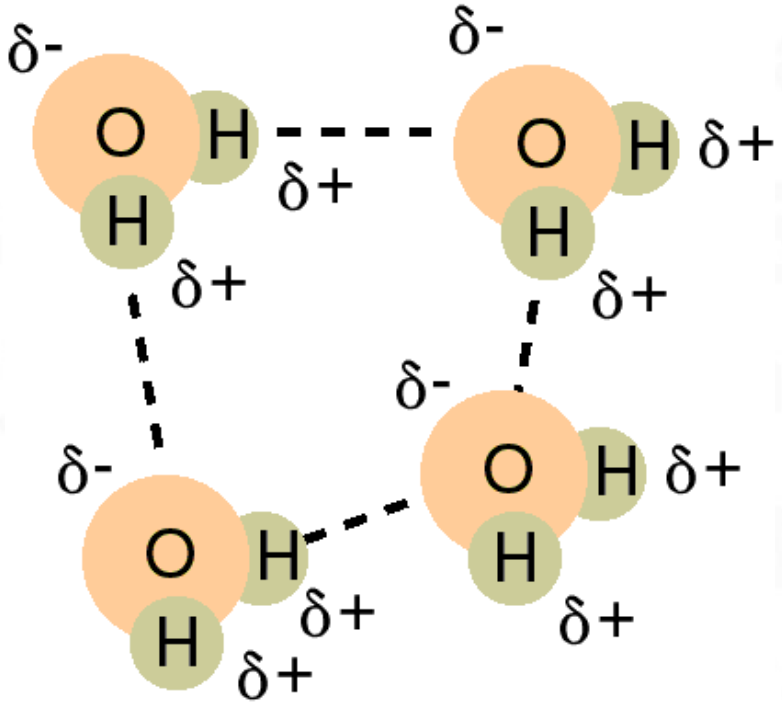
Specificity comes from **hydrogen bonding**

The first clue for the base pairing came by Erwin Chargaff

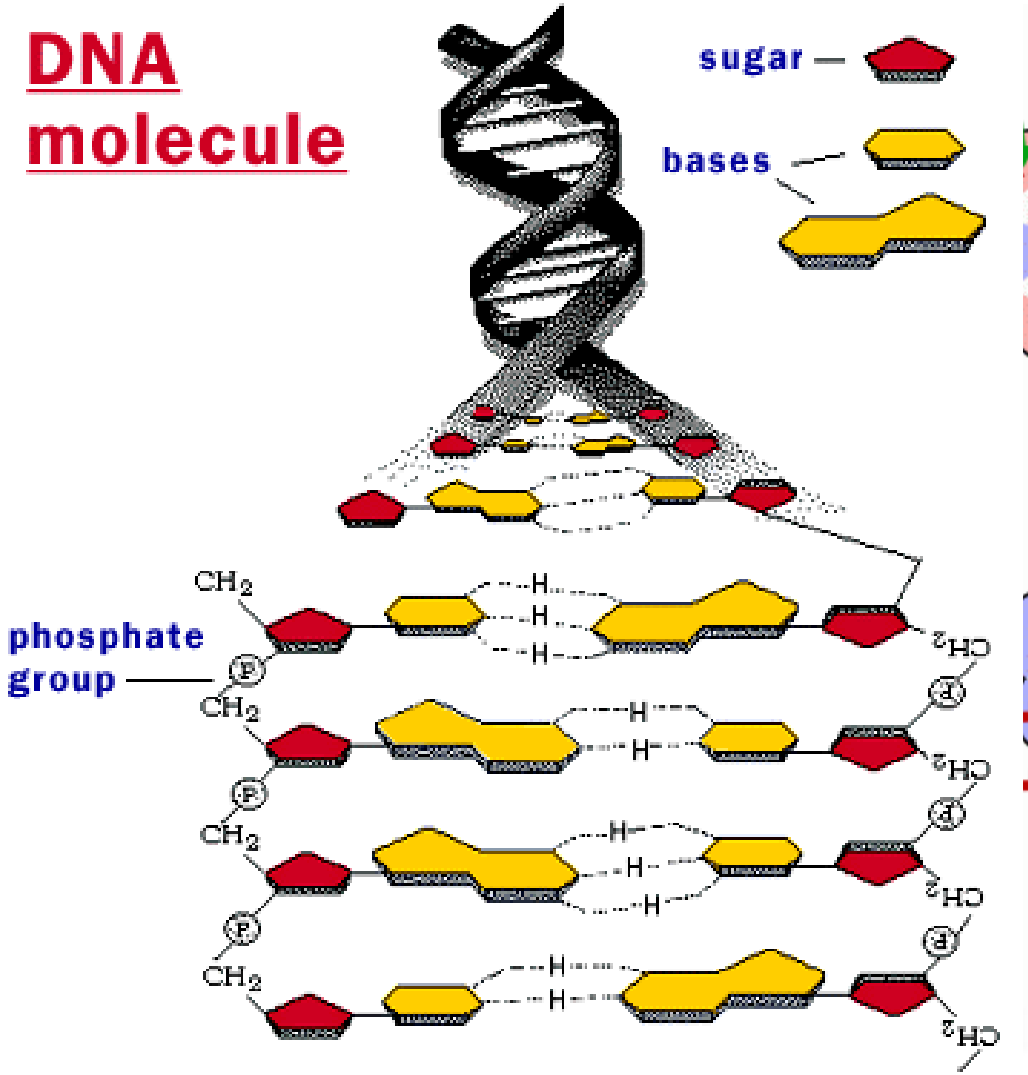
Nitrogenous Bases

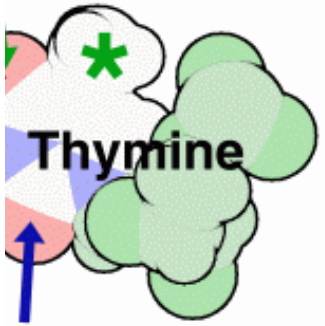
A hydrogen bond is the electromagnetic attractive interaction between polar molecules in which hydrogen (H) is bound to a highly electronegative atom, such as nitrogen (N), oxygen (O) or fluorine (F).

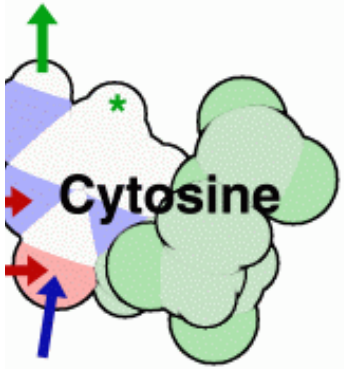




Nitrogenous Bases







Overview of DNA

- a thread-like molecule (twisted ladder-like)
- the DNA isolated from different cells consists of two polynucleotide strands wound together to form a long, slender, helical molecule, the DNA double helix.
- each DNA strand consists of four types of nucleotides: adenine (A), cytosine (C), guanine (G), and thymine (T)
- the strands run in the opposite directions, that is, they are antiparallel
- the strands are held together in the double helical structure through inter-chain hydrogen bonds
- the H-bonds pair the bases of nucleotides in one chain to complementary bases in the other (so-called base pairing)

Virus

Sub microscopic entity consisting of a single nucleic acid surrounded by a protein coat and capable of replication only within the living cells of bacteria, animals or plants.

When it comes into contact with a host cell, a virus can insert its genetic material into its host, literally taking over the host's functions.

Virus

Influenza A virus (A/Ahvaz/106899/2015(H1N1))

```
   1 atgaaggcaa tactagtagt tctgctatat acatttgcaa ccgcaaatgc agacacatta
  61 tgtataggtt atcatgcgaa caattcaaca gacactgtag acacagtact agaaaagaat
 121 gtaacagtaa cacactctgt taaccttcta gaagacaagc ataacgggaa actatgcaaa
 181 ctaagagggg tagccccatt gcatttgggt aaatgtaaca ttgctggctg gatcctggga
 241 aatccg
 301 agttca
 361 caattg
 421 cccaat
 481 ttctac
 541 tcctac
 601 actact
 661 tcaaga
 721 gaaggg
 781 gcaact
 841 ggtatt
 901 ggtgct
 961 ccaaag
1021 attcaa
1081 atggta
1141 gacctg
1201 gaaaag
1261 atagag
1321 gaactgctgg ttctattgga aaatgaaaga actttggact accacgattc aaatgtgaag
1381 aacttgtatg aaaaggtaag aaaccagtta aaaaacaatg ccaaggaaat tggaaacggc
1441 tgctttgaat tttaccacaa atgcgacaac acgtgcatgg aaagtgtcaa aaatgggact
1501 tatgactacc caaaatactc aggggaagca aaattaaaca gagaaaaaat agatggggta
1561 aagctggaat caacaagaat ttaccagatt ttggcgatct attcaactgt cgccagttca
1621 ttggtactgg tagtctccct gggggcaatc agcttctgga tgtgctctaa tgggtctcta
1681 cagtgtagaa tatgtattta a
```

```
/translation="MKAILVVLLYTFATANADTLCIGYHANNSTDTVDTVLEKNVTVT
HSVNLLEDKHNGKLCKLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVETSSS
DNGTCYPGDFINYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAACPHAGAKS
FYKNLIWLVKKGNSYPKLSQSYINDKGKEVLVLWGIHHPSTTADQQSLYQNADAYVFV
GTSRYSKKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFTMER
NAGSGIIISDTPVHDCNTTCQTPEGAINTSLPFQNVHPITIGKCPKYVKSTKLRLATG
LRNVPSIQSRGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQGSGYAADLKSTQNAIDKI
TNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLLENER
TLDYHDSNVKNLYEKVRNQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSG
EAKLNREKIDGVKLESTRIYQILAIYSTVASSLVLVVSLGAISFWMCSNGSLQCRICI
```