In the name of God The compassionate the merciful

Nucleic Acids

For the Students of Medicine and Dentistry

By

Dr. A. Seghatoleslam

Professor of Cell & Molecular Biology

Shiraz Medical School

بسم الله الرحمن الرحيم

اللهم اخرجنا من ظلمات الوهـــم

و اكرمنا بنور الفهـــم

اللهــم افتــح علينا ابواب رحمتك

و انشرعلينا خزاين علومك

برحمتك يا ارحم الراحمـــين

CHEMICAL COMPONENTS OF LIFE (Macromolecules)

1. CARBOHYDRATES
2. NUCLEIC ACIDS:
   - DNA
   - RNA
3. PROTEINS
4. LIPIDS

NUCLEIC ACIDS

- **Genetic material of all known organisms:**

  - **DNA: deoxyribonucleic acid**
  - **RNA: ribonucleic acid (i.e. some viruses: retrovirus, hepatitis C virus,..)**

- **High Molecular weight polymers, informational macromolecules.**

- **5-15% of dry weight of all cells.**

NUCLEIC ACIDS

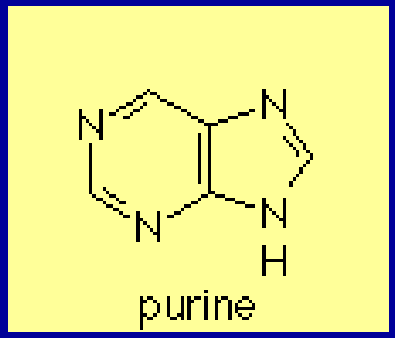
- **Consist of chemically linked sequences of nucleotides:**
  - **A heterocyclic nitrogenous base.**
  - **Pentose- 5-carbon sugar (ribose or deoxyribose).**
  - **Phosphate group.**
- **The sequence of bases provides the genetic information.**

  (Genetic code)

Bases

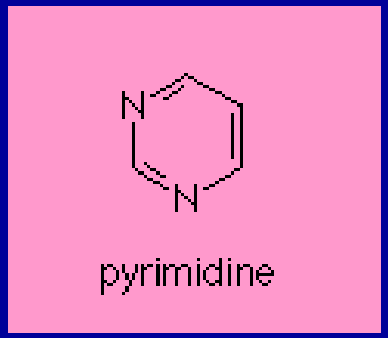
- Two types of bases:

- <u>Purines</u> are fused five- and six-membered rings:
  - Adenine A DNA RNA
  - Guanine G DNA RNA



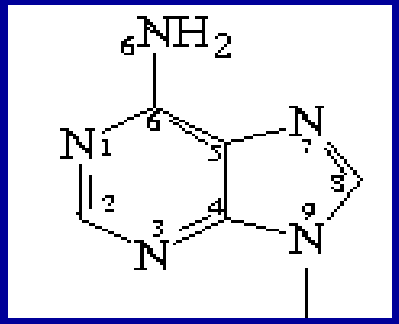
purine

- <u>Pyrimidines</u> are six-membered rings:
  - Cytosine C DNA RNA
  - Thymine T DNA
  - Uracil U RNA

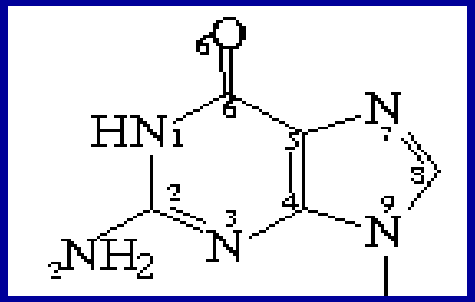


pyrimidine

Purines

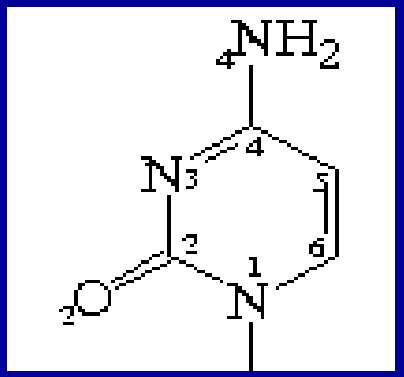


Adenine (A)
(6-amino purine)

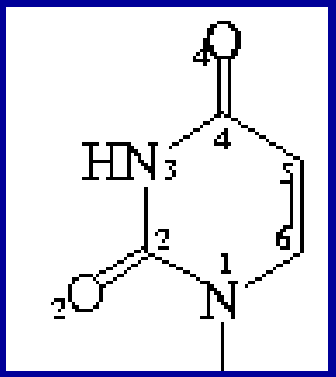


Guanine (G)
(2-amino-6-oxy purine)

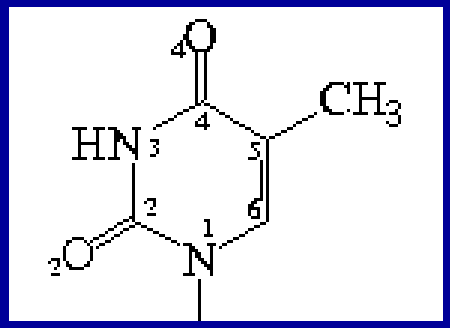
Pyrimidine



Cytosine (C)
(2-oxy-4-amino pyr.)



Uracil (U)
(2,4-dioxy pyr.)



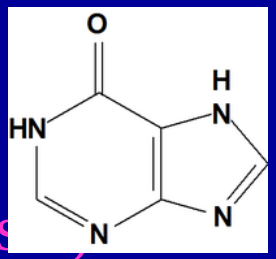
Thymine (T)
(2,4-dioxy-5-methyl pyr.)
(methyl uracil)

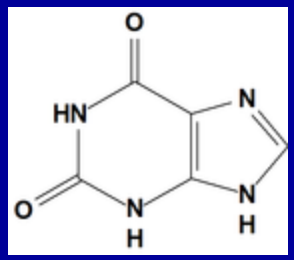
Some of minor bases

- Various derivatives of bases:
  - **5-methyl cytosine** (bacterial & human DNA)
  - **N6-methyl adenosine** (mRNA)

- Bases present in the free state in cells:
  - **Hypoxanthin** (6-oxypurine)
  - **Xanthin** (2,6-dioxy purine)

  (Intermediates of A & G metabolis)

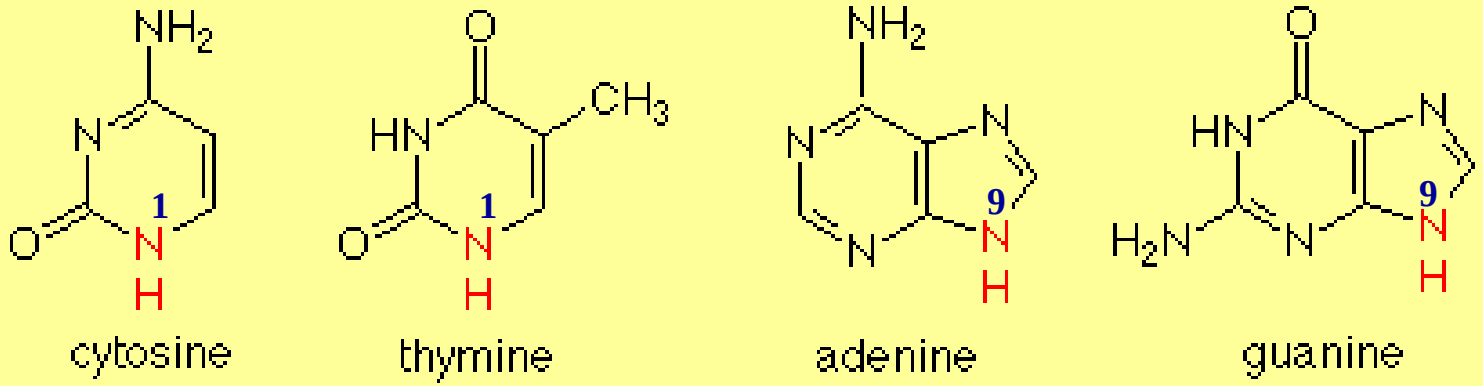




- Methylated purines of plants have pharmacologic properties:
  - ***Caffeine** (coffee): 1,3,7-trimethyl xanthine
  - **Theophylline** (tea): 1,3-dimethyl xanthine
  - **Theobromine** (cocoa): 3,7-dimethyl xanthine

* Adenosine / receptor

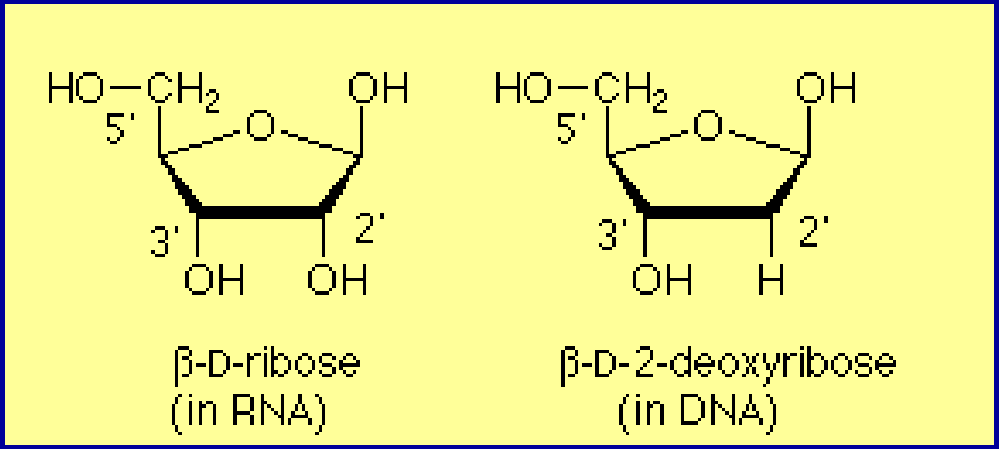
Nitrogenous bases:



In each of these bases there is a secondary amine whose nitrogen forms a bond to the anomeric carbon of a deoxyribose in the DNA backbone.

Sugars

- **2'-deoxy-D-ribose**
- **D-ribose**



(The "primes" ' are there to differentiate atoms in the sugar ring from those in the bases).

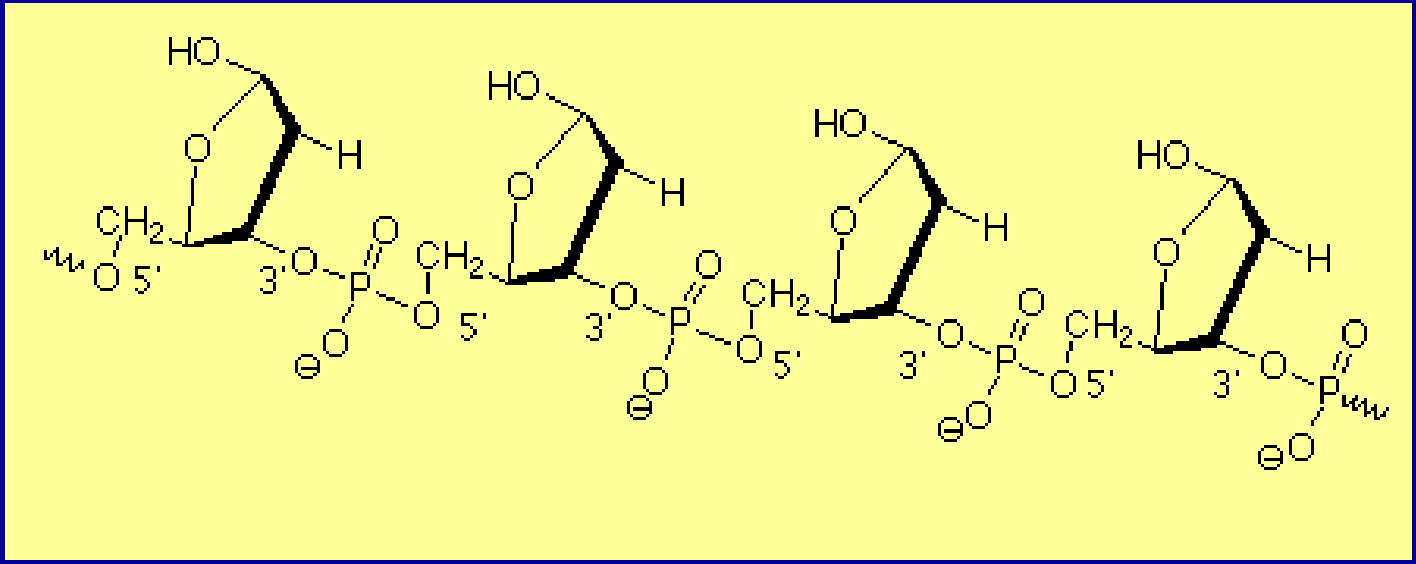
The pentose linked via a **β-N-glycosidic bond** to a ring nitrogen of a Purine or Pyrimidine base:

C1 (anomeric carbon) sugar → N9 Purine (A & G)

C1 (anomeric carbon) sugar → N1 Pyrimidine (T, C, U)

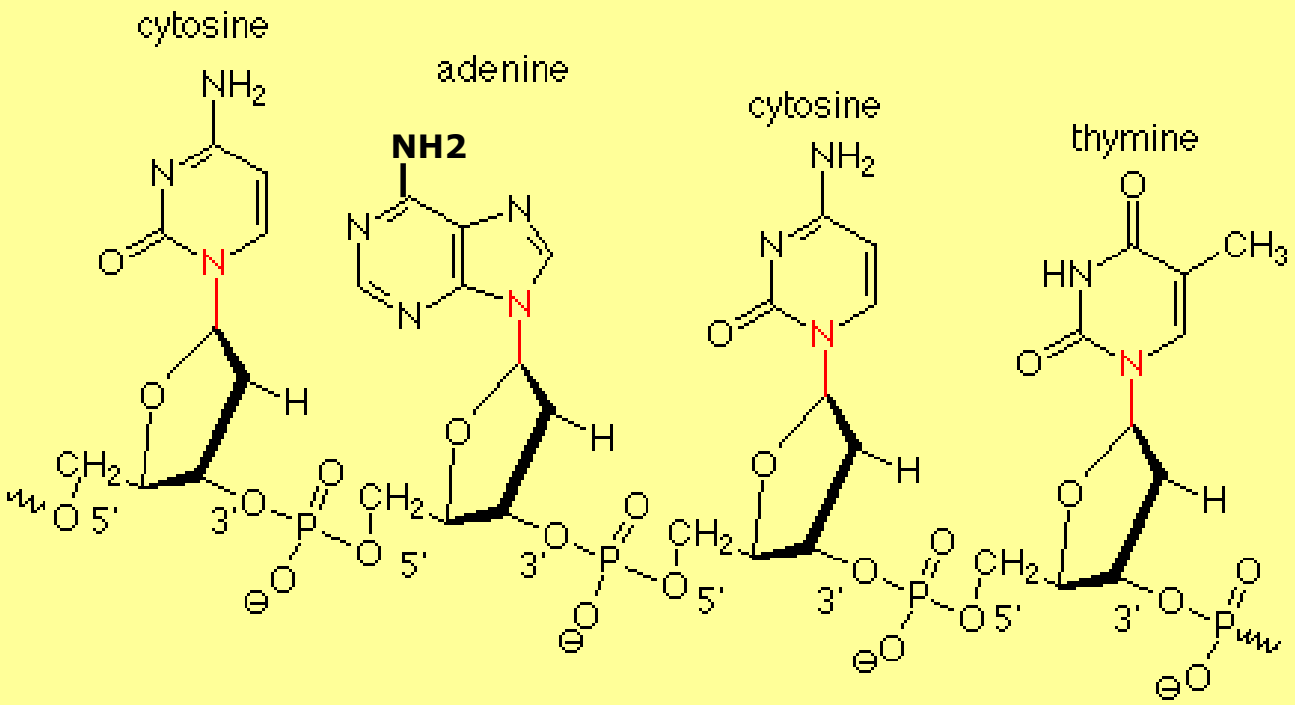
Backbone of the polymer Sugar & phosphate:

- 2-deoxyribose (or ribose) units are linked together by phosphate esters (phosphodiester bond) which link the 3' oxygen of one sugar with the 5' oxygen of the next.

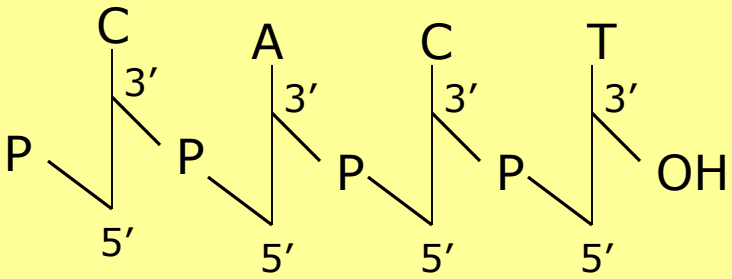


Part of nucleic acid chain

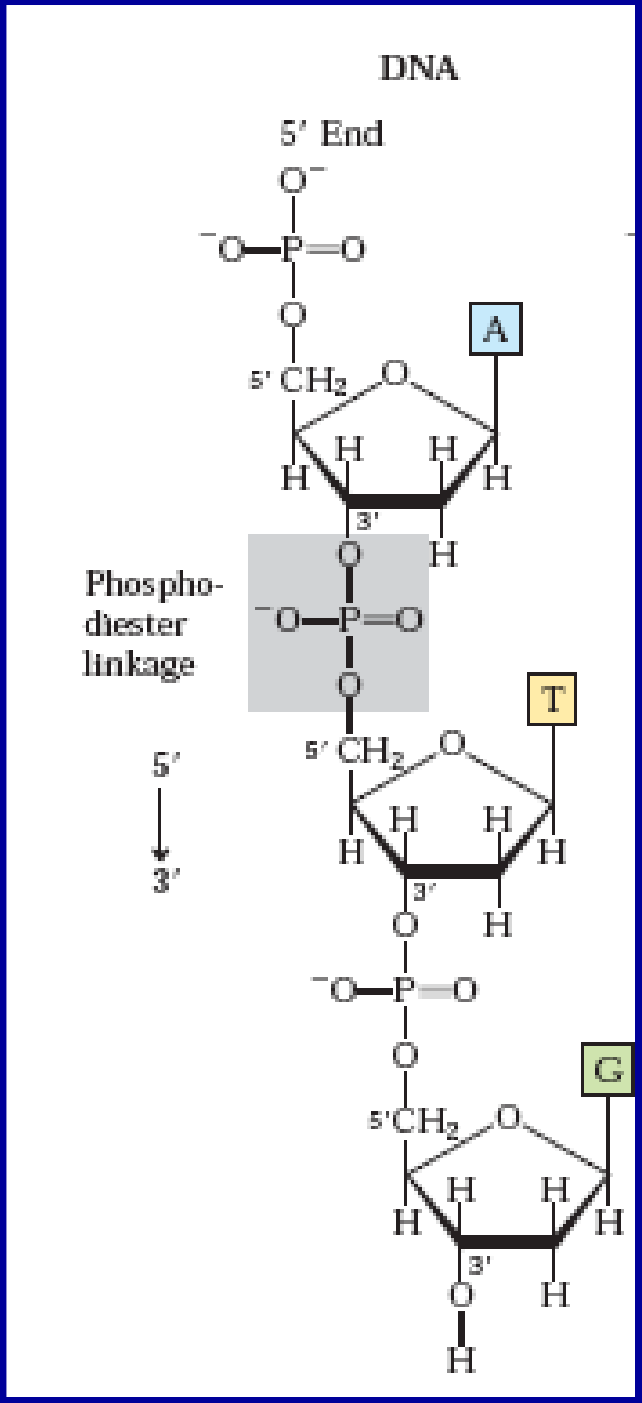
- CACT (~ pCpApCpT~)



- Shorthand notation for structure of oligonucleotides



N nucleotides(n-1 phosphodiester bond) polynucleotides (DNA & RNA)


DNA
5′ End
A
T
G
Phospho-
diester
linkage
5′
3′

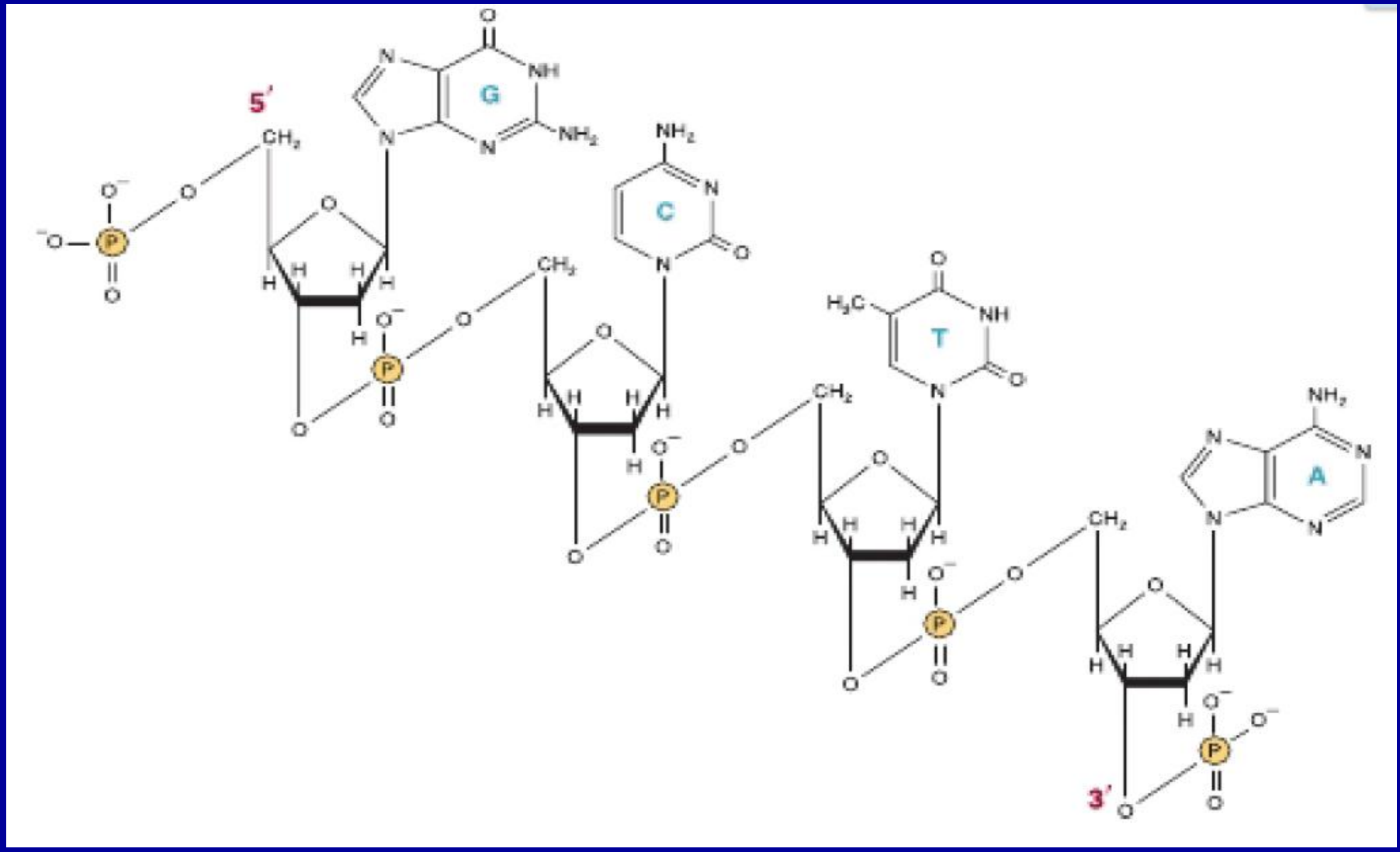


FIGURE 34–1, Harper's Illustrated Biochemistry

Nomenclature

Base	***Nucleoside***	***Nucleotide***	***Abbreviation***
Adenine	(Deoxy) adenosine	(Deoxy) adenosine monophosphate	(d) AMP
Guanine	guanosine	guanosine monophosphate	GMP
Uracil	uridine	uridine monophosphate	UMP
Cytosine	cytidine	cytidine monophosphate	CMP
Thymine	thymidine	thymidine monophosphate	TMP

Major biochemical function of purine and pyrimidine nucleotides

- **Energy carrier (i.e. ATP & other nucleoside-3-phosphate).**

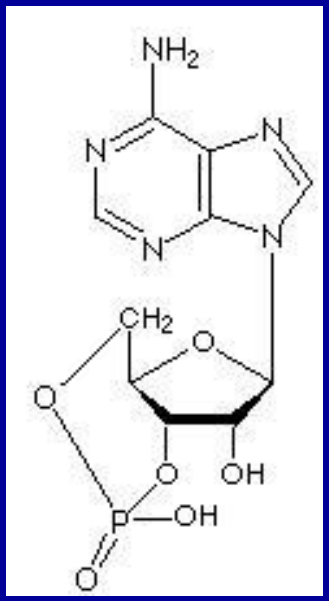
- **Carrier in biosynthetic reaction as "high energy intermediate" for covalent bond synthesis** (UDP-Glc & UDP-Gal: carbohydrate biosynthesis, CDP-acyl glycerol: lipid biosynthesis).

……Major biochemical function

- **A Portion of coenzyme structure such as FAD, NAD+, NADP+, CoA & s-adenosyl methionine.**
- **"Second messenger" for activating certain enzymatic cascade (cyclic AMP/ cAMP & cGMP).**
- **NTP serve as the monomer units precursor of DNA & RNA.**

Cyclic adenosine monophosphate (cAMP)

- ATP —(Adenylate cyclase)→ cAMP + PPi



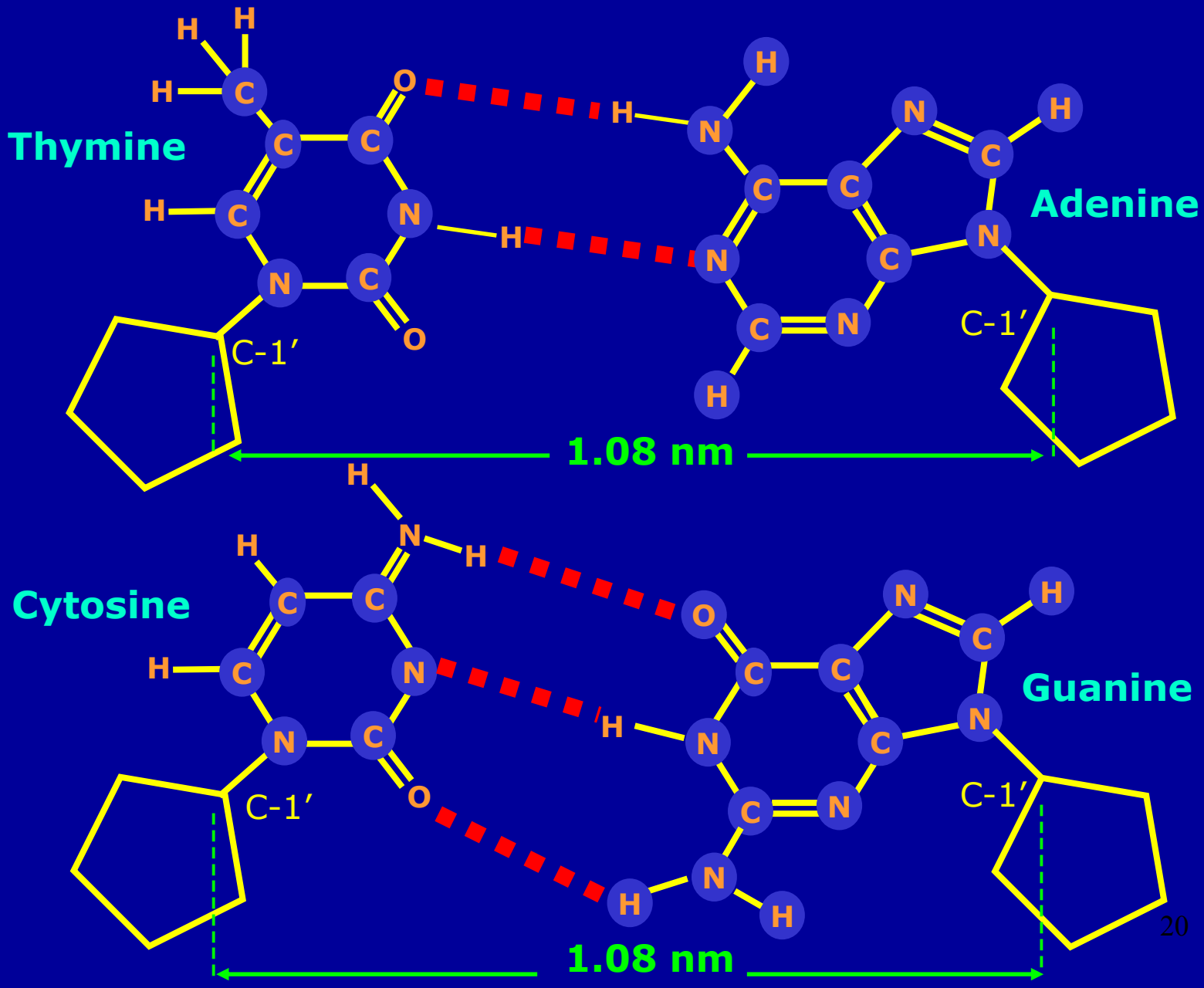
•**Phosphodiester bond are very much stable (fossils), but in the presence of catalysts, when NA are digested, this hydrolysis is rapid.**

•Catalysts: **Phosphodiesterase**

Base-pairing

- **In the DNA double helix, purines and pyrimidines face each other.**
- **The two polynucleotide chains in the double helix are connected by hydrogen bonds (relatively weak bonds compared to covalent bonds) between the bases.**
- **Chargaff's Rules (1949):**
  - **The No. of A residues is equal to the No. of T residues.**
  - **The No. of G residues is equal to the No. of C residues.**
  - **In human : Ratio (A+T)/(G+C) = 1.52**
- **Watson-Crick base-pairing rules:**
  - **A═T**
  - **G≡C**

Hydrogen Bonds



Complementarity: Base-pairing

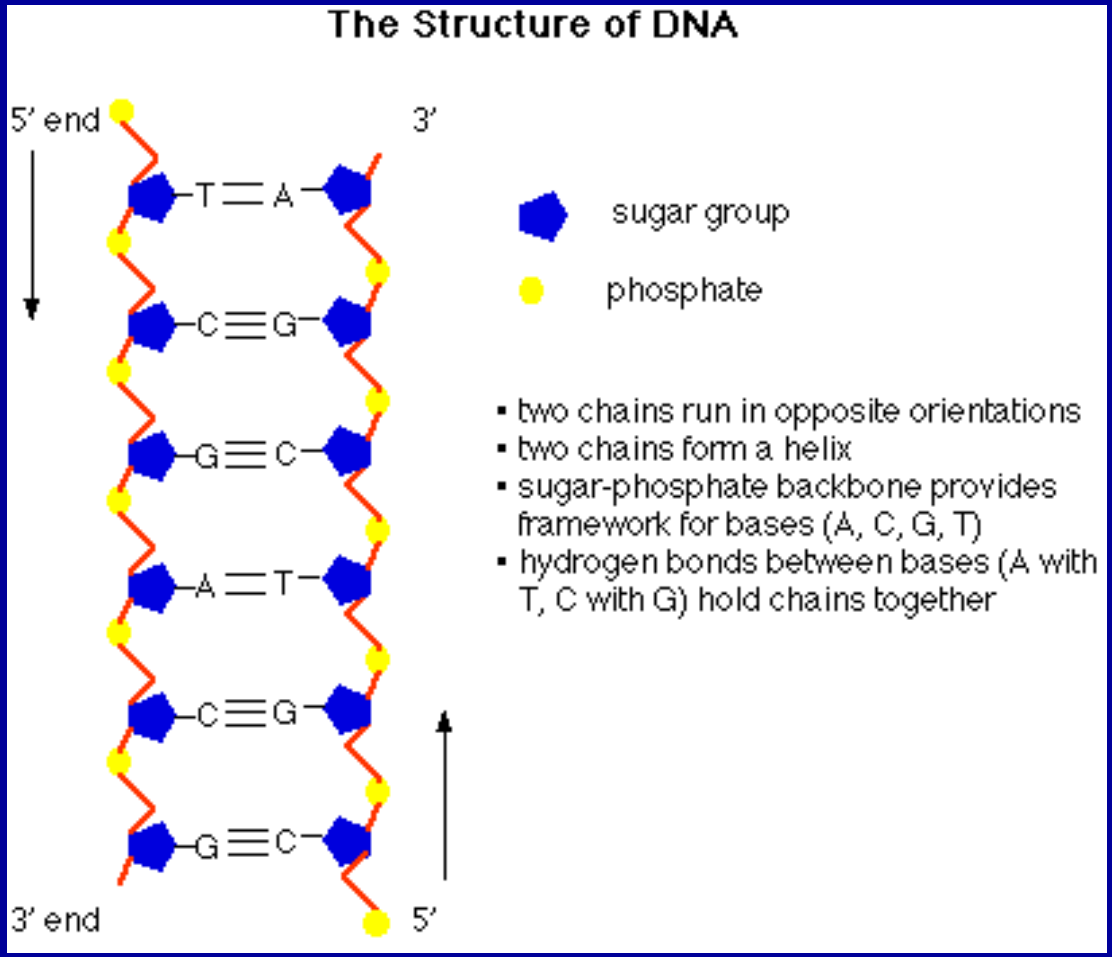
- **GC base pairs(bps)have more energy than AT bps.**
- **Since one strand of DNA is complementary to the other, genetic material can be accurately reproduced, therefore :**

  Each strand serves as the template for the synthesis of the other:

  The process of DNA replication / duplication.

The structure of DNA

- **Nitrogenous bases are in the middle.**
- **Sugar & phosphate are in the external parts.**
- **The molecule has negative charge.**



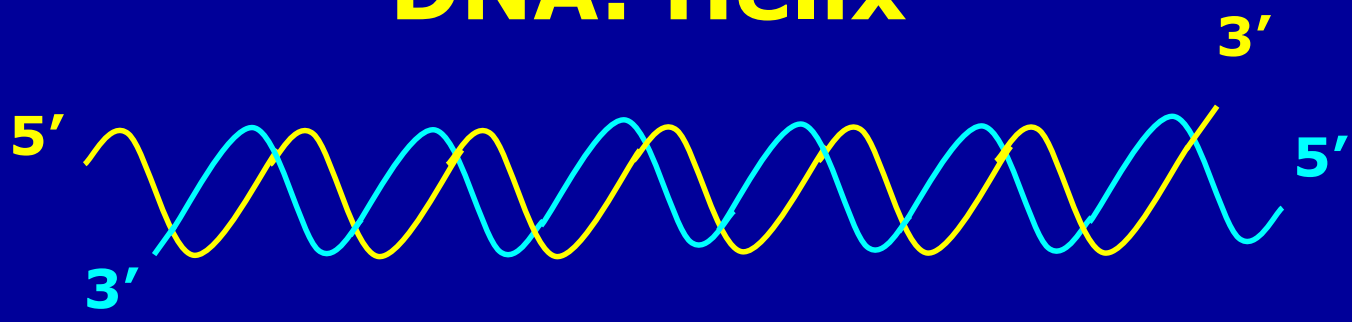
Antiparallel Chains

- Two strands of the DNA double helix are antiparallel and complementary to each other.
- In eukaryotic cells:
  - all DNA molecules in nucleus interact with basic protein (histons) with ionic interaction.

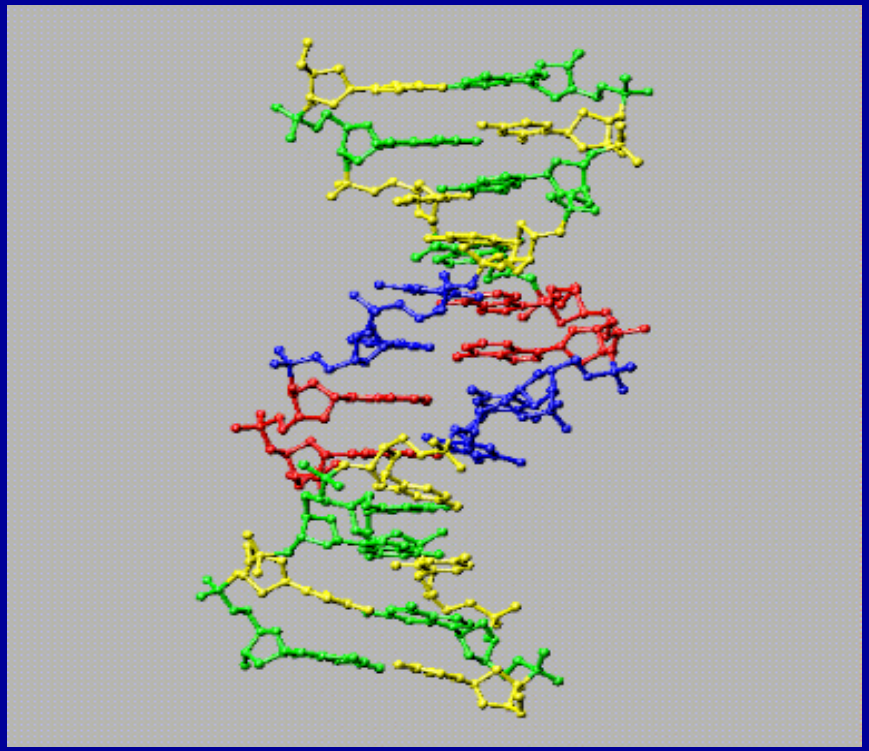
(DNA+ Histon: chromatin)

- Small amount of DNA in Mitochondria and Chloroplasts.

DNA: Helix

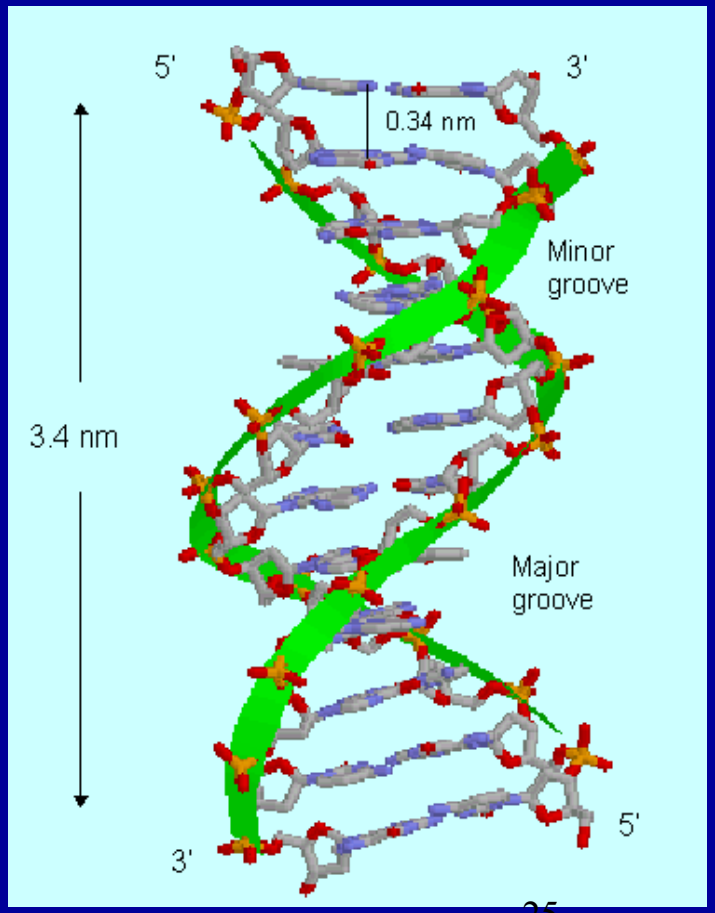


- **In general, The two strands form a "double helix" structure, which was first discovered by James D. Watson and Francis Crick in 1953.**



Conformations of Double-helical DNA

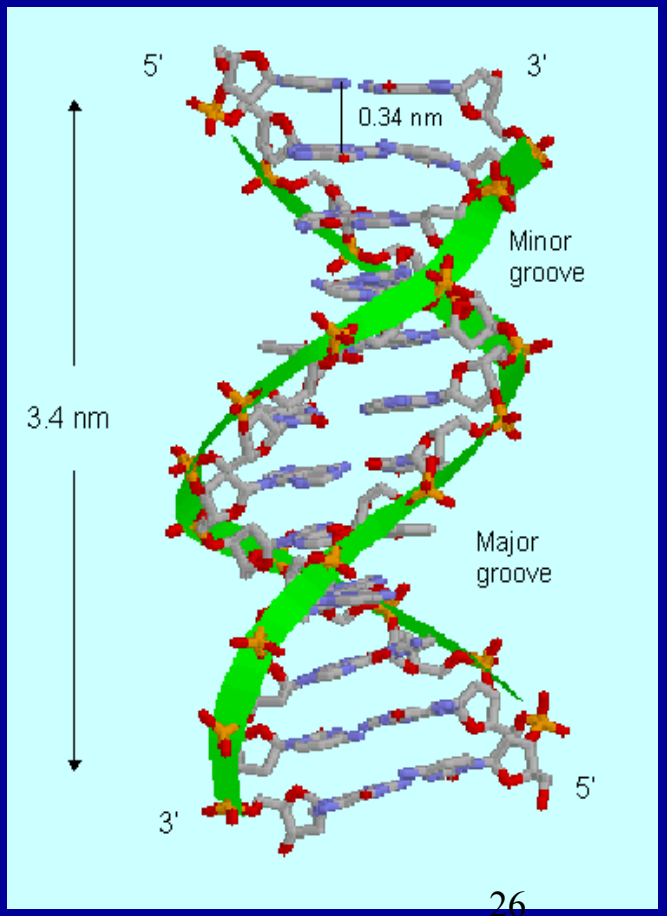
- **B-DNA (**the basis of Watson-Crick structure**):**
  - The great majority of the genome is organized into the duplex B-form of DNA:
    - High humidity/low salt
    - Double-stranded
    - Right handed helix
    - ~10 base pairs/turn of the helix
    - The diameter of the helix is 2 nm.



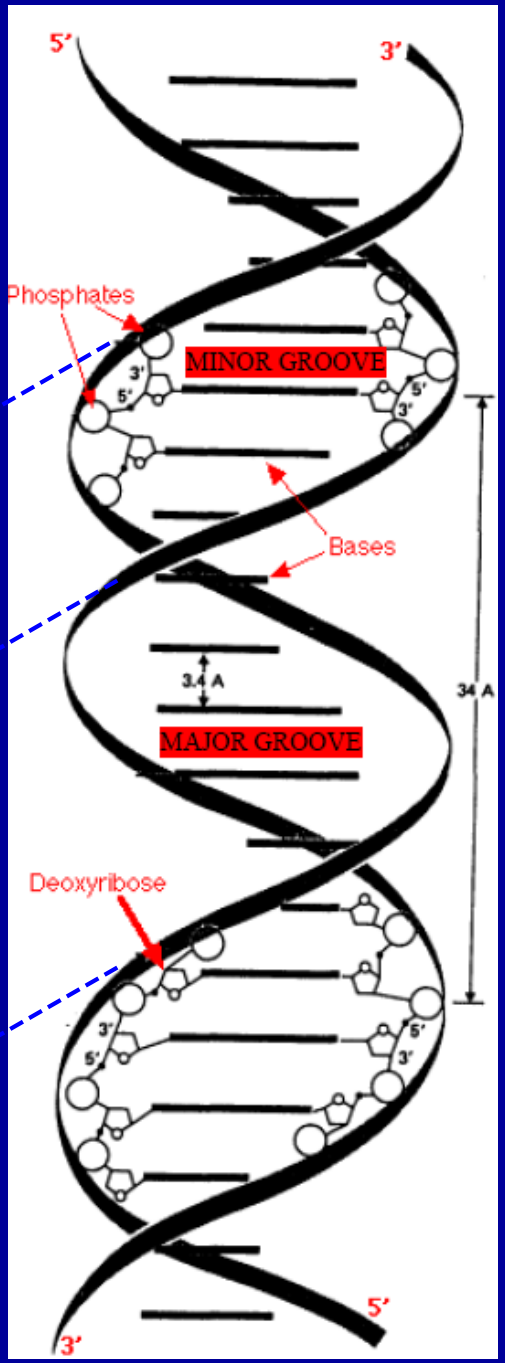
(Refer to table 2.3, page 40)

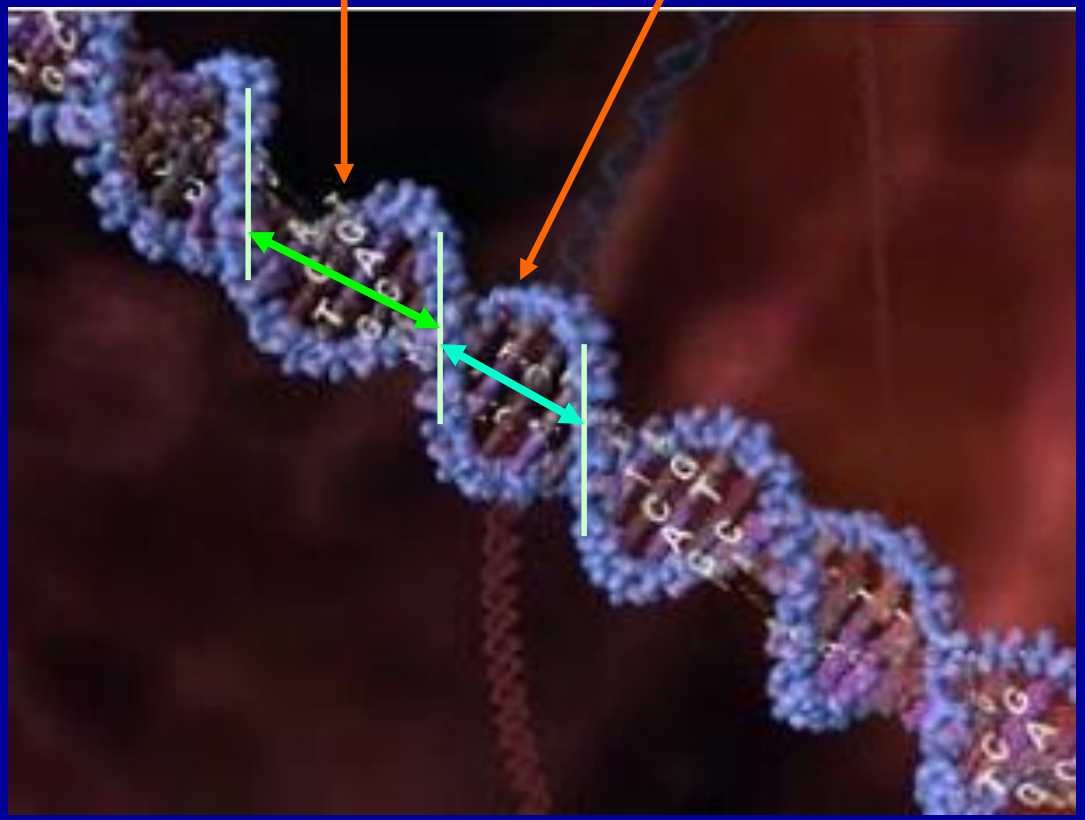
Conformations of Double-helical DNA

- **B-DNA:**
  - The major and minor grooves are roughly the same depth
  - The minor groove is relatively narrow.
  - Base pairs are nearly perpendicular to the helical axis.
  - Adjacent bases are separated by 0.34 nm along the helix axis.



Major groove & minor groove





(Refer to Biochemistry, J.David Rawn, p.669)

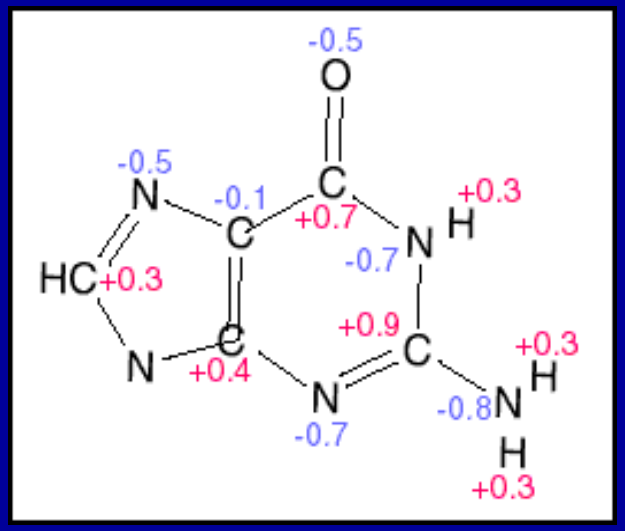
Polynucleotide conformations

- -NH2, =N-, =O groups:
  - Interaction with polar group & water
  - H-bonding between complementary bases
- Base stacking

- Important factors in stabilization:
  - pH
  - Salt concentration
  - Temperature

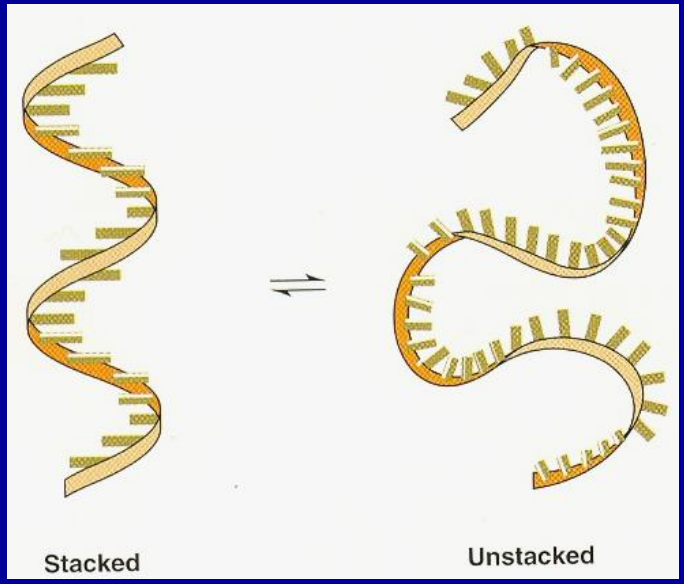
Base stacking

- Most of the bonds in the heterocyclic bases have a significant polar character due to the partial electrostatic charges on the atoms.
- The faces of the rings, however cannot participate in such interactions and tend to avoid contact with water.
- Stacking interactions, is a combination of:
  - Van der waals interactions
  - Hydrophobic forces



Base stacking

- Distance between two bases <5.5A°
- Estimation: generation of 4-15 kcal/mol of stabilization energy for each adjacent pair of stacked nucleobases.



Structures of DNA

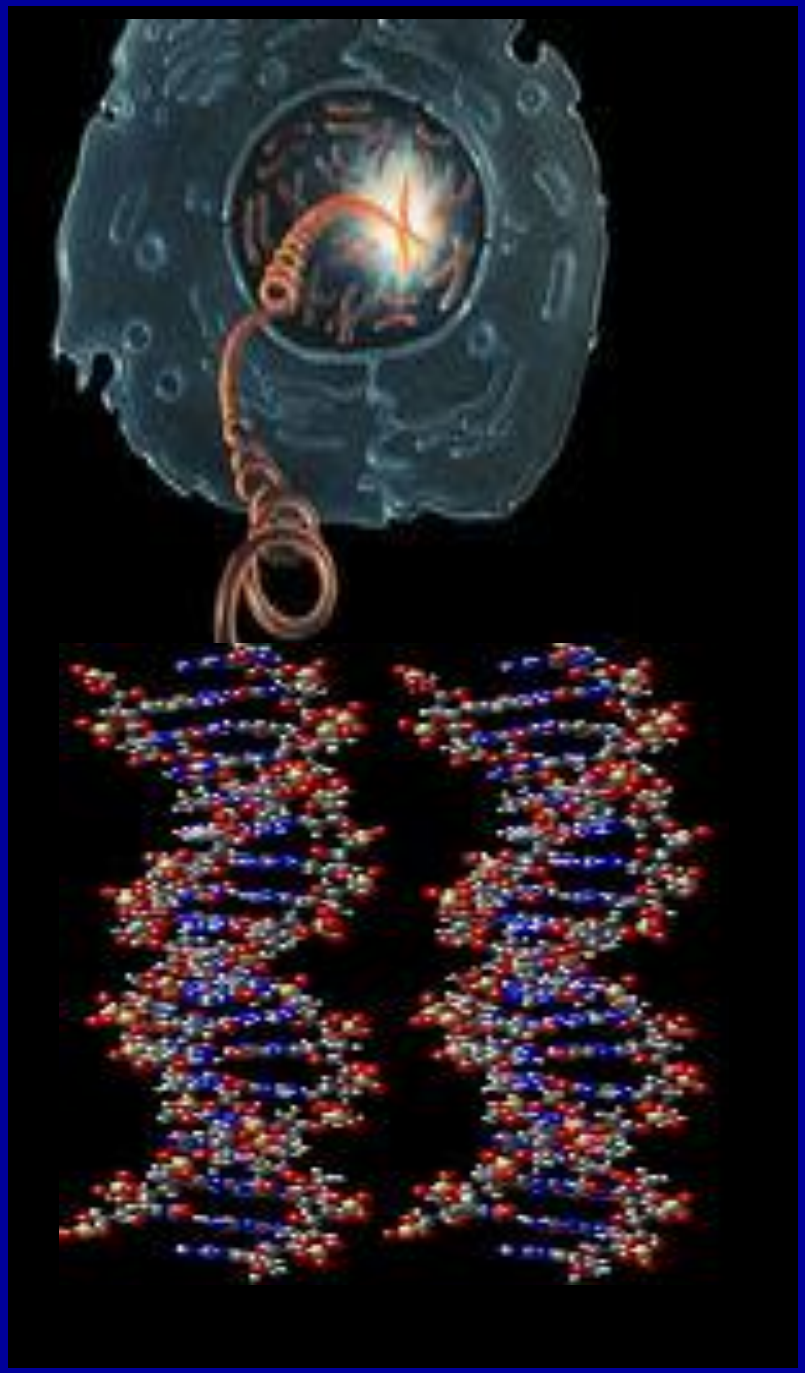
- **dsDNA has:**
  - **Primary structure (linear sequence)**
  - **Secondary structure (right handed double helix)**
  - **Tertiary/Quaternary structure**

  (it is folded and packed in the cell)

DNA PACKAGING

- A single diploid human cell contain $6x10^9$ bp of chromosomal DNA packaged in 46 chromosomes.
- So the total extended length of DNA is nearly 2 m, but this DNA must fit into a nucleus with a diameter of about 11-22 μm.

Mitchel, Campbell Reece. *Biology Concept and Connections. California, 1997.*	"At actual size, a human cell's DNA totals about 3 meters in length."	3.0 m
McGraw Hill Encyclopedia of Science and Technology. New York: McGraw Hill, 1997.	"If stretched out, would form very thin thread, about 6 feet (2 meters) long."	2.0 m
Matthews, Harry R. *DNA Structure Prerequisite Information. 1997.*	"The length is (length of 1 bp)(number of bp per cell) which is (0.34 nm)(6×10^9)"	2.0 m
Leltninger, Albert L. *Biochemistry. New York: Worth, 1975.*	"Chromosome 13 contains a DNA molecule about 3.2 cm long."	1.5 m
"Cell." *The World Book Encyclopedia. Chicago: Field Enterprises, 1996.*	"On the average, a single human chromosome consists of DNA molecule that is about 2 inches long."	2.3 m

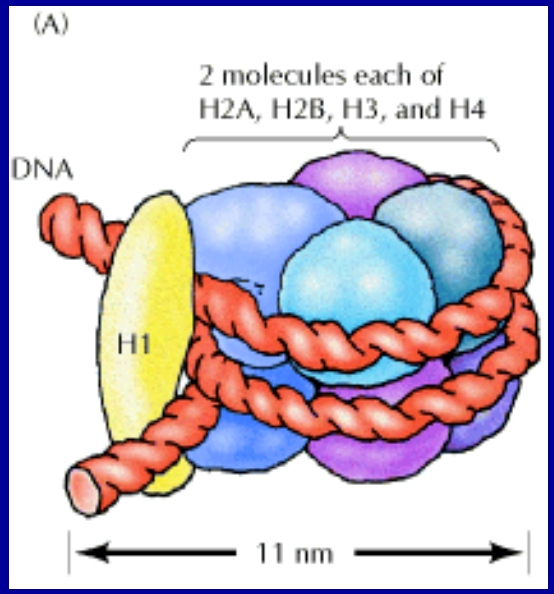


DNA PACKAGING:

Structure of a chromatosome

- The **nucleosome** core particle consists of **146 base pairs** of DNA wrapped **1.65 turns** around a **histone octamer** consisting of two molecules each of **H2A**, **H2B**, **H3**, and **H4**.
- A **chromatosome** contains **two full turns** of DNA (**166 base pairs**) locked in place by one molecule of **H1**.

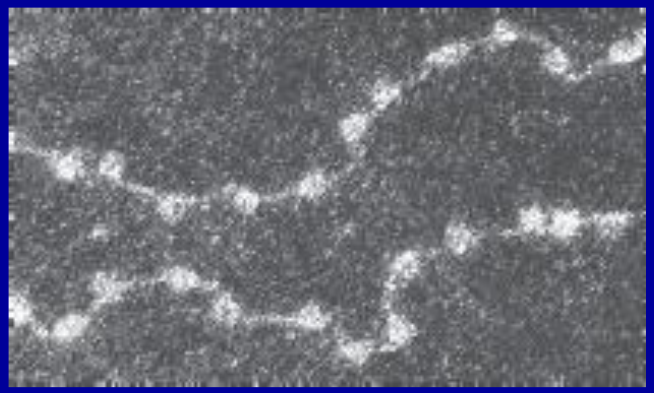
The DNA is wrapped around histones in nucleosome core particles and sealed by histone H1.



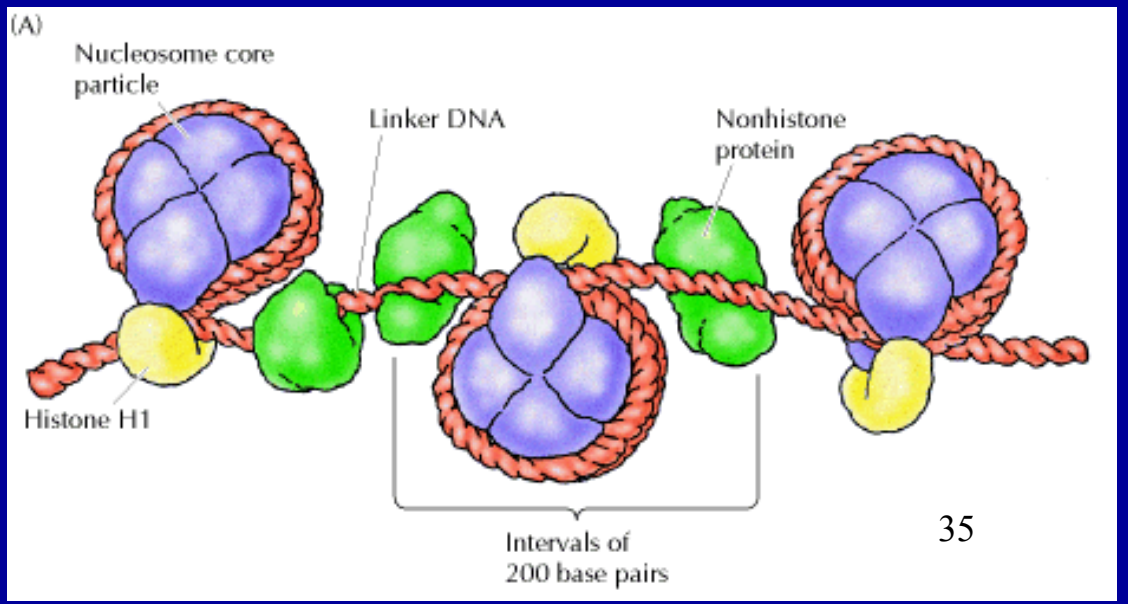
The organization of chromatin in nucleosomes

- Nonhistone proteins bind to the linker DNA between nucleosome core particles.

The packaging of DNA into nucleosomes yields a chromatin fiber approximately 10 nm in diameter.

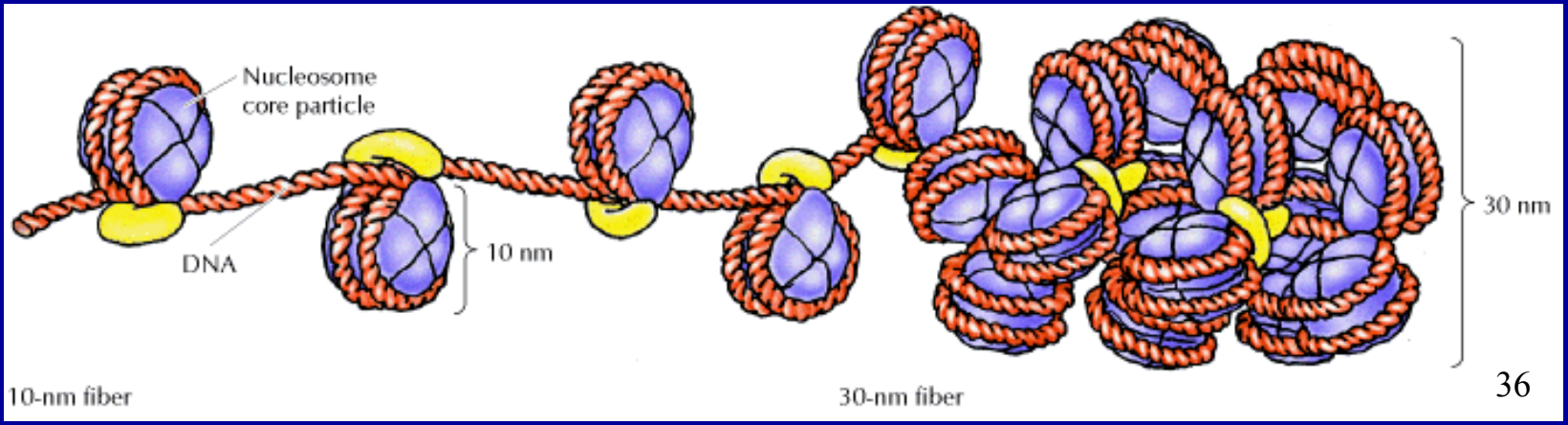


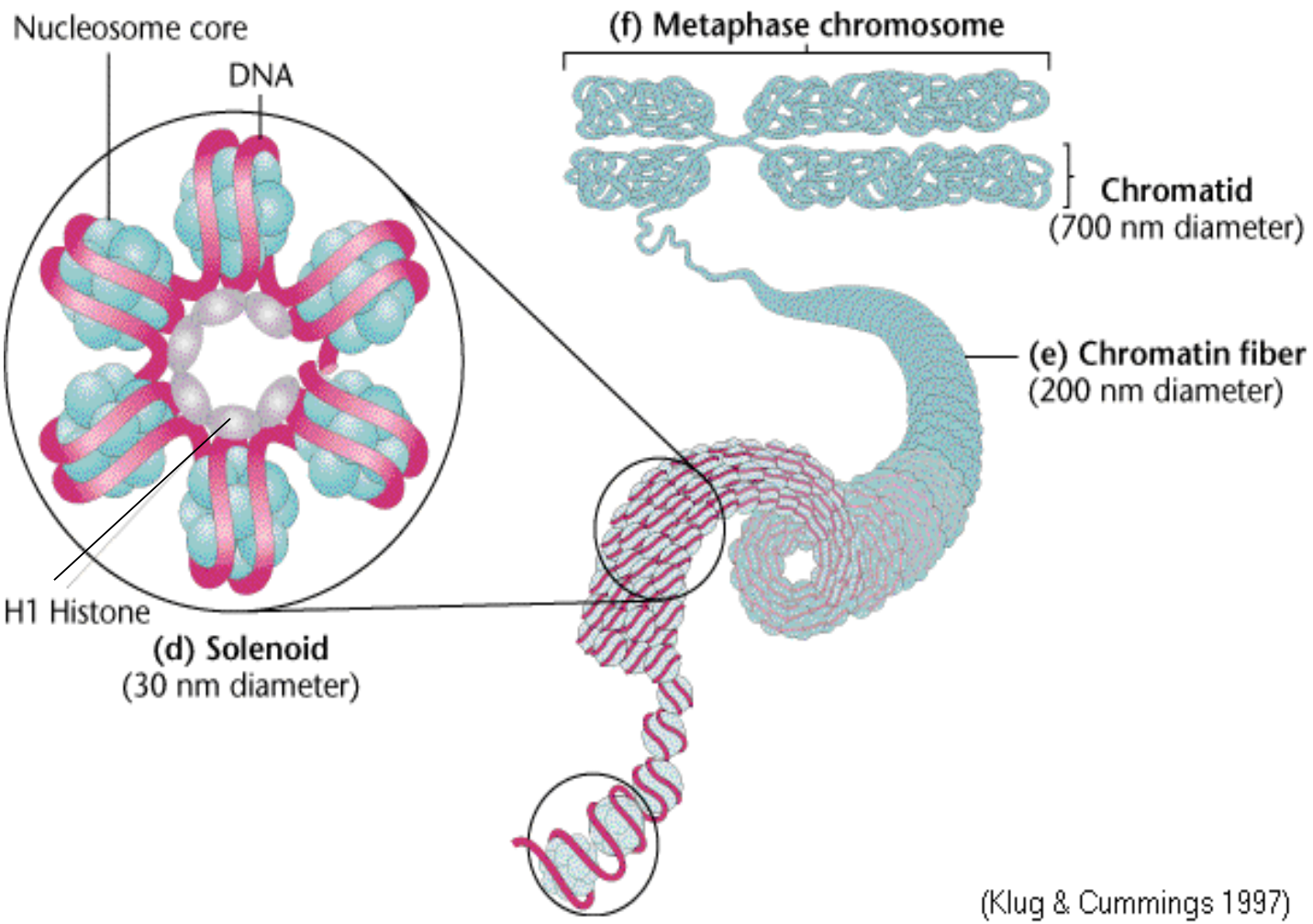
Electron microscopy photo

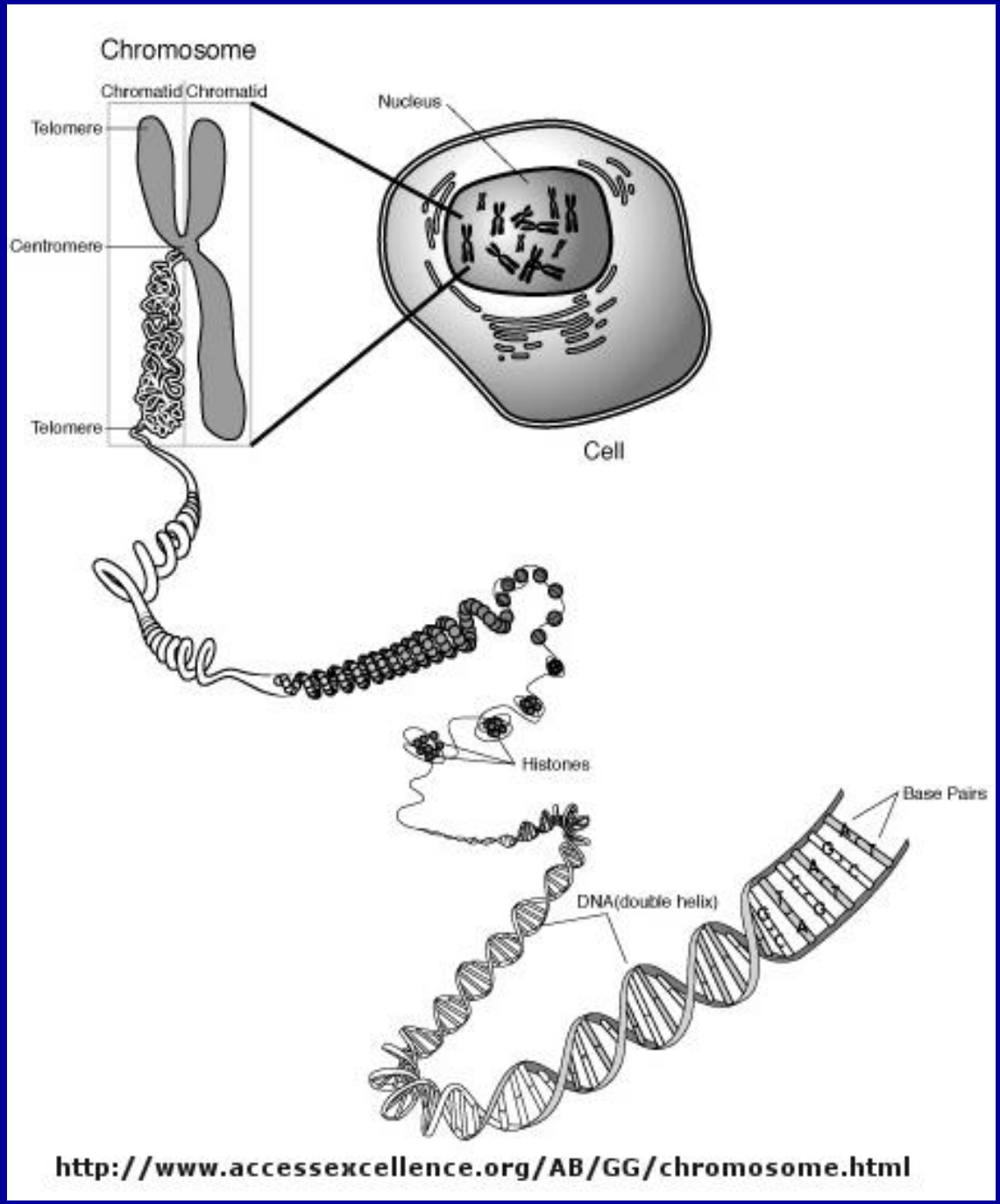


Chromatin fibers

- The chromatin is further condensed by coiling into a 30-nm fiber, containing about six nucleosomes per turn.




Nucleosome core
DNA
H1 Histone
(d) Solenoid
(30 nm diameter)
(f) Metaphase chromosome
Chromatid
(700 nm diameter)
(e) Chromatin fiber
(200 nm diameter)
(Klug & Cummings 1997)


Chromosome
Chromatid
Chromatid
Telomere
Centromere
Telomere
Nucleus
Cell
Histones
Base Pairs
DNA(double helix)
http://www.accessexcellence.org/AB/GG/chromosome.html

- Nucleosome - achieved by winding DNA around a protein core to produce a "bead-like" structure; packing ratio *6.

- 30 nm fiber - a coil of nucleosomes in a helical structure; found in both interphase and mitotic chromosomes; increases the packing ratio to *40.

- 700 nm structure and scaffolding - organization of the fibers in loops and scaffolds; gives a final packing ratio of about 1000 in interphase chromosomes and *10,000 in mitotic chromosomes.

Circular DNA

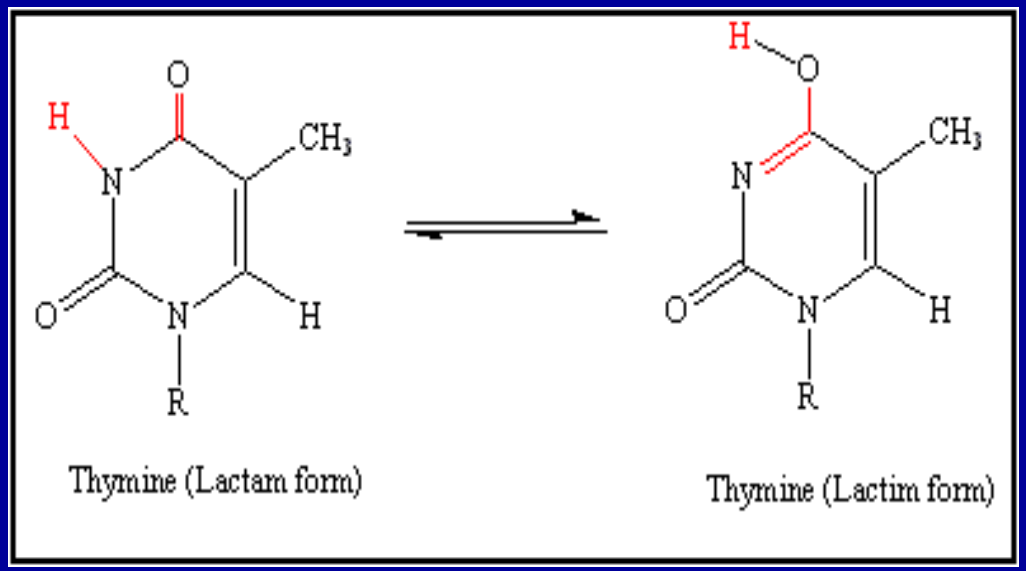
- In mitochondria (and most prokaryotes) DNA is in circular form, including genes encoding for:
  - tRNA (22 types)
  - rRNA (2 types)
  - Protein (13 types)

Physicochemical properties of nucleobases and polynucleotides

- Aqueous Solubility:
  - Nucleotides > Nucleosides> Nucleobases
- Thymine dimer formation
- Methylation of nucleobases
- Spontaneous deamination (C⇨U, G⇨X, A⇨HX)
- Tautomerism
- Denaturation and renaturation
- Pu and Pyr bases strongly absorb UV light (at 260 nm)

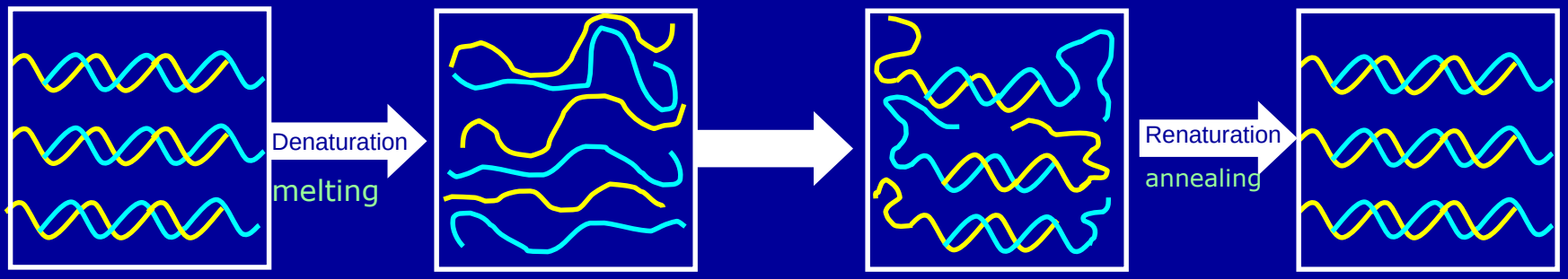
TAUTOMERIC FORMS OF BASES

- Two possibilities
  - KETO (Lactam)
  - ENOL (Lactim)



- Important in order **to specify hydrogen bonding** relationships
- The keto form predominates

Denaturation - Renaturation



Double-stranded DNA

Single-stranded DNA

Initial Base pairing

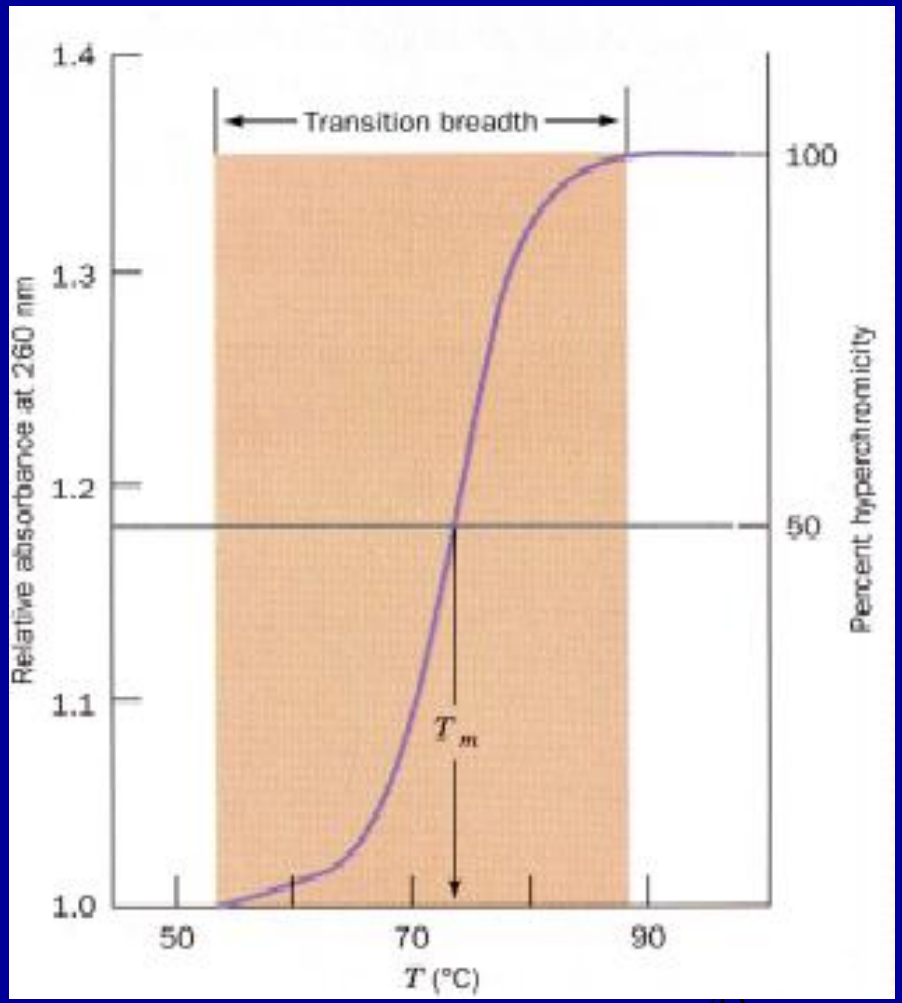
Renatured DNA

Monitoring DNA denaturation and renaturation

- Study the change of A260.
- Temperature at which the change in A260 is half maximal is called Tm.
- single-stranded DNA is hyperchromic (higher A260 than double-stranded DNA)
- Double-stranded DNA with a high GC-content has a higher Tm than DNA with a lower GC-content .

Temperature optical density profile for DNA (Melting curve)

- When DNA is heated, the absorbance at 260 nm increases with rising temperature.
- When native (double strand) DNA is heated above its “melting” temperature, it is denatured (separates into single strands).



Conditions favoring denaturation

- High temperature (breaks weak hydrogen bonds holding base pairs together, but **NOT** the strong phosphodiester bonds).
- Low salt concentrations (DNA is a polyanionic molecule. High salt concentrations "shield" the negative charges on each phosphate. When the charges are **NOT** shielded in low salt concentrations, the electrostatic repulsion of the negatively charged strands makes it energetically more favorable to separate the strands).
- High (basic) pH (also breaks hydrogen bonds).

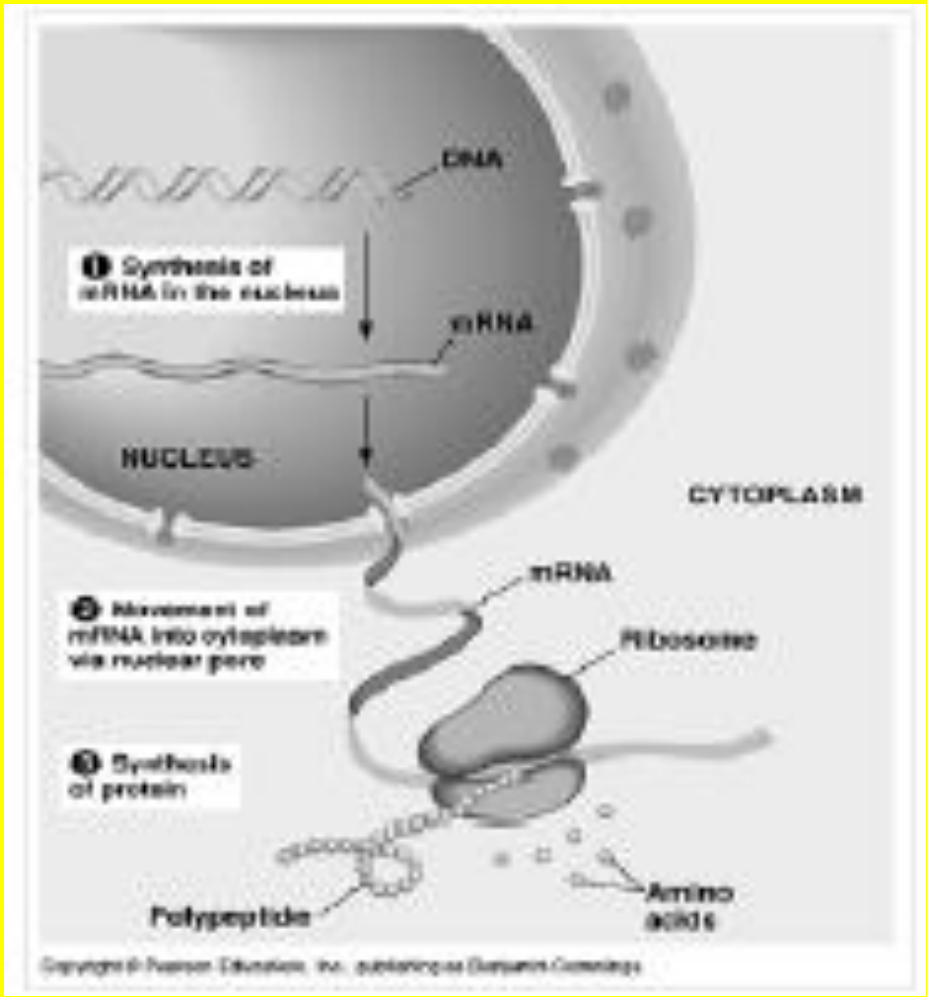
Conditions favoring renaturation

- Temperature lower than the Tm (allows hydrogen bonds to form between base pairs).
- Higher salt concentration (cations shield negatively charged phosphates, eliminating the repulsive force between the two antiparallel strands).
- Neutral pH (also allows hydrogen bonds to form).

…..Monitoring DNA de/re-naturation

- The melting temperature also depends on the salt concentration:
  - in low salt, a given DNA will melt at a lower temperature than in a higher salt concentration.

Central Dogma of Flow of Genetic Information



Crick 1958

- **Transcription of DNA to RNA to protein**

Nature, vol. 227, pp. 561-563, 1970

The **Central Dogma** of Molecular Biology

- This dogma forms the backbone of molecular biology and is represented by four major stages:
- 1. The DNA replicates its information in a process that involves many enzymes: replication.
- 2. The DNA codes for the production of messenger RNA (mRNA) during transcription.

…..The Central Dogma

- 3. In eucaryotic cells, the mRNA is processed (essentially by splicing) and migrates from the nucleus to the cytoplasm.
- 4. Messenger RNA carries coded information to ribosomes. The ribosomes "read" this information and use it for protein synthesis. This process is called translation.
- Proteins do not code for the production of protein, RNA or DNA:

Proteins are involved in almost all biological activities, structural or enzymatic.

Ribonucleic acids (RNA)

- There are three common types of ribonucleic acids on the base of their function (in protein production):

	in Ecoli
• mRNA (messenger RNA)	2%
• tRNA (transfer RNA)	16%
• rRNA (ribosomal RNA)	82%

* snRNAs and regulatory ncRNAs

Ribonucleic acids (RNA)

Small RNA

- Most of these molecules are complexed with proteins to form ribonucleoproteins in the nucleus or cytoplasm or both.
- They range in size from 20 to 1000 nucleotides
- 100,000 to 1,000,000 copies per cell,
- Collectively : ≤5% of cellular RNA.

Ribonucleic acids (RNA)

Small RNA:

-Small Nuclear RNAs (SnRNA, …)

- Large & Small Noncoding Regulatory RNAs:

- Micro-RNAs (miRNAs),
- Silencing RNAs (siRNAs),
- Long Noncoding RNAs (lncRNAs),
- Circular RNAs (circRNAs)

Varieties of RNA

- Messenger RNA:
  - The structure that carries the genetic code instructions for new proteins and enzymes production.
  - mRNA provides the physical link between the DNA and the actual polypeptide chains.

…..m RNA

- Every three nitrogenous bases in the molecule makes up a unit in this code called a codon, and each codon, represents a specific amino acid.
- Because there are four different nitrogenous bases and three bases in a codon, there are 64 (4x4x4=64) possible units in the genetic code system.

…..m RNA

- In eukaryotes mRNA is not created ready to work, it must go through "posttranscriptional modification" to purify its information.
- It contains many noncoding sequences called introns that must be removed by enzymes then the relevant portions (exons) remain.

……mRNA (Eukaryotes)

- **hnRNA (heterogeneous nuclear RNA): Eukaryotic mRNA primary transcripts** (in nucleus) **whose introns have not yet been excised (pre-mRNA).**

- **snRNP (small nuclear Ribo-Nucleo-Protein): a complex of (usually) a single snRNA (small nuclear RNA) and several (~10) nuclear proteins.**
  - **It binds to the RNA that is to be spliced, catalyzes the excision of the introns and covalent joining of the exons.**

……mRNA (Eukaryotes)

The final mRNA is arranged as follows:

- **a leading segment, called the 5’ “cap”** (7-methyl guanosine 3-p; functions in recognition by ribosome).
- **The actual gene, or cistron.**
- **The end of the molecule is attached to many adenine bases (poly-A tail) in 3’-end** (several hundred A nucleotides, functions in stabilization of mRNA).

……mRNA (prokaryotes)

- In prokaryotes mRNA is transcribed from the DNA in its finished form.
- No posttranscriptional modification is necessary.
  - mRNA lacks a cap and a poly-A tail.
  - mRNA is polycistronic.
    - It means that many of these strands contain more than one gene.

    (many different proteins are created within the same small area).

Varieties of RNA

- **Transfer RNA:**
  - **It transfer the genetic messages from mRNA into polypeptide chains in the ribosomes.**
  - **For each amino acid at least one tRNA exists.**
  - **Normally it has 73-95 nucleotides units.**
  - **They are diverse in their primary structure, but have remarkable similarity in their secondary and tertiary structure.**

…tRNA: Secondary structure

- **Clover leaf:**

• It's due to four base-paired stems.

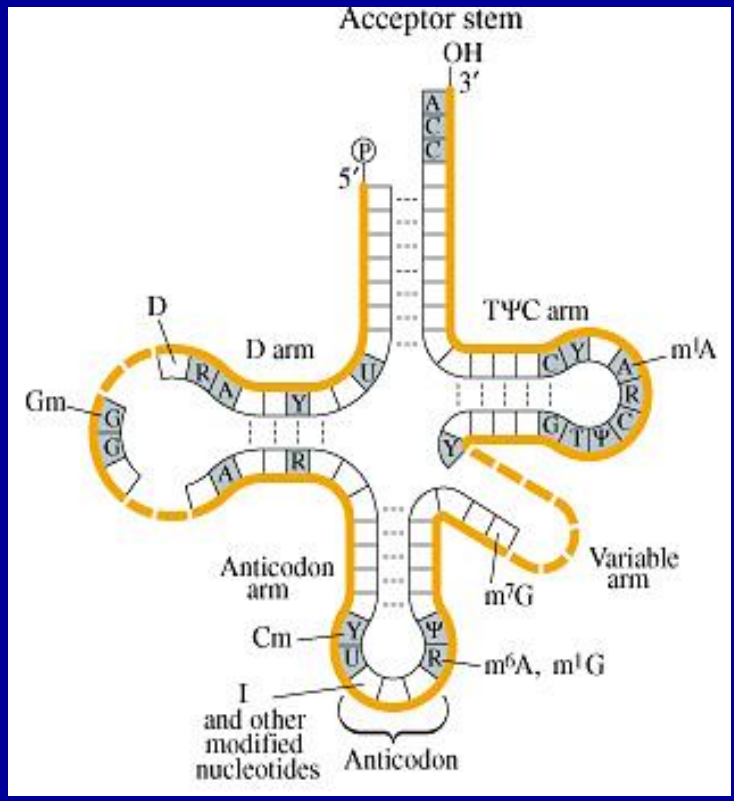
• It contains three non-base-paired loops:

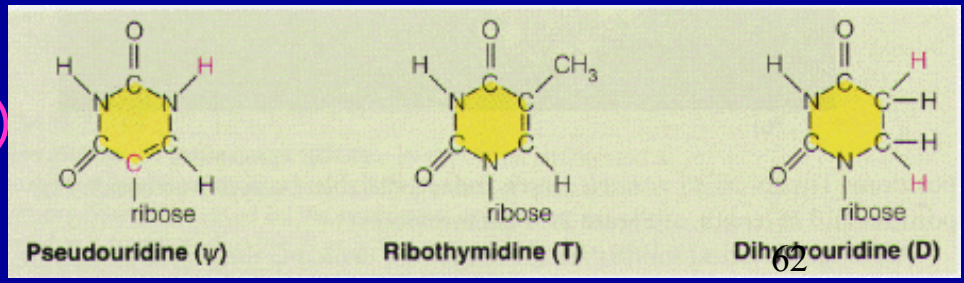
- **D loop**

(D:dihydrouridine)

- **Anticodon loop**
- **TpsiC loop (TψC)**

(T:ribothymidine;
Ψ:pseudouridine: R- C5)





…tRNA

- 3' end *always terminates* with the sequence CCA, with the 3‘OH of the ribose of the terminal A being the point of covalent attachment of the amino acid.

- It contains a number (7-15%) of unique/modified bases. These are post-transcriptionally modified after synthesis by RNA polymerase.

…tRNA

- ***Wobble Theory*-**1966:

(Some amino acids are being coded by more than one codon, i.e.. Leu: CUU,CUC, CUA, CUG)

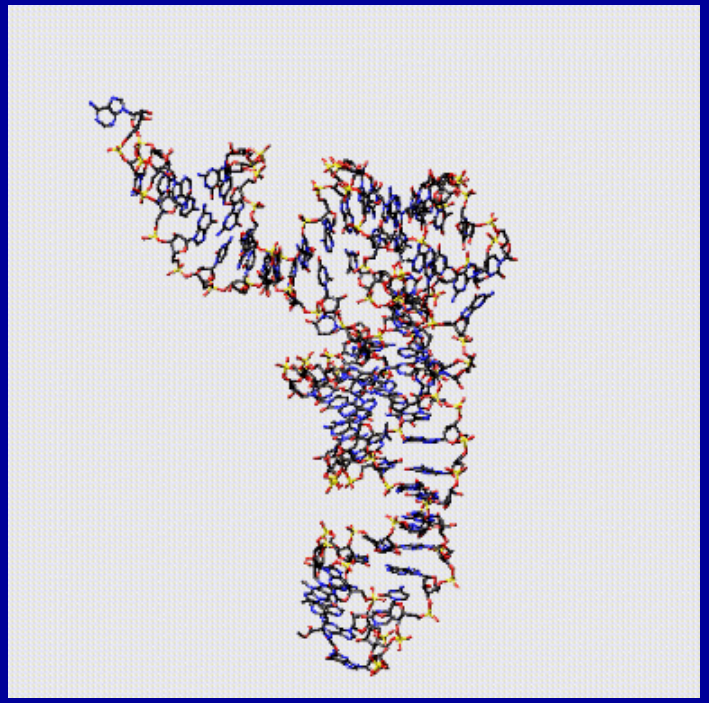
- **adenosine (A) in first or 5' position of the anticodon** (corresponding to the third or 3' position of the codon) **is always modified to inosine (*I).**
- **Inosine can base-pair with A, G, U or C, therefore a single tRNA could recognize several different codons.**

* nucleosidic form of hypoxanthin

…tRNA: Tertiary structure

- *L* shape:

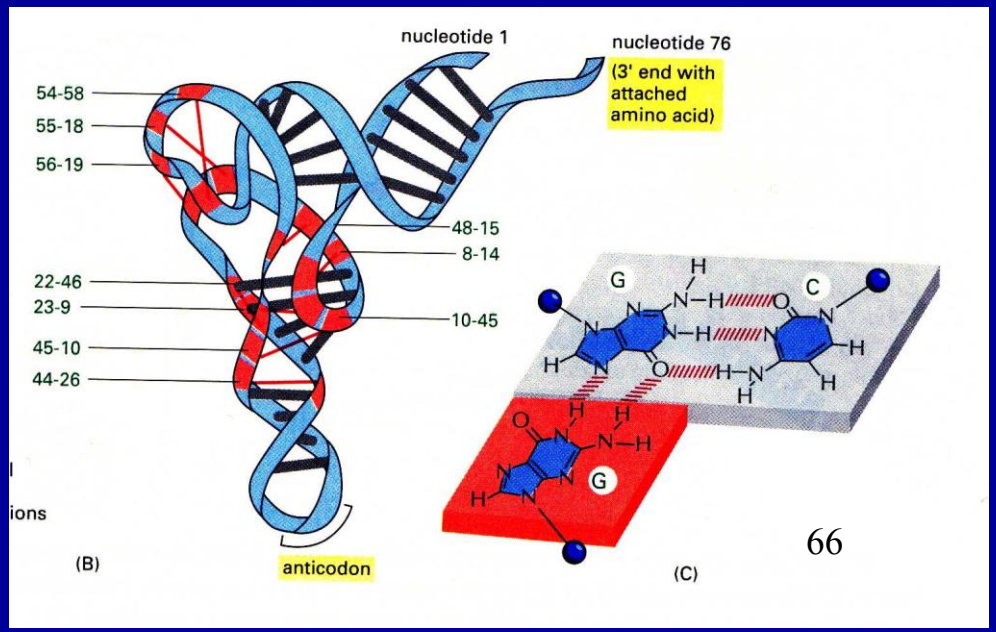
- The *anticodon* is a single-stranded loop which base-pairs with the triplet codon.
- The amino acid is attached to the terminal A on the upper right.
- The active sites (*anticodon* and *amino acid*) are maximally separated.



…tRNA

- As in proteins, the tertiary structure is dictated by the primary sequence.
- The tertiary structure is stabilized by base pairing and base stacking.

- Bases sometimes make hydrogen bonds with more than one partner.



Varieties of RNA

- Ribosomal RNA:
  - plays a crucial role in the creation of ribosome (RNA+protein), Possibly have enzymatic activity.
  - It is transcribed in the nucleolus (all other types are not), from DNA segments that are looped into structures called nucleolar organizer regions.

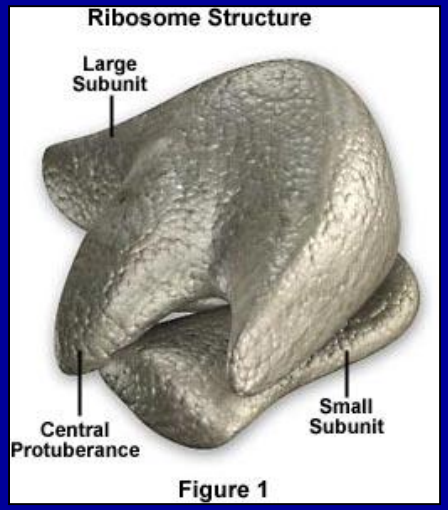
…..r RNA

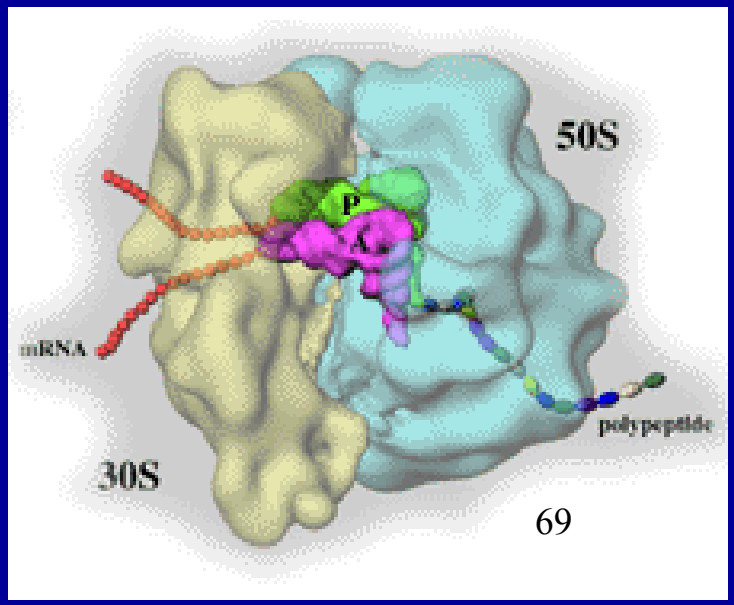
- Because rRNA is often in high demand in a cell, the data necessary for their creation is often repeated a number of times on the DNA strand.

- This repetition allows thousands of identical molecules to be transcribed simultaneously.

Ribosomal RNA

- **Ribosome** – multimolecular structures that act as factories for protein synthesis.
- Ribosomes composed of 2 subunits





Structure of ribosomes

- Ribosomes composed of 2 subunits:
  - Prokaryotes (E. coli) = 50S + 30S
  - Eukaryotes = 60S + 40S

Svedberg unit

- Ribosome pieces measured in S units (Svedberg unit)
- Measured by sedimentation analysis
  - Measures rate at which particle sediments through dense solution.
  - Depends on mass and shape
  - Larger S = Larger macromolecule

Structure of ribosomes

Prokaryotes: 50S + 30S

- 50S
  - 2 rRNA molecules (*23S + 5S)
  - 34 polypeptides
- 30S
  - one rRNA molecule (16S)
  - 21 polypeptides

*the active site for peptide bond synthesis

Structure of ribosomes

Eukaryotes: 60S + 40S

- 60S
  - 3 rRNA molecules (28S +5.8S + 5S)
  - 50 polypeptides
- 40S
  - one rRNA molecule (18S)
  - 33 polypeptides

*** The rRNAs (28S + 18S) represent roughly 70% of total cellular RNA.**
• large rRNA component performs the peptidyl transferase activity and thus is a ribozyme.

Ribosomal RNA

- For efficiency, each rRNA should be equally synthesized.
  - Solved by gene clustering:
    - All transcribed in one RNA strand (pre-rRNA)

    (The 45S precursor rRNA gives rise to the 28S, 18S and 5.8S rRNAs, cleaved and trimmed to final product).

*5S rRNA is not transcribed in the nucleolus (from a separate gene within the nucleoplasm)