- A **nucleic acid** is a complex, high-molecular-weight biochemical macromolecule found in all living cells and viruses.

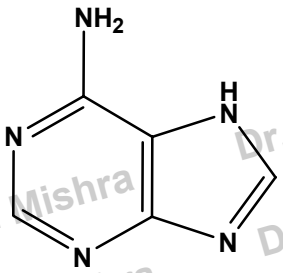
- They are composed of nucleotide chains that convey genetic information.

- They are responsible for the storage and transmission of genetic material

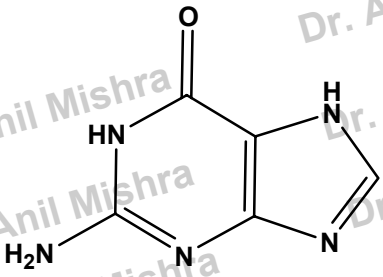
- The most common nucleic acids are
  - **Deoxyribonucleic acid** (DNA)
  - **Ribonucleic acid** (RNA).
- Are biopolymers having a alternating phosphate and sugar moieties

STRUCTURE

- Comprises of three moieties
  - **Heterocyclic bases**
    - **Purine**

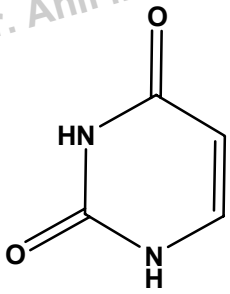


Adenine

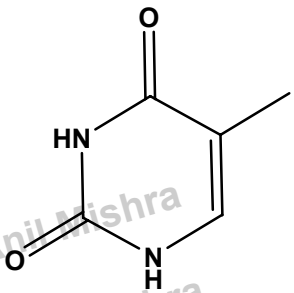


Guanine

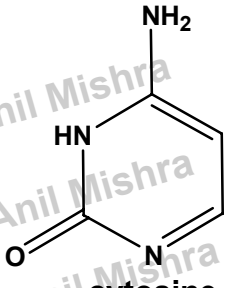
    - **Pyrimidine**



uracil



Thymine



cytosine

  - **Sugar**
    - Ribose
    - Deoxyribose
  - **Phosphate**

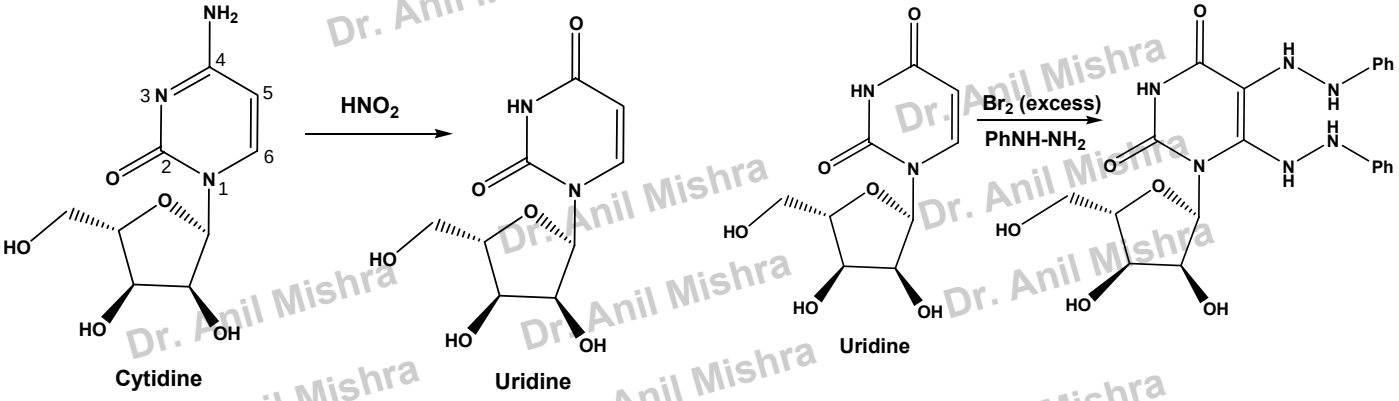
- Heterocyclic bases are attached to the sugar moiety
- Sugars and phosphate linked through **phospohdiester** bonds

STRUCTURE OF NUCLEOSIDES

- Hydrolysis of nucleotides give
  - Nucleoside
  - Phosphoric Acid

- Hydrolysis of nucleosides give
  - Heterocyclic Base
  - Sugar

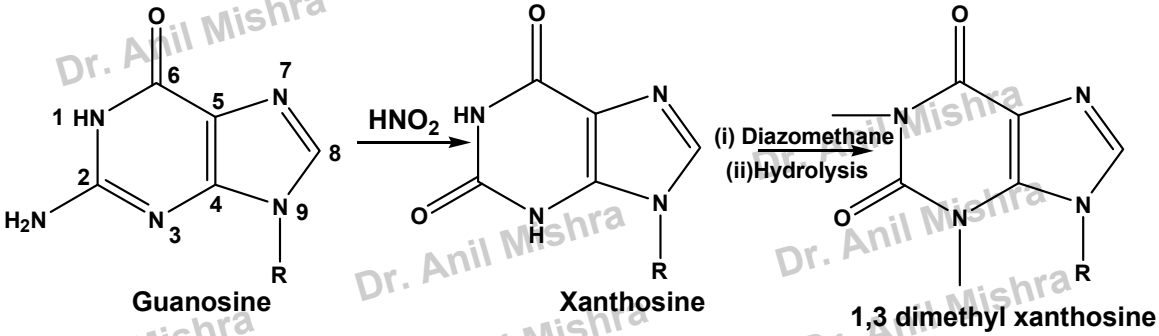
POINT OF LINKAGE OF SUGAR IN PYRIMIDINES



- This indicates that the point of linkage of sugar is the same in both the nucleosides
- Point of linkage cannot be 3 or 4 as cytidine has a free amino group at position 4 and consequently there cannot be hydrogen atom at position 3
- Uridine form a 5-bromo derivative therefore position 5 must be free
- Uridine on treatment with excess bromine followed by reaction with phenyl hydrazine gives a compound having two phenyl hydrazine groups indication that position the sugar cannot be linked to position 6 as well

POINT OF LINKAGE OF SUGAR IN PURINES

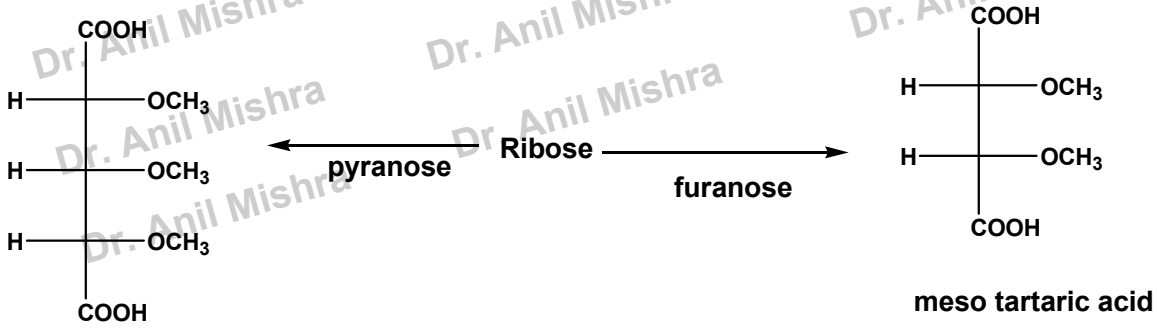
- Adenosine has a free amino group at position 6 therefore the sugar cannot be at this position or at position1.
- Guanosine has a free amino group at position 2 therefore the point of linkage cannot be 2 or 3.
  - The only positions where the sugar can be linked is 7, 8 or 9.
- This reaction proves this



- This reaction also indicates that Position 8 is free because:
  - Trimethyl compound should have been formed
  - Sugar at position 8 will give a C-nucleoside which is difficult to hydrolyze
- The UV spectra of guanosine was identical to that of 9-methyl guanine while it was different from that of 7 methyl guanine
- **This confirms that the sugar is attached to position 9 in purine nucleosides**

NATURE OF SUGAR MOIETY

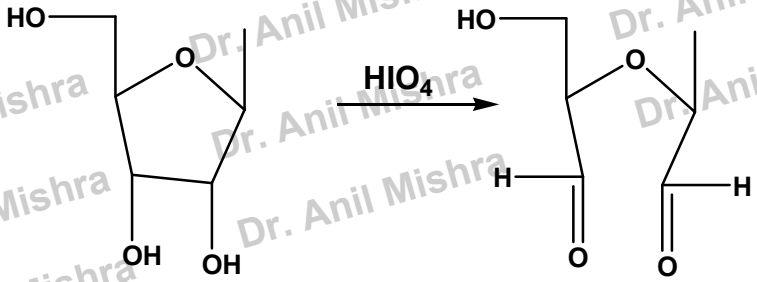
- Degradative experiments indicate that the sugar is present in the furanose form



trimethoxy glutaric acid

meso tartaric acid

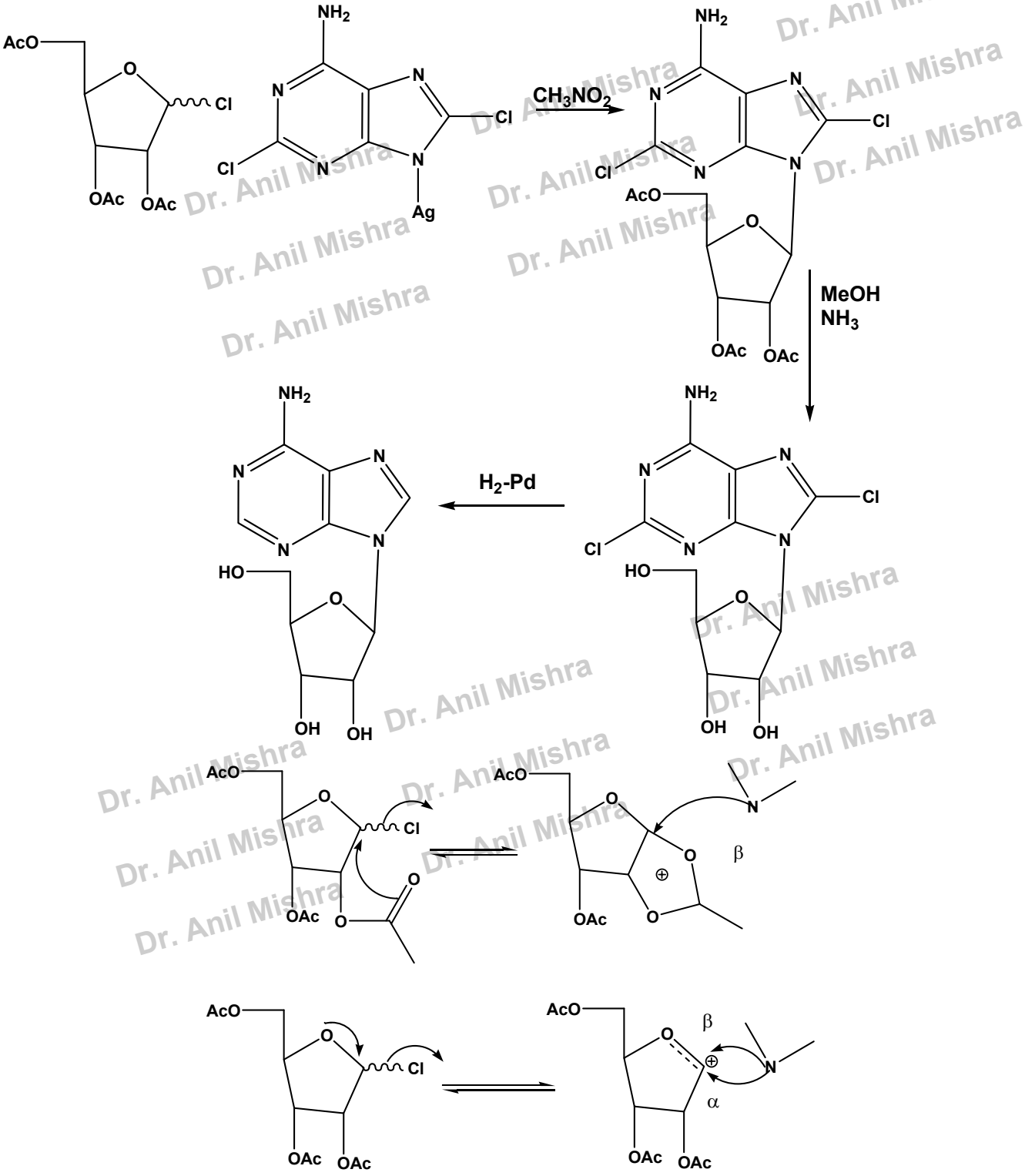
- Oxidation with periodic acid forms a dialdehyde with no loss of carbon



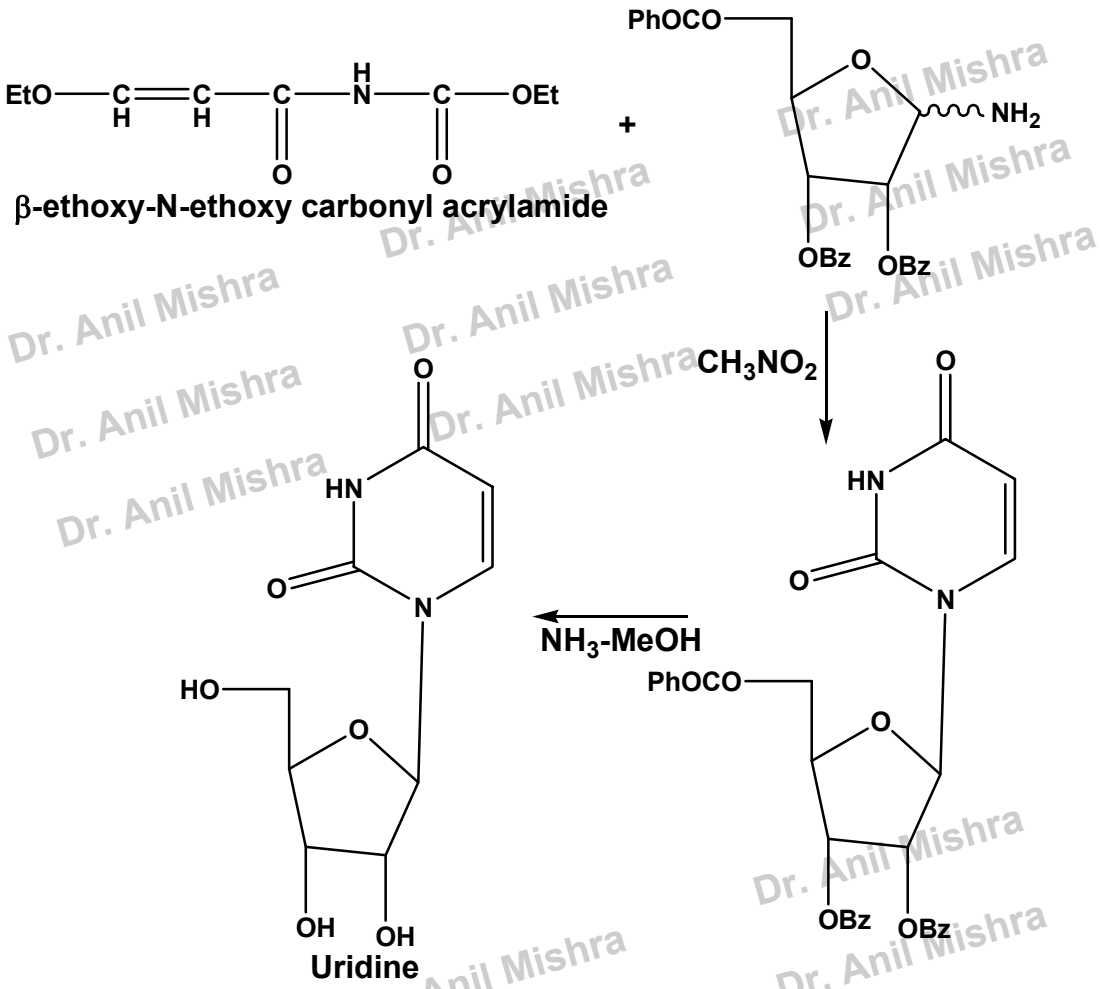
- The nucleosides of natural origin have the sugar in the **β** orientation

SYNTHESIS OF PURINE NUCLEOSIDES

- Silver Salt method (Todd et.al.)

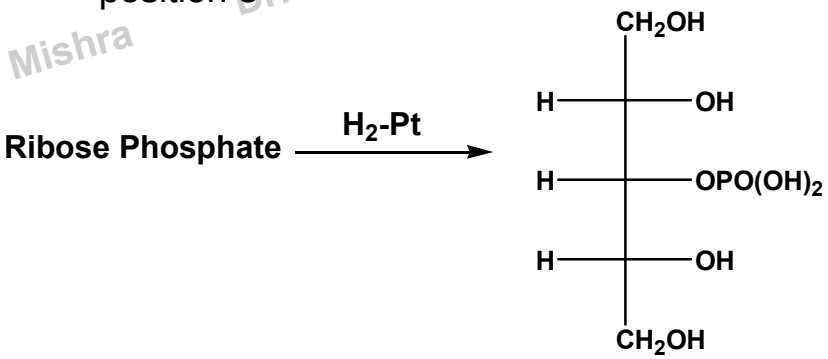


SYNTHESIS OF PYRIMIDINE NUCLEOSIDES

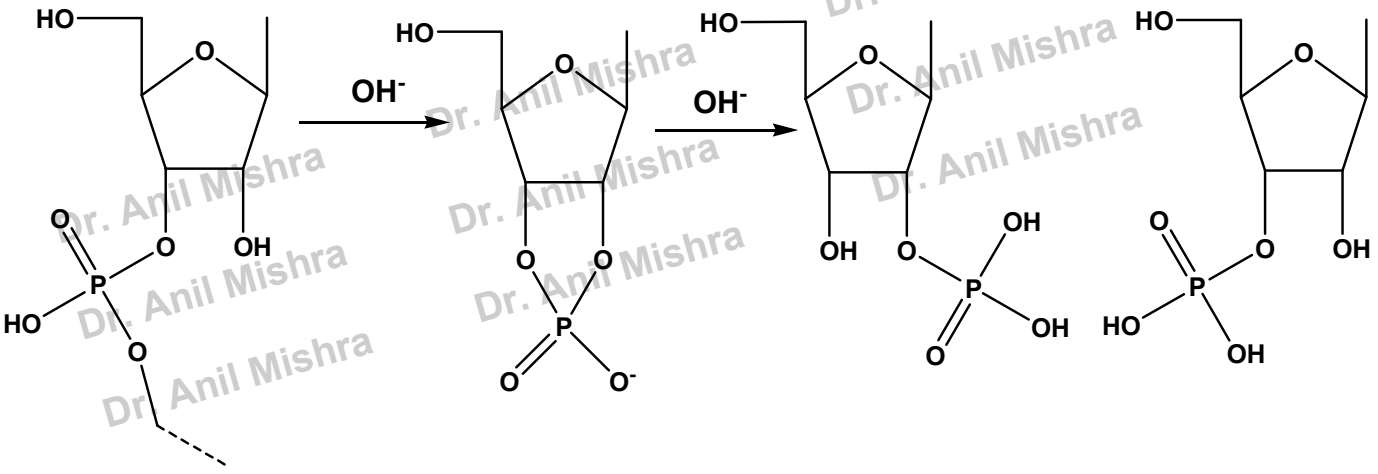


Structure of Nucleotides

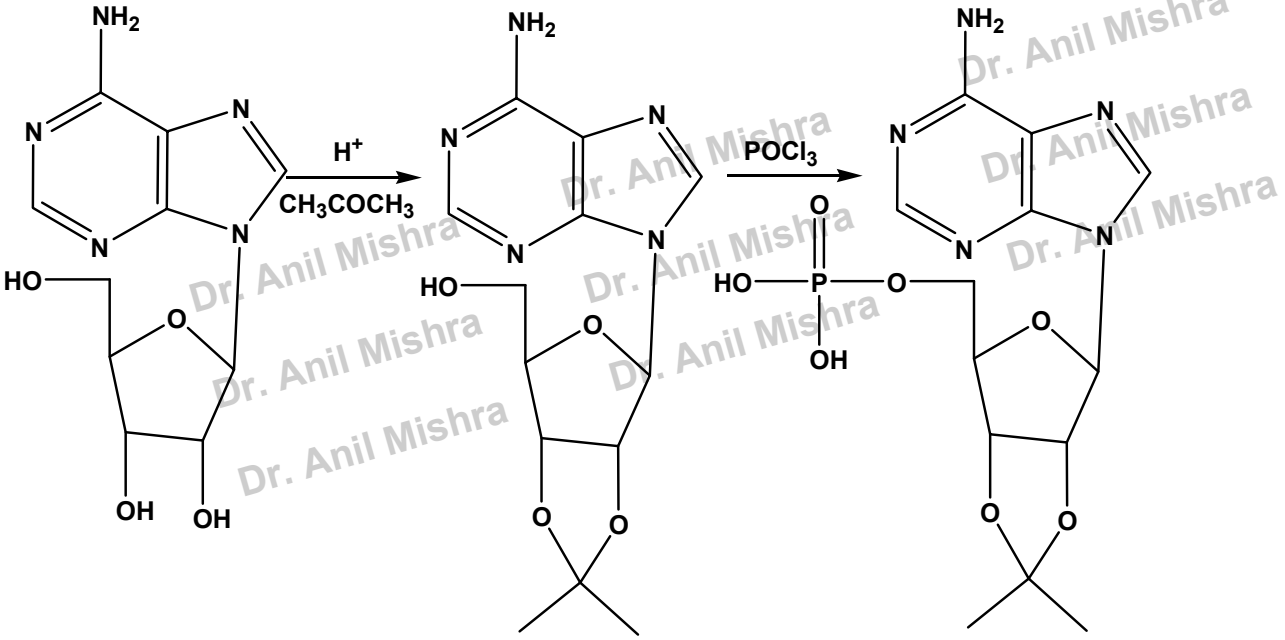
- These are phosphate esters of nucleosides
- Careful hydrolysis gives ribose phosphate
  - Phosphate is attached to the sugar moiety
- Point of attachment of phosphate could be at 2, 3or 5 in ribose and 3or 5 in deoxy ribose
- Reduction of ribose phosphate gives optically inactive **phosphoribitol**
  - Optical inactivity only possible when phosphate residue is attached to center hydroxyl group i.e. position 3



- Enzymatic hydrolysis of nucleotides give 2, 3 and 5 phosphates in RNA
- Mixture o f 2 and 3 phosphate can be explained as



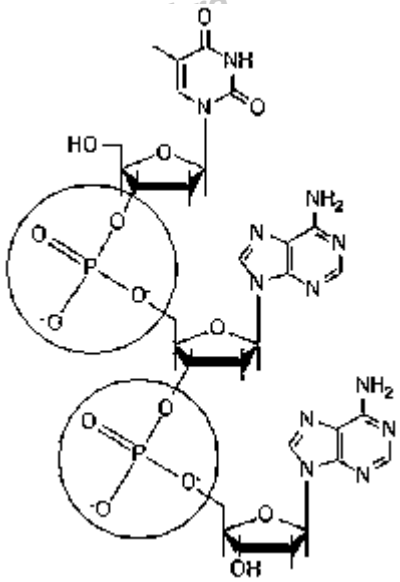
SYNTHESIS OF NUCLEOTIDES



STRUCTURE OF NUCLEIC ACIDS

- The most common nucleic acids are
  - **Deoxyribonucleic acid** (DNA)
  - **Ribonucleic acid** (RNA)
- Are biopolymers having a alternating phosphate and sugar moieties
- Each nucleotide sub-unit consists of a
  - Phosphate
  - Deoxyribose sugar (ribose in RNA)
  - One of the 4 nitrogenous nucleotide bases
    - Purine
    - Pyrimidine.
- The purine bases **adenine** (**A**) and **guanine** (**G**) are larger and consist of two aromatic rings.
- The pyrimidine bases **cytosine** (**C**) and **thymine** (**T**) are smaller and only consist of one aromatic ring.
- In RNA however, **thymine** (**T**) is substituted by **uracil** (**U**) and the **deoxyribose** is substituted by **ribose**.

PHOSPHODIESTER BOND

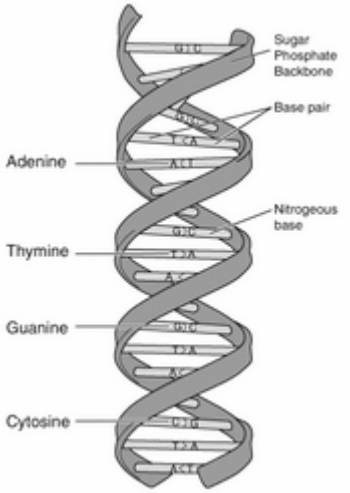


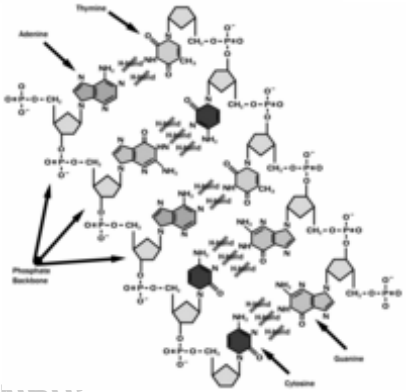
- A **phosphodiester bond** is a group of strong covalent bonds between the phosphorus atom in a phosphate group and two other molecules over two ester bonds.
- Phosphodiester bonds are central to all life on Earth, as they make up the backbone of the strands of DNA. In DNA and RNA, the phosphodiester bond is the linkage between the 3' Carbon atom and the 5' Carbon of the ribose sugar.
- The phosphate groups in the phosphodiester bond are very negatively charged. Because the phosphate groups are so negatively charged, there is a large repulsion, which forces the phosphates to take opposite sides of the DNA strands.
- In order for the phosphodiester bond to be formed and the nucleotides to be joined, the tri-phosphate or di-phosphate forms of the nucleotide building blocks are broken apart to give off energy required to drive the enzyme-catalyzed reaction.
  - When a single phosphate or two phosphates known as pyrophosphates break away and catalyze the reaction, the phosphodiester bond is formed.
- Phosphodiester bonds can be catalyzed by the action of phosphodiesterases, which play an important role in repairing DNA, sequences.
- In biological systems, the phosphodiester bond between two ribonucleotides can be broken by alkaline hydrolysis because of the free 2' hydroxyl group.

- Genes can be loosely viewed as the organism's "cookbook" or "blueprint";
- A strand of DNA (deoxyribonucleic acid) contains genes, areas that regulate genes, and areas that either have no function, (or function we do not (yet) know)
- DNA differs from RNA by having the sugar 2-deoxyribose instead of ribose in its backbone. This is the basic chemical distinction between RNA and DNA.
- DNA is organized as two complementary strands, head-to-toe, with bonds between them that can be "unzipped" like a zipper, separating the strands;
- DNA is a chain of chemical "building blocks", called "bases", of which there are four types: these can be abbreviated A, T, C, and G, with U rarely replacing T
- Because each strand of DNA has a directionality, the sequence order does matter: A+T is not the same as T+A, just as C+G is not the same as G+C
- For each given base, there is just one possible complementary base, so naming the bases on the conventionally chosen side of the strand is enough to describe the entire double-strand sequence
- The order of the bases along the length of the DNA is the genetic material, the sequence itself is the description for genes

WATSON-CRICK MODEL

(James Watson and Francis Crick, 1953)





- Although sometimes called "**the molecule of heredity**", pieces of DNA as people typically think of them are not single molecules.
- Rather, they are pairs of molecules, which entwine like vines to form a **double helix**
- Each vine-like molecule is a strand of DNA: a chemically linked chain of nucleotides, each of which consists of a sugar, a phosphate and one of five kinds of **nucleobases** ("bases").
- When using the twisted ladder analogy, think of the sugar-phosphate backbones as the two sides of the ladder and the bases in the middle as the rungs of the ladder.
- The two strands of the DNA double helix run in opposite directions, one in the 5' to 3' direction, the other in the 3' to 5' direction. The term that describes how the two strands relate to each other is known as ***antiparallel***.
- The diversity of the bases means that there are five kinds of nucleotides,
- In a DNA double helix, two polynucleotide strands can associate through the **hydrophobic effect** and **pi stacking.** Specificity of which strands stay associated is determined by complementary pairing. Each base forms hydrogen bonds readily to only one other -- **A to T** and **C to G** -- so that the identity of the base on one strand dictates the strength of the association; the more complementary bases exist, the stronger and longer-lasting the association.
- Because pairing causes the nucleotide bases to face the helical axis, the sugar and phosphate groups of the nucleotides run along the outside; the two chains they form are sometimes called the "**backbones**" of the helix. In fact, it is chemical bonds between the phosphates and the sugars that link one nucleotide to the next in the DNA strand.

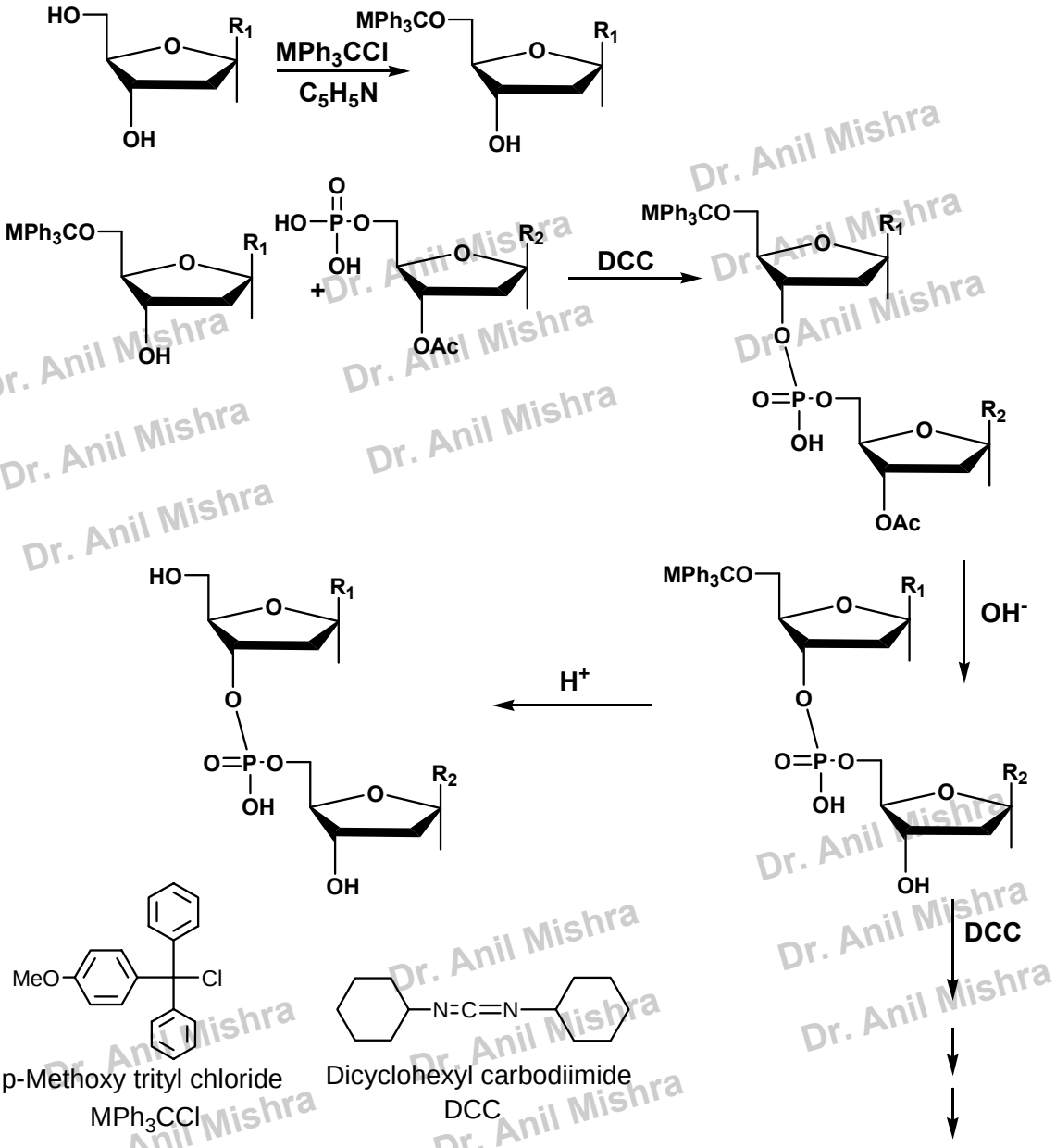
Nature of Internucleotide Bond in DNA

1. ***Pancreatic deoxyribo nuclease*** converts DNA into a mixture of oligonucleotides containing 5’ phosphate residue and 3’ hydroxyl is free

2. Action of ***spleen phosphodiesterase*** on DNA results in the formation of deoxyribo nucleotide-3’-phosphate

    - Thus DNA’s have a structure with the nucleotide units linked by 3’-5’ phosphodiester bonds

Chemical synthesis of DNA
(Khorana et.al.)



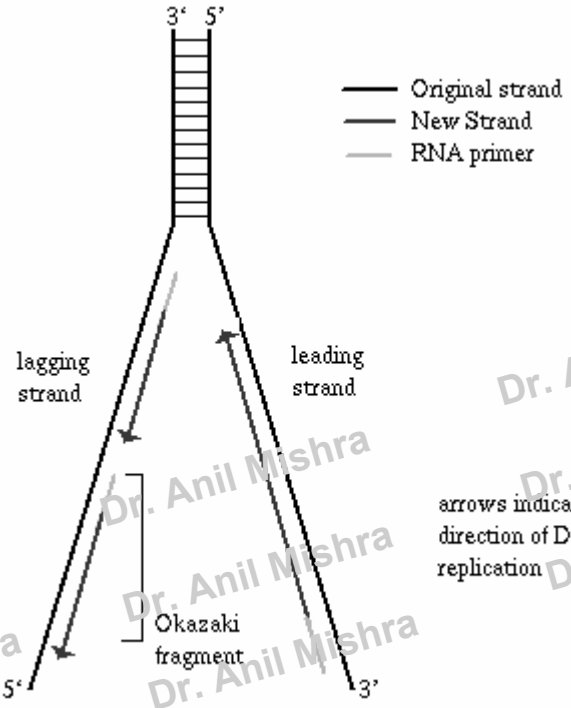
p-Methoxy trityl chloride
MPh_3CCl

Dicyclohexyl carbodiimide
DCC

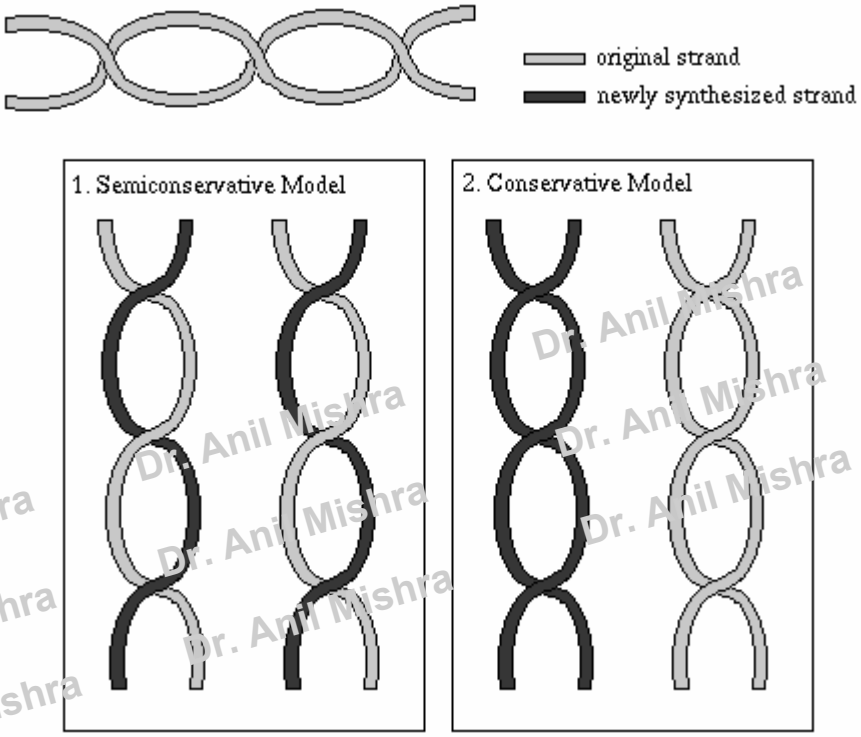
DNA Replication

- DNA replication begins with a partial unwinding of the double helix at an area known as the ***replication fork***. This unwinding is accomplished by an enzyme known as ***DNA helicase***.
- As the two DNA strands separate ("unzip") and the bases are exposed, the enzyme ***DNA polymerase*** moves into position at the point where synthesis will begin.
- But where does the DNA polymerase enzyme know where to begin synthesis? Is there some sort of marker, a start point?
- YES; the start point for DNA polymerase is a short segment of RNA known as an ***RNA primer***.
- The very term "primer" is indicative of its role, which is to "prime" or start DNA synthesis at certain points. The primer is "laid down" complementary to the DNA template by an enzyme known as ***RNA polymerase*** or ***Primase***.
- The DNA polymerase then adds nucleotides one by one in an exactly complementary manner, A to T and G to C.
- How does the polymerase "know" which base to add?
- DNA polymerase is described as being "template dependent" in that it will **"read"** the sequence of bases on the template strand and then **"synthesize"** the complementary strand.
- The template strand is **ALWAYS** read in the 3' to 5' direction (that is, starting from the 3' end of the template and reading the nucleotides in order toward the 5' end of the template).
- The new DNA strand (since it is complementary) **MUST BE SYNTHESIZED** in the 5' to 3' direction (remember that both strands of a DNA molecule are described as being antiparallel).
- DNA polymerase catalyzes the formation of the hydrogen bonds between each arriving nucleotide and the nucleotides on the template strand.

- Because the original DNA strands are complementary and run antiparallel, only one new strand can begin at the 3' end of the template DNA and grow continuously as the point of replication (***the replication fork***) moves along the template DNA.
- The other strand must grow in the opposite direction because it is complementary, not identical to the template strand. The result of this side's discontiguous replication is the production of a series of short sections of new DNA called ***Okazaki fragments*.**
- To make sure that this new strand of short segments is made into a continuous strand, the sections are joined by the action of an enzyme called ***DNA ligase***, which LIGATES the pieces together by forming the missing phosphodiester bonds!
- The last step is for an enzyme to come along and remove the existing RNA primers and then fill in the gaps with DNA. This is the job of yet another type of DNA polymerase which has the ability to chew up the primers (dismantle them) and replace them with the deoxynucleotides that make up DNA.



- Since each new strand is complementary to its old template strand, two **identical** new copies of the DNA double helix are produced during replication.
- In each new helix, one strand is the old template and the other is newly synthesized, a result described by saying that the replication is ***semi-conservative***.
- Crick described the DNA replication process and the fitting together of two DNA strands as being like a hand in a glove. The hand and glove separate, a new hand forms inside the old glove, and a new glove forms around the old hand. As a result, two identical copies now exist.



STRUCTURE OF RNA

Both nucleic acids are sugar-phosphate polymers and both have nitrogen bases attached to the sugars of the backbone- but there are several important differences.

- They differ in composition:
  - The sugar in RNA is ribose, not the deoxyribose in DNA
  - The base uracil is present in RNA instead of thymine.
- They also differ in size and structure:
  - RNA molecules are smaller (shorter) than DNA molecules,
  - RNA is single-stranded, not double-stranded like DNA.
- Another difference between RNA and DNA is in function. DNA has only one function- ***STORING GENETIC INFORMATION*** in its sequence of nucleotide bases.

Nature of Internucleotide Bond in RNA

1. Hydrogen ion titration on purified RNA showed that secondary phosphate ionizations are absent.
   - ♦ There are no free phosphate groups
   - ♦ Individual ribonucleotides are linked through phosphodiester bonds
2. Attachment of phosphate in mononucleotide is at position 3'
   - ♦ Possible internuclear bonds are 2'-3' or 3'-5'

3. The enzyme ***spleen phosphodiesterase*** (specific for C-5'-O-P bonds) converts RNA's into a mixture of ribonucleotide-3'-phosphates

4. The enzyme ***snake venom phosphodiesterase*** (specific for C-3'-O-P bonds) converts RNA's into a mixture of ribonucleotide-5'-phosphates
   - ♦ There is phosphodiester bond between 3'-5'

Classification of RNA

There are three main kinds of ribonucleic acid, each of which has a specific job to do.

- **Ribosomal RNA's**-exist outside the nucleus in the cytoplasm of a cell in structures called **ribosome**.
  - Ribosomes are small, granular structures where protein synthesis takes place. Each ribosome is a complex consisting of about 60% ribosomal RNA **(rRNA)** and 40% protein.

- **Messenger RNA's**-are the nucleic acids that "record" information from DNA in the cell nucleus and carry it to the ribosomes and are known as messenger RNA's **(mRNA).**

- **Transfer RNA's**-The function of transfer RNA's **(tRNA)** is to deliver amino acids one by one to protein chains growing at ribosomes.

GENETIC CODE

- The **genetic code** is a set of rules, which maps DNA sequences to proteins in the living cell, and is employed in the process of protein synthesis. Nearly all-living things use the same genetic code, called the **standard genetic code**, although a few organisms use minor variations of the standard code.
- The genetic information carried by an organism - its **genome** - is inscribed in one or more DNA molecules.
  - Each functional portion of a DNA molecule is referred to as a **gene**.
  - Each gene is **transcribed** into a short template molecule of the related polymer RNA, which is better suited for protein synthesis.
  - This in turn is **translated**, by mediation of a machinery consisting of ribosomes and a set of transfer RNA's and associated enzymes, into an amino acid chain (polypeptide), which will then be folded into a protein.
- The gene sequence inscribed in DNA, and in RNA, is composed of tri-nucleotide units called **codons**, each coding for a single amino acid.
  - Overall, there are $4^3 = 64$ different codon combinations.
- A series of three nucleotide bases on a DNA molecule is called a **TRIPLET**
- A set of three nucleotide bases on an mRNA molecule is called a **CODON**
- A set of three nucleotide bases on a tRNA molecule is called an **ANTICODON**
- For example, the RNA sequence UUUAAACCC contains the codons UUU, AAA and CCC, each of which specifies one amino acid. So, this RNA sequence represents a protein sequence, three amino acids long. (DNA is also a sequence of nucleotide bases, but there thymine takes the place of uracil.)

♦

Table showing the 64 codons and the amino acid each codon codes for

		2nd base			
		U	**C**	**A**	**G**
1st base	**U**	UUU (Phe/F)Phenylalanine UUC (Phe/F)Phenylalanine UUA (Leu/L)Leucine UUG (Leu/L)Leucine, *Start*	UCU (Ser/S)Serine UCC (Ser/S)Serine UCA (Ser/S)Serine UCG (Ser/S)Serine	UAU (Tyr/Y)Tyrosine UAC (Tyr/Y)Tyrosine UAA Ochre (*Stop*) UAG Amber (*Stop*)	UGU (Cys/C)Cysteine UGC (Cys/C)Cysteine UGA Opal (*Stop*) UGG (Trp/W)Tryptophan
	C	CUU (Leu/L)Leucine CUC (Leu/L)Leucine CUA (Leu/L)Leucine CUG (Leu/L)Leucine, *Start*	CCU (Pro/P)Proline CCC (Pro/P)Proline CCA (Pro/P)Proline CCG (Pro/P)Proline	CAU (His/H)Histidine CAC (His/H)Histidine CAA (Gln/Q)Glutamine CAG (Gln/Q)Glutamine	CGU (Arg/R)Arginine CGC (Arg/R)Arginine CGA (Arg/R)Arginine CGG (Arg/R)Arginine
	A	AUU (Ile/I)Isoleucine, *Start* AUC (Ile/I)Isoleucine AUA (Ile/I)Isoleucine AUG (Met/M)Methionine, *Start*	ACU (Thr/T)Threonine ACC (Thr/T)Threonine ACA (Thr/T)Threonine ACG (Thr/T)Threonine	AAU (Asn/N)Asparagine AAC (Asn/N)Asparagine AAA (Lys/K)Lysine AAG (Lys/K)Lysine	AGU (Ser/S)Serine AGC (Ser/S)Serine AGA (Arg/R)Arginine AGG (Arg/R)Arginine
	G	GUU (Val/V)Valine GUC (Val/V)Valine GUA (Val/V)Valine GUG (Val/V)Valine, *Start*	GCU (Ala/A)Alanine GCC (Ala/A)Alanine GCA (Ala/A)Alanine GCG (Ala/A)Alanine	GAU (Asp/D)Aspartic acid GAC (Asp/D)Aspartic acid GAA (Glu/E)Glutamic acid GAG (Glu/E)Glutamic acid	GGU (Gly/G)Glycine GGC (Gly/G)Glycine GGA (Gly/G)Glycine GGG (Gly/G)Glycine

Reverse codon table

Table showing the 20 standard amino acids used in proteins, and the codons that code for each amino acid.

Ala	A	GCU, GCC, GCA, GCG	**Leu**	L	UUA, UUG, CUU, CUC, CUA, CUG
Arg	R	CGU, CGC, CGA, CGG, AGA, AGG	**Lys**	K	AAA, AAG
Asn	N	AAU, AAC	**Met**	M	AUG
Asp	D	GAU, GAC	**Phe**	F	UUU, UUC
Cys	C	UGU, UGC	**Pro**	P	CCU, CCC, CCA, CCG
Gln	Q	CAA, CAG	**Ser**	S	UCU, UCC, UCA, UCG, AGU,AGC
Glu	E	GAA, GAG	**Thr**	T	ACU, ACC, ACA, ACG
Gly	G	GGU, GGC, GGA, GGG	**Trp**	W	UGG
His	H	CAU, CAC	**Tyr**	Y	UAU, UAC
Ile	I	AUU, AUC, AUA	**Val**	V	GUU, GUC, GUA, GUG
Start		AUG, GUG	***Stop***		UAG, UGA, UAA

- A series of three nucleotide bases on a DNA molecule is called a **TRIPLET**
- A set of three nucleotide bases on an mRNA molecule is called a **CODON**
- A set of three nucleotide bases on a tRNA molecule is called an **ANTICODON**

TRANSFER OF GENETIC INFORMATION

Three fundamental processes take place in the transfer and use of genetic information:

1. **Replication** is the process by which a replica, or identical copy, of DNA is made. Replication occurs every time a cell divides so that information can be preserved and handed down to offspring. This is similar to making a copy of a file onto a disk so you can take that file to a different computer.
2. **Transcription** is the process by which the genetic messages contained in DNA are "read" or transcribed. The product of transcription, known as messenger RNA (mRNA), leaves the cell nucleus and carries the message to the sites of protein synthesis. This tutorial explains later why this step is necessary in organisms with a nucleus!
3. **Translation** is the process by which the genetic messages carried by mRNA are decoded and used to build proteins.

STEP I OF PROTEIN SYNTHESIS:

INITIATION

- Protein synthesis is initiated when an mRNA, a ribosome, and the first tRNA molecule (carrying its Methionine amino acid) come together.
- The ribosome is inactive when it exists as two subunits (a large one and a small one) before it contacts a mRNA. The small unit of the ribosome will initiate the process of translation when it encounters a mRNA in the cytoplasm.
- The first **A-U-G** codon on the 5' end of the mRNA acts as a "start" signal for the translation machinery and codes for the introduction of a methionine amino acid. **THIS CODON AND, THUS, AMINO ACID WILL ALWAYS BE THE FIRST IN ANY AND ALL mRNA MOLECULES!!**
- Initiation is complete when the methionine tRNA occupies one of the **two binding sites** on the ribosome. Since this first site is the site where the growing peptide (another word for protein) will reside, it's known as the ***P site***. This is where the growing ***P***rotein will be. There is another site just to the 3' direction of the P site; it is known as the ***A site***. This is where the incoming tRNA will ***A***ttach itself.
  - Even though every protein begins with the Methionine amino acid, not all proteins will ultimately have methionine at one end. If the "start" methionine is not needed, it is removed before the new protein goes to work (either inside the cell or outside the cell, depending on the type of protein synthesized)

STEP II IN PROTEIN SYNTHESIS:
ELONGATION

- The incoming tRNA will bind to the A site (next to where the tRNA with the methionine attached is on the P site). ALL available tRNAs will approach the site and try to attach, but the only tRNA which will successfully attach is the one whose anticodon IS COMPLEMENTARY to the codon of the A site on the mRNA.

STEP III IN PROTEIN SYNTHESIS:
TERMINATION

- The elongation procedure continues until the proper protein is completed. A "stop" codon (U-A-A, U-G-A, or U-A-G) signals the end of the process. There is no tRNA that is complementary to the Stop Codon, so the process of building the protein stops.
- An enzyme called the ***releasing factor*** then frees the newly made polypeptide chain, also known as the PROTEIN, from the last tRNA.
- The mRNA molecule is released from the ribosome as the small and large subunits fall apart.
- The mRNA can then be re-translated or it may be degraded, depending on how much of that particular protein is needed.
- All mRNA messages are eventually degraded when the protein no longer needs to be made.