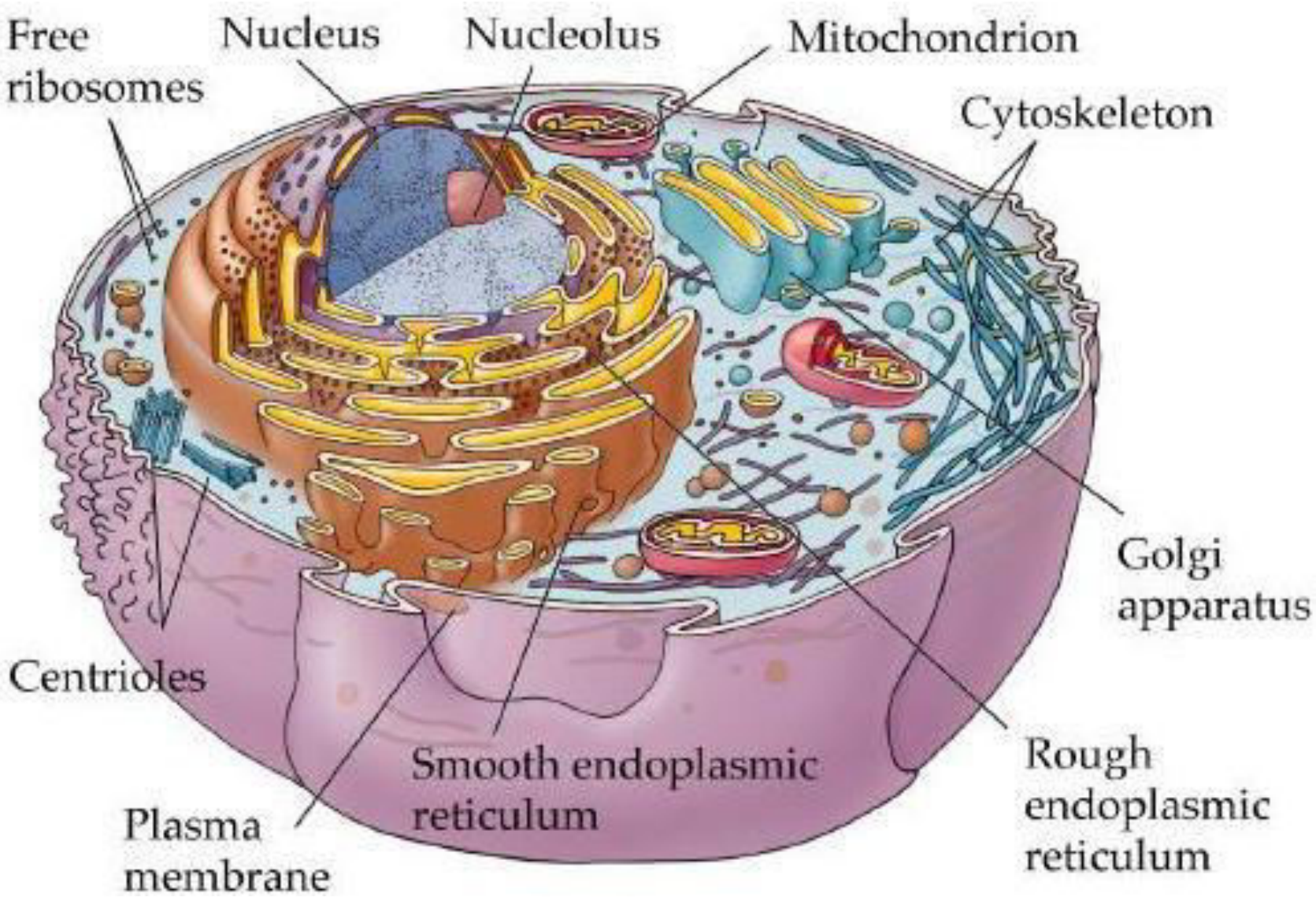
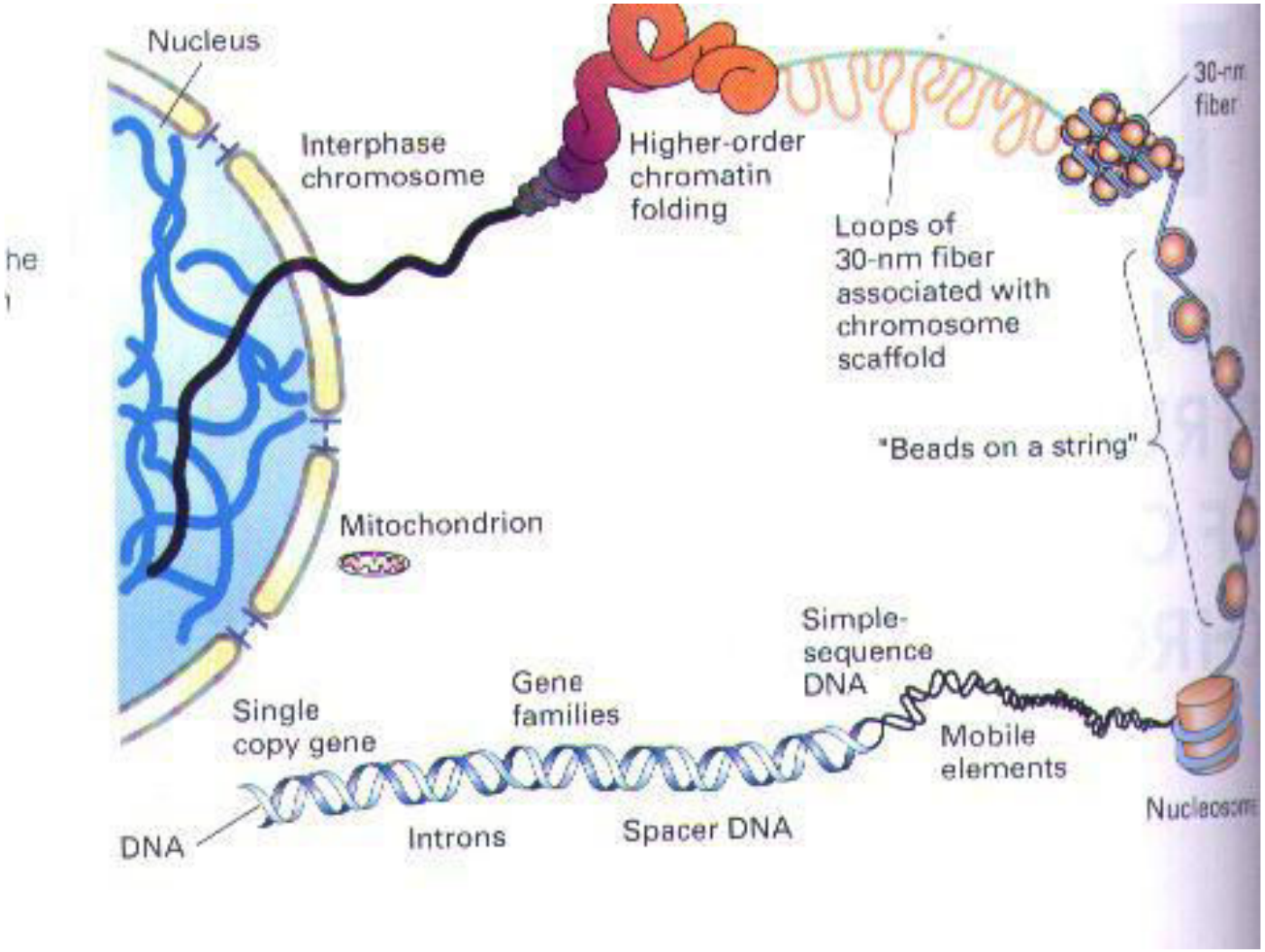
Basics of DNA Structure & Function

Dr. Monisha Banerjee
Professor
Molecular & Human Genetics Lab
Department of Zoology
University of Lucknow
Lucknow

Introduction



Cell-Chromosome-DNA



What's So Special About DNA?

DNA is one of the most boring macromolecules imaginable - its made of only four building blocks and has a perfectly monotonous structure.

Worse yet, DNA just sits there - it doesn't catalyze reactions or build the cell or organism.

So, what's so good about DNA?

The answer lies in DNA's ability to store and copy information.

How Can DNA Store and Copy Information?

Key properties that allow these neat tricks are that DNA is a:

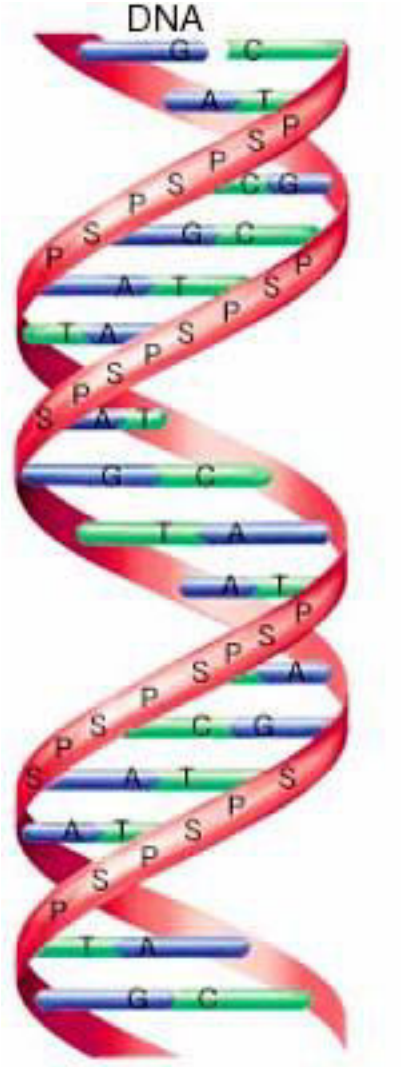
double stranded molecule ….

…. held together by complementary bases …..

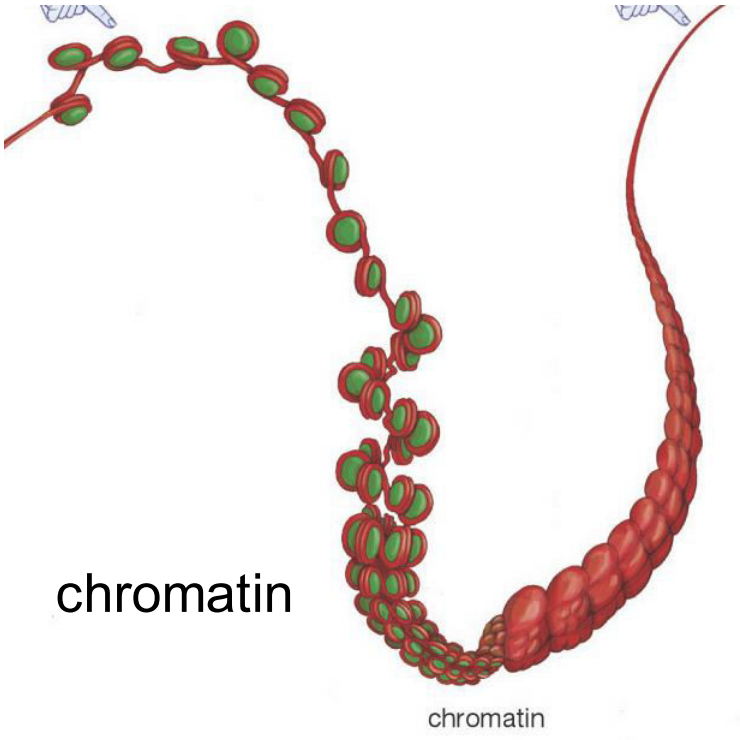
…… that pair through simple rules.

DNA is also capable of occasional change, and occasionally, change is good.

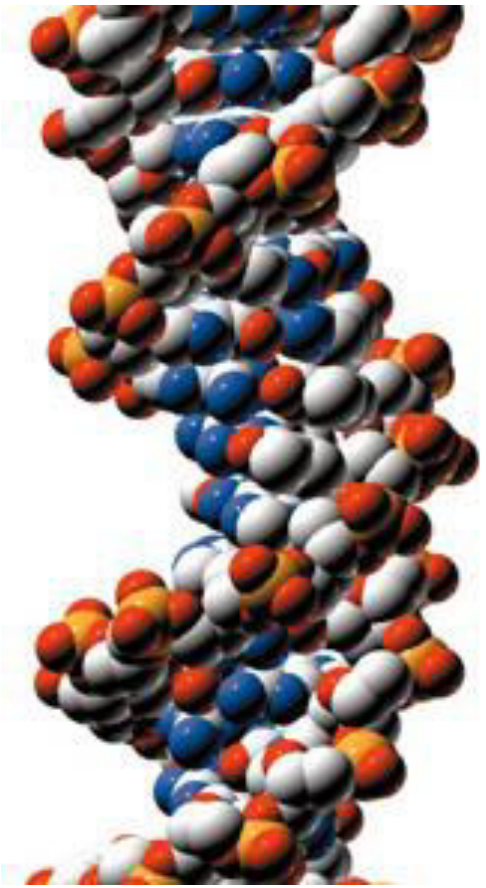
Two Views of the Double Helix



DNA is Almost Always Wrapped Around Proteins



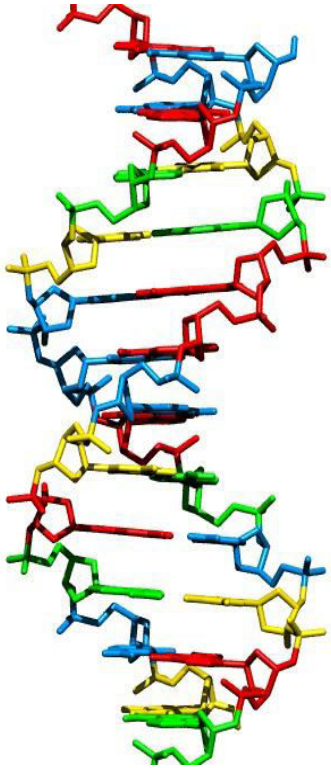
This ….

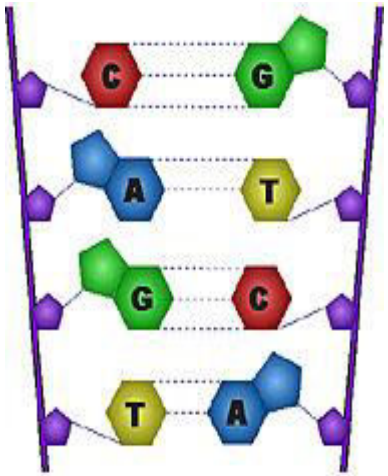


…. not this is what's found in the cell.

DNA

Sequence of nucleotides

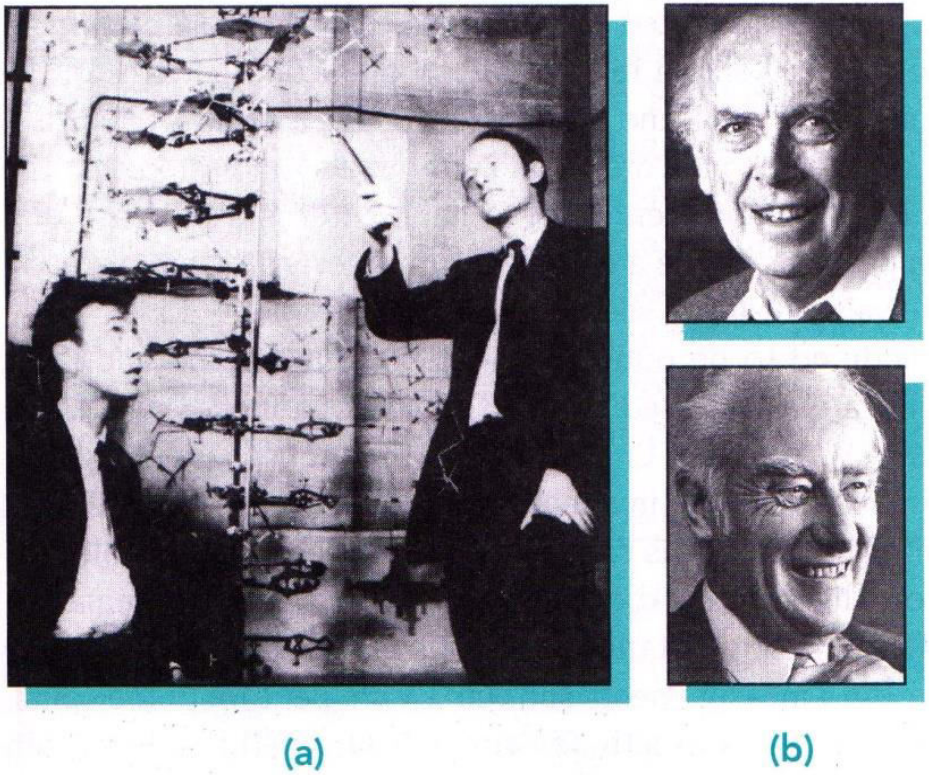




5′…ACGTGACTGAGGACCGTG
CGACTGAGACTGACTGGGT
CTAGCTAGACTACGTTTTA
TATATATATACGTCGTCGT
ACTGATGACTAGATTACAG
ACTGATTTAGATACCTGAC
TGATTTTAAAAAAATATT…3′

Winners of the Race to Learn DNA's Structure – Watson and Crick 53 Years Ago

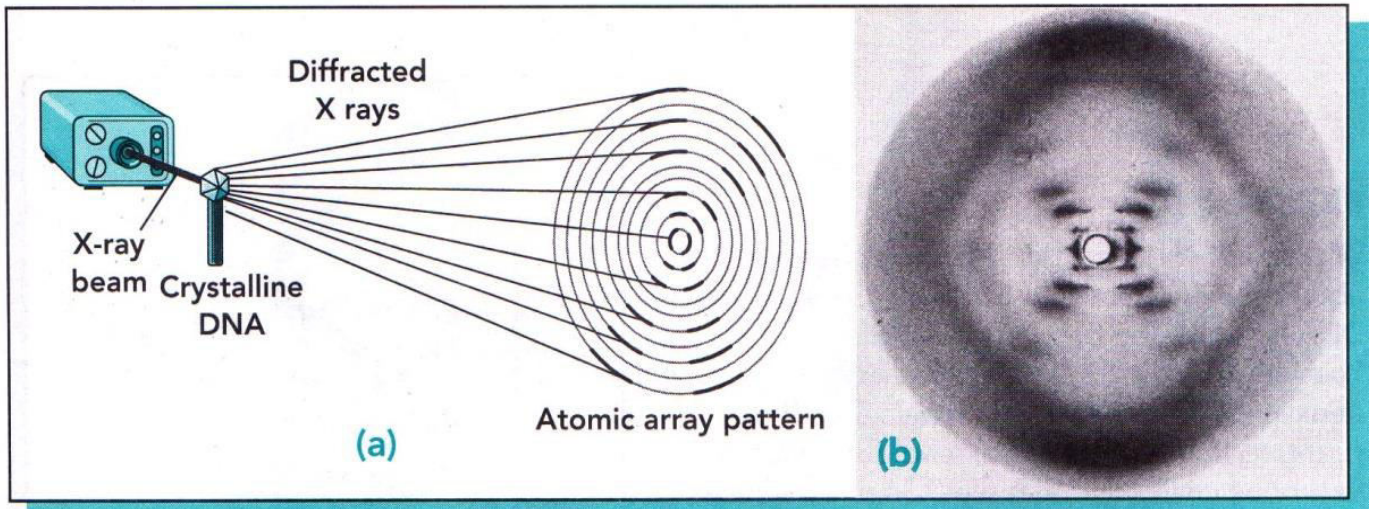
Discovery of DNA structure by X-ray Diffraction Technology



(a) (b)

FIGURE 2.5

James D. Watson and Francis H. C. Crick, the American graduate student and the British biochemist, who correctly explained the structure of DNA. **(a)** The scientists as they appeared in 1952, when the structure of DNA was formulated. **(b)** Photographs of more recent vintage.



Landmark Publication

NATURE

No. 4356 April 25, 1953

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A structure for Deoxyribose Nucleic Acid

We wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.

This figure is purely diagrammatic. The Two ribbons symbolize the two phosphate—sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions, namely, that each chain consists of phosphate diester groups joining β-D-deoxyribofuranose residues with 3′, 5′ linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right-handed gelices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Furberg's[2] model No. 1; that is the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's 'standard configuration', the sugar being roughly perpendicular to the attached base. There is a residue on each chain every 3·4. A. in the *z*-direction. We have assumed an angle of 36° between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 A. The distance of a phosphorus atom from the fibre axis is 10 A. As the phosphates are on the outside, cations have easy access to them.

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical *z*-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrmidine position 6.

If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

It has been found experimentally[3,4] that the ratio of the amounts of adenine to thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

The previously published X-ray data[5,6] on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on inter-atomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. F. Wilkins, Dr. R. E. Franklin and their co-workers at King's College, London. One of us (J. D. W.) has been aided by a fellowship from the National Foundation for Infantile Paralysis.

J. D. Watson
F. H. C. Crick

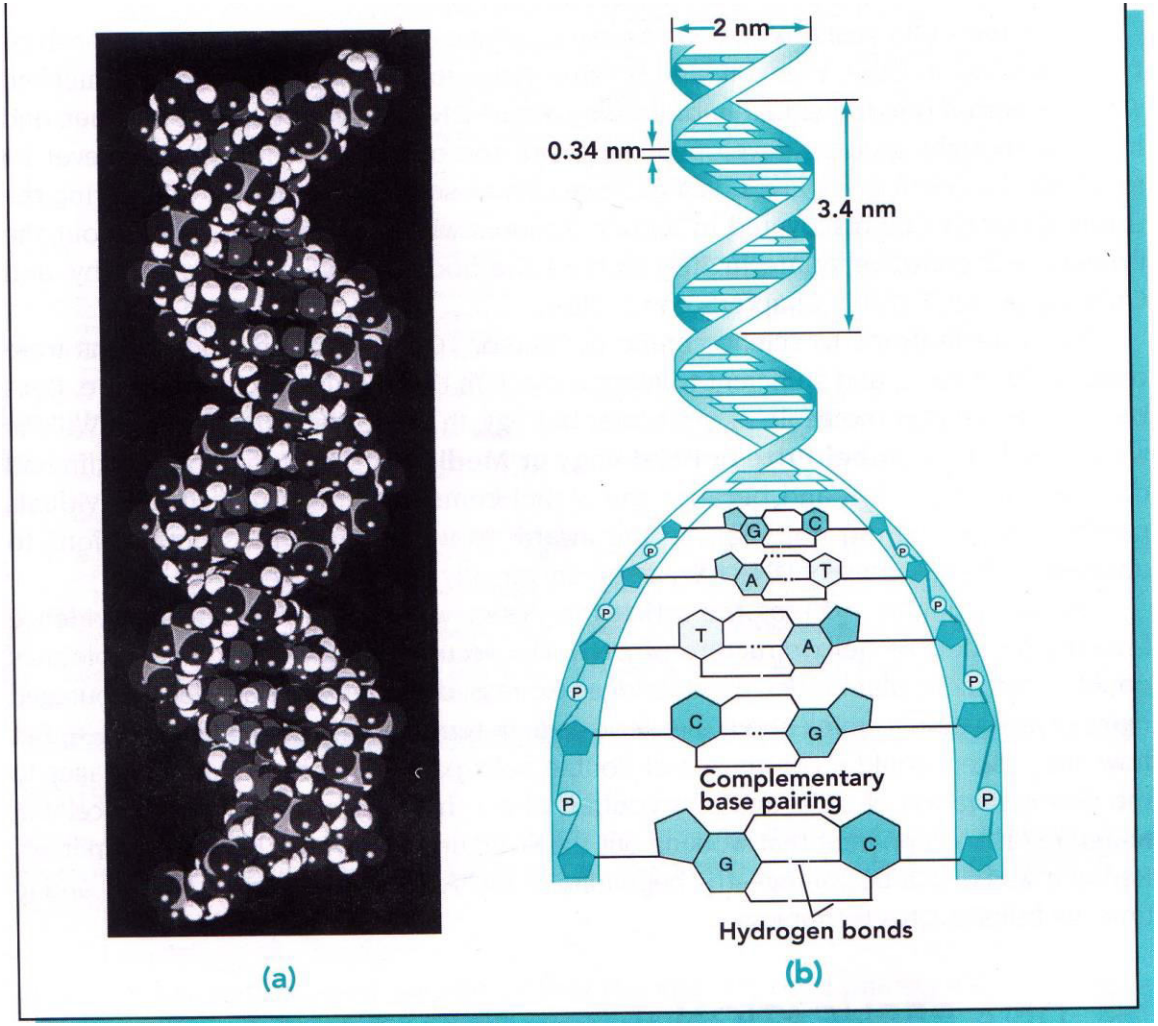
Medical Research Council Unit for the Study of the Molecular Structure of Biological Systems,
Cavendish Laboratory, Cambridge.
April 2.

[1]Pauling, L., and Corey, R. B., *Nature*, 171, 346 (1953); *Proc. U.S. Nat. Acad. Sci.*, 39, 84 (1953).
[2]Furberg, S., *Acta Chem. Scand.*, 6, 634 (1952).
[3]Chargaff, E., for references see Zamenhof, S., Brawerman, G., and Chargaff, E., *Biochim. et Biophys. Acta*, 9, 402 (1952).
[4]Wyatt, G. R., *J. Gen. Physiol.*, 36, 201 (1952).
[5]Astbury, W. T., Symp. Soc. Exp. Biol. 1, Nucleic Acid, 66 (Camb. Univ. Press, 1947).
[6]Wilkins, M. H. F., and Randall, J. T., *Biochim. et Biophys. Acta*, 10, 192 (1953).

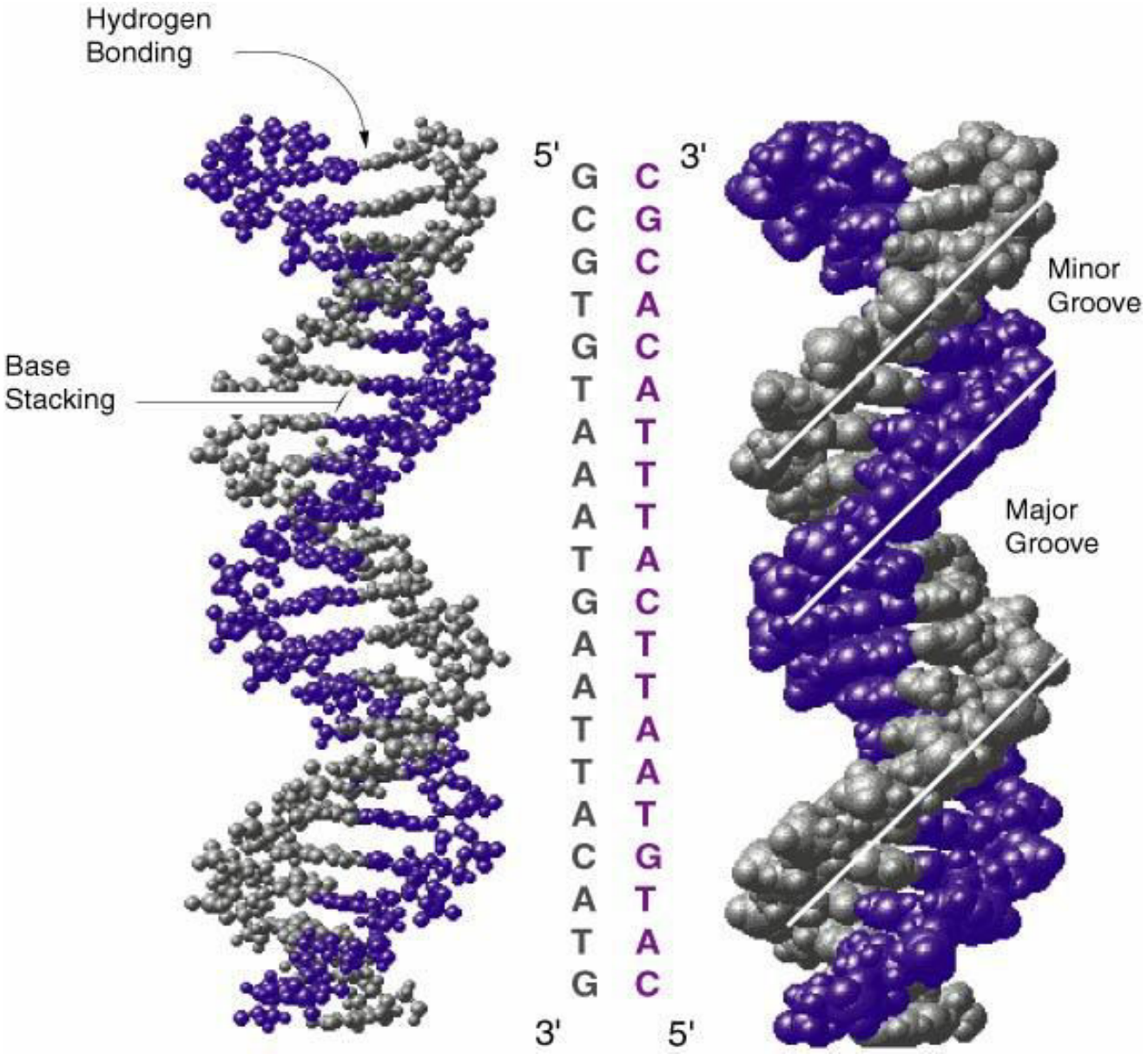
DNA structure

- DNA is a double stranded molecule consists of 2 polynucleotide chains running in opposite directions.
- Both strands are complementary to each other.
- The bases are on the inside of the molecules and the 2 chains are joined together by double H-bond between A and T and triple H-bond between C and G.
- The base pairing is very specific which make the 2 strands complementary to each other.
- So each strand contain all the required information for synthesis (replication) of a new copy to its complementary.

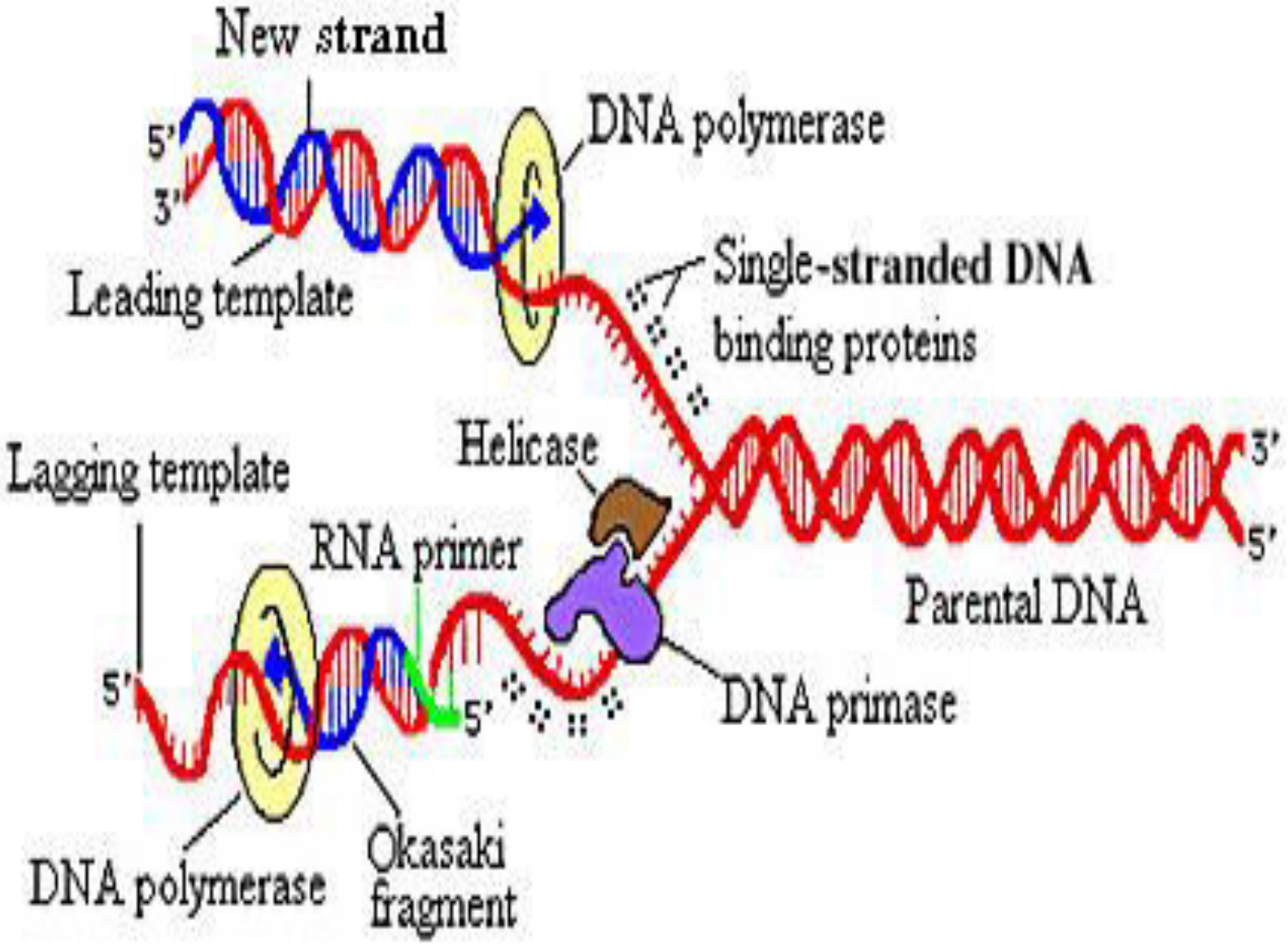
Structure of DNA



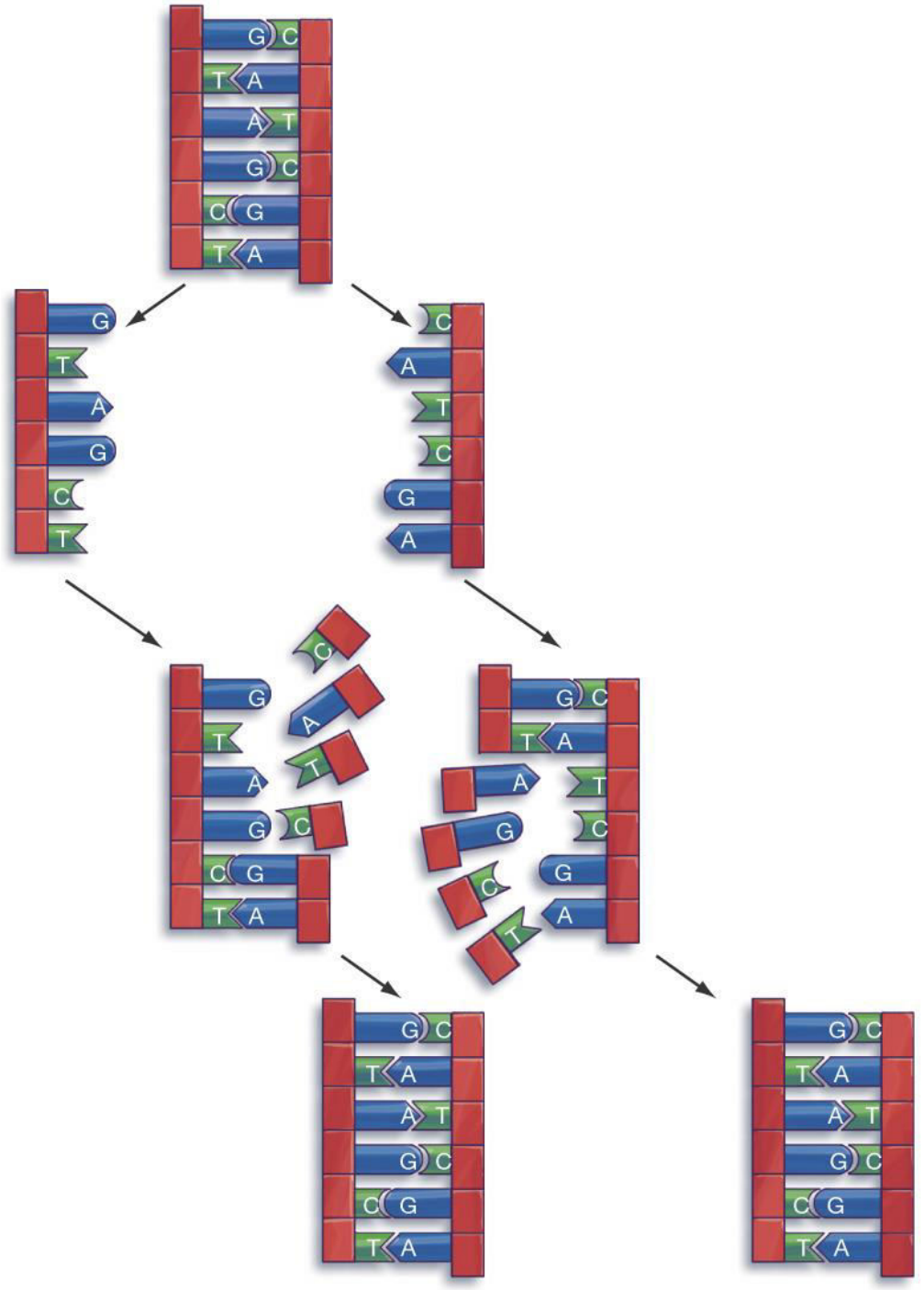
How Can DNA Store and Copy Information?



DNA Replication



Complementary Base Pairing Allows Each Strand of DNA to Serve as a Template for DNA Replication

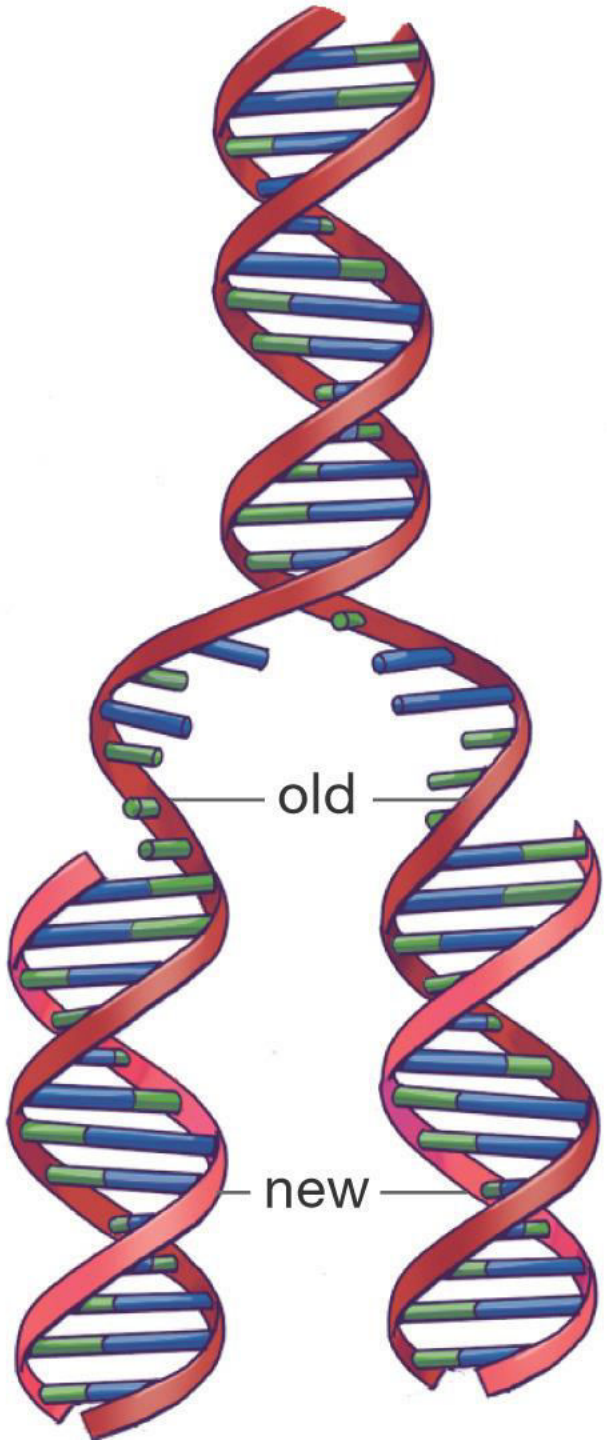


DNA is a perfect illustration of function following form (structure dictates function).

DNA Replication – Something Old and Something New In Each Daughter Molecule

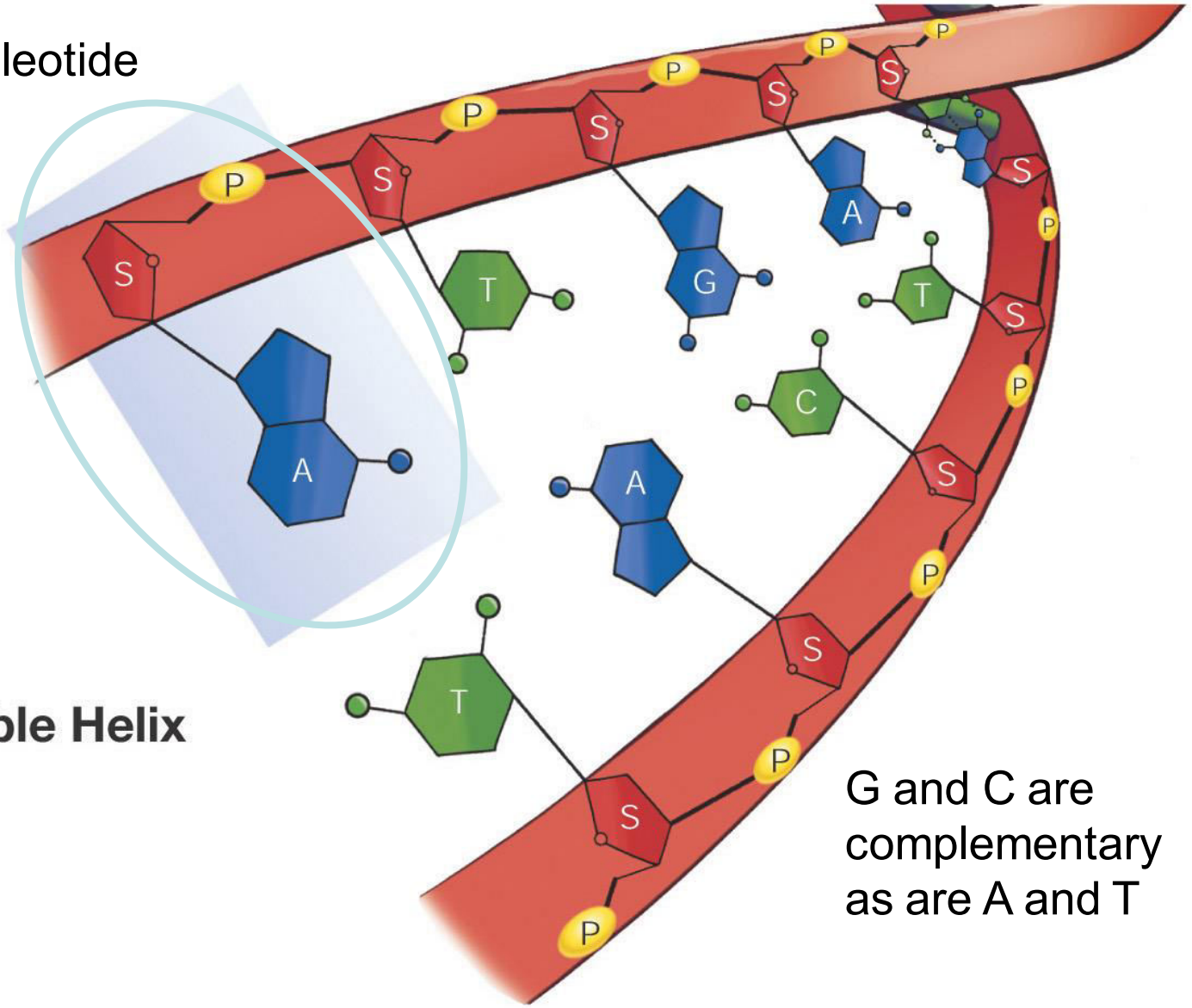
Simple as it is in Principle, DNA Replication Requires Many Enzymes That Work Coordinately

DNA polymerases are the first and foremost of the replication enzymes.



DNA is Made of Two Long Chains of Nucleotides Joined by Hydrogen Bonds

A Nucleotide



The Double Helix

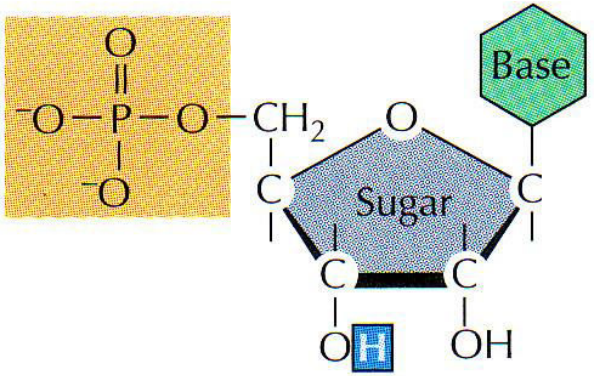
G and C are complementary as are A and T

Nucleic acids

- Principle information molecule in the cell.
- All the genetic codes are carried out on the nucleic acids.
- Nucleic acid is a linear polymer of nucleotides

Nucleotides

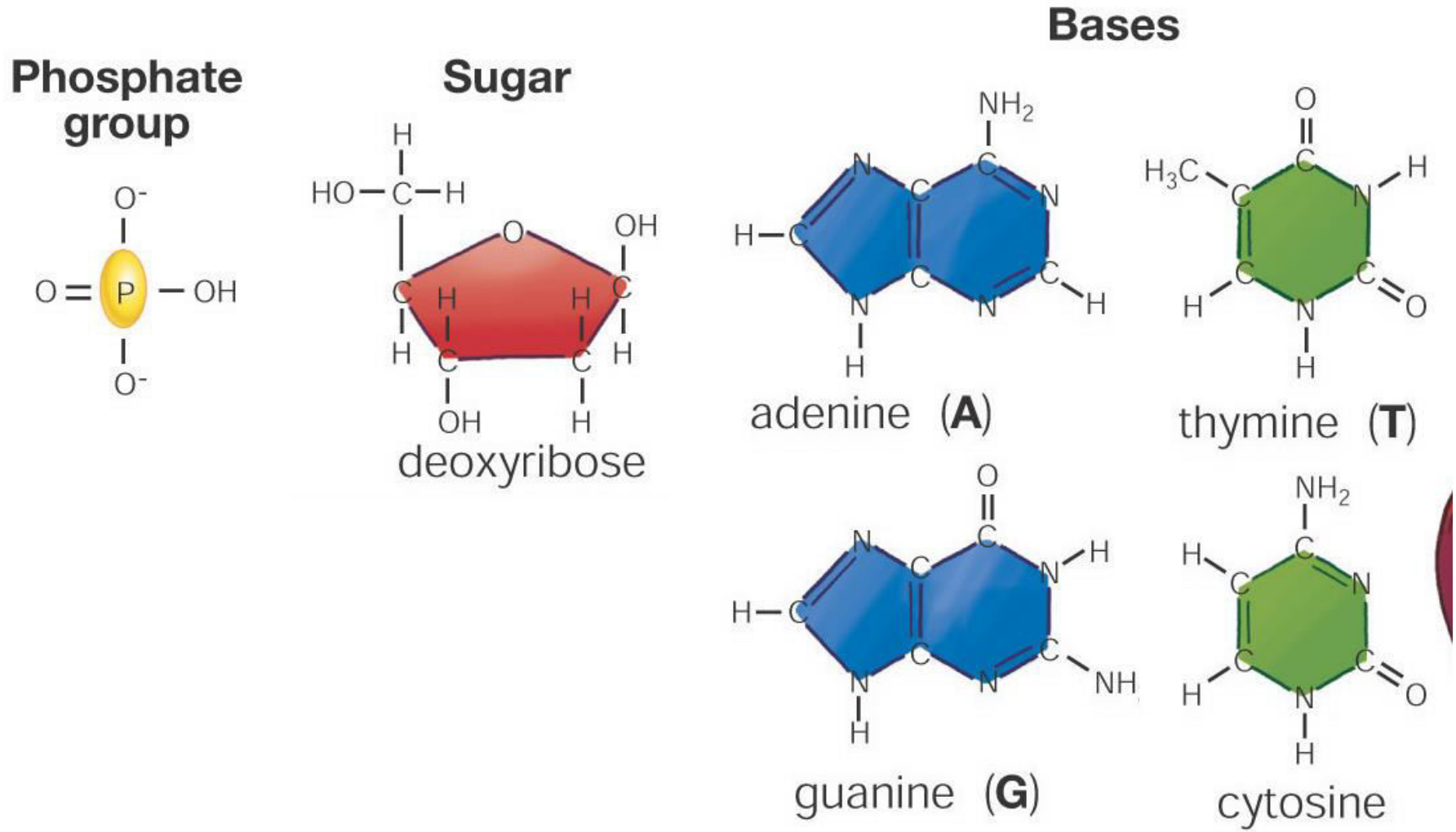
- **Nucleotides are the unit structure of nucleic acids.**
- **Nucleotides composed of 3 components:**
  - **Nitrogenous base (A, C, G, T or U)**
  - **Pentose sugar**
  - **Phosphate**



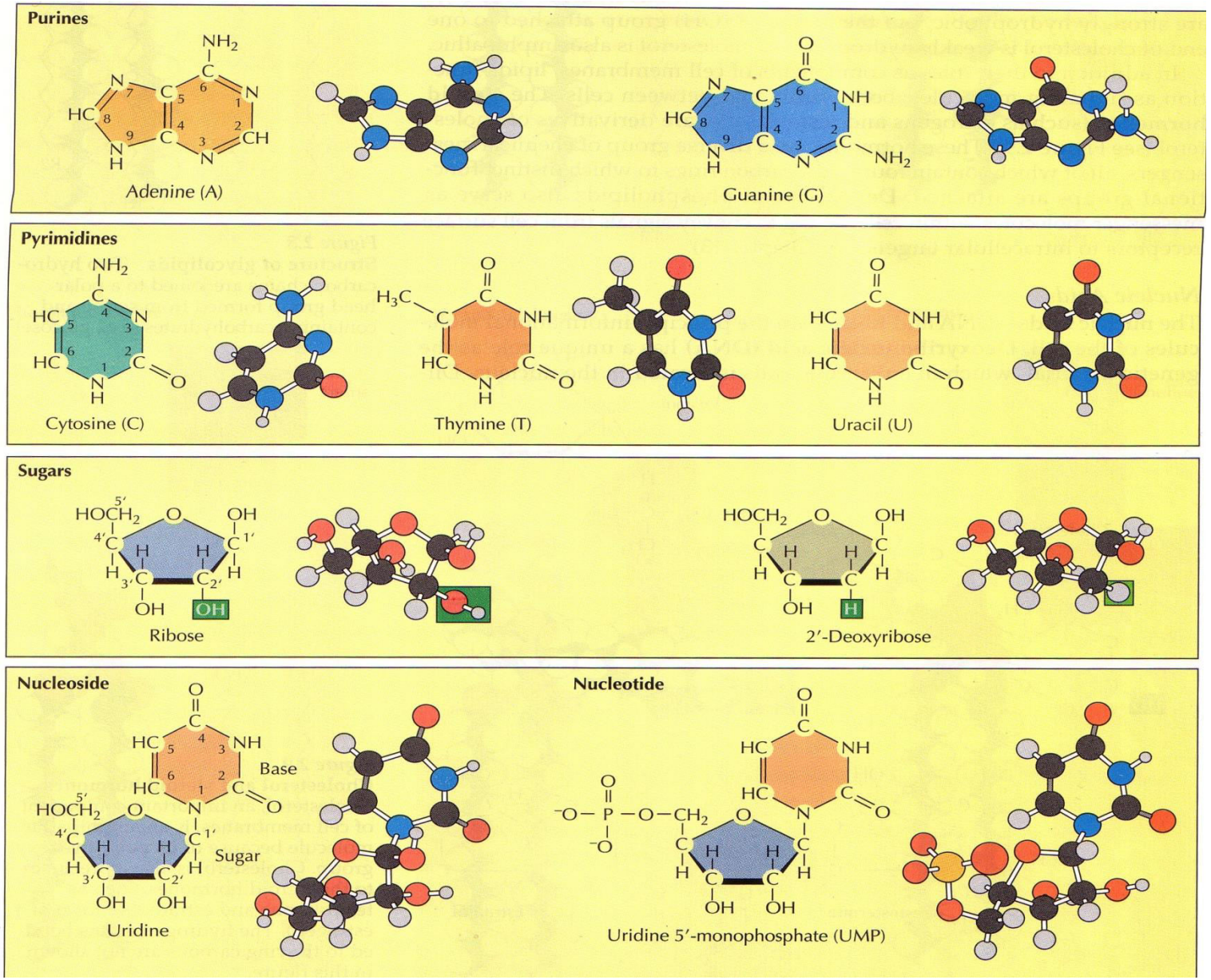
Nitrogenous bases

- **There are 2 types:**
  - **Purines:**
    - **Two ring structure**
    - **Adenine (A) and Guanine (G)**
  - **Pyrimidines:**
    - **Single ring structure**
    - **Cytosine (C) and Thymine (T) or Uracil (U).**

Building DNA Building Blocks

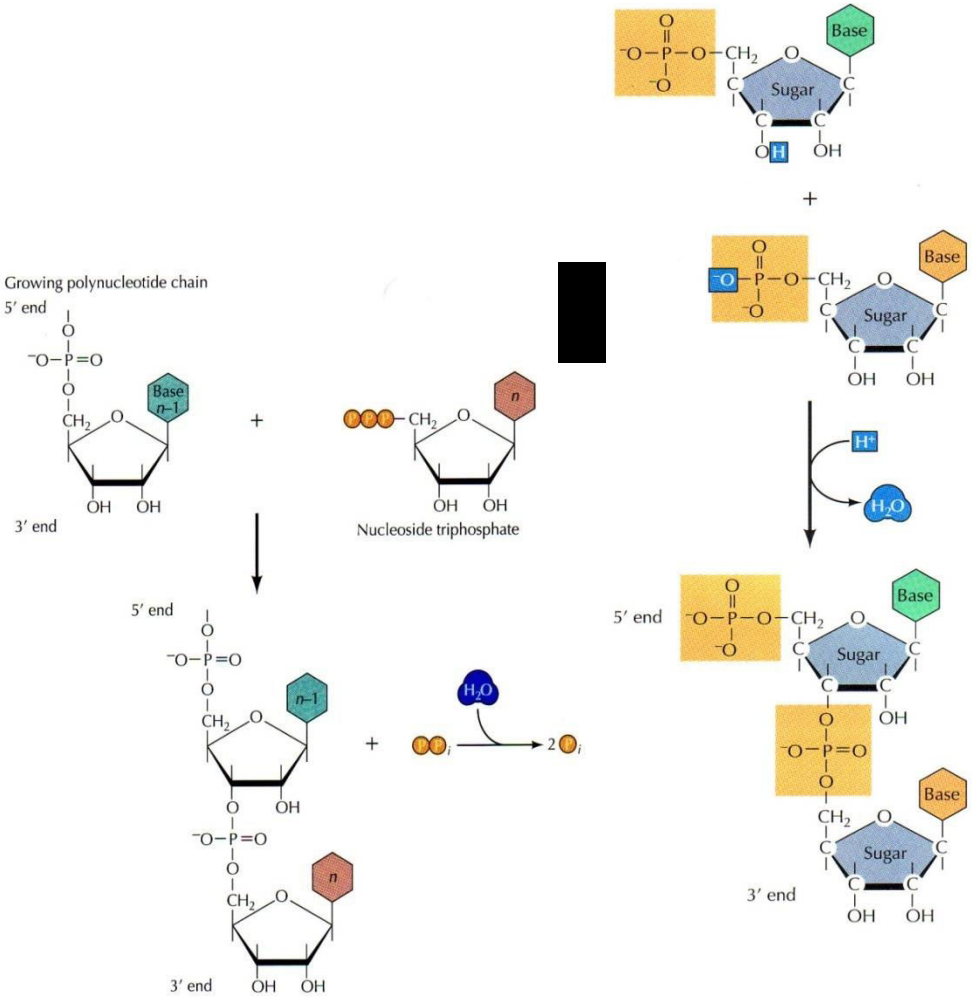


Nucleotide bases



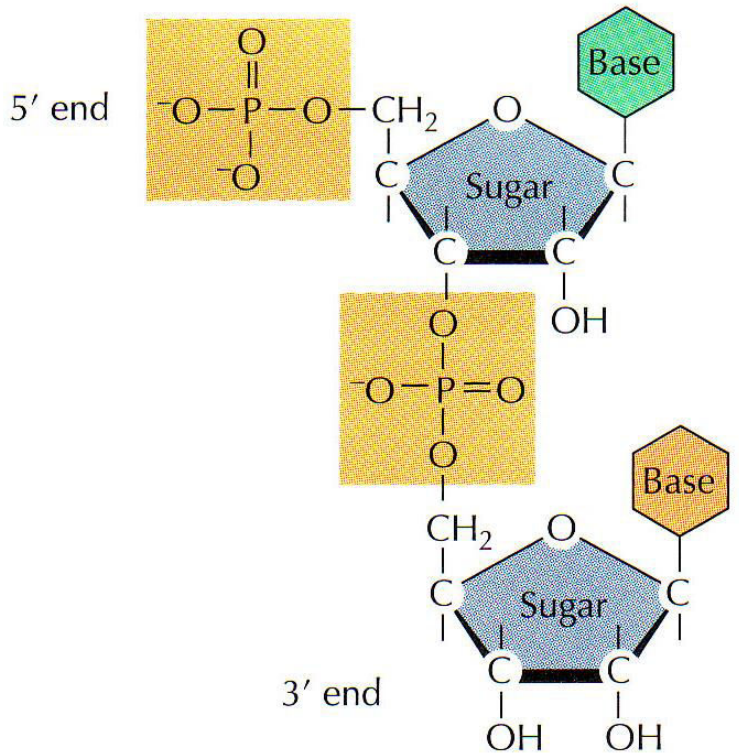
Linear Polymerization of Nucleotides

- Nucleic acids are formed of nucleotide polymers.
- Nucleotides polymerize together by *phospho-diester bonds* via condensation reaction.
- The phospho-diester bond is formed between:
  - Hydroxyl (OH) group of the sugar of one nucleotide.
  - Phosphate group of other nucleotide.



Polymerization of Nucleotides

- The formed polynucleotide chain is formed of:
  - Negative (-ve) charged Sugar-Phosphate backbone.
    - Free 5’ phosphate on one end (5’ end)
    - Free 3’ hydroxyl on other end (3’ end)
  - Nitrogenous bases are not in the backbone
    - Attached to the backbone
    - Free to pair with nitrogenous bases of other polynucleotide chain



Polymerization of Nucleotides

- Nucleic acids are polymers of nucleotides.
- The nucleotides formed of purine or pyrimedine bases linked to *phosphorylated sugars* (nucleotide back bone).
- The bases are linked to the pentose sugar to form *Nucleoside*.
- The nucleotides contain one phosphate group linked to the 5’ carbon of the nucleoside.

Nucleotide = Nucleoside + Phosphate group

Summary

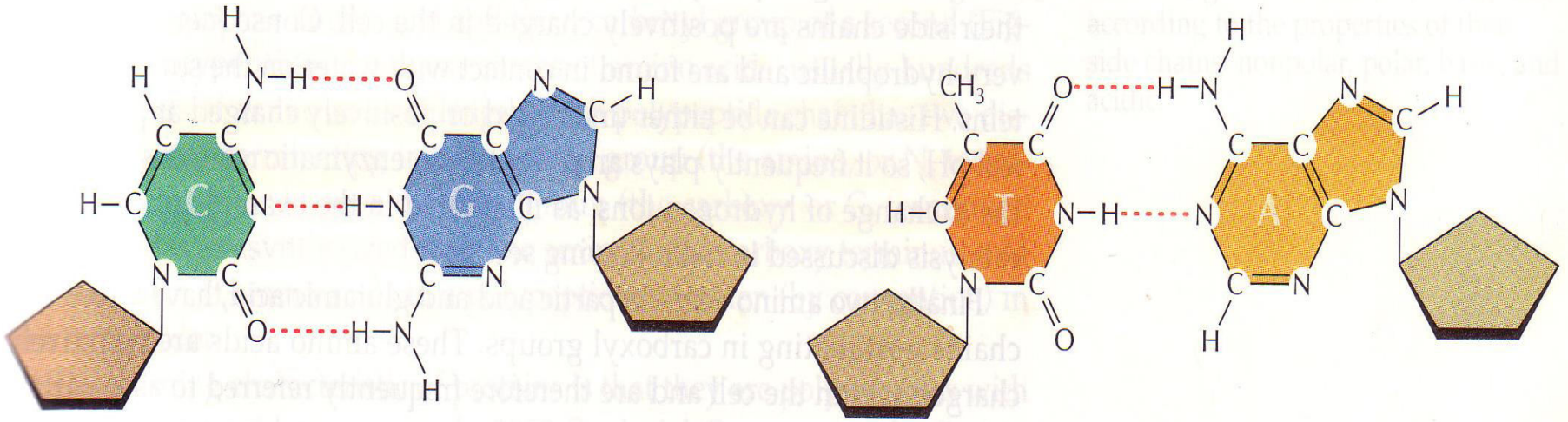
- The polymerization of nucleotides to form nucleic acids occur by condensation reaction by making phospho-diester bond between 5’ phosphate group of one nucleotide and 3’ hydroxyl group of another nucleotide.

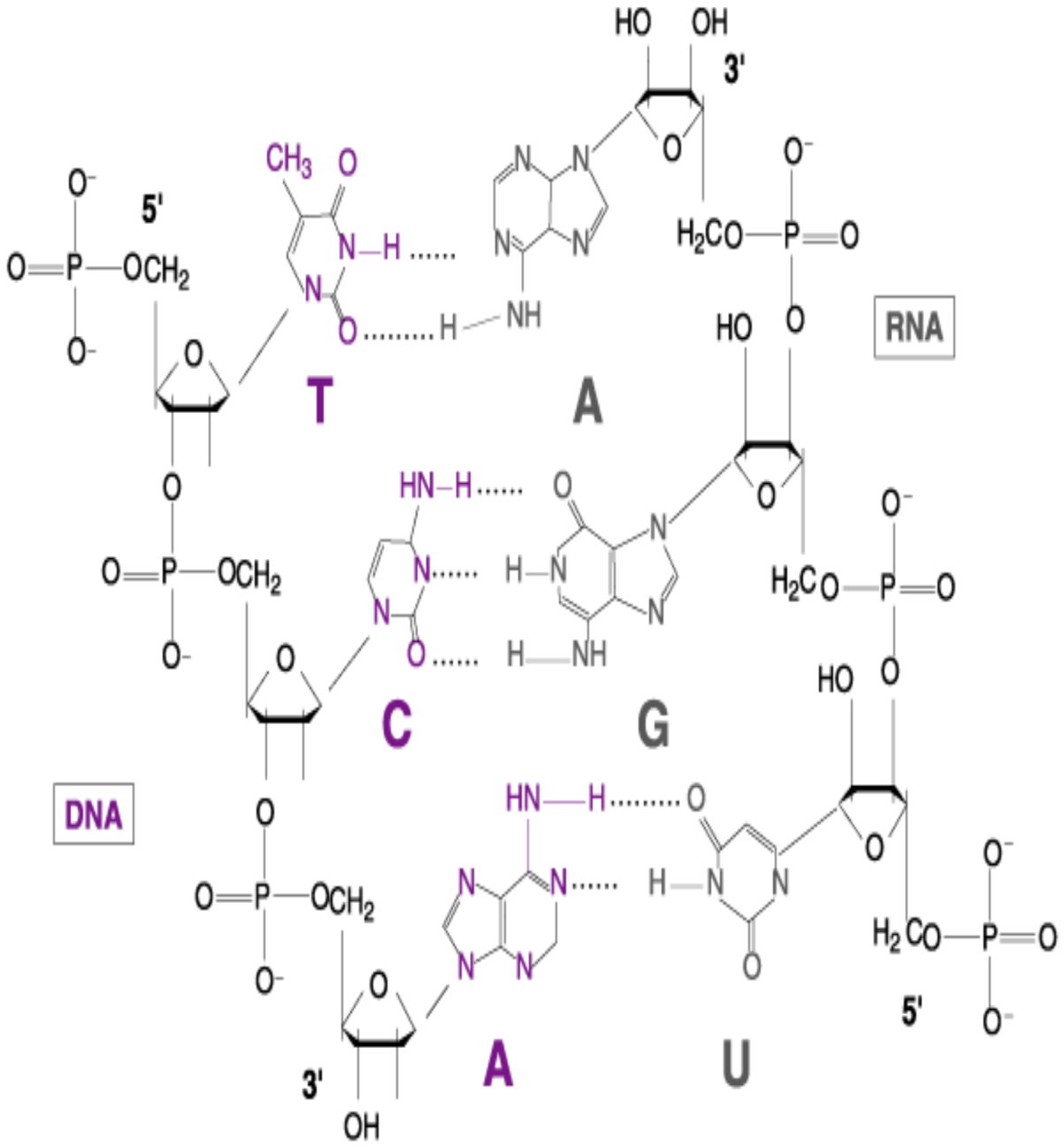
- Polynucleotide chains are always synthesized in the 5’ to 3’ direction, with a free nucleotide being added to the 3’ OH group of a growing chain.

Complementary base pairing

- It is the most important structural feature of nucleic acids.
- It connects bases of one polynucleotide chain (nucleotide polymer) with complementary bases of other chain.
- Complementary bases are bonded together via:
  - Double hydrogen bond between A and T (DNA), A and U (RNA) (A═T or A═U)
  - Triple H-bond between G and C in both DNA or RNA (G≡C)

Base pairing




HO
OH
3'
5'
RNA
DNA
T
A
C
G
A
U
3'
OH
5'

Significance of complementary base pairing

- The importance of such complementary base pairing is that each strand of DNA can act as template to direct the synthesis of other strand similar to its complementary one.

- Thus *nucleic acids are uniquely capable of directing their own self replication*.

- The information carried by DNA and RNA direct the synthesis of specific proteins which control most cellular activities.

Forms of DNA

1- B-form helix

- It is the most common form of DNA in cells.
  - Right-handed helix
  - Turn every 3.4 nm.
  - Each turn contain 10 base pairs (the distance between each 2 successive bases is 0.34 nm)
  - Contain 2 grooves;
    - Major groove (wide): provide easy access to bases
    - Minor groove (narrow): provide poor access.

2- A-form DNA

- Less common form of DNA , more common in RNA
  - Right handed helix
  - Each turn contain 11 b.p/turn
  - Contain 2 different grooves:
    - Major groove: very deep and narrow
    - Minor groove: very shallow and wide (binding site for RNA)

3- Z-form DNA

- Radical change of B-form
  - Left handed helix, very extended
  - It is GC rich DNA regions.
  - The sugar base backbone form Zig-Zag shape
  - The B to Z transition of DNA molecule may play a role in gene regulation.

Denaturing and Annealing of DNA

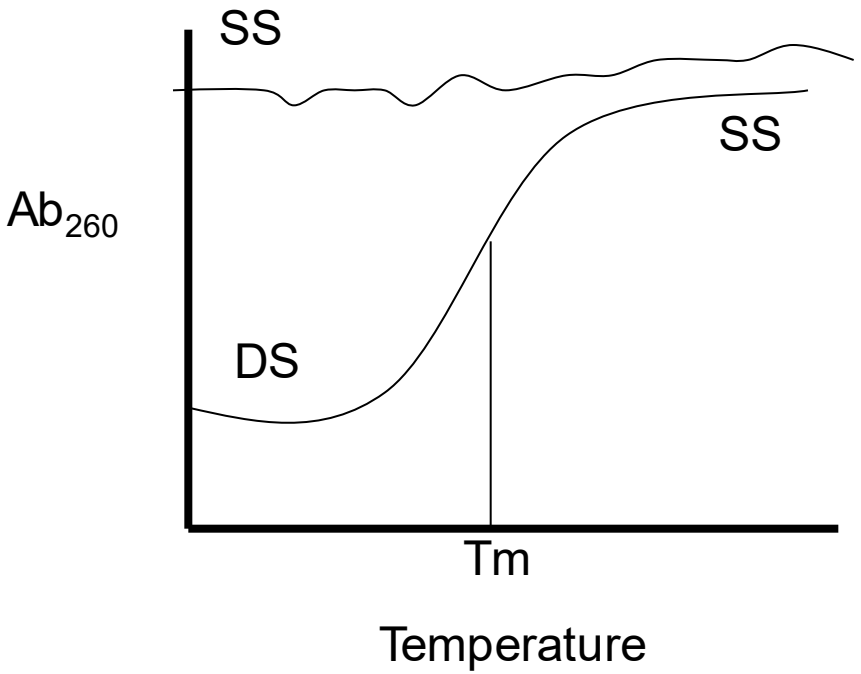
- The DNA double strands can denatured if heated (95ºC) or treated with chemicals.
  - AT regions denature first (2 H bonds)
  - GC regions denature last (3 H bonds)
- DNA denaturation is a reversible process, as denatured strands can re-annealed again if cooled.
- This process can be monitored using the hyperchromicity (melting profile).

Hyperchromicity (melting profile)

- It is used to monitor the DNA denaturation and annealing.
- It is based on the fact that single stranded (SS) DNA gives higher absorbtion reading than double stranded (DS) at wavelength 260º.
- Using melting profile we can differentiate between single stranded and double stranded DNA.

Hyperchromicity (melting profile)

Using melting profile we can differentiate between SS DNA and DS DNA



Tm (*melting temp.*): temp. at which 50% of DS DNA denatured to SS

- Heating of SS DNA: little rise of Ab reading
- Heating of DS DNA: high rise of Ab reading

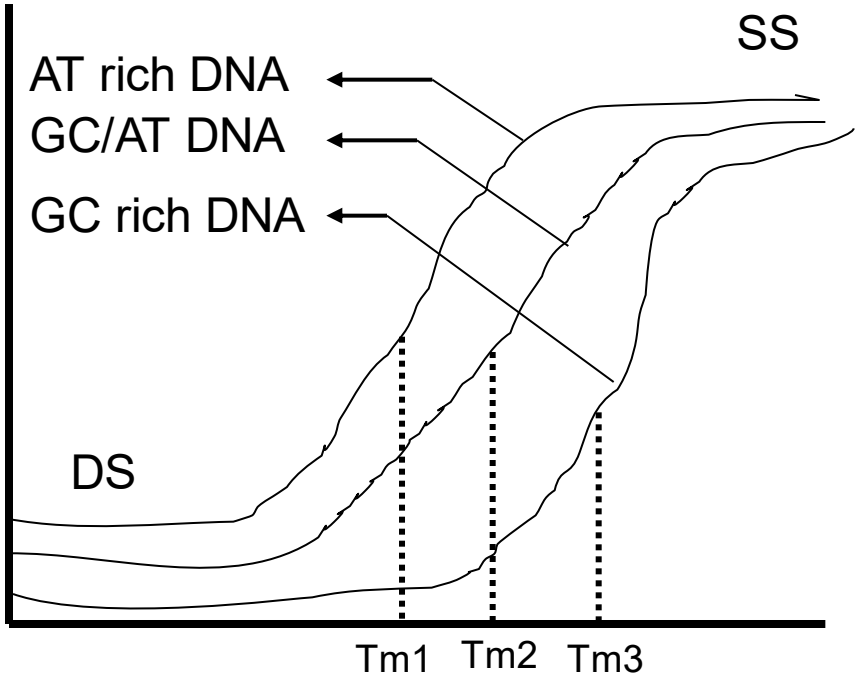
Melting profile *continue…..*

- Melting profile can be also used to give an idea about the type of base pair rich areas using the fact that:
  - A═T rich regions: denatured first (low melting point)
  - G≡C rich regions: denatured last (higher melting point)

Tm1: Small melting temp. of AT rich DNA

Tm2: higher melting temp. of AT/GC equal DNA

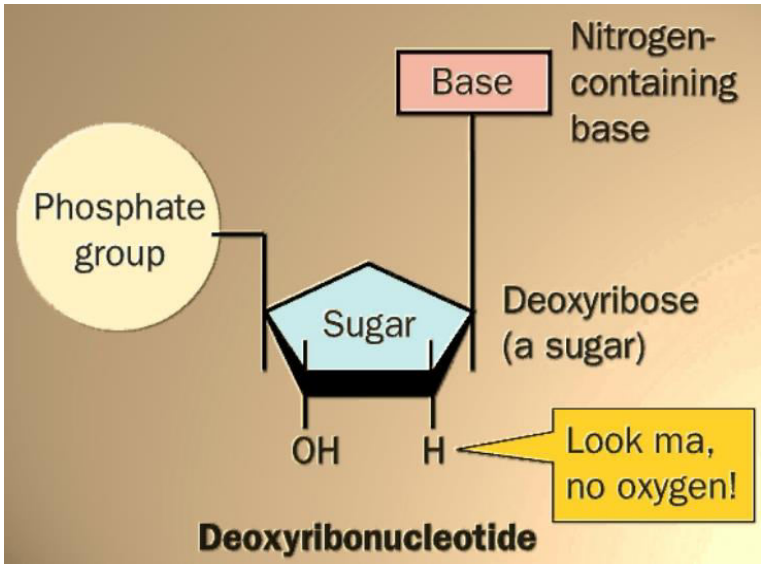
Tm3: Highest melting temp. of GC rich DNA

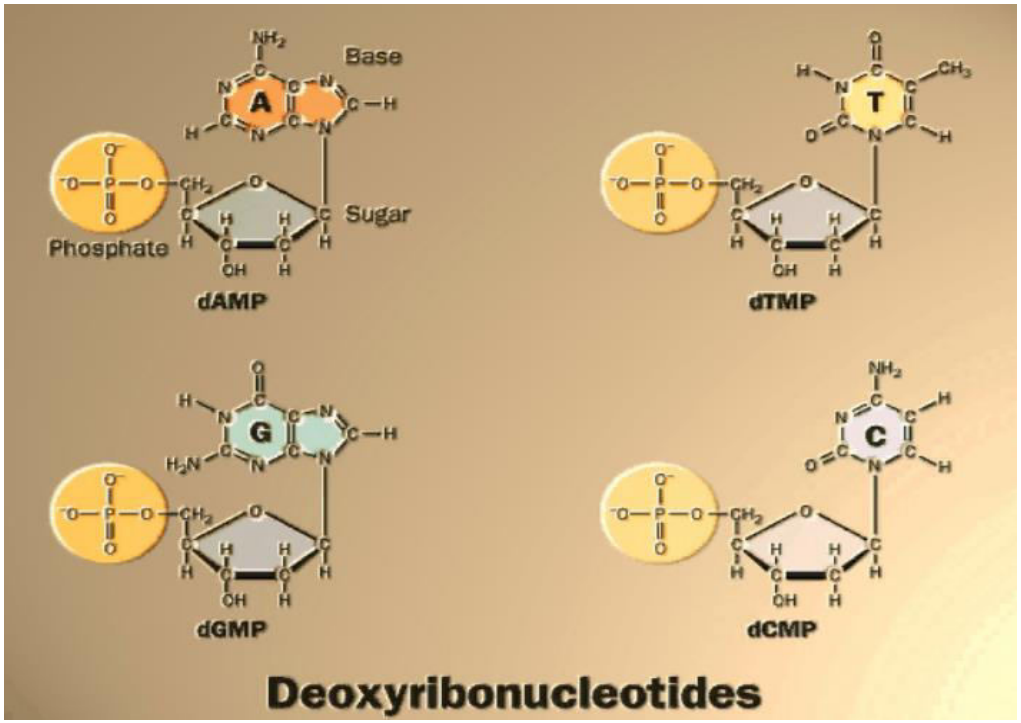


Types of Nucleic acids

There are 2 types of nucleic acids:

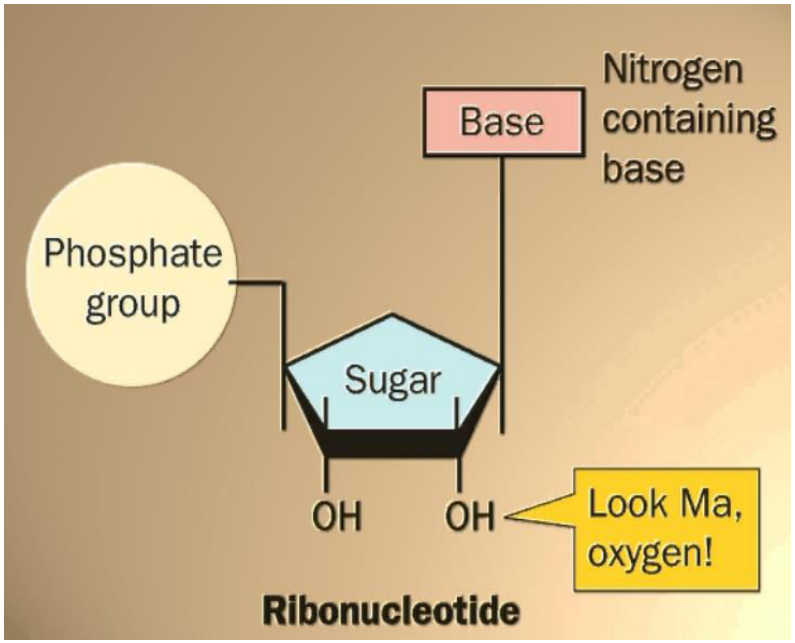
1. **Deoxy-ribonucleic acid (DNA)**
   - **Pentose Sugar is deoxyribose (no OH at 2' position)**
   - **Bases are Purines (A, G) and Pyrimidine (C, T).**

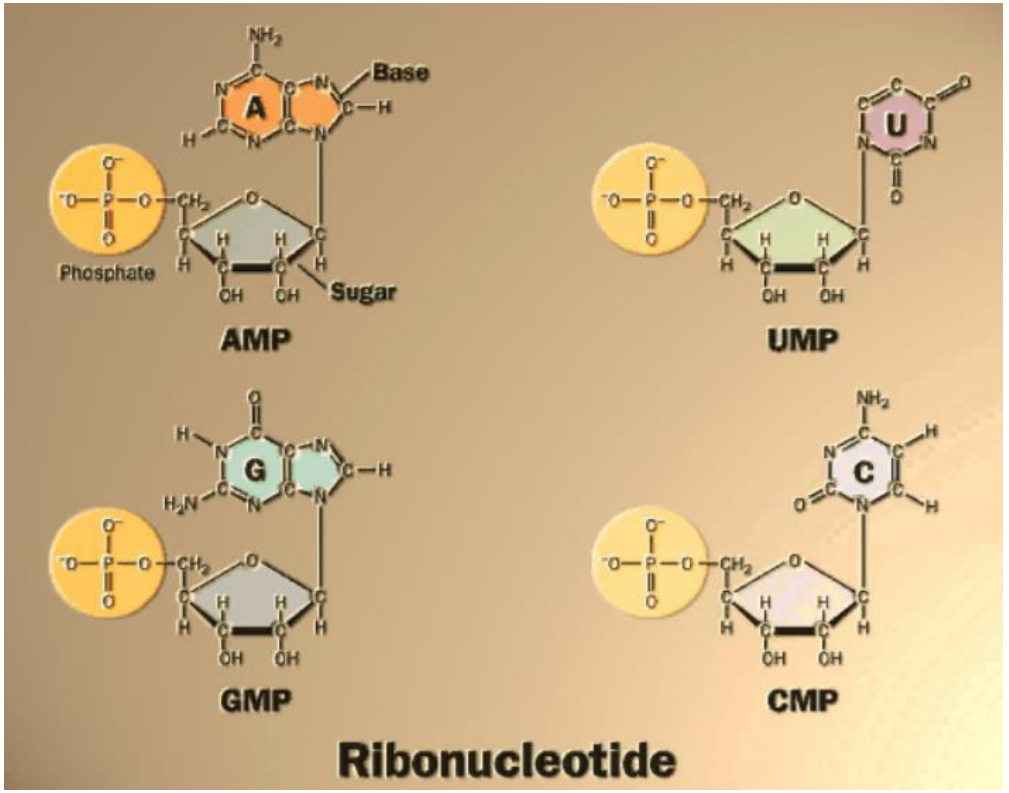




2. Ribonucleic acid (RNA)

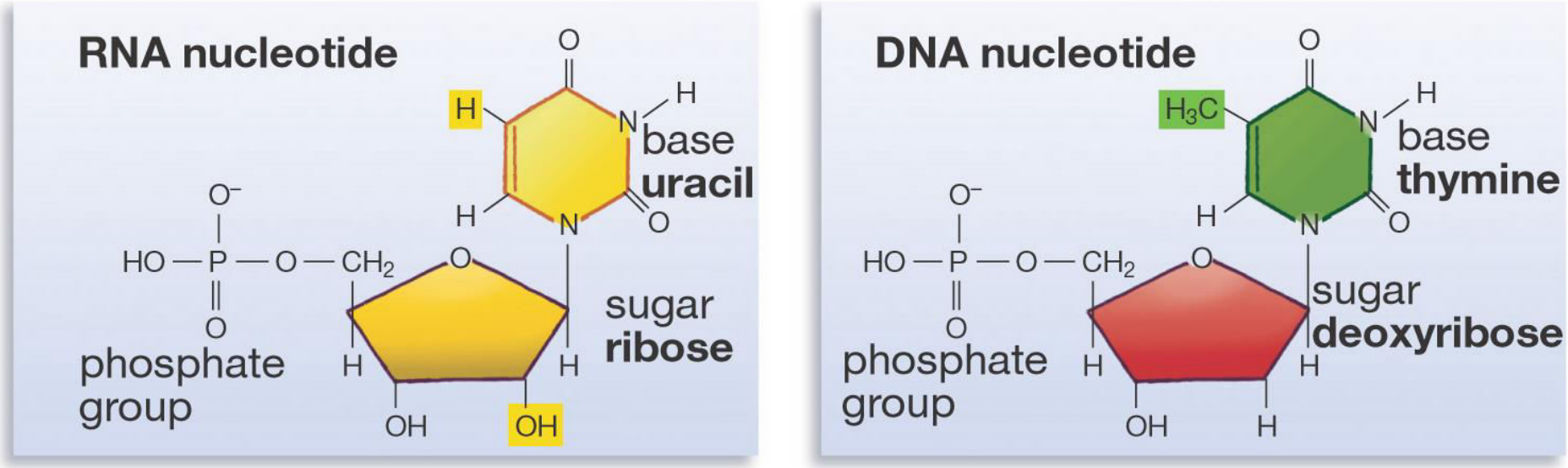
- Pentose Sugar is Ribose.
- Bases are Purines (A, G) and Pyrimidines (C, U).





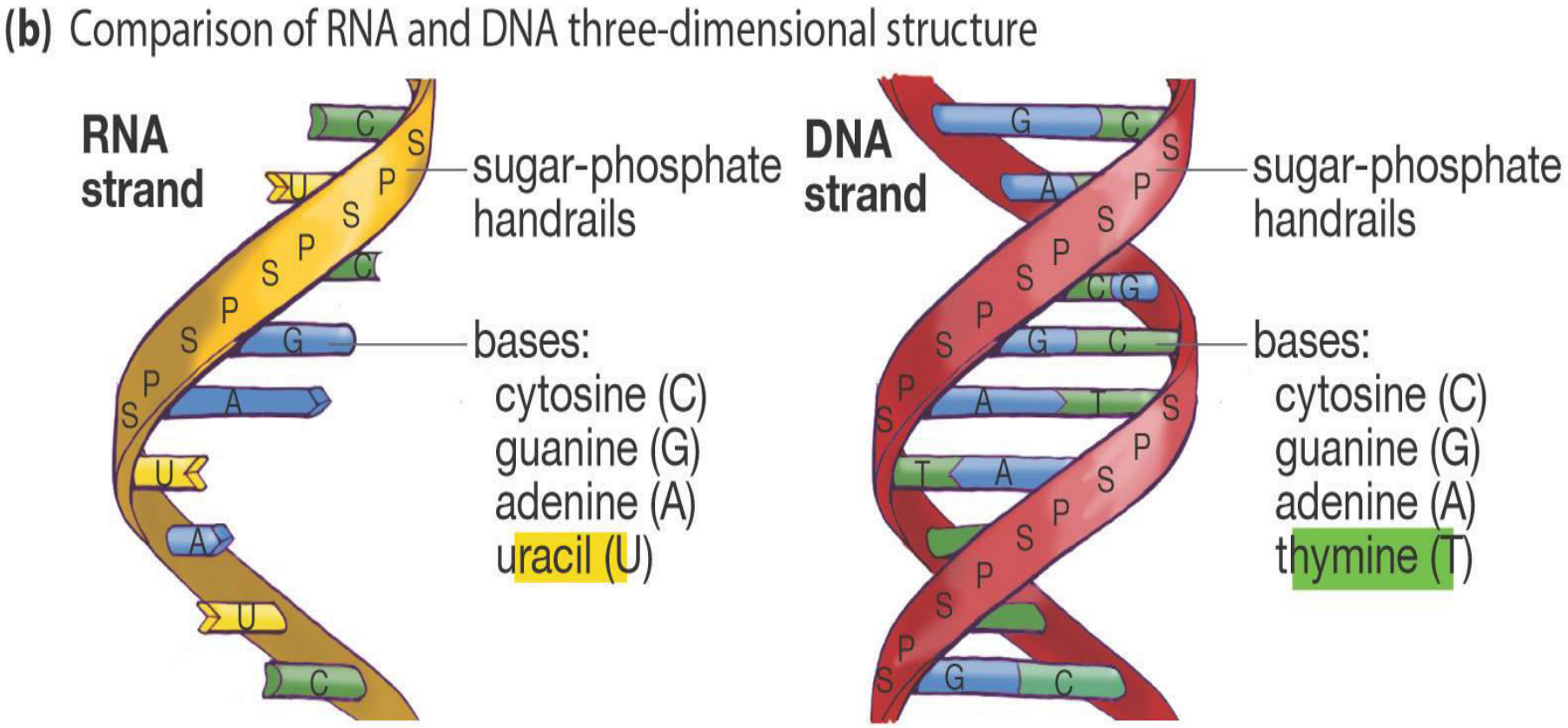
RNA & DNA Nucleotide

(a) Comparison of RNA and DNA nucleotides



RNA is a nucleic acid polymer that uses a slightly different sugar than DNA and the base uracil (U) in place of thymine (T).

RNA & DNA


(b) Comparison of RNA and DNA three-dimensional structure
RNA strand
sugar-phosphate handrails
bases:
cytosine (C)
guanine (G)
adenine (A)
uracil (U)
DNA strand
sugar-phosphate handrails
bases:
cytosine (C)
guanine (G)
adenine (A)
thymine (T)

Function of DNA

DNA in Action

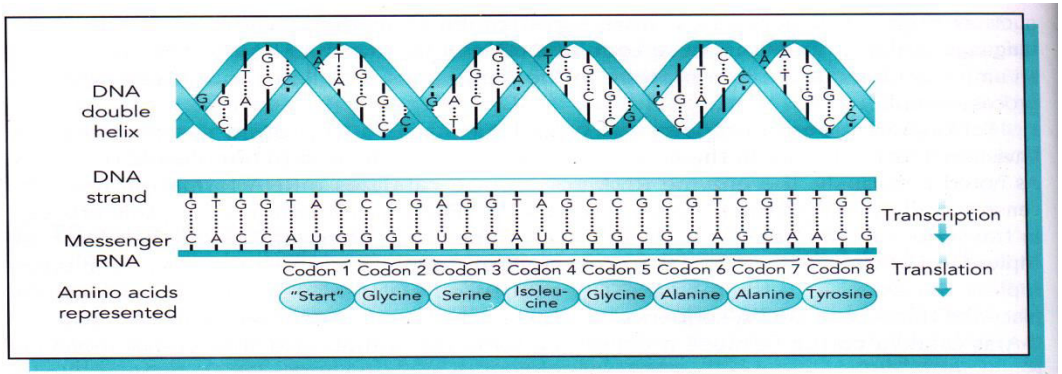
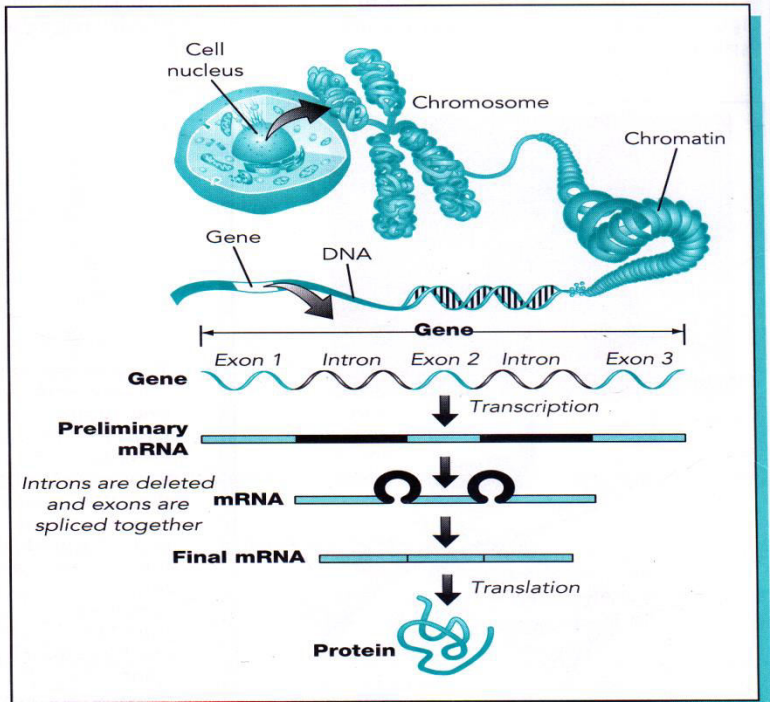


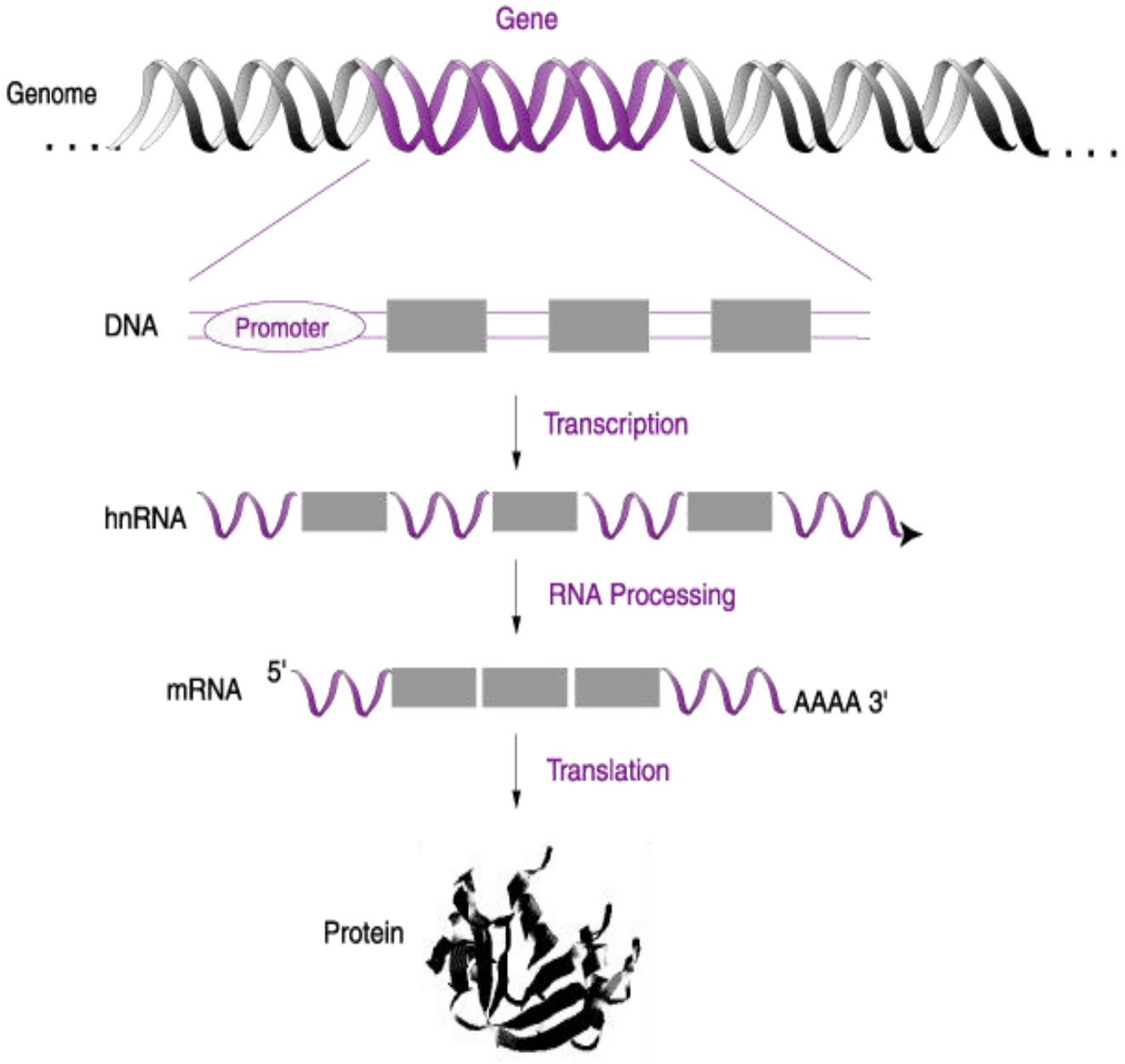
FIGURE 3.14

A broad view of protein synthesis. The DNA molecule unwinds, and the coding strand of DNA is transcribed to messenger RNA. The mRNA then operates as a series of codons, each codon containing three bases. During translation, a particular codon specifies a certain amino acid for placement in the protein chain during translation. Because codons 6 and 7 are identical, alanine molecules occur next to one another in the protein. Codons 2 and 5 (identical) encode glycine.

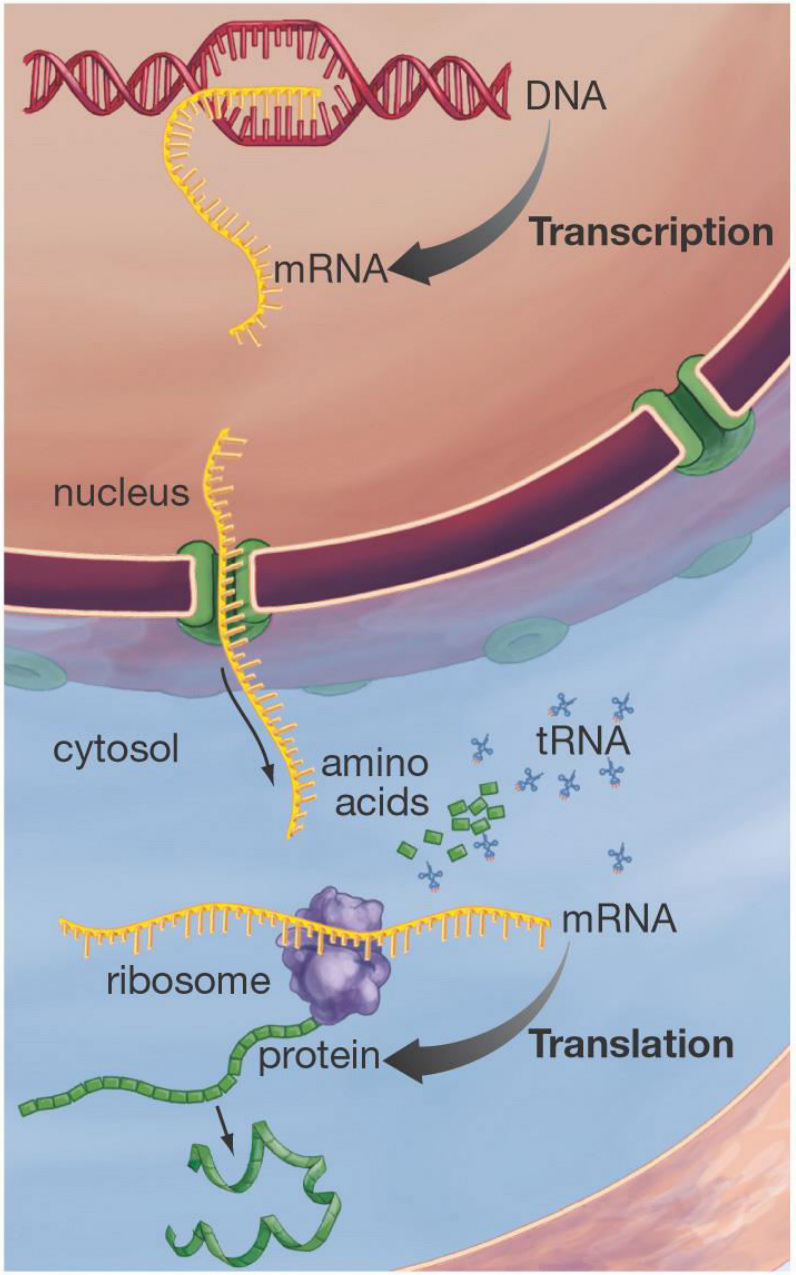
FIGURE 3.15

The formation of mRNA. A gene consists of exons, the parts of the gene expressed as protein, and introns, the intervening sequences between the exons. In the formation of mRNA, the gene is transcribed to a preliminary mRNA molecule. Then the introns are removed biochemically and the exons are spliced together. This activity results in the functional mRNA molecule, which is then ready for translation. This type of processing does not occur in mRNA production in prokaryotic cells such as bacterial cells; it occurs only in eukaryotic cells such as plant animal, and human cells.




Gene
Genome
DNA
Promoter
Transcription
hnRNA
RNA Processing
mRNA
5'
AAAA 3'
Translation
Protein

Molecular Genetics - From DNA to Trait



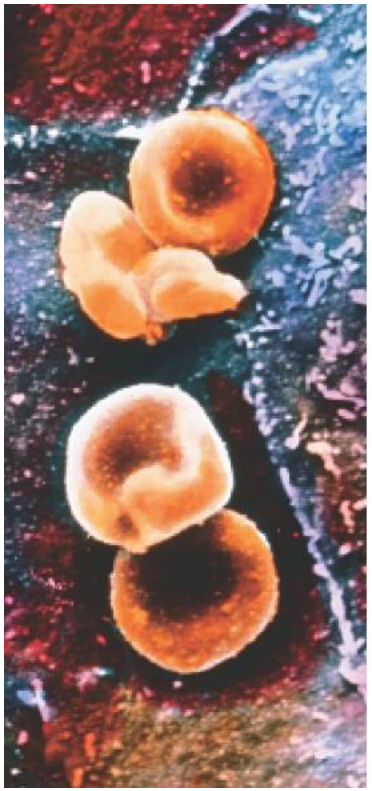
How Are Different Types of Cells Created and Maintained?

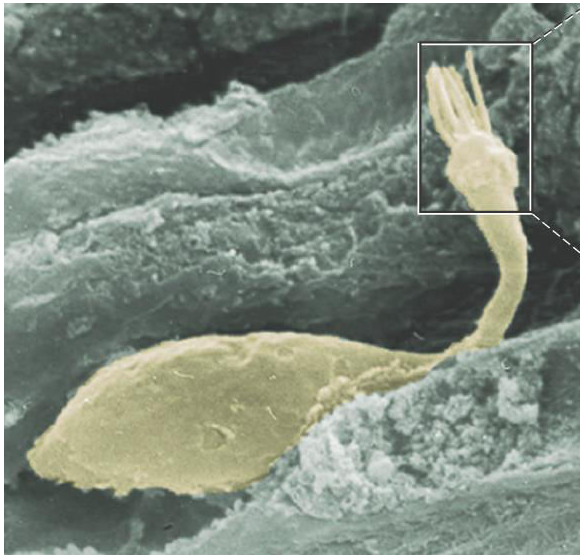
By differential gene expression.

The same genetic information is in all 100 trillion cells of any one person. Different cells use the same blueprint in different ways.

How?

In essence, the control of gene expression occurs by regulating the flow of information from DNA to protein.





Basic Genetic Mechanisms are Universal

The storage of genetic information in DNA, the use of an RNA intermediate that is read in three letter words, and the mechanism of protein synthesis are essentially the same in all organisms.

Among other things, this means cancer can be studied productively in flies or yeast.

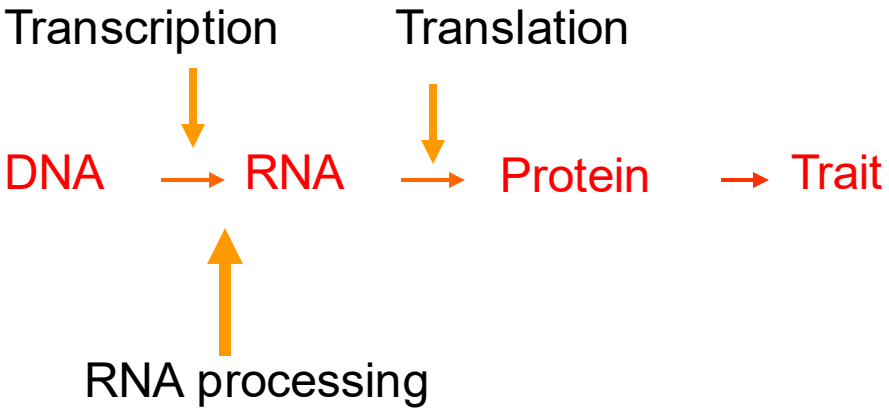
It also means that human genes can be expressed in a plant or mouse genes in a yeast.

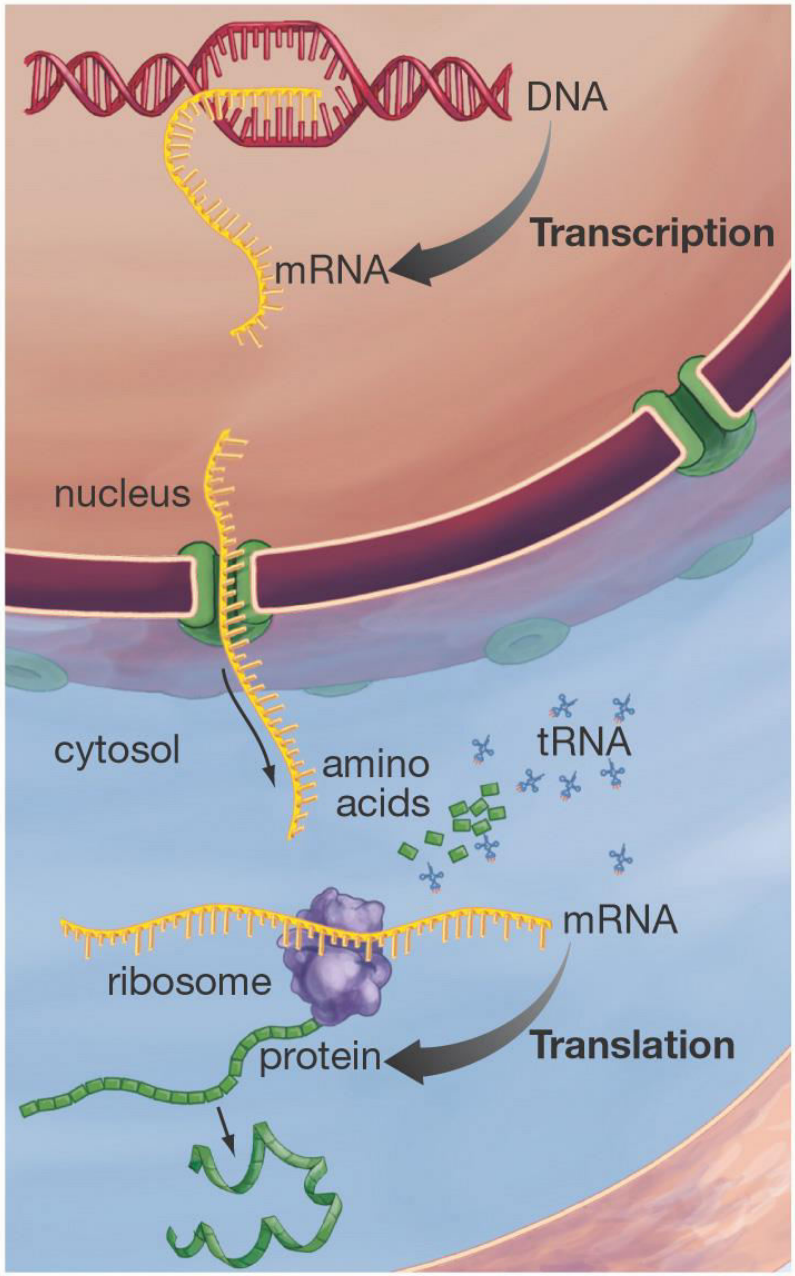
Central Dogma

DNA ----------→ RNA----------→Protein

- This unidirectional flow equation represents the Central Dogma (fundamental law) of molecular biology.
- This is the mechanism whereby inherited information is used to create actual objects, namely enzymes and structural proteins.
- An exception to the central dogma is that certain viruses (retroviruses) make DNA from RNA using the enzyme reverse transcriptase.

The "Central Dogma" of Molecular Genetics

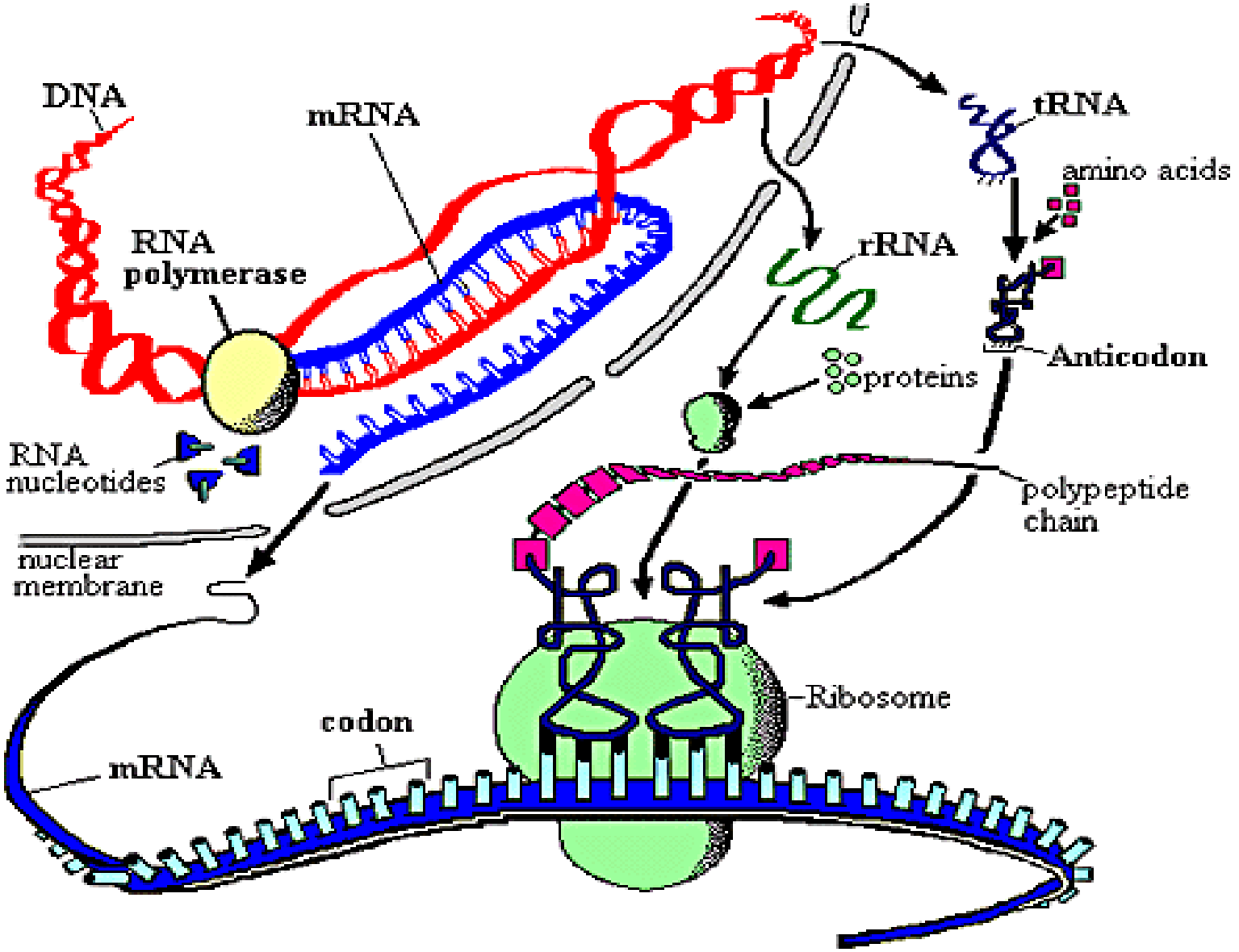




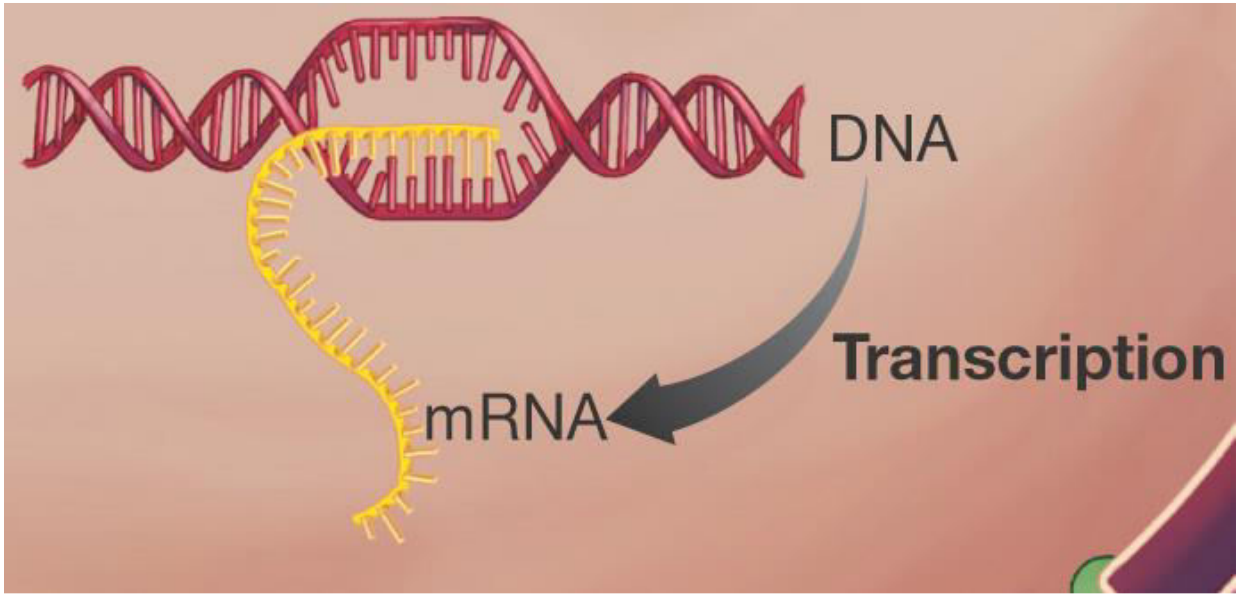
Gene Expression

- Genes are DNA sequences that encode proteins (the gene product).
- Gene expression refers to the process whereby the information contained in genes begins to have effects in the cell.
- DNA encodes and transmits the genetic information passed down from parents to offspring.

RNA & protein Synthesis

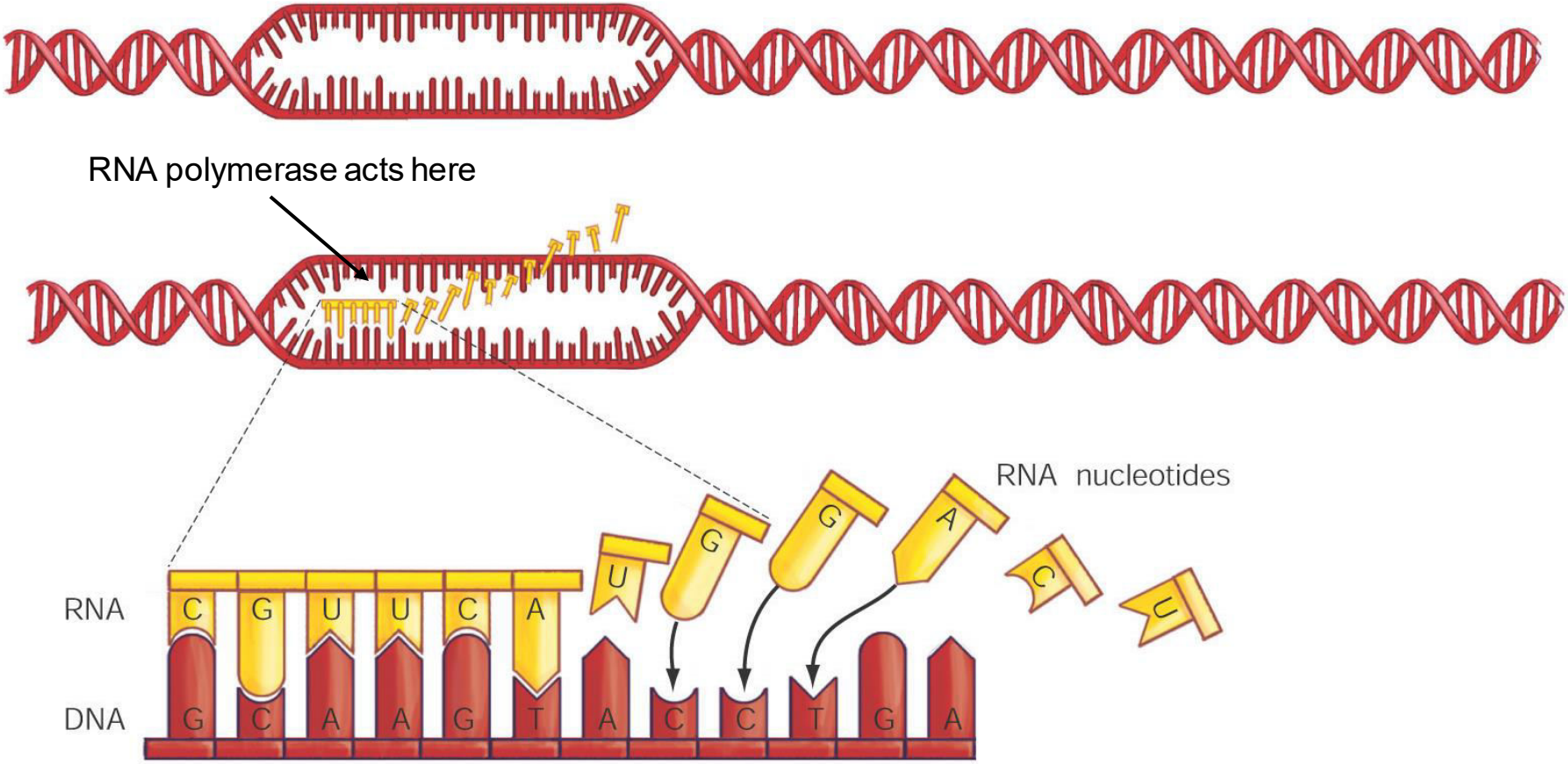


Transcription is a Key Step in Gene Expression



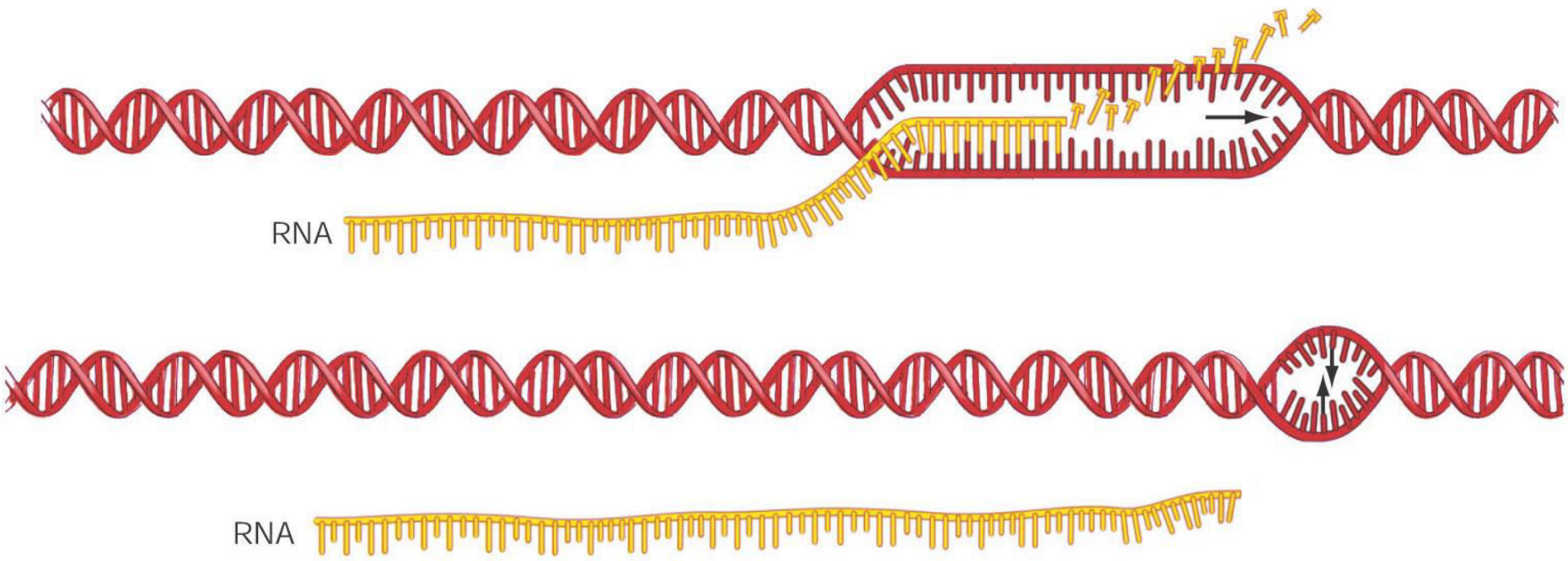
Transcription makes an RNA copy of DNA.

Transcription



The enzyme RNA polymerase opens the DNA strands and synthesizes an RNA complementary to only one of the DNA strands.

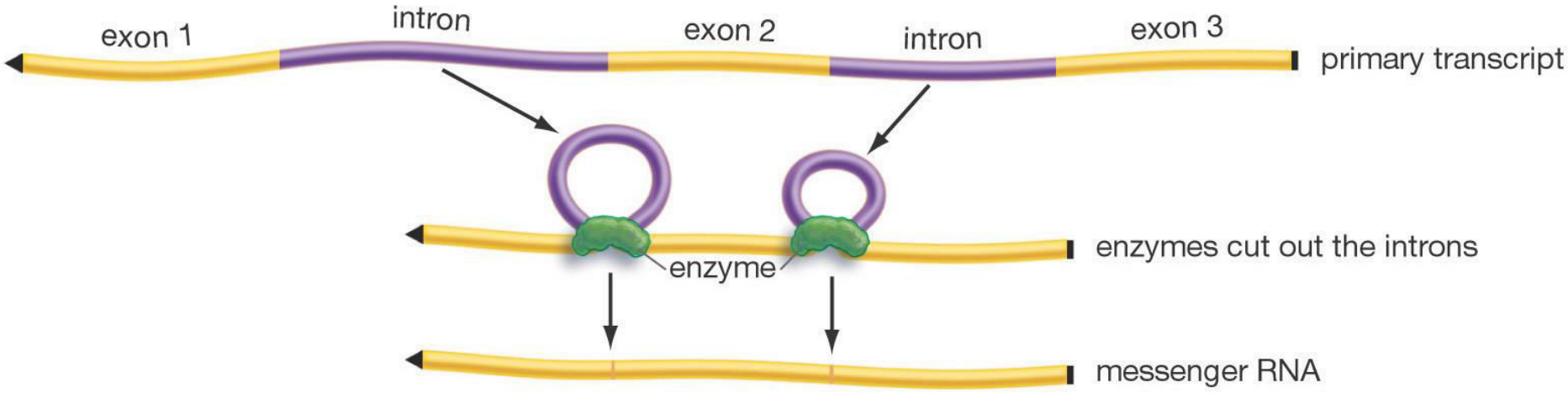
Transcription



The decision to transcribe a gene is the most important step in the control of gene expression.

Transcription starts and stops at distinct sites at the ends of a gene.

Eukaryotic Genes are Segmented

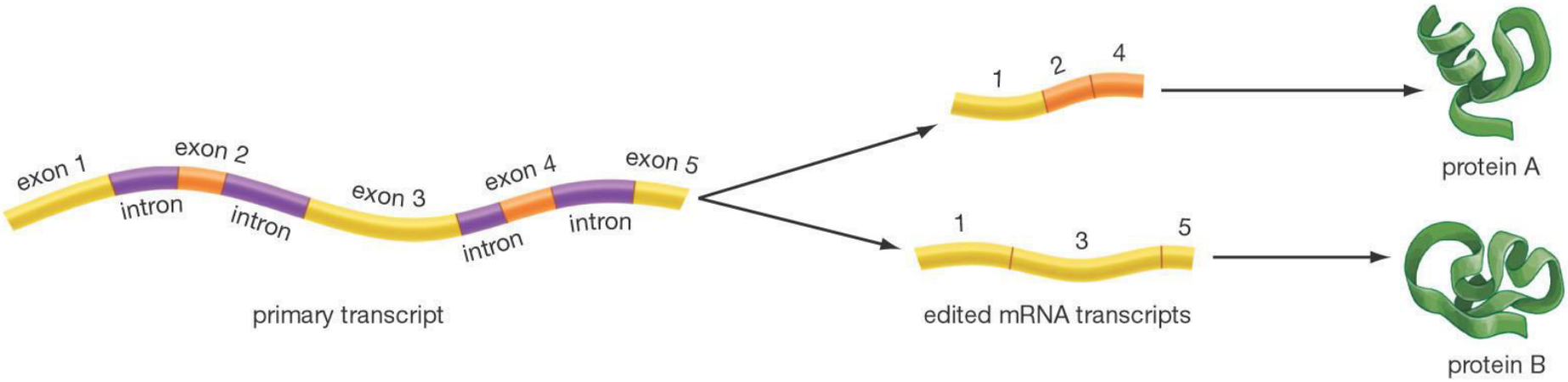


Genes are made of parts represented in the mRNA (exons) and parts that are transcribed but not present in the mRNA (introns).

Introns are removed from the primary transcript and exons are spliced together to make mRNA.

In some genes more than 90% of the pre-mRNA is destroyed, never to appear in the mRNA.

Alternative Splicing

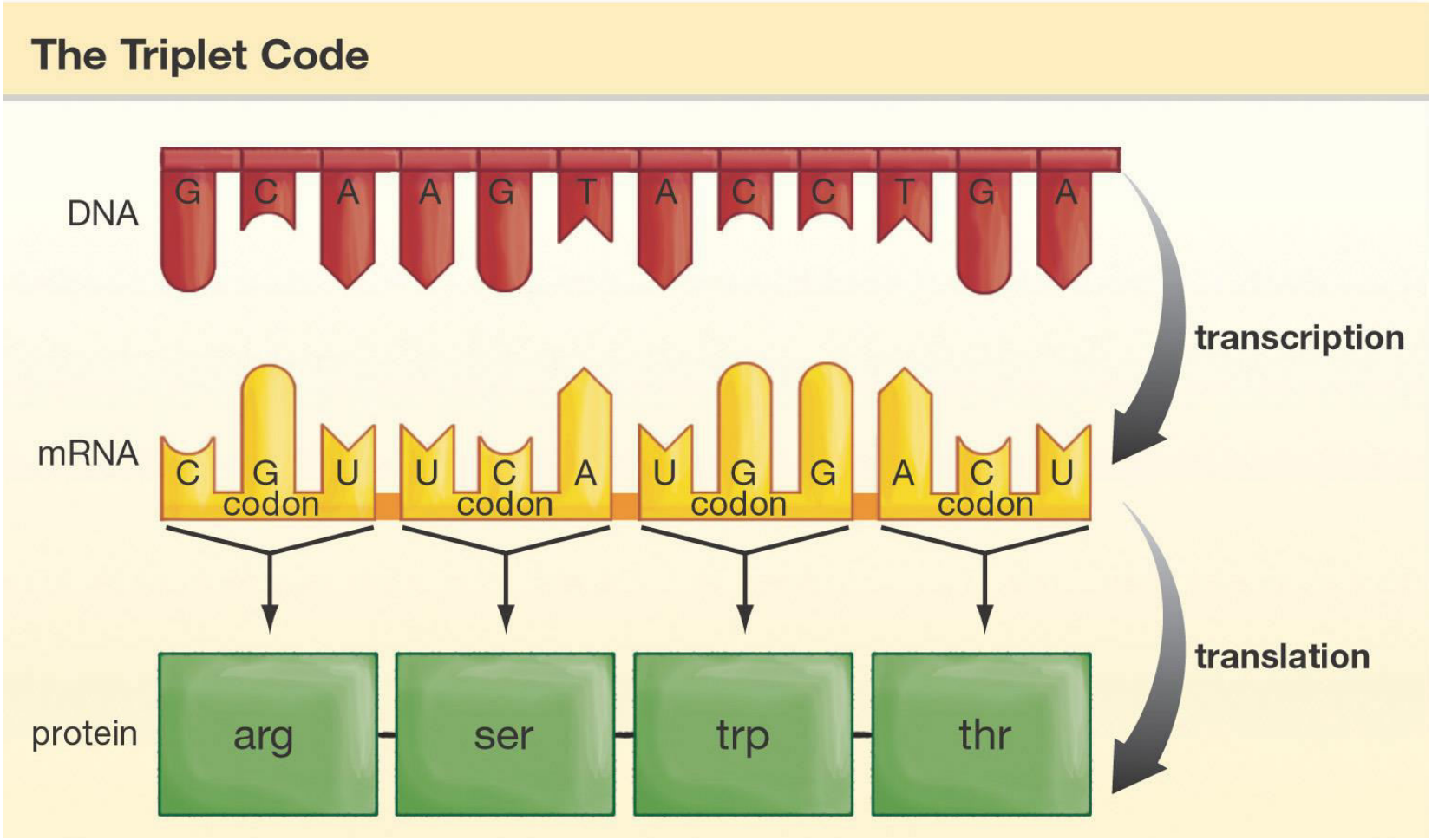


This has the consequence that the count of our genes (~20,000) seriously underestimates the count of our different proteins.

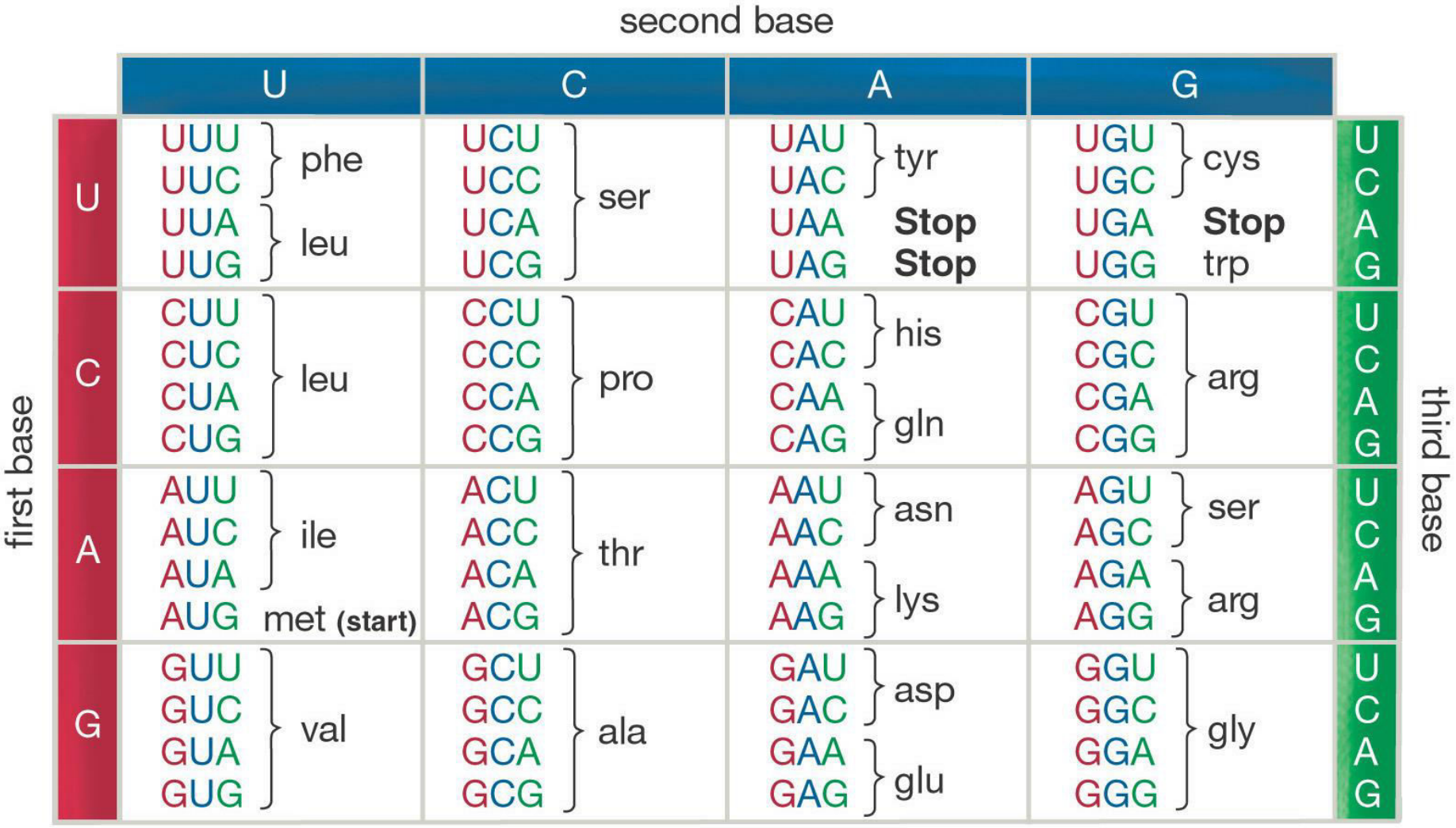
Genetic code

- The alphabet of the genetic code contains only four letters (A,T,G,C).
- A number of experiments confirmed that the genetic code is written in 3-letter words, each of which codes for particular amino acid.
- A nucleic acid word (3 nucleotide letters) is referred to as a *codon.*

The Genetic Language Uses 4 Letters Written Into 3-Letter Words



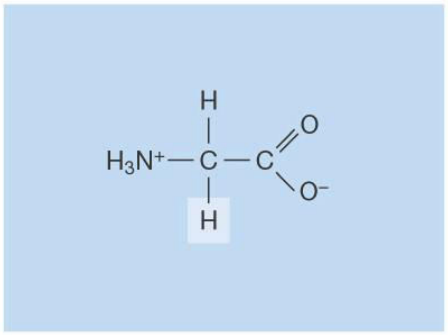
The Genetic Code is Biology's Rosetta Stone



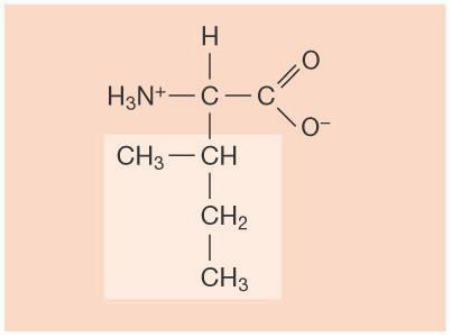
These are the words of the genetic language

Amino Acids – What the Genetic Code Specifies

glycine **(gly)**



isoleucine **(ile)**



Two examples

There are 20 different amino acids

Table 14.1 Amino Acids

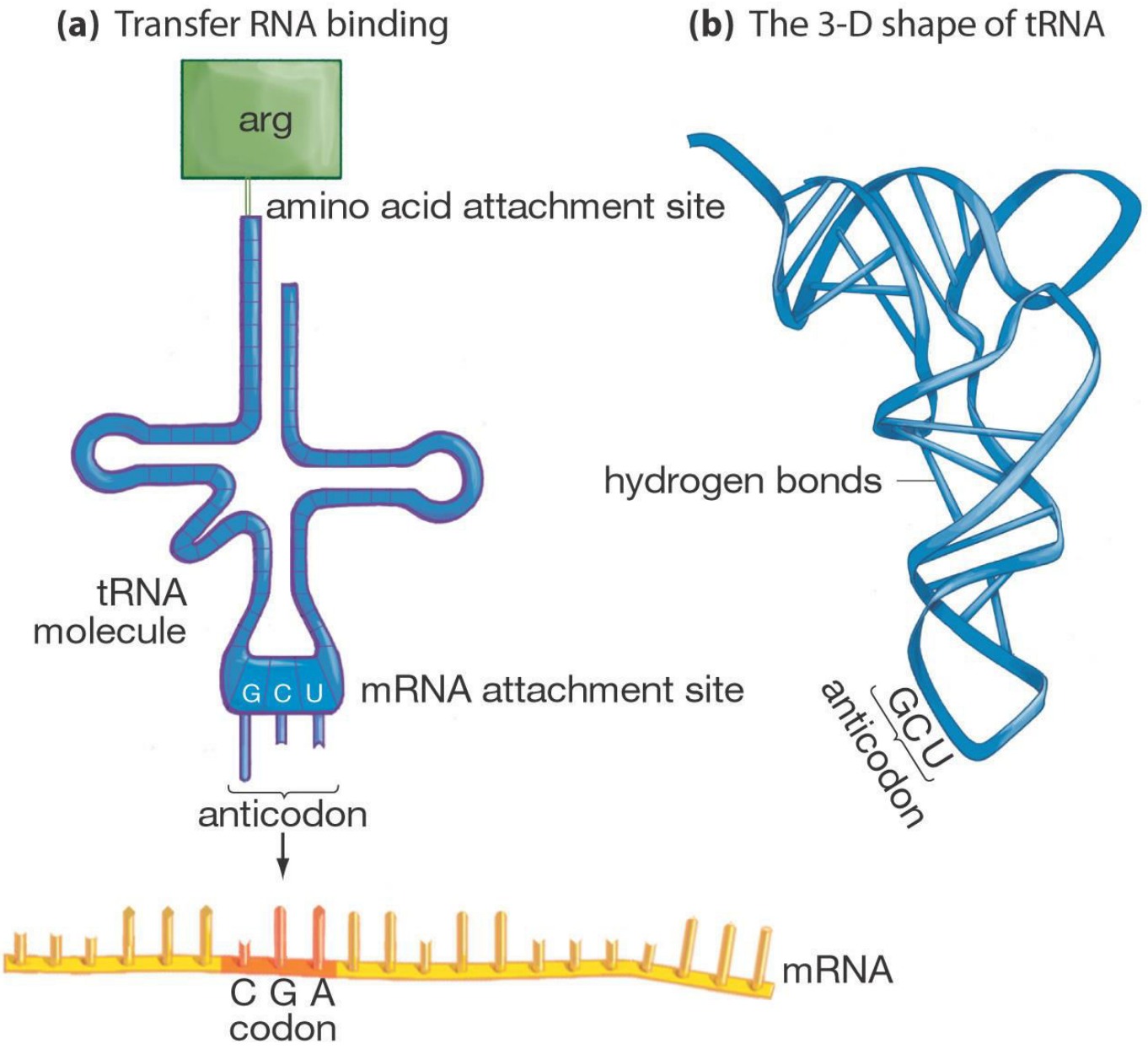
Amino Acid	Abbreviation
Alanine	ala
Arginine	arg
Asparagine	asn
Aspartic acid	asp
Cysteine	cys
Glutamine	gln
Glutamic acid	glu
Glycine	gly
Histidine	his
Isoleucine	ile
Leucine	leu
Lysine	lys
Methionine	met
Phenylalanine	phe
Proline	pro
Serine	ser
Threonine	thr
Tryptophan	trp
Tyrosine	tyr
Valine	val

There are Different RNAs with Distinct Functions

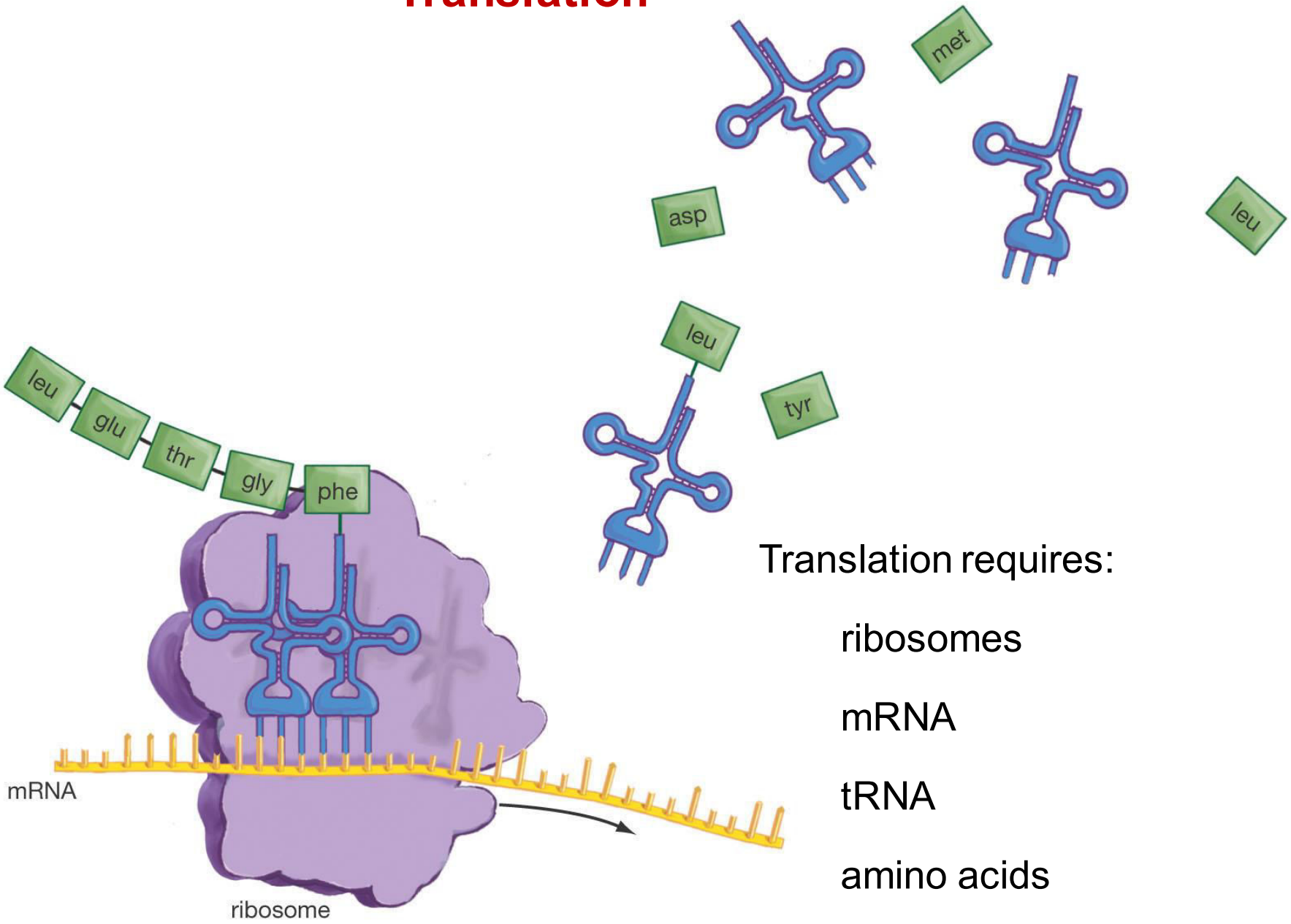
Type of RNA	Functions in	Function
Messenger RNA (mRNA) 	Nucleus, migrates to ribosomes in cytoplasm	Carries DNA sequence information to ribosomes
Transfer RNA (tRNA) 	Cytoplasm	Provides linkage between mRNA and amino acids; transfers amino acids to ribosomes
Ribosomal RNA (rRNA) 	Cytoplasm	Structural component of ribosomes

Recently, a new class of RNA, microRNA and small interfering RNA (siRNA, has been shown to regulate gene expression.

tRNA Is An Adpator That Couples Codons and Amino Acids



Translation



Translation requires:

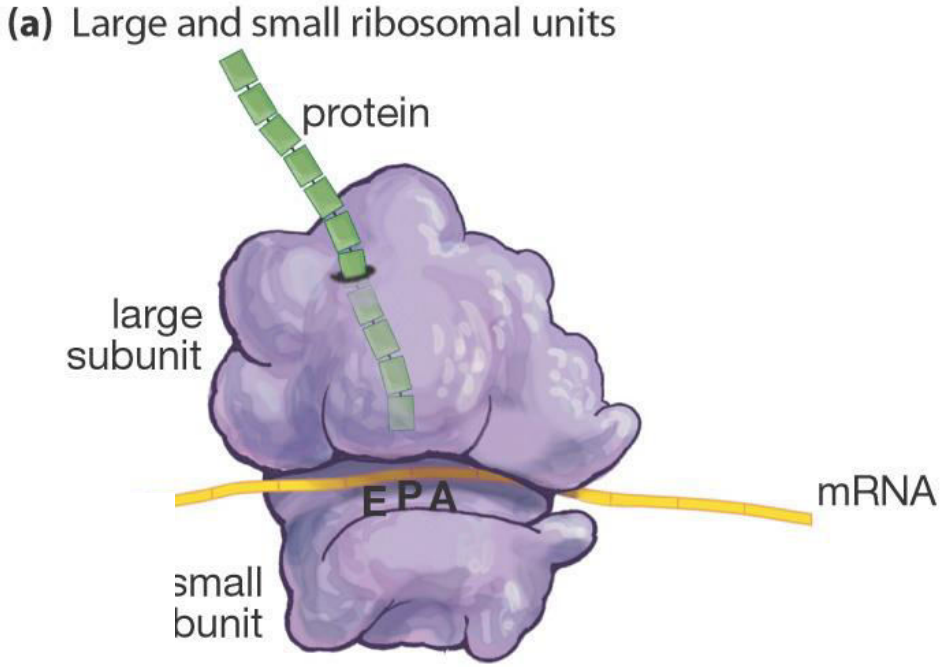
ribosomes

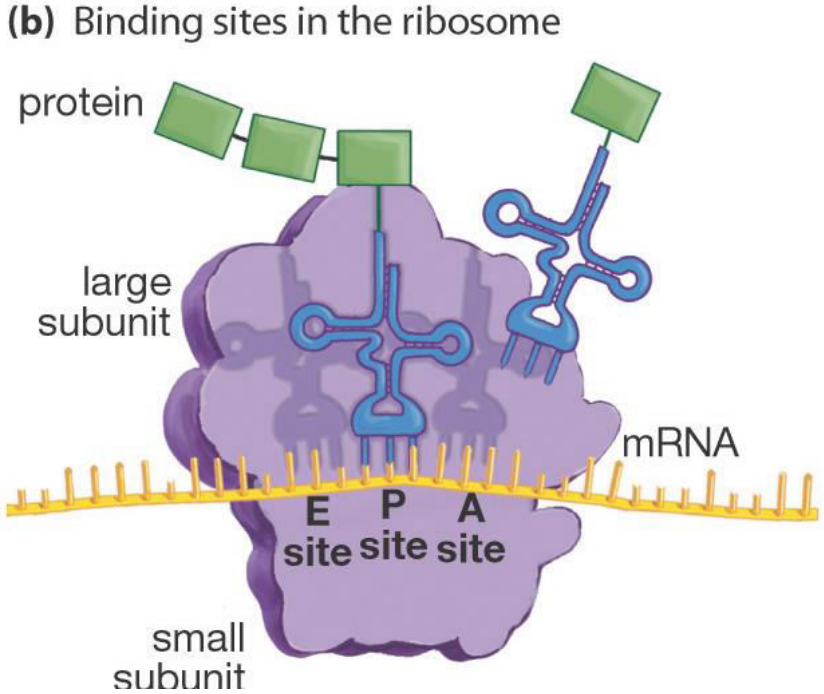
mRNA

tRNA

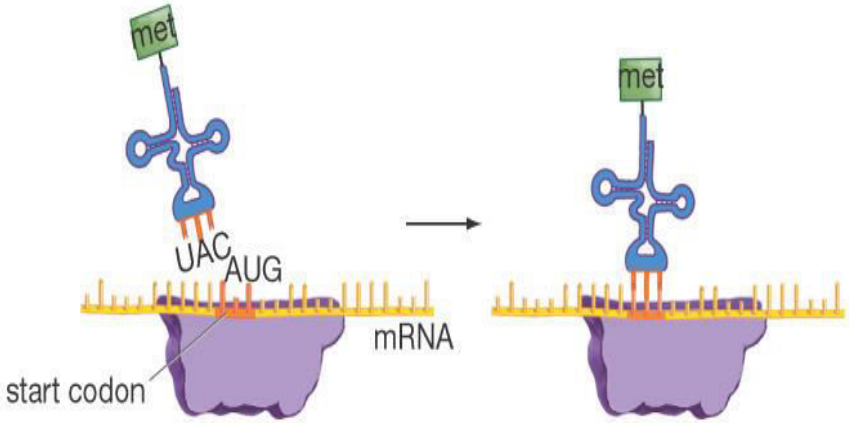
amino acids

Ribosomes are Complicated Protein Synthesizing Machines

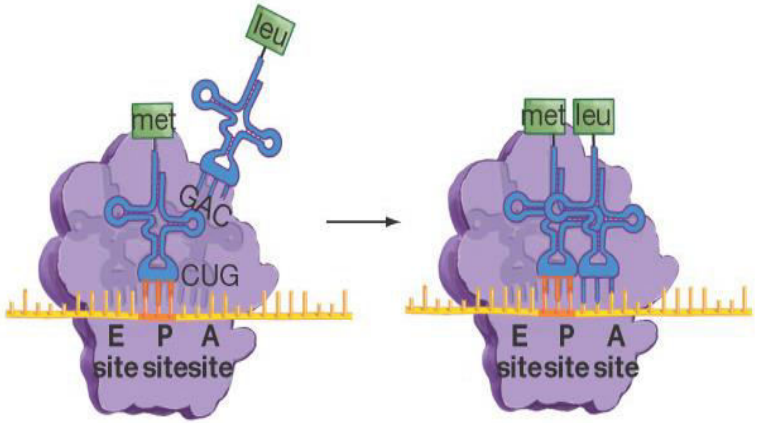




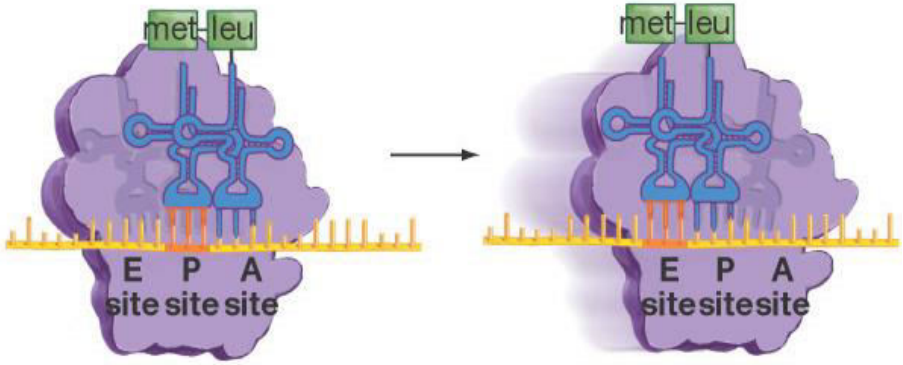
Translation is a Cyclic, Multistep Process



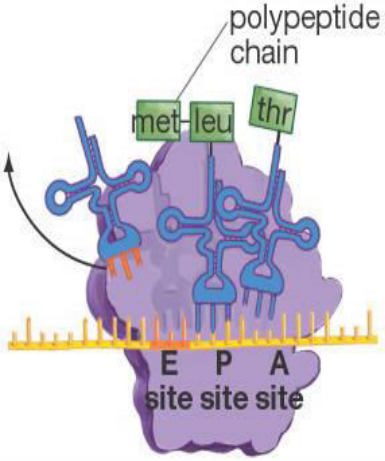
Step 1



Step 2

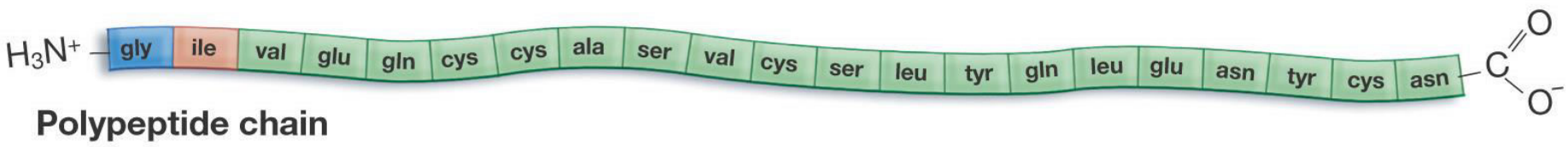


Step 3

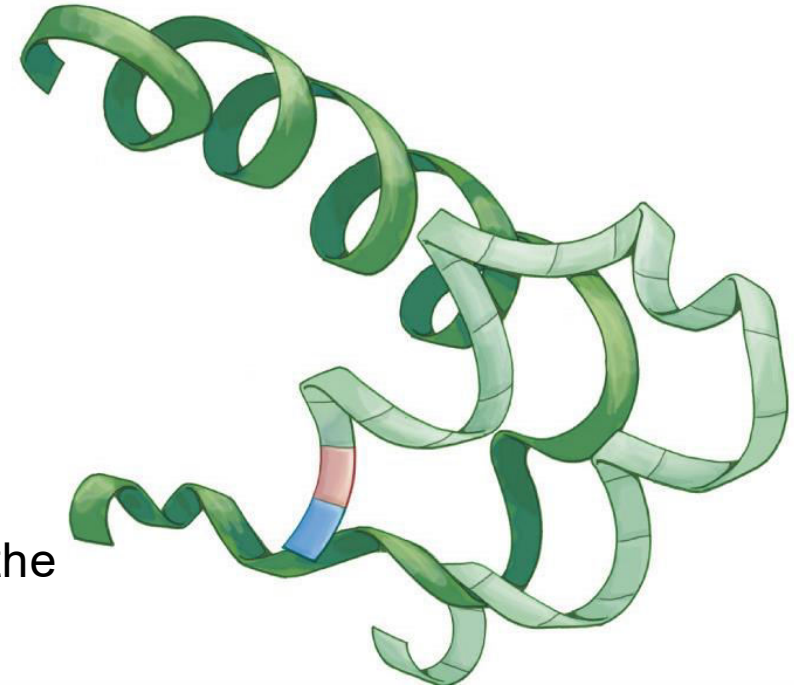


Step 4

What Translation Accomplishes



Polypeptide chain



Protein

The sequence of amino acids determines the structure, and therefore the function, of a protein.

In translation, information present in the mRNA is read by the ribosome to synthesize a polypeptide.

Thank you