APPLICATION CHEMISTRY:

Chemistry of Life

(Genetic Information)

By: Mr. Chan M.H., Lucas

DNA

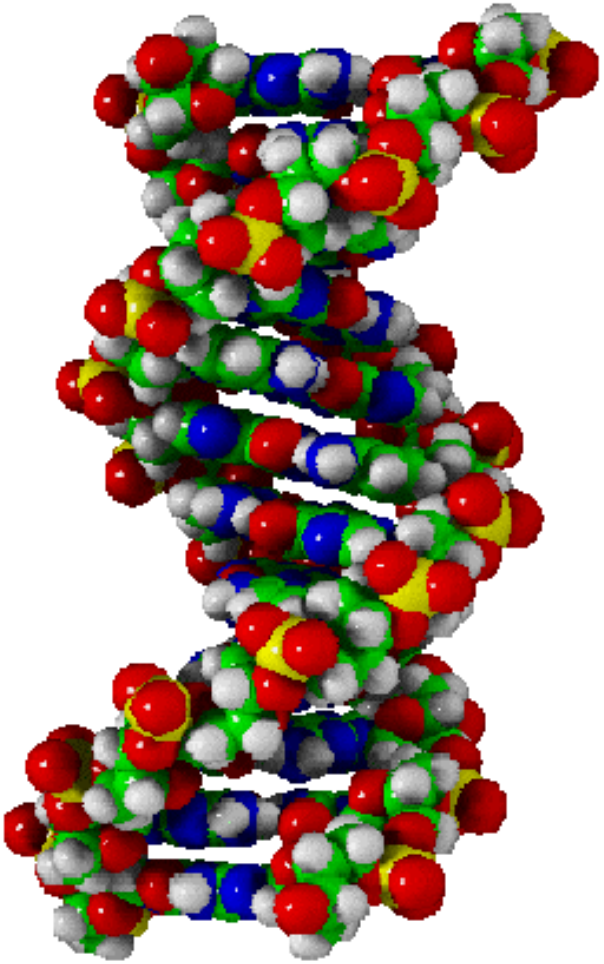


Figure 1.30 – Computer-generated picture of the double helix

DNA

- **Two** deoxyribonucleic acid, **DNA strands, running in opposite directions**, are **linked together** in a ladder-like molecule – but a **twisted ladder** – a right-handed helix (Figure 1.30).

- Deoxyribonucleic acid (DNA) controls heredity on a molecular level:

  – it is a self-replicating molecule capable of passing genetic information from one generation to the next;

  – it contains in its base sequence the genetic code used to synthesise proteins.

DNA

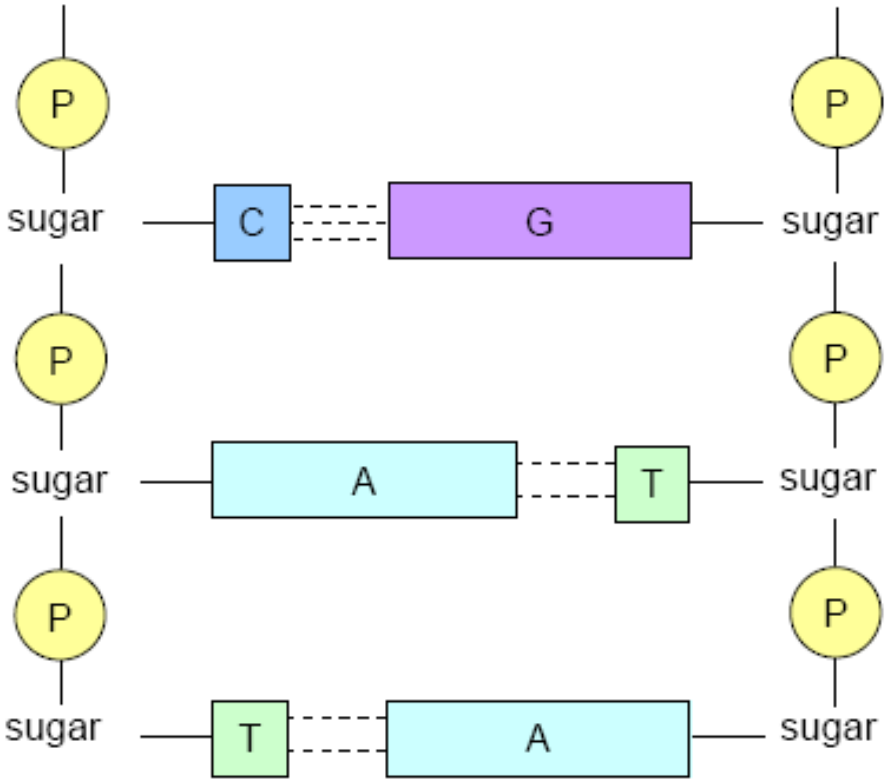


Figure 1.32 – The sugar-phosphate-base structure of DNA

Structure of DNA

- A strand of DNA is a **condensation polymer** of units called **nucleotides**.

- Nucleotides (Figure 1.33) are themselves made from three components:

  (1) a **sugar** - deoxyribose (a **pentose sugar** with a five membered ring).

  (2) a **phosphate group** - attached by an **ester link to** the **deoxyribose**.

  (3) a **nitrogen-containing organic base**
  - adenine (**A**),
  - guanine (**G**),
  - thymine (**T**),
  - cytosine (**C**).

Structure of DNA

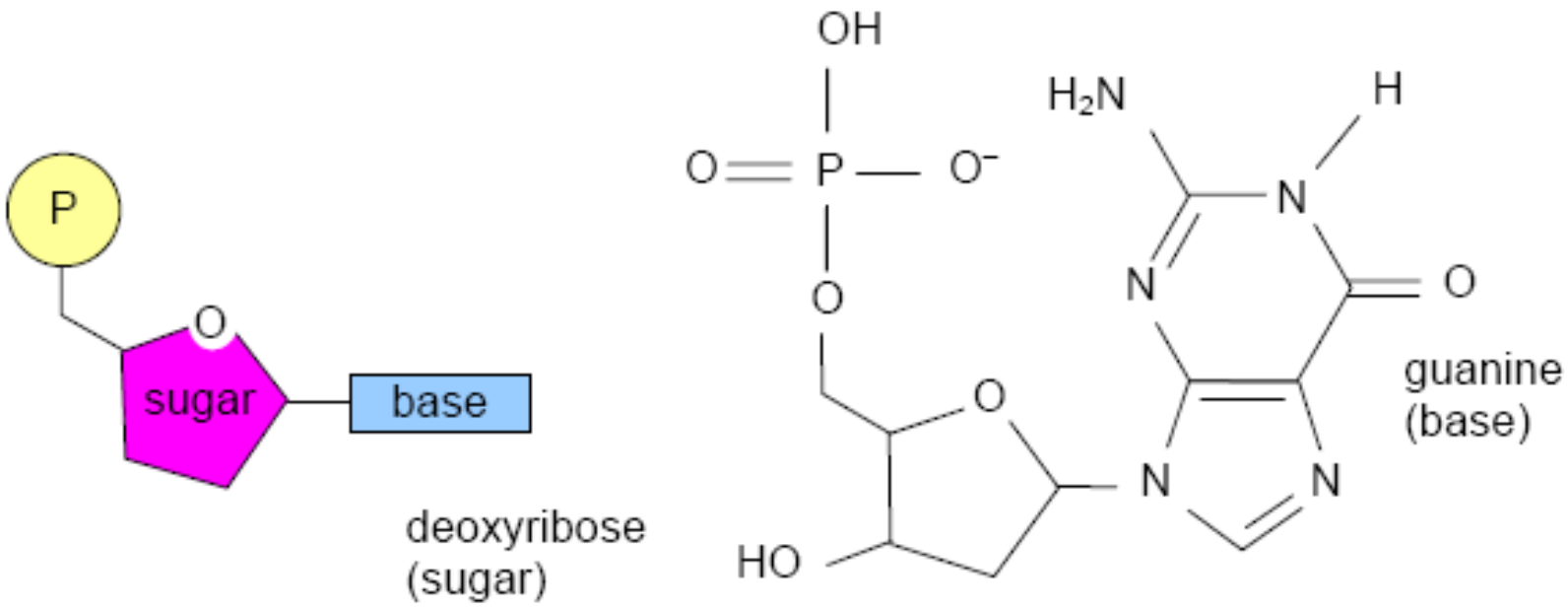


Figure 1.33 – the three components that make up a nucleotide

Structure of DNA

- The **sugar** molecule in the nucleotides that make up DNA is deoxyribose (a **pentose sugar with a five membered ring**).
- The **phosphate group** is **attached by an ester link to the deoxyribose**.
- The **final components of the nucleotides** in DNA are the **four different bases** (all of which are cyclic compounds formed from carbon, nitrogen and hydrogen):
  - **adenine (A),**
  - **guanine (G),**
  - **thymine (T),**
  - **cytosine (C).**

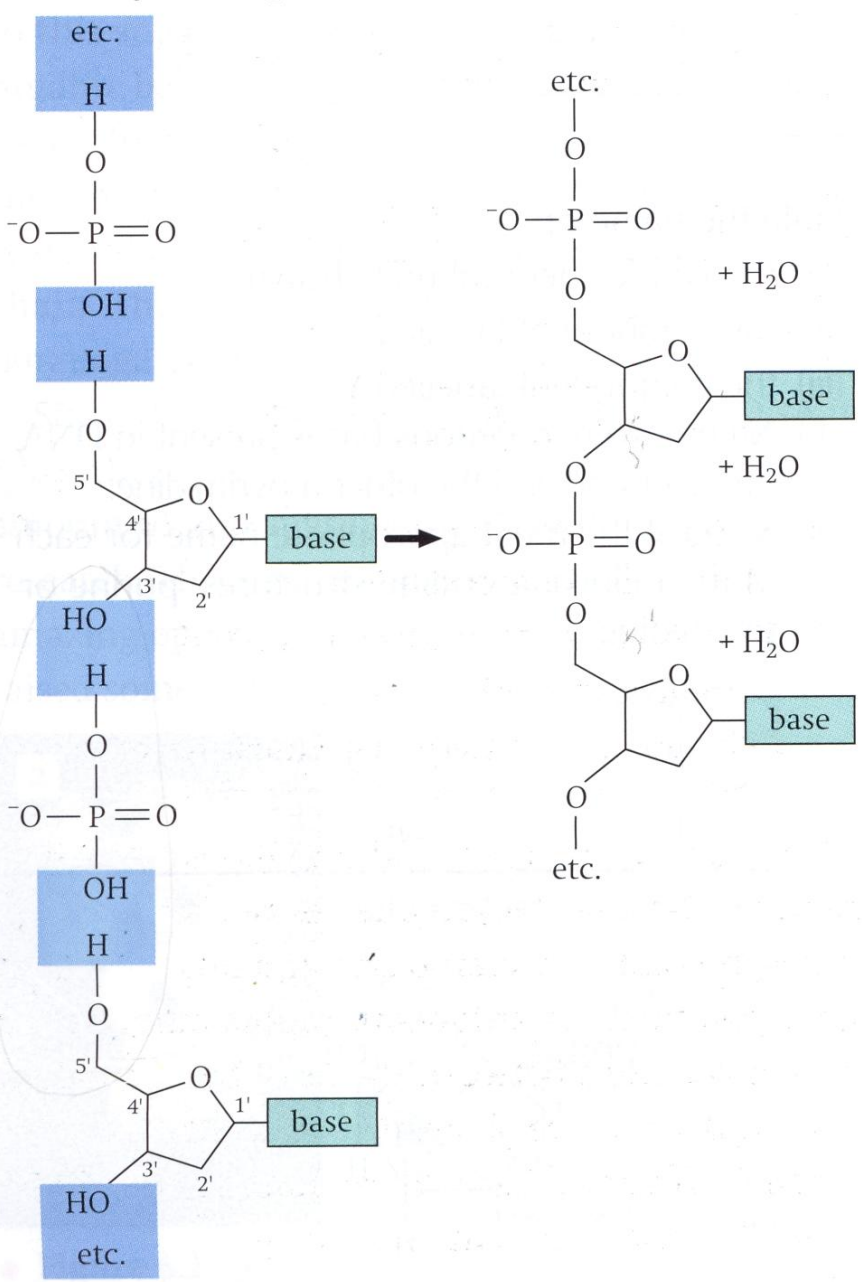
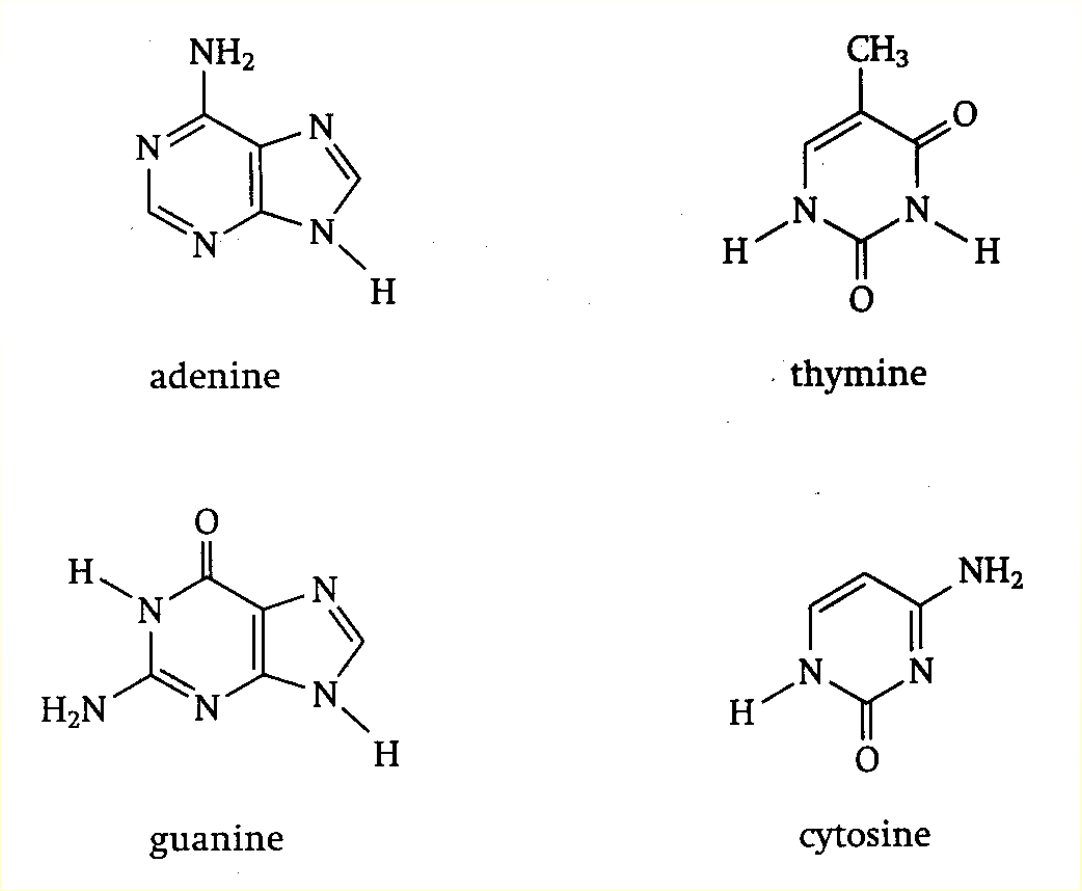


Figure 6.7 The sugar–phosphate backbone of DNA is formed by condensation reactions that produce a polyester.

Structure of DNA

- **adenine(A)** and **guanine(G),** have **planar two-ring** structure **structures** [they are **purines**].
- **thymine(T)** and **cytosine(C),** are **planar single-ring** structure molecules [they are **pyrimidines**].



DNA double helix

- DNA molecule consists of two strands.
- **Linked together by hydrogen-bonding between the bases**

Figure 1.34 – the DNA double helix

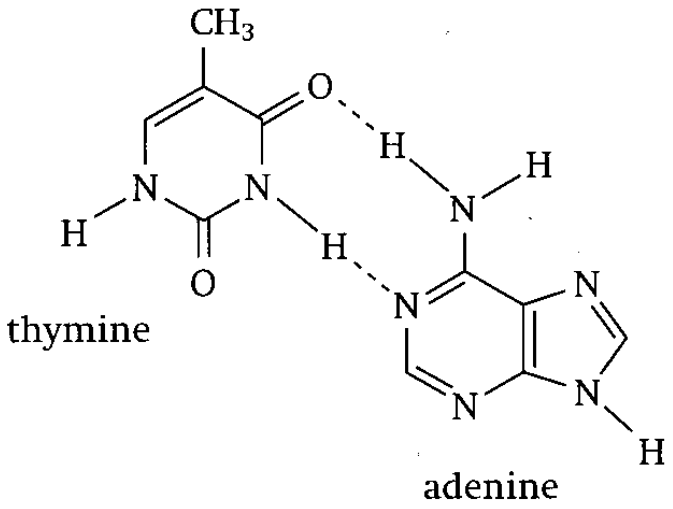
DNA double helix

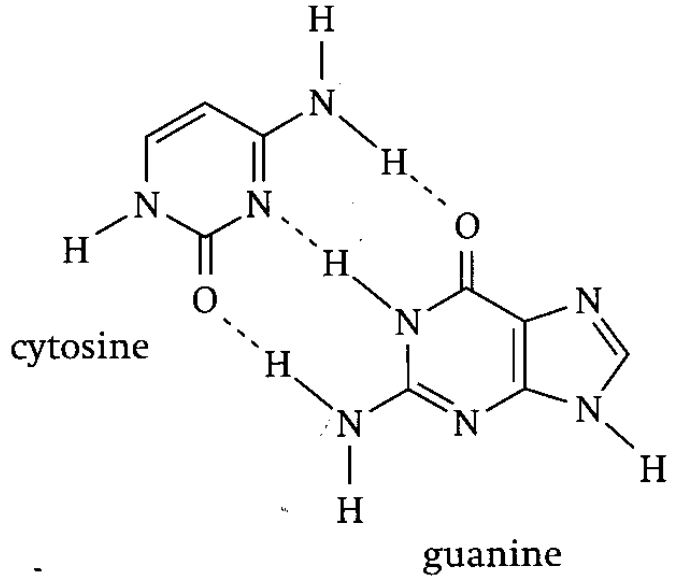
- These **two anti-parallel strands** are **twisted together in a double helix**.

- The **bases positioned between the two chains lie at right angles to the backbone**, filling the space between the strands.

DNA double helix

- **Two hydrogen bonds form** between each adenine-thymine pair **(A = T).**
- **Three hydrogen bonds are formed** between a guanine-cytosine pair **(G ≡ C).**
- This gives rise to the specificity of the pairing of the bases in DNA.
- The bases always pair up as follows:
  - **adenine is always paired with thymine**;
  - **guanine is always paired with cytosine**.





- **Figure 6.9** Complementary base pairing between the bases in DNA. Thymine pairs with adenine by two hydrogen bonds and cytosine pairs with guanine by three hydrogen bonds.

DNA double helix

- This is **known as complementary base pairing**.
- **Hydrogen bonding** and **van der Waals' forces between the stacks of bases are responsible for holding the chains together**.
- The precise sequence of the bases carries the genetic information.

DNA double helix

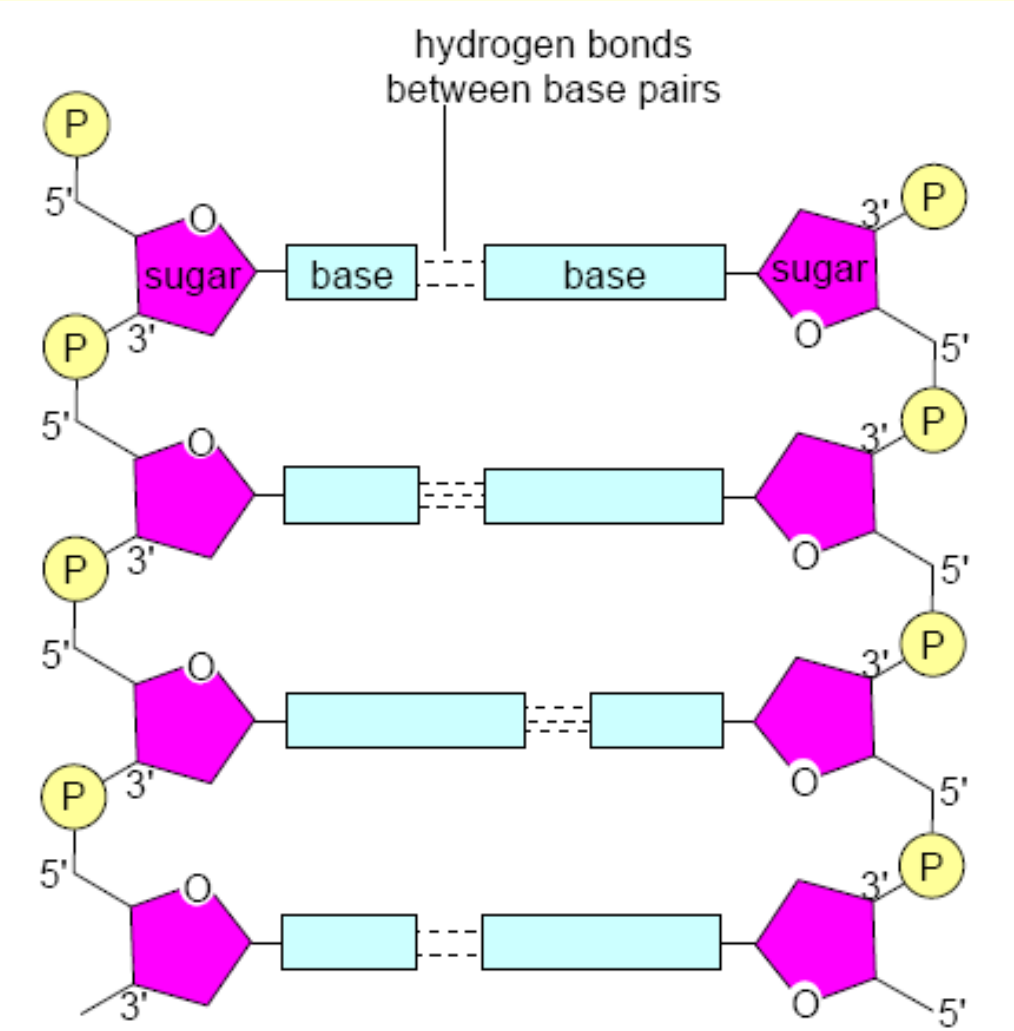


Figure 1.35 – the outline structure of DNA

Structure of RNA

- In order to express the genetic message Nature has devised a system of '**information transfer'** that **involves** various forms of **ribonucleic acid (RNA)** as well.

- The major different forms of RNA in gene expression are:

  – messenger RNA (**mRNA**);
  – ribosomal RNA (**rRNA**);
  – transfer RNA (**tRNA**).

Structure of RNA

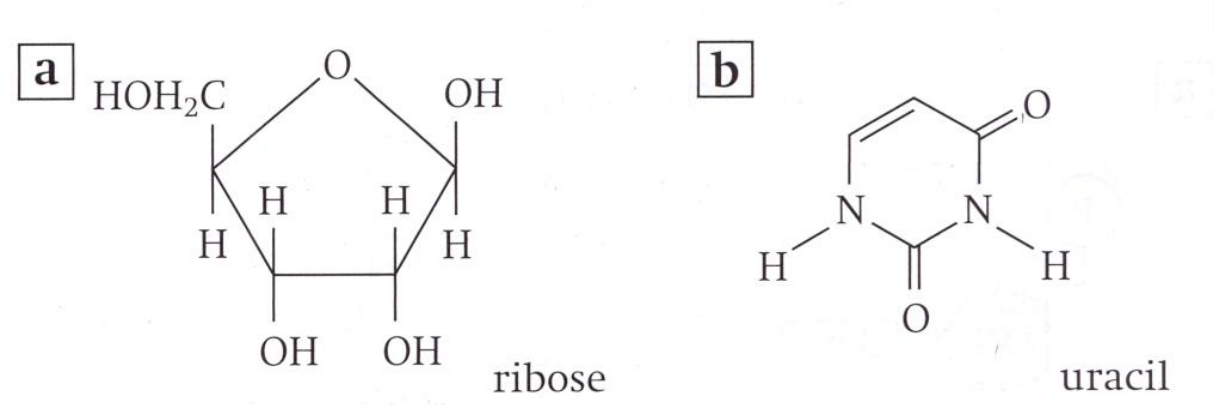
Table 1.3: Comparison of the structures of DNA and RNA

factor	deoxyribonucleic acid, DNA	ribonucleic acid, RNA
sugar	pentose sugar present is deoxyribose	pentose sugar present is ribose
bases	• adenine • cytosine • guanine • thymine	• adenine • cytosine • guanine • uracil
structure	a double helix made of two anti-parallel strands	single stranded, though the chain can fold on itself to form helical loops

• uracil (U), replaces thymine (T) in RNA

•uracil is a single-ring structure and can form a complementary base pair with adenine.

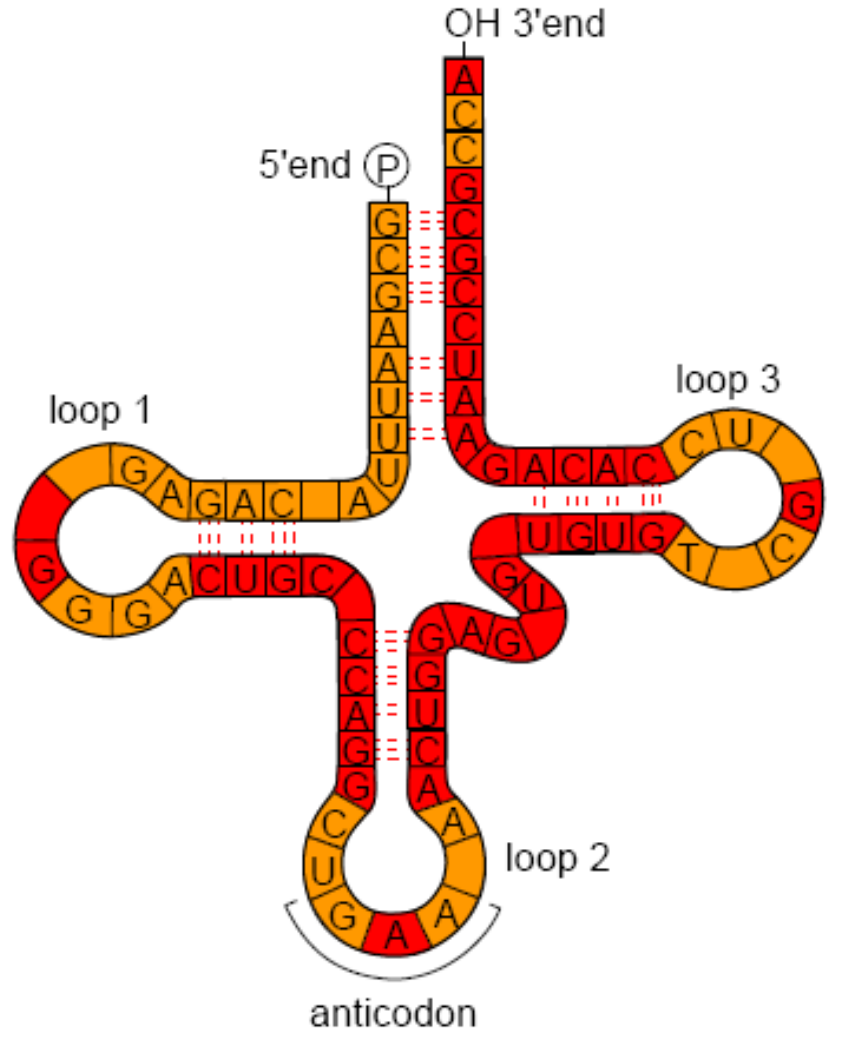
Structure of RNA



- Although an RNA molecule is a single long chain it **can bend back on itself to form hairpin loops**.

- These **loops are hydrogen-bonded** and are **important features of the structure of ribosomal RNA (rRNA) and transfer RNA (tRNA) molecules**.

Structure of t-RNA



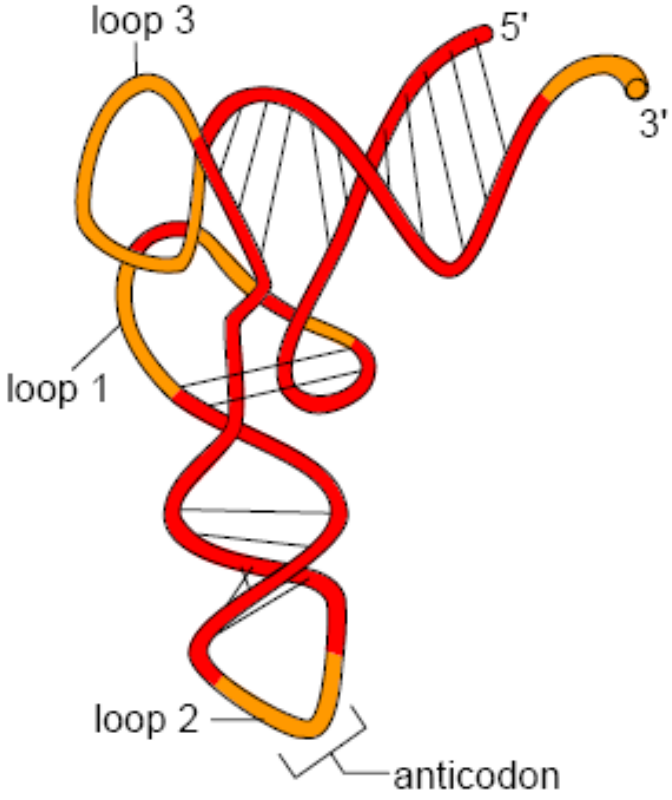


Figure 1.36 – the cloverleaf structure of tRNA molecules

Gene Expression

- Broadly speaking the amino acid sequence of each polypeptide chain is encoded in a specific stretch of DNA, or gene.

- The **code in a gene is used to make copies of a particular polypeptide chain through a two-stage process**.

Gene Expression

(1) **Transcription:** The DNA template is first ***copied***, or ***transcribed***, into an intermediary nucleic acid molecule, messenger ribonucleic acid (***mRNA***).

(2) **Translation:** mRNA molecules copied from the gene programme the assembly of the polypeptide chain. The translation process involves ribosomes attaching to, and moving along, the mRNA as the chain is synthesised.

- These two processes of transcription and translation is ultimately **responsible for the nature of all the proteins synthesised by cells**.

Gene Expression

- The double helix of DNA controls heredity on the molecular level. **DNA both preserves the genetic information (replication),** and **uses it to direct the synthesis of proteins (transcription and translation).**

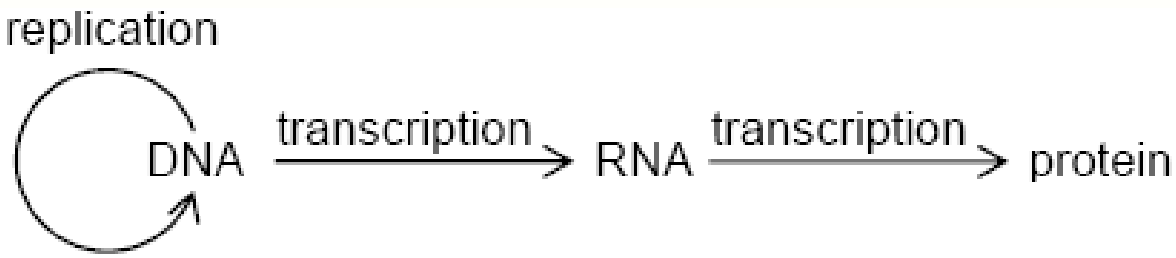


Figure 1.37 – summary diagram of replication, transcription

Replication

- The process of formation of new DNA strands is catalysed by the enzyme **DNA polymerase**.
- The new nucleotide units are fed into the reaction process in the form of nucleotide triphosphate.
- The breakdown of the triphosphates into monophosphate form is exothermic.
- The energy released by this breakdown drives the addition of the next nucleotide unit to the growing DNA copy.

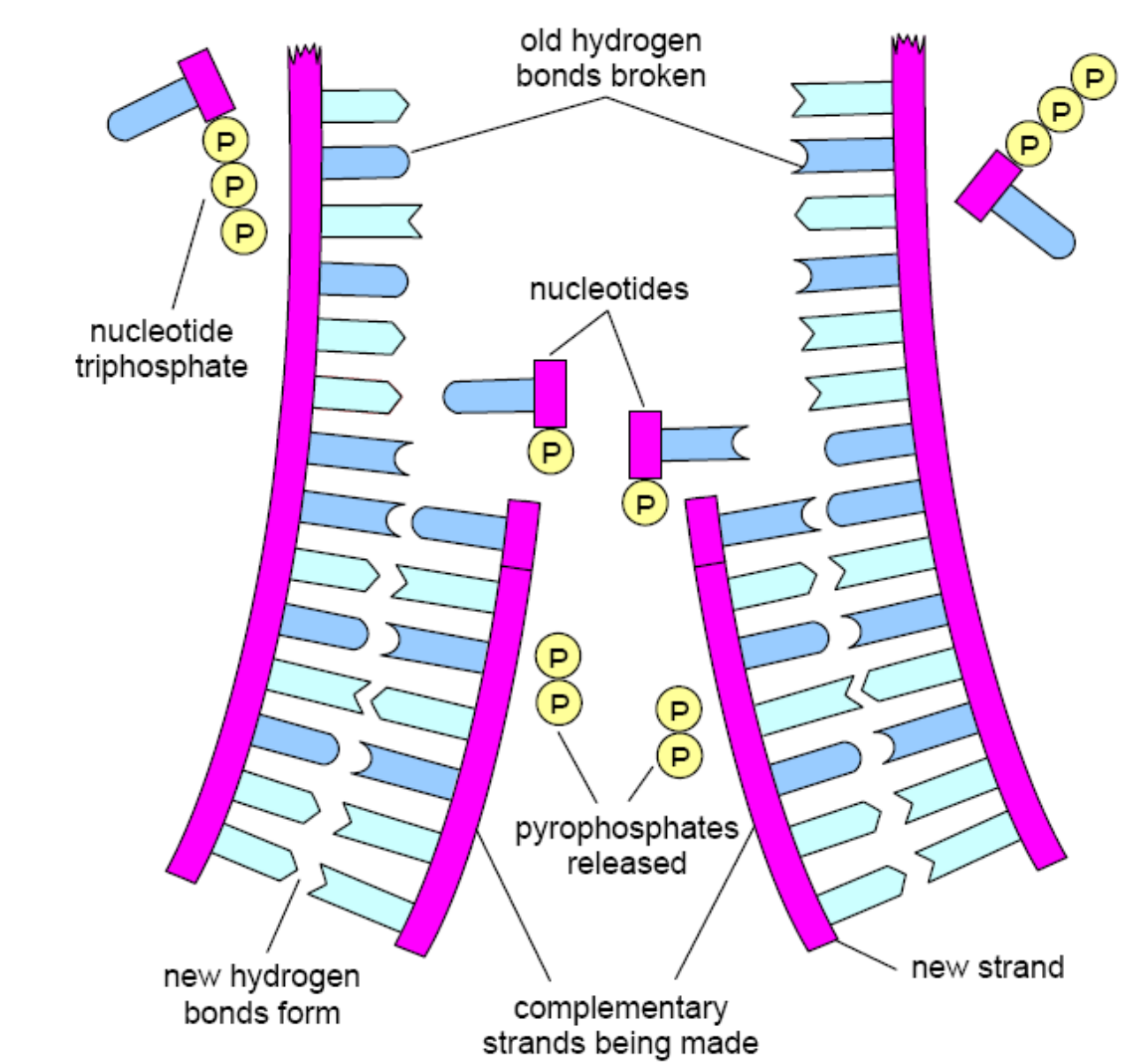


Figure 1.38 – strand of DNA acting as template for replication

Semi-conservative Replication

- During replication, the **hydrogen bonds** and van der Waals' forces between the base pairs in the double helix are **broken**.

- Two new strands are formed using the original strands as templates for the synthesis.

- Each new strand contains a complementary sequence of bases as dictated by the order of the bases in the original strand.

Semi-conservative Replication

- **Hydrogen bonds** and Van der Waals' interactions **then form** between the **original** and new **strands** creating a stable helical structure.

- Thus two daughter molecules are formed from the parent double helix.

- Each daughter molecule contains one new strand and one original strand.

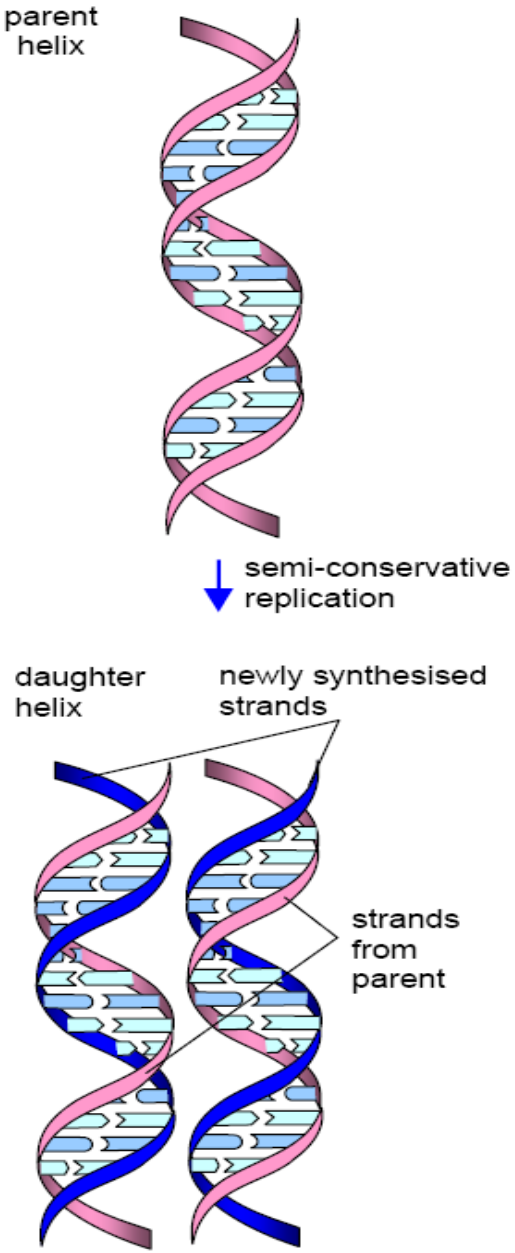


Figure 1.39 – semi-conservative replication

Semi-conservative Replication

- Gene sequences are always written in the 5'→ 3'direction, for example,

5'-ATGCCGTTAGACCGT _ _ _ _ _ _ _ GT-3'.

- The DNA in almost every cell in our bodies (the exceptions are certain white blood cells and sex cells) should be an identical copy of the DNA in the fertilised egg.

- Different genes have different sequences of these four nucleotides and so code for different polypeptide chains.

The Role of RNA

- Gene expression through transcription and translation processes involves several different ribonucleic acid (RNA) molecules.

- Each of the three main types of RNA has a different role within the complex mechanism of translation:

  - **Ribosomal RNA (rRNA)** – rRNA's form part of the structure of the ribosomes. Ribosomes are the **small organelles where protein synthesis takes place within the cell**.

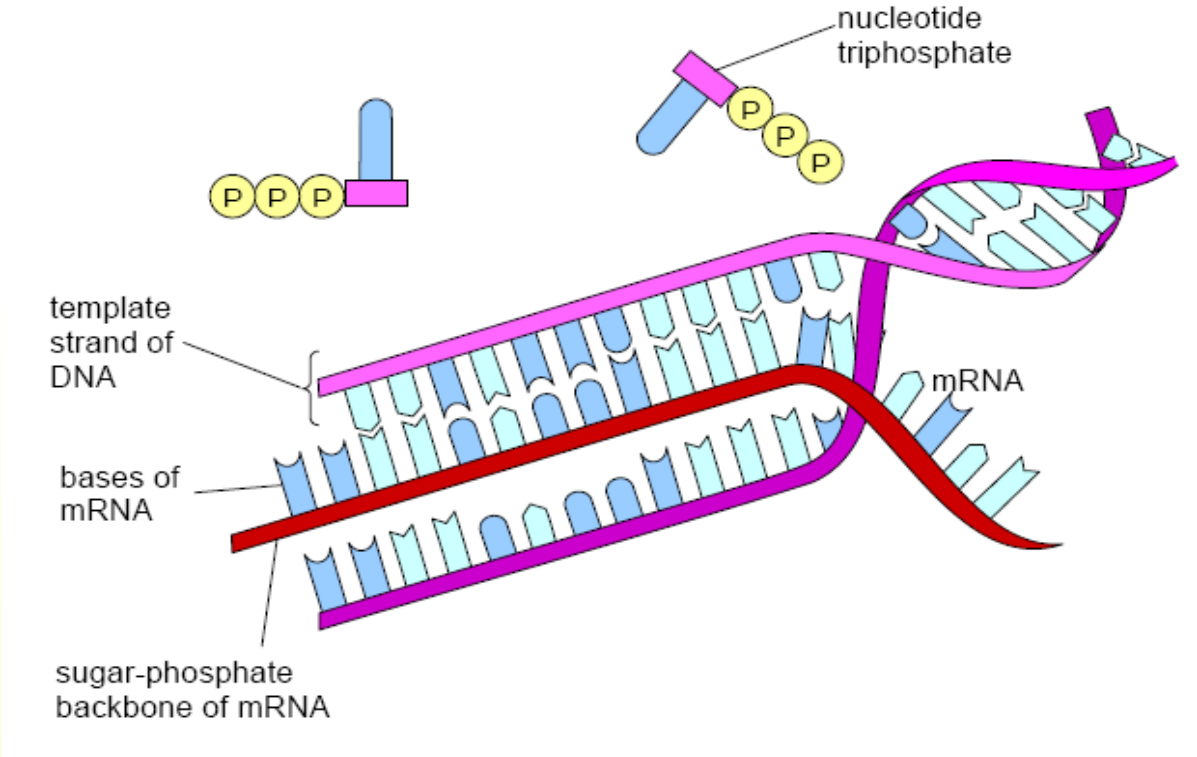
  - **Transfer RNA (tRNA**) – '**carry' the amino acids to the ribosomes for protein synthesis.** Each tRNA recognises the coding sequence for a particular amino acid in the messenger RNA.

  - **Messenger RNA (mRNA)** – the **RNA copied from the DNA gene** sequence for a particular polypeptide chain. The 'message' encoded in the mRNA molecule is translated into the **primary sequence of a polypeptide chain.**

Transcription

- First the code is transcribed into the mRNA for the protein chain by the enzyme RNA polymerase.

- Part of the DNA double helix unravels and an RNA copy of the gene is synthesised using the appropriate nucleotides.



Transcription

- The **mRNA molecule is synthesised from the 5'end to the 3'end**.

- This is also the direction in which the ‘message’ will subsequently be translated on the ribosomes.

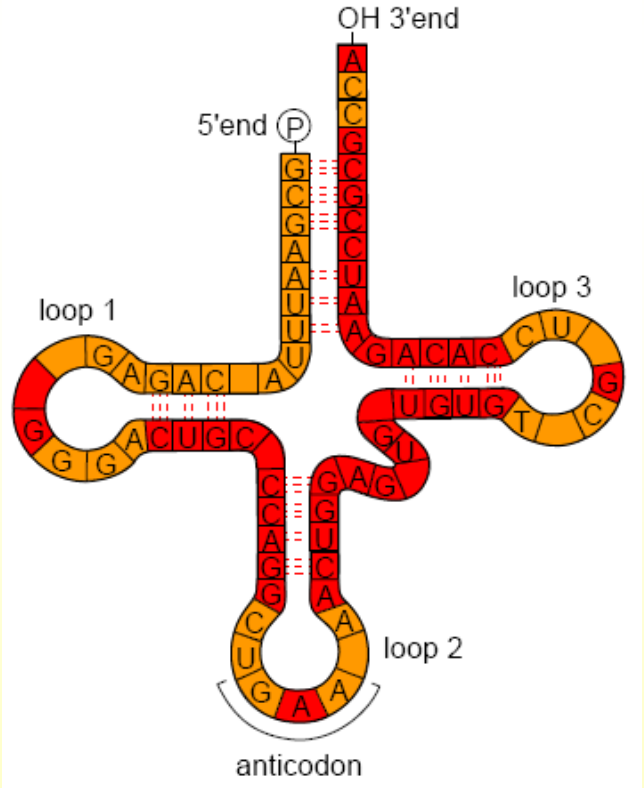
Translation

- Ribosomes are the cellular ‘machines’ that synthesise protein chains.
- During translation several ribosomes can attach to a particular mRNA molecule at any one time.
- As the ribosomes move along the mRNA the sequence of bases directs the bringing together of amino acids in the correct order to produce proteins.
- **Amino acids on their own cannot bind to mRNA**. **Transfer RNA molecules (tRNA’s) act as the vehicles** for these interactions.

Translation

- Each **tRNA binds a specific amino acid at one end** of the molecule.

- At the **other end** it has a specific triplet of bases (the **anticodon**) which can bind to the codon triplet on the mRNA.



Translation

- Each tRNA, carrying its specific amino acid, can interact with the ribosome and the correct codon on the messenger RNA to continue the process of translation.

- The translation process is a complex one involving three steps – **initiation, elongation, and termination**

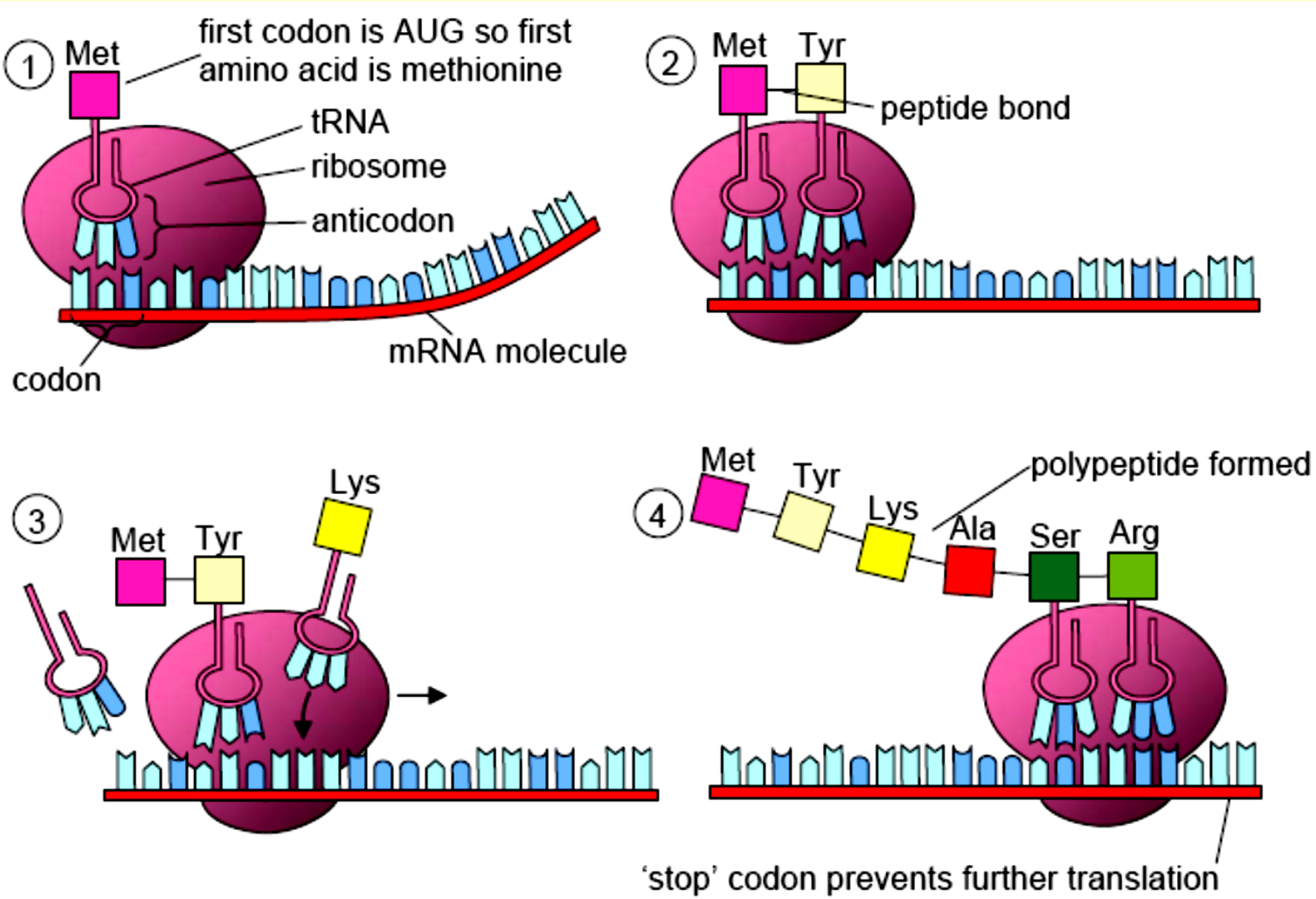


Figure 1.42 – the process of translation

The Genetic Code

- There are 20 amino acids used in making proteins.

- A **three-base (or triplet) code** (known as **codon**) would provide 64 possible combinations (4^3). This would allow coding for 64 different amino acids if all the codes were unique – each triplet coding for just one amino acid.

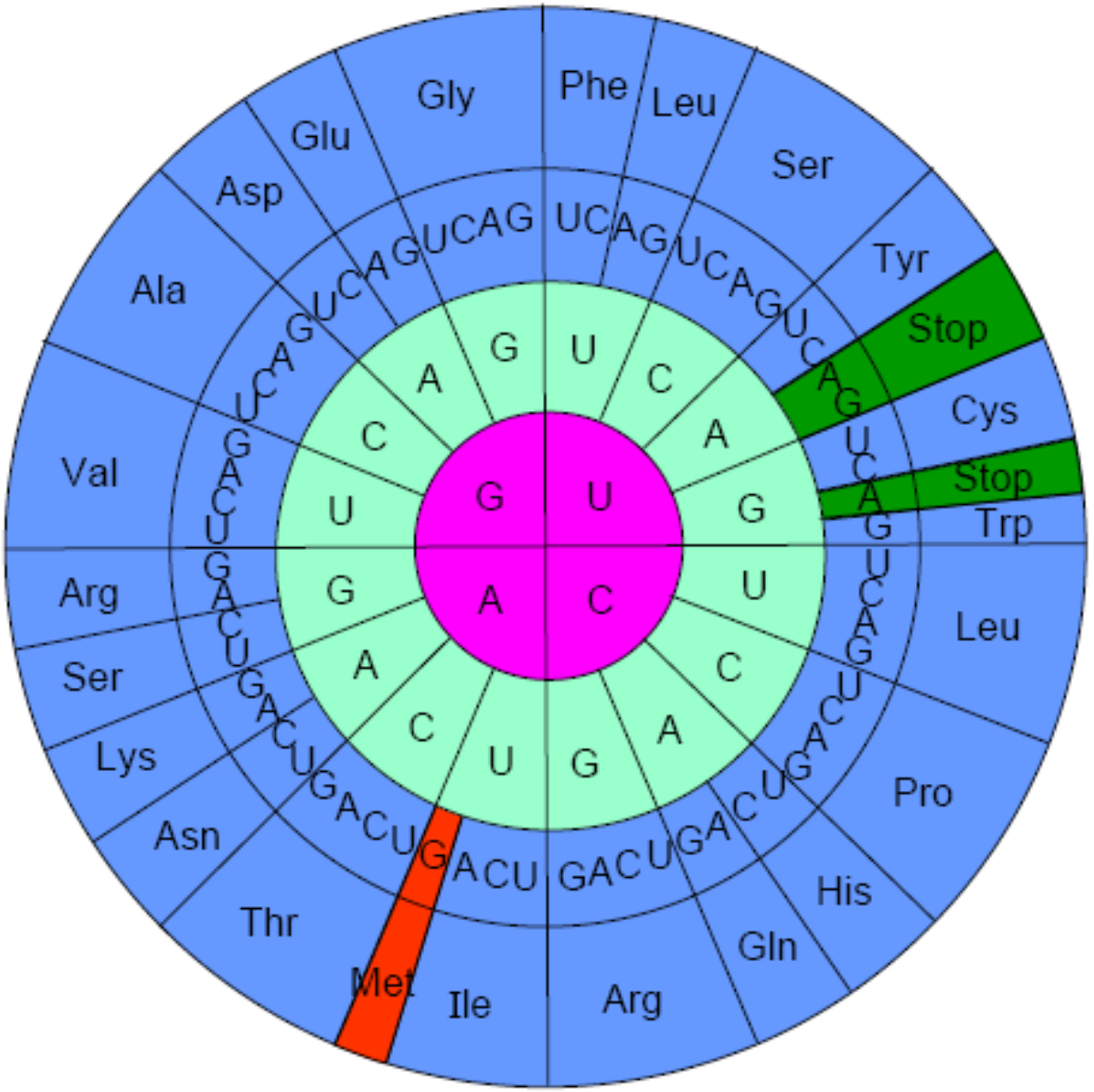


Figure 1.41 – the genetic code

The Genetic Code

- Most amino acids are coded for by more than one triplet codon.

- For all amino acids **except methionine and tryptophan**, **more than one base is allowed in the third position** of the combination.

- This arrangement offers some **protection from mutations** since a base change in the third position of a codon will often still mean that it still codes for the same amino acid.

The Genetic Code

- This ‘START’ signal is **5´-AUG-3´,** which codes for methionine.

- The **‘start’ signal ensures** that the **series of triplet codons is read in the correct groups of three**.

- There are also three codons that do not code for any amino acid - act as **‘STOP’** signals to end the assembly of a polypeptide chain.

Mutations

- Changes from the original DNA are known as mutations.
- Apart from errors in the replication process, mutations can also be caused by any process that damages DNA.
- UV light, cigarette smoke and many other chemical compounds can cause mutations.
- A single change in the base sequence has no effect on the protein that is being produced.
- This is because most amino acids have several codons, and a change in the DNA from a CAA sequence to a CAG, for instance, will still produce a protein containing valine in the correct place.

Mutations

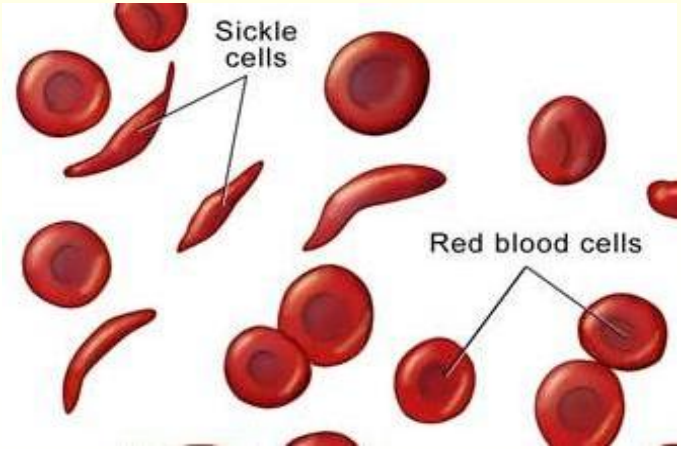
- Mutations which result in the **deletion of a base** would alter the way the message is read and produce a different sequence of amino acids in the protein chain.

- Such mutations together with those that remove a start or stop codon may have serious consequences.

- A crucial protein may not be produced or may be so changed that it is unable to function properly. Such a situation may result in a genetically based condition such as **sickle cell anaemia** or **cystic fibrosis**.

Sickle cell anemia

- The red blood cells of these patients do not have the normal disc shape, but have a **crescent moon** (or sickle) shape.



- People with sickle cell anemia have sickle haemoglobin (HbS) which is different from the normal haemoglobin (HbA).

- Arises from a single mutation in the DNA of the gene for one of the haemoglobin chains.

Sickle cell anemia

- Resulting an abnormal amino acid sequence in one of the protein chains in haemoglobin (the β-chain).

Normal β-chain	Val	His	Leu	Thr	Pro	**Glu**	Glu
Sickle cell β-chain	Val	His	Leu	Thr	Pro	**Val**	Glu

- When sickle haemoglobin gives up its oxygen to the tissues, it sticks together to form long rods inside the red blood cells, making these cells rigid and sickle-shaped.

- Because of their shape, sickle-shaped red blood cells cannot squeeze through small blood vessels as easily as the almost doughnut-shaped normal cells.

- This can lead to these small blood vessels getting blocked, stopping oxygen from getting through to where it is needed.

Cystic Fibrosis

- The condition affects the **lungs, pancreas, gut and sweat glands**. Instead of the normal fluid secretions a thick sticky mucous forms. This viscous mucous blocks and damages the intestines and lungs.

- Because the supply of digestive enzymes from the pancreas is blocked, nutrients cannot be absorbed and babies fail to thrive.

- These babies have repeated chest infections and in particular can get intestinal obstruction. The malfunctioning of the sweat glands results in abnormally salty sweat,

Cystic Fibrosis

- No cure, but the faulty gene has been identified.
- Affects the cells that line the cavities and tubes inside organs such as the lungs.
- Membranes of these cells have a mechanism for pumping chloride ions into the cells from the blood supply.
- In lungs the chloride ions normally diffuse out of the cells through channels in the cell membrane lining the airways.
- This is part of the process for keeping a runny layer of watery mucus on the surface of the cells.

Cystic Fibrosis

- Chloride ions diffuse out of the cell through a channel created by a protein.
- The protein channel is only open in the presence of ATP.
- The name of this membrane protein is CFTR protein (short for cystic fibrosis transmembrane regulatory protein).
- In a person with cystic fibrosis the CFTR protein may be missing or, if present, it does not work properly.
- It does not allow chloride ions which are being pumped into the cell to leave (Figure 1.43). The chloride ion concentration in the cell builds up.

Cystic Fibrosis

- The high solute concentration in the cell causes water to move into the cell instead of out of it by osmosis. As a result the mucus covering the cells lining the airways becomes thick and sticky.

- In some cases ATP is unable to bind to it so the channel cannot open.

- In other cases the channel opens but in a way that does not let the chloride ions escape.

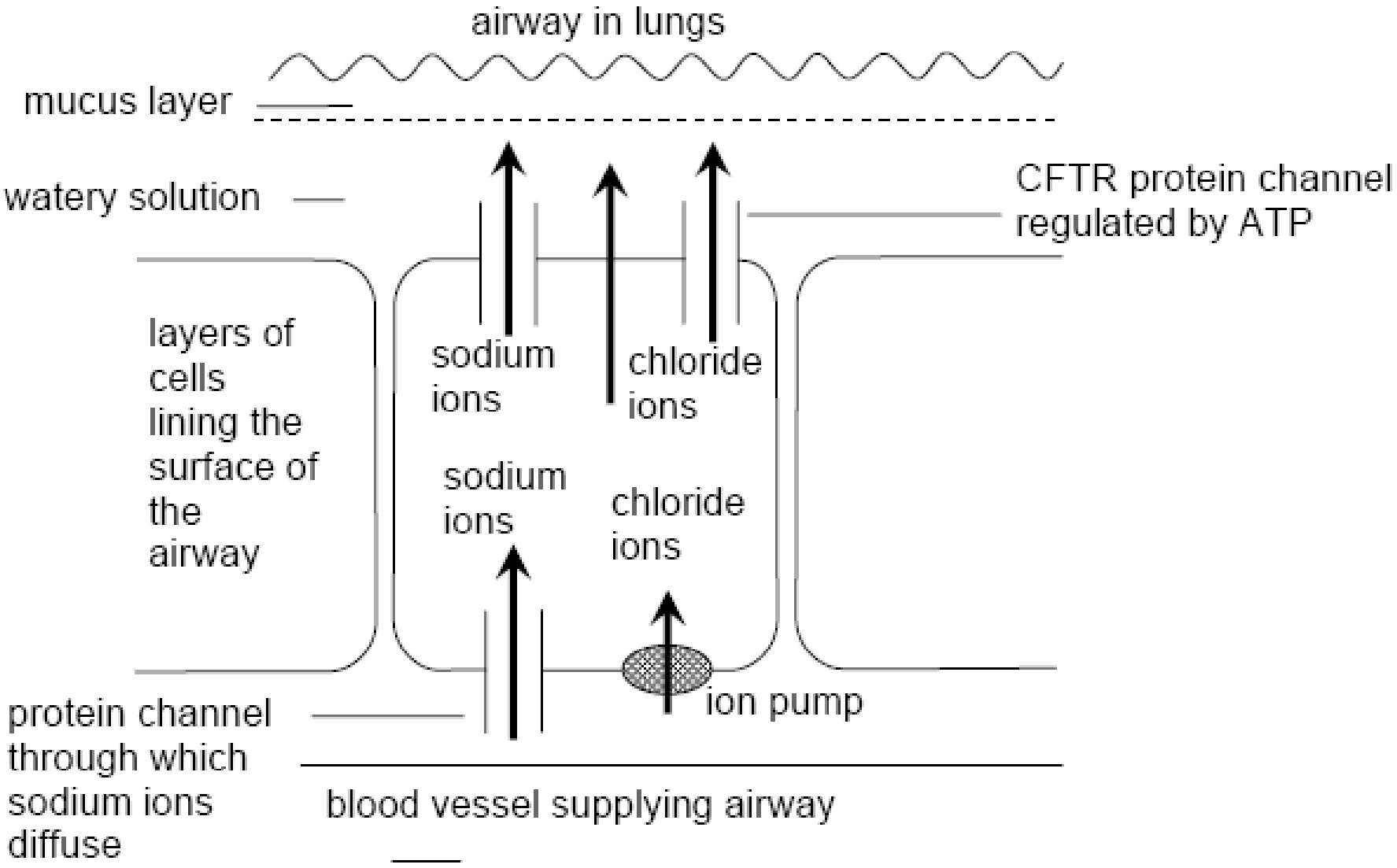


Figure 1.43 – the movement of ions across cell membranes in the lungs

Cystic Fibrosis

- The commonest mutation is the deletion of three nucleotides which result in the loss of phenylalanine, the 508th amino acid in the structure of the protein.