

Nucleic acids and protein synthesis

A-Level Biology

Nucleotides —the building blocks

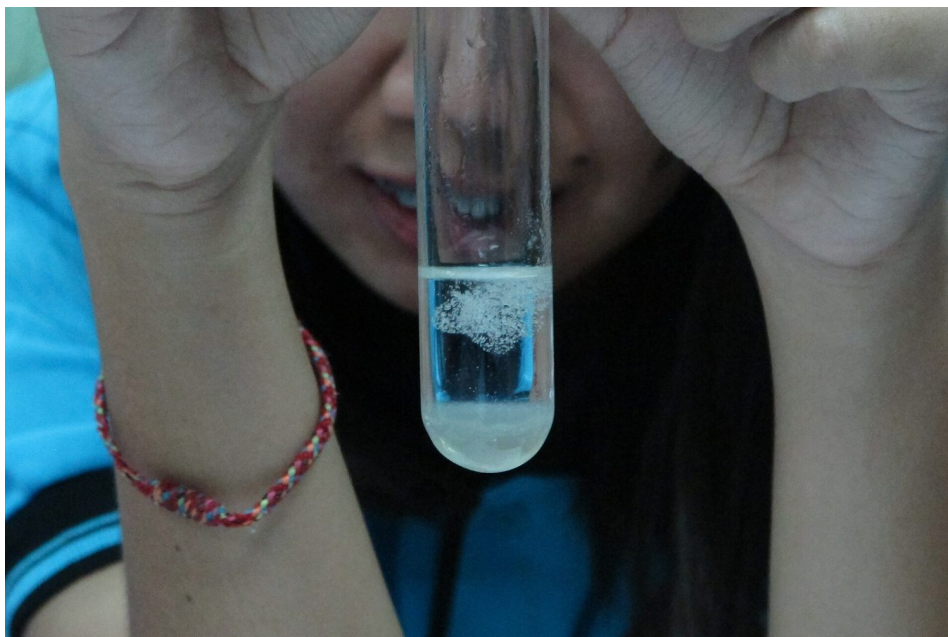
Nucleic acids 核酸 (DNA and RNA) are polymers of small units called **nucleotides** 核苷酸. Each nucleotide is made of three parts joined together:

- a **phosphate** 磷酸 group,
- a sugar (a 5-carbon sugar),
- a nitrogen-containing **base** 碱基.

ATP is a special nucleotide. It has the base **adenine** 腺嘌呤, the sugar **ribose** 核糖, and **three** phosphate groups. Breaking off the last phosphate releases **energy** 能量 for the cell.



A nucleotide is a phosphate, a sugar and a base; ATP is a nucleotide with three phosphates

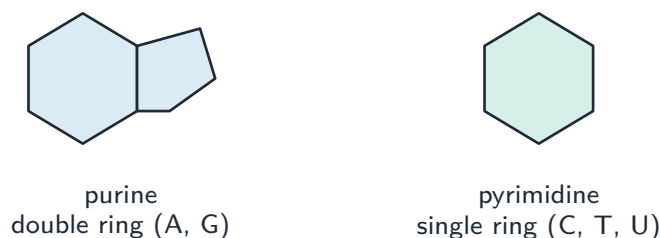


DNA extracted from cells appears as pale, stringy strands of these nucleotide polymers

Image: Joo Nath, CC BY-SA 4.0 (commons.wikimedia.org)

There are five bases, in two groups:

- **purines** 嘌呤 have a **double ring** (two rings): adenine and **guanine** 鸟嘌呤.
- **pyrimidines** 嘧啶 have a **single ring** (one ring): **cytosine** 胞嘧啶, **thymine** 胸腺嘧啶 and **uracil** 尿嘧啶.



The two groups differ in shape: a purine is a double ring, a pyrimidine a single ring — which is why a purine always pairs with a pyrimidine

The structure of DNA

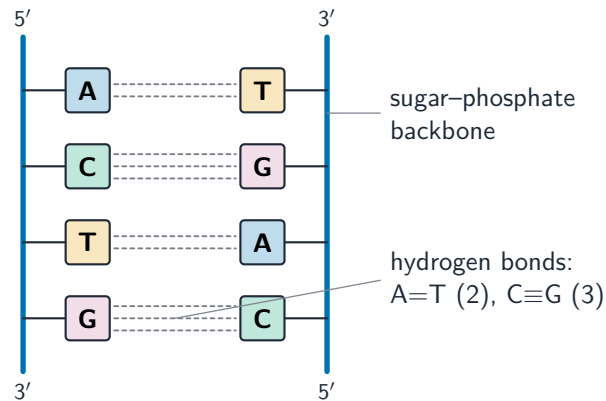
A DNA molecule is two strands twisted together into a **double helix** 双螺旋.

Each **strand** 链 has a backbone of alternating sugar (here the sugar is **deoxyribose** 脱氧核糖) and phosphate. The sugar of one nucleotide is joined to the phosphate of the next by a **phosphodiester bond** 磷酸二酯键.

The two strands are held together by their bases, which meet in the middle. The pairing is exact — this is **complementary base pairing** 碱基互补配对:

- A always pairs with T, held by **two hydrogen bonds** 氢键.
- C always pairs with G, held by **three hydrogen bonds** (so a C–G **base pair** 碱基对 is harder to separate).

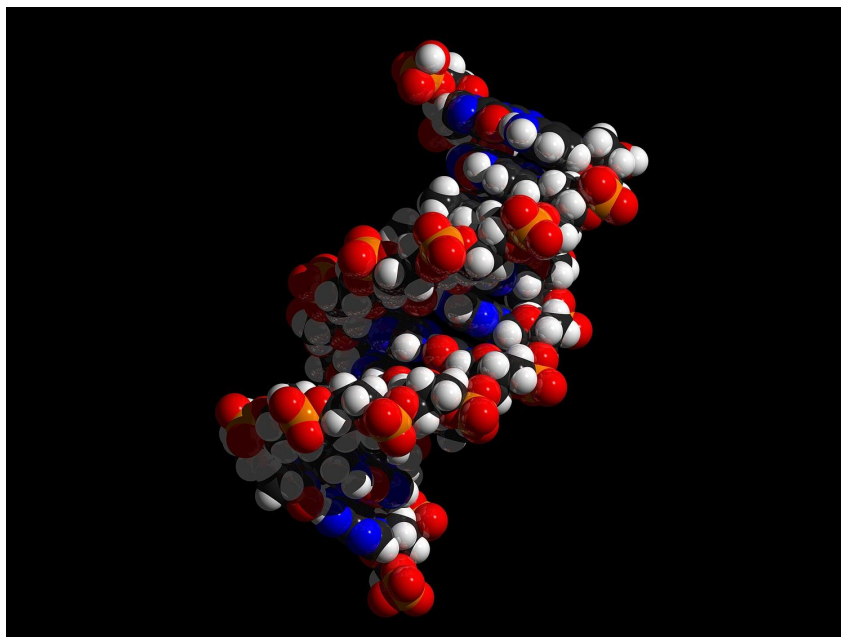
The two strands run in opposite directions: one goes 5' to 3' while the other goes 3' to 5'. We say they are **antiparallel** 反平行.



the two strands run in opposite directions (antiparallel)

Complementary base pairing: A pairs with T (2 hydrogen bonds), C with G (3); the strands are antiparallel

The flat ladder above is twisted into a spiral. This space-filling model, where every atom is a ball, shows the real shape of the **double helix**:



*A space-filling model of DNA: the two strands twist around each other into the **double helix** —the ladder of the diagram, coiled up*

Image: Ude, Public domain (commons.wikimedia.org)

DNA replication

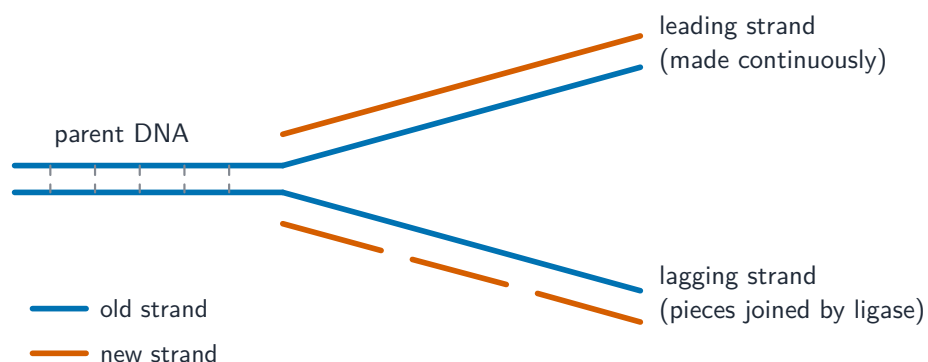
DNA **replication** 复制 (copying) happens during the S phase of the cell cycle. It is **semi-conservative** 半保留复制: each new molecule keeps one old strand and one new strand. The steps are:

1. the double helix unwinds and the hydrogen bonds break, so the two strands separate.

2. each old strand acts as a template. Free nucleotides pair with the exposed bases by complementary base pairing.
3. the **enzyme 酶 DNA polymerase 聚合酶** joins the new nucleotides into a strand. It can only add nucleotides in the 5' to 3' direction.

Because of that 5' to 3' rule, the two new strands are made differently:

- the **leading strand 前导链** is built continuously, following the unwinding.
- the **lagging strand 后随链** is built in short pieces, working away from the unwinding point. The enzyme **DNA ligase 连接酶** then joins these pieces together.



each new molecule = one old strand + one new strand (semi-conservative)

Replication is semi-conservative: each new molecule keeps one old strand (blue) and one new strand (orange)

RNA

RNA is also made of nucleotides, but it is a single strand, its sugar is ribose, and it uses uracil in place of thymine. The most important type here is **messenger RNA (mRNA)**, which carries a copy of a gene's instructions out of the nucleus to be used.

The genetic code

A **gene 基因** is a sequence of DNA nucleotides that codes for one **polypeptide 多肽**.

The code is read in **triplets 三联体**—groups of three bases. Each triplet either codes for one specific **amino acid 氨基酸**, or acts as a start or stop signal. The code is **universal**: nearly all living things use the same triplets for the same amino acids.

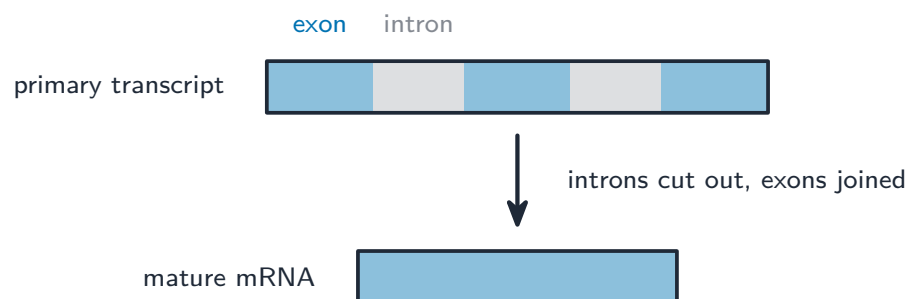
There are 64 possible triplets but only 20 common amino acids, so most amino acids are coded by more than one triplet.

Protein synthesis: transcription and translation

Transcription (in the nucleus)

The DNA gene is copied into mRNA. This is **transcription** 转录.

- the strand of DNA that is copied is the **template strand** 模板链; the partner strand is the **non-transcribed strand** 非转录链.
- the enzyme **RNA polymerase** joins RNA nucleotides that pair with the template bases (with uracil pairing to adenine).
- in eukaryotes the first RNA made (the **primary transcript** 初级转录本) contains coding parts called **exons** 外显子 and non-coding parts called **introns** 内含子. The introns are cut out and the exons joined together to form the finished mRNA.

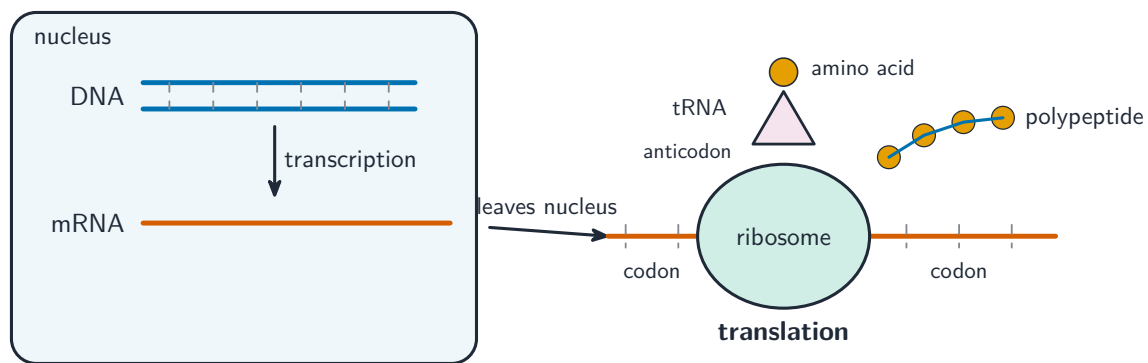


Splicing: the non-coding introns are cut out and the coding exons joined to make the finished mRNA

Translation (at the ribosome)

The mRNA leaves the nucleus and attaches to a **ribosome** 核糖体. Building the polypeptide from the mRNA code is **translation** 翻译.

- the mRNA is read in **codons** 密码子 (each codon is one triplet of mRNA bases).
- molecules of **transfer RNA (tRNA)** bring amino acids to the ribosome. Each tRNA has an **anticodon** 反密码子 that pairs with a matching codon.
- as the codons are read in order, the ribosome joins the amino acids with **peptide bonds** 肽键, building the polypeptide.



Transcription copies DNA into mRNA in the nucleus; translation at the ribosome builds the polypeptide

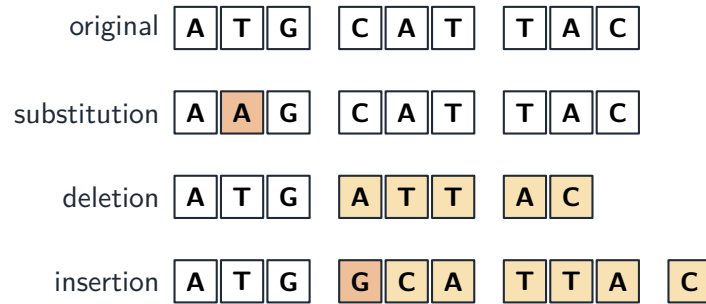
Worked example. A DNA template strand reads TAC GGA CTT. Give the mRNA codons, and say how long the polypeptide is. Transcribe by **complementary base pairing**, remembering that RNA uses **uracil** in place of thymine: TAC gives AUG, GGA gives CCU, CTT gives GAA. So the mRNA reads AUG CCU GAA. AUG is the **start** codon (methionine), so the polypeptide is **three** amino acids long - two once the start methionine is removed. Two errors cost most of the marks here: pairing A with **T** instead of **U** when writing mRNA, and transcribing the **coding** strand instead of the **template** strand. The mRNA is complementary to the template, and identical to the coding strand apart from U replacing T.

Gene mutations

A gene **mutation** 突变 is a change in the base sequence of a DNA molecule. It may change the polypeptide made. There are three types:

- **substitution** 替换—one base is swapped for a different base. This changes at most one amino acid, and sometimes none (because most amino acids have more than one triplet).
- **deletion** 缺失—a base is removed.
- **insertion** 插入—an extra base is added.

A deletion or insertion shifts how every later triplet is read, so it usually changes many amino acids after that point and has a large effect on the polypeptide.



red = changed base; shaded = shifted by a frameshift (deletion or insertion)

A substitution changes one triplet; a deletion or insertion shifts every later triplet (a frameshift)

Exam tips

- DNA is **antiparallel**, held by hydrogen bonds (A=T two, C G three); replication is **semi-conservative** (each product keeps one old strand).
- State **where** and **what**: transcription (nucleus → mRNA), translation (ribosome, mRNA + tRNA → polypeptide).
- Describe the code with all four terms: **triplet, non-overlapping, degenerate, universal**.
- For a **gene mutation** (substitution, insertion, deletion) explain the effect on the protein —and why a substitution can be silent (degeneracy).