Exam Papers Practice

Boost your performance and confidence with these topic-based exam questions

Practice questions created by actual examiners and assessment experts

Detailed mark schemes

Suitable for all boards

Designed to test your ability and thoroughly prepare you

Nucleic Acids

IB Biology - Revision Notes

DNA & RNA Structure

Genetic Material of Life

DNA as the genetic material of living organisms

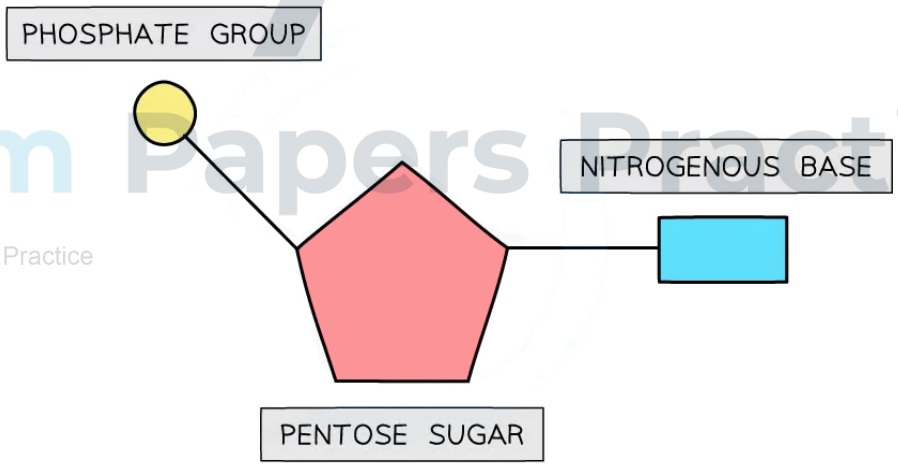
- **Deoxyribose nucleic acid** (DNA) carries the **genetic code** in all living organisms
- This is the reason why the genetic code is said to be **universal**, it applies to all forms of life
- DNA is **mainly found in the nucleus** where it forms chromosomes
  - It is also found in **chloroplasts** and **mitochondria** of eukaryotic cells
- **Ribonucleic acid** (RNA) is another type of nucleic acid which is the **main component of ribosomes**, which play an important role in protein synthesis
  - Some RNA is also found in the **nucleus** and **cytoplasm**
- Certain viruses (such as SARS-CoV-2) contain **RNA** as their genetic material instead of DNA
- These viruses cause a variety of different diseases, such as COVID-19, Ebola, mumps and influenza
- Viruses are **not considered to be living organisms**, since they are unable to replicate by themselves
- They are **dependent on other living cells** for replication and survival
- Viruses also **lack a cellular structure**, which is another reason they are not considered to be living

Nucleotide Components

Components of a nucleotide

- Both DNA and RNA are **polymers** that are made up of **many repeating units** called **nucleotides**
- Each nucleotide is formed from:
  - A **pentose sugar** (a sugar with 5 carbon atoms)
  - A nitrogen-containing **organic base** (with either 1 or 2 rings of atoms)
  - A **phosphate group** (this is acidic and negatively charged)
- The base and phosphate group are both **covalently bonded** to the sugar
- The **nitrogenous bases** in DNA are:
  - Adenine (A)
  - Guanine (G)
  - Cytosine (C)
  - Thymine (T)
- RNA share the same nitrogenous bases as DNA except thymine, which is replaced by **uracil (U) in RNA**
- The nitrogenous bases can be grouped as either **purine** or **pyrimidine** bases:
  - **Adenine** and **guanine** are purine bases
  - **Cytosine**, **thymine** (in DNA) and **uracil** (in RNA) are pyrimidine bases

Nucleotide structure diagram



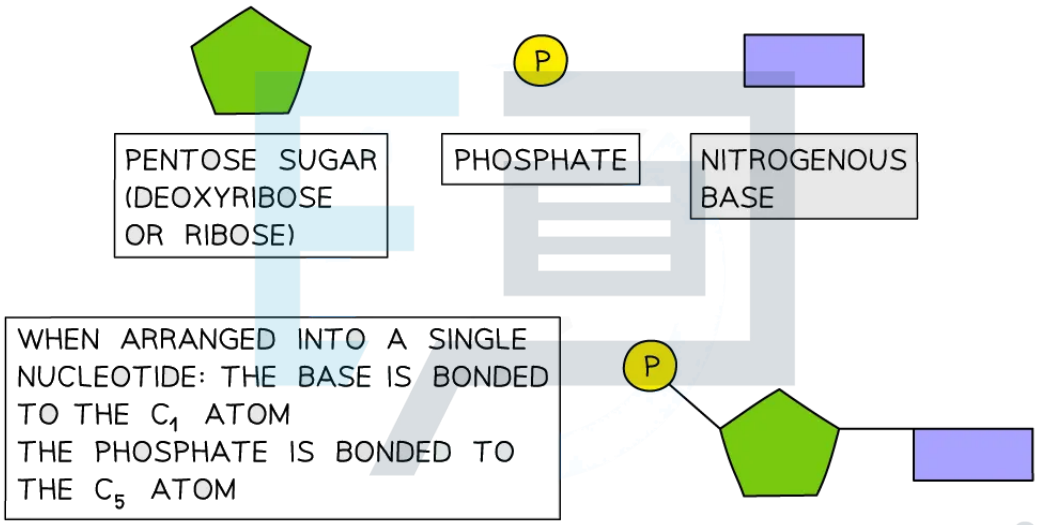
The basic structure of a nucleotide

Drawing simple diagrams of the structure of single nucleotides of DNA and RNA

- **Simple shapes** can be used to draw the **main building blocks of nucleotides** and the DNA double helix

- Advanced drawing skills are not required!
- **Pentagons** can represent **pentose sugars**
- **Circles** can represent **phosphates**
  - Often shown as a circle with the letter P inside: Ⓟ
- **Rectangles** can represent **bases**
- **Covalent bonds** can be shown with **solid lines**
- **Hydrogen bonds** can be shown with **dashed lines**
  - Or with complementary shapes that fit together (see diagrams)

Components of a nucleotide diagram



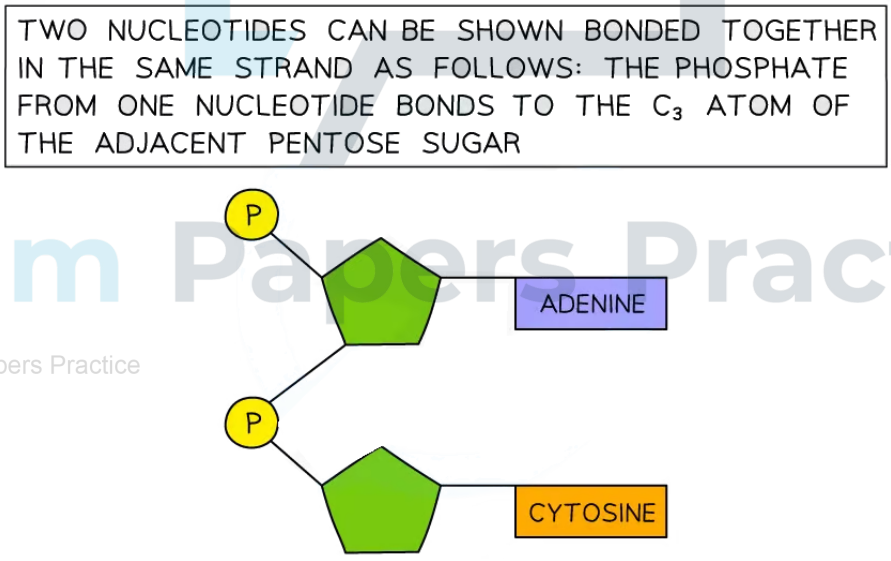
Simple shapes can be used to represent parts of nucleotide molecules

Linking Nucleotides

Forming the sugar-phosphate backbone

- Nucleotides join together in chains to form DNA or RNA strands
- The **phosphate group of one nucleotide** forms a covalent bond to the **pentose sugar of the next one**
  - This carries on to form a **large polymer**
  - These polymers form **nucleic acids**, which are also known as **polynucleotides**
- The phosphate group of one nucleotide is linked to the pentose sugar of the next one by **condensation reactions**
  - This means a **molecule of water is released** during the formation of each covalent bond
- This forms a '**sugar-phosphate backbone**' with a base linked to each sugar
- The polymer of nucleotides is known as a **strand**
- DNA is double-stranded, RNA is usually single-stranded
- There are just **4 separate bases** that can be joined in **any combination/sequence**
  - Because the **sugar and phosphate are the same** in every nucleotide

Linking nucleotides together diagram



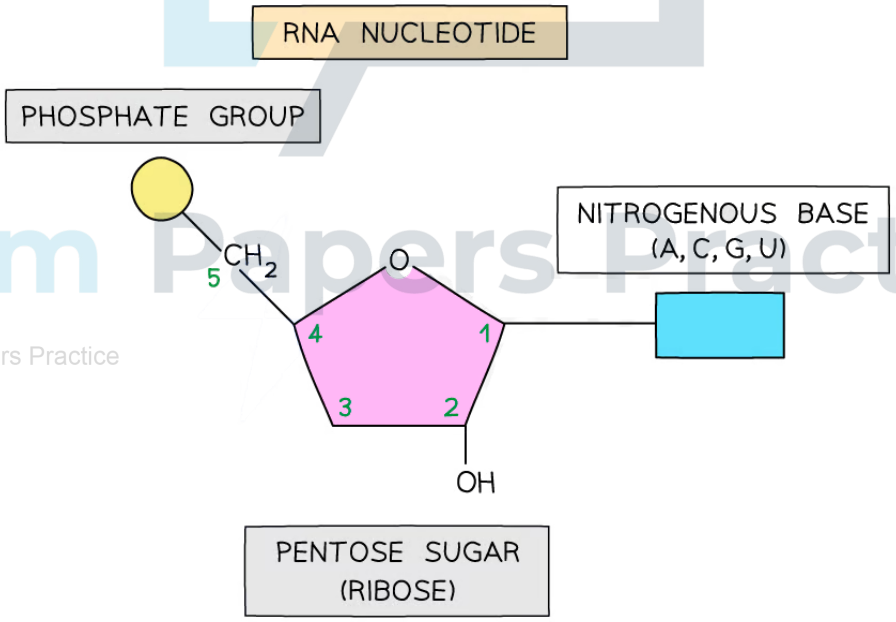
Two nucleotides shown bonded together covalently within a strand

RNA Structure

RNA structure

- Unlike DNA, RNA molecules are **relatively short** with lengths of between a hundred to a few thousand nucleotides
- It usually forms a **single-stranded polynucleotide** with **ribose** as the pentose sugar in each nucleotide
- RNA nucleotides contain the following **nitrogenous bases**:
  - Adenine
  - Guanine
  - Cytosine
  - Uracil (instead of thymine in DNA)
- The carbon atoms in nucleotides are numbered **from the right in a clockwise direction**
  - This makes it easier to **identify the bonds** in the sugar-phosphate backbone of polynucleotides
  - It also indicates the **orientation of the polynucleotide**

RNA nucleotide diagram



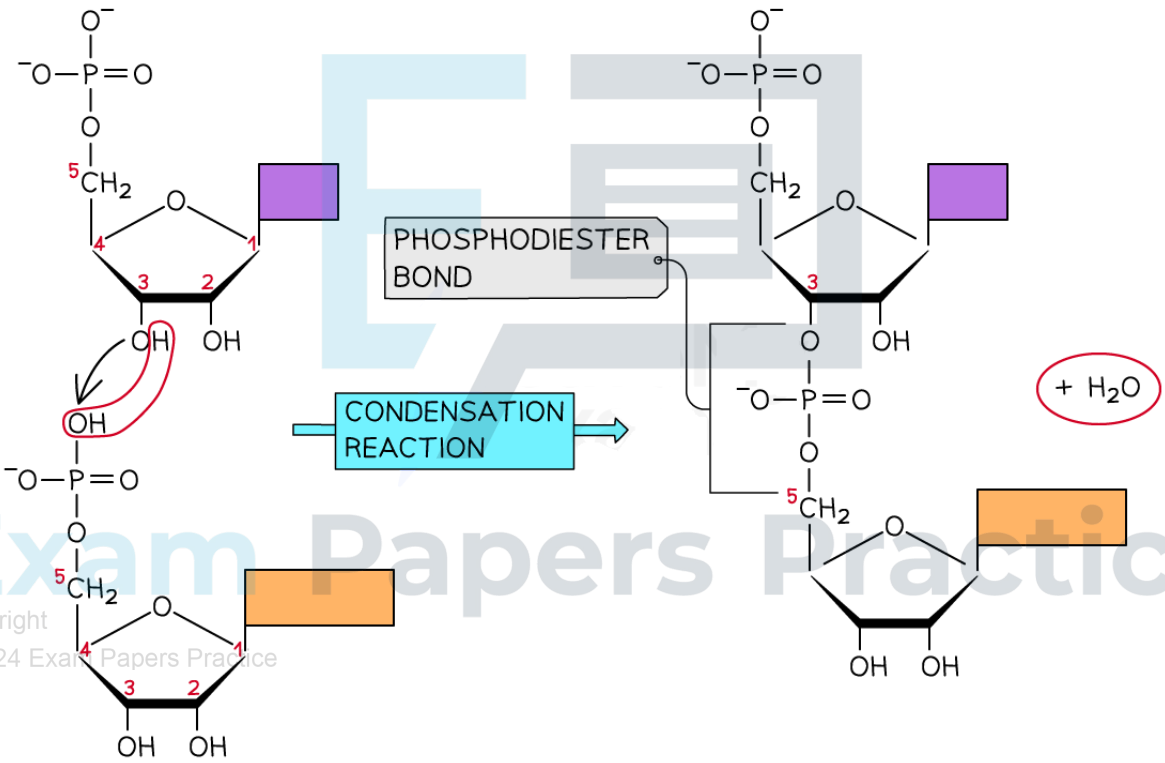
The structure of an RNA nucleotide

- Different types of RNA are found in the cells of living organisms:

- **messenger RNA** (mRNA), which is formed in the nucleus and transported to the ribosomes in the cytoplasm
- **transfer RNA** (tRNA), which is responsible for transporting amino acids to ribosomes during protein synthesis
- **ribosomal RNA** (rRNA), which forms part of ribosomes

- Adjacent RNA nucleotides are linked together by **condensation reactions**, during which a molecule of water is released
- This forms a **phosphodiester bond** between the pentose sugar of one nucleotide and the phosphate group of the next nucleotide

The formation of an RNA polymer diagram



Linking RNA nucleotides together by condensation reactions will result in the formation of phosphodiester bonds

Exam Tip

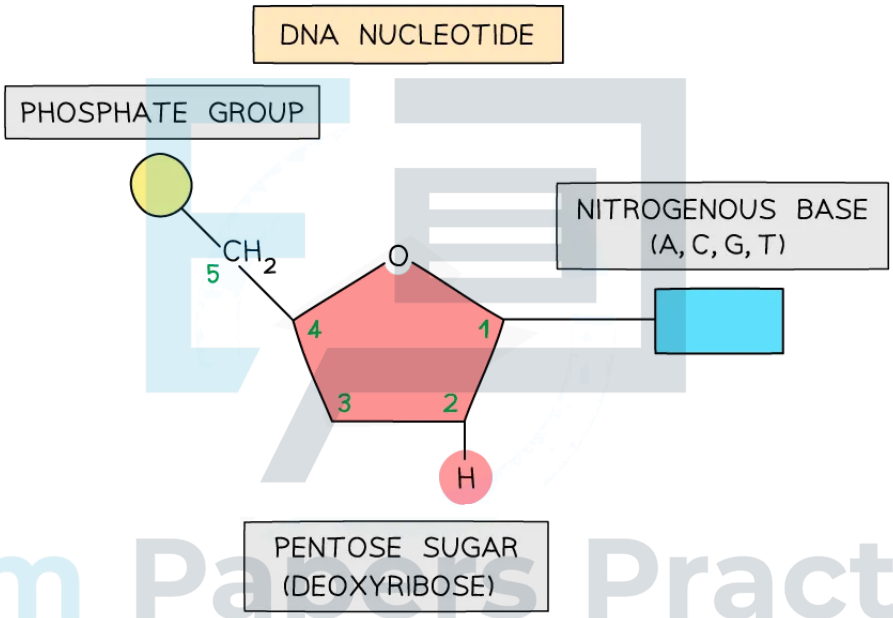
Ensure that you are able to draw and recognise diagrams of a single RNA nucleotide, as well as RNA polymers

DNA Structure

DNA structure

- DNA is a **double helix** made of two antiparallel strands of nucleotides linked by hydrogen bonding between complementary base pairs
- The nucleic acid DNA is a **polynucleotide** – it is made up of **many nucleotides** bonded together in a **long chain**

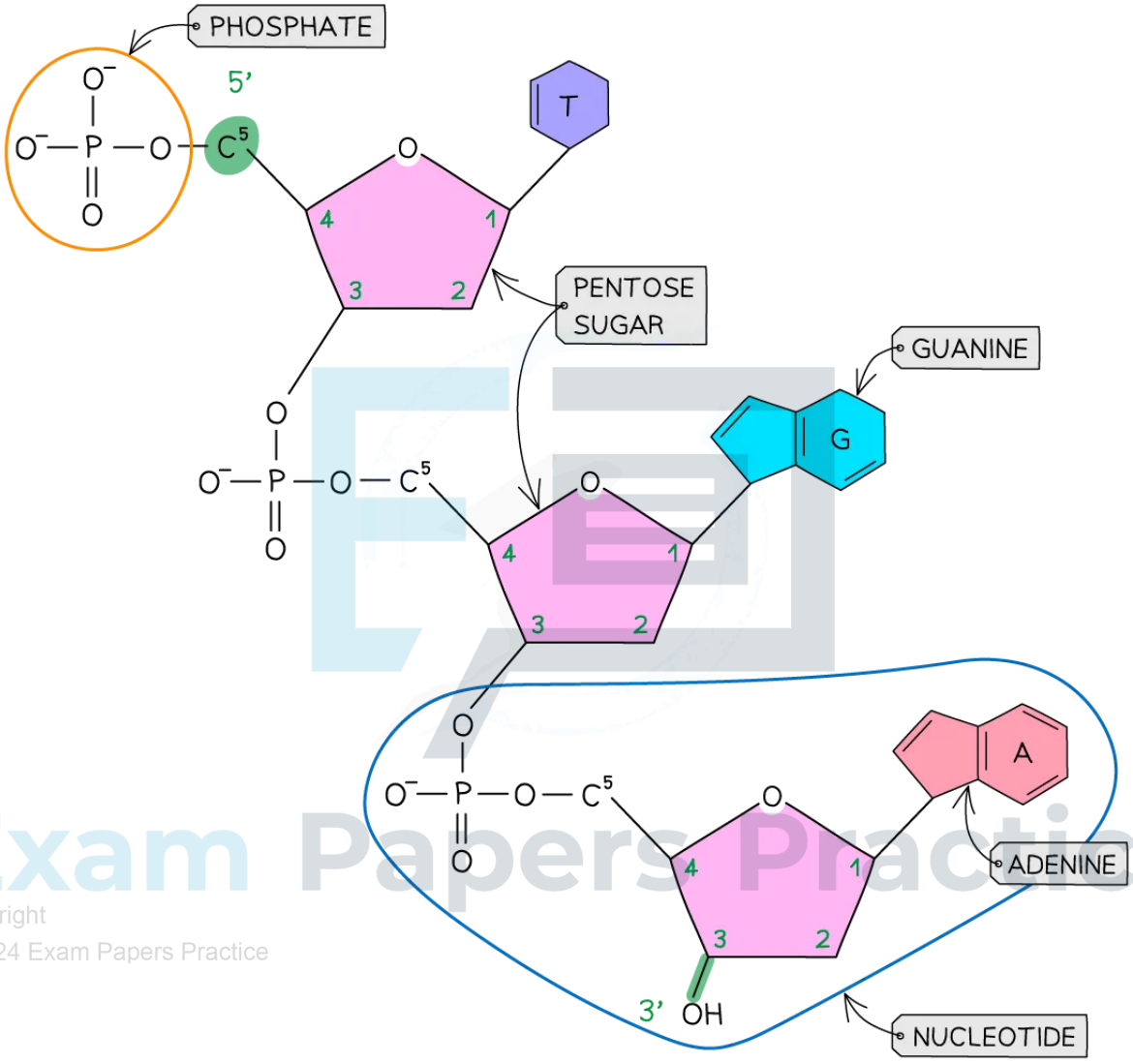
DNA nucleotide diagram



A DNA nucleotide

- DNA molecules are made up of **two polynucleotide strands** lying side by side, running in opposite directions – the strands are said to be **antiparallel**
- Each DNA polynucleotide strand is made up of **alternating deoxyribose sugars and phosphate groups bonded together** to form the **sugar-phosphate backbone**
- Each DNA polynucleotide strand is said to have a **3' end and a 5' end** (these numbers relate to which carbon atom on the pentose sugar could be bonded with another nucleotide)
- Because the strands run in opposite directions (they are **antiparallel**), one is known as the **5' to 3' strand** and the other is known as the **3' to 5' strand**
- The nitrogenous bases of each nucleotide project out from the backbone **towards the interior** of the double-stranded DNA molecule

A single DNA polynucleotide strand diagram



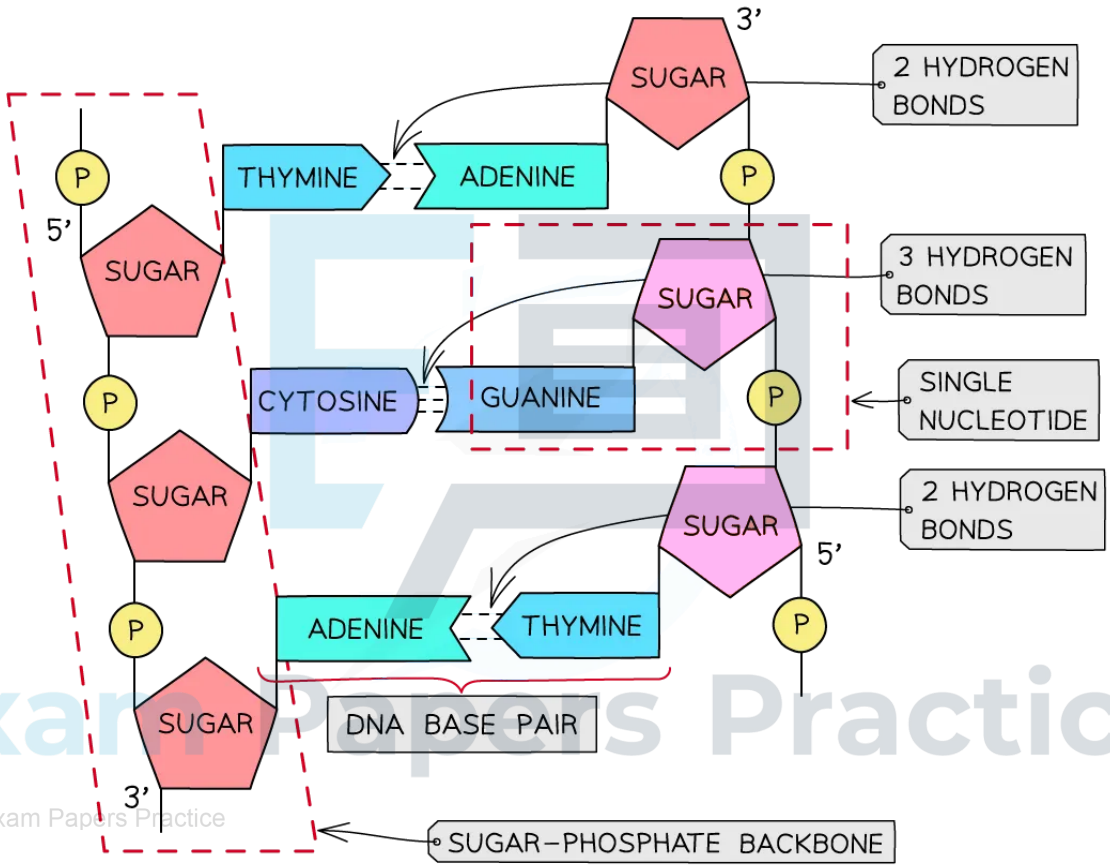
A single DNA polynucleotide strand showing 3 nucleotides in a sequence

Hydrogen bonding

- The two antiparallel DNA polynucleotide strands that make up the DNA molecule are **held together by hydrogen bonds** between the nitrogenous bases
- These hydrogen bonds always occur between the **same pairs of bases**:
  - The purine **adenine** (A) always pairs with the pyrimidine **thymine** (T) – two hydrogen bonds are formed between these bases

- The purine **guanine** (G) always pairs with the pyrimidine **cytosine** (C) – three hydrogen bonds are formed between these bases
- This is known as **complementary base pairing**
- These pairs are known as **DNA base pairs**

DNA molecule with hydrogen bonding diagram

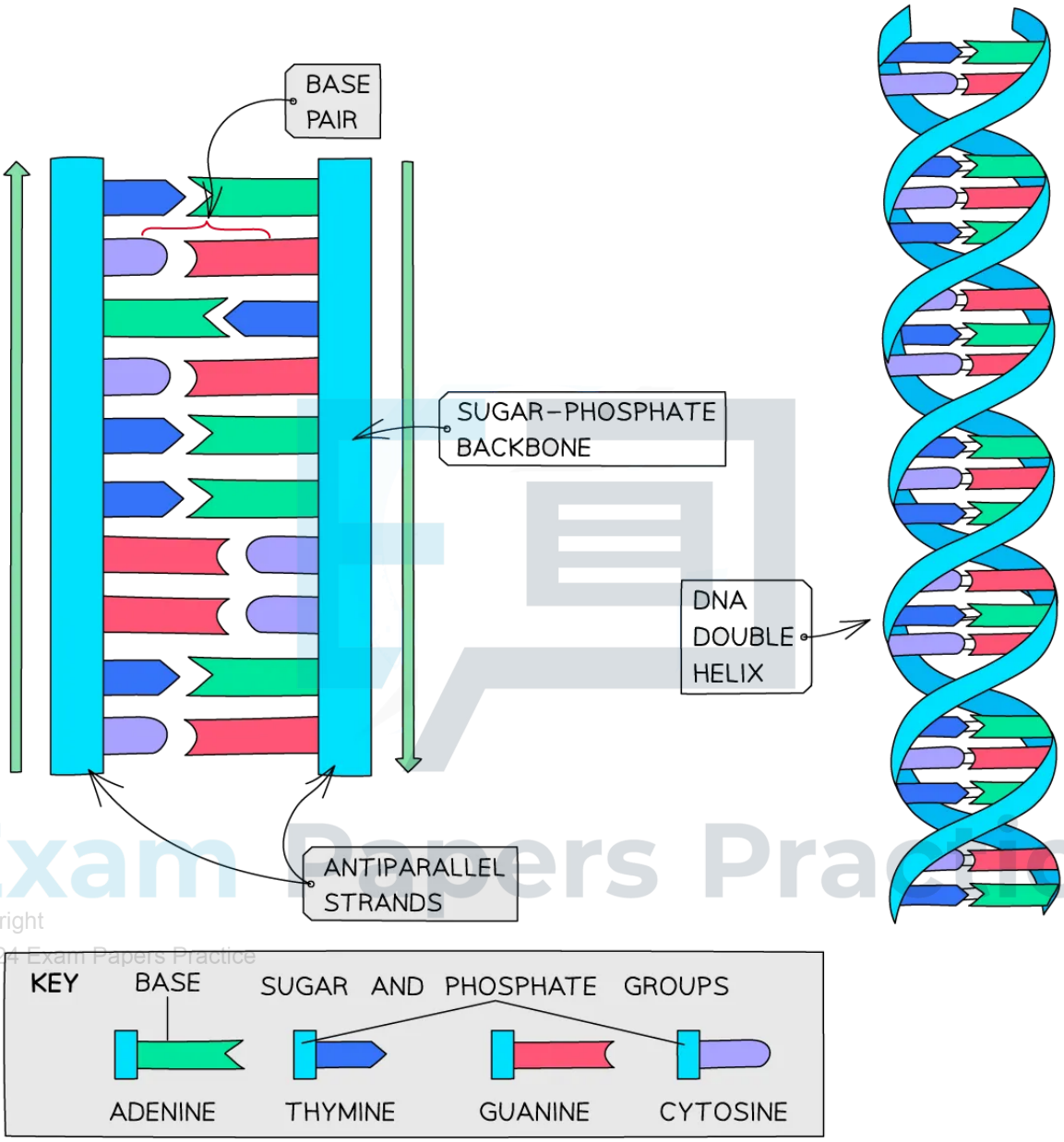


A section of DNA – two antiparallel DNA polynucleotide strands held together by hydrogen bonds

Double helix

- DNA is not two-dimensional as shown in the diagram above
- DNA is described as a double helix
- This refers to the **three-dimensional shape** that DNA molecules form

DNA double helix formation diagram

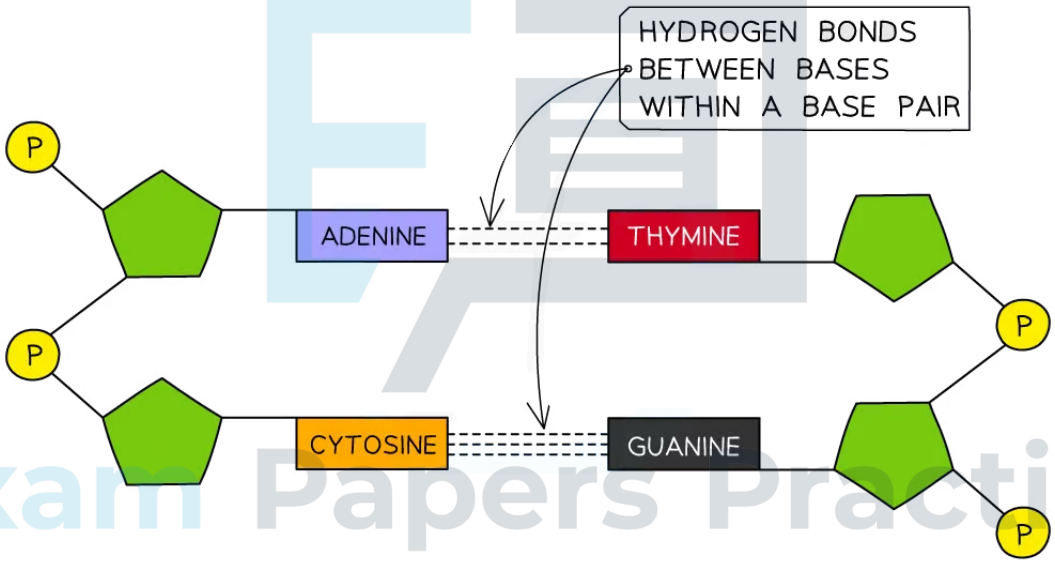


DNA molecules form a three-dimensional structure known as a DNA double helix

Exam Tip

Make sure you can name the different components of a DNA molecule (sugar-phosphate backbone, nucleotide, complementary base pairs, hydrogen bonds) and make sure you are able to locate these on a diagram. Remember that covalent bonds join the nucleotides in the sugar-phosphate backbone, and hydrogen bonds join the bases of the two complementary strands together. Remember that the bases are complementary, so the number of A = T and C = G. You could be asked to determine how many bases are present in a DNA molecule if given the number of just one of the bases.

Drawing base-pairing in a DNA molecule



When drawing the base pairing, the opposite strand should be antiparallel to the first. The presence of hydrogen bonding is shown, but the numbers/lengths of bonds is not required

Exam Tip

Simple, hand-drawn shapes will suffice in an exam. Expert tip - a **large** drawing is always easier for an examiner to read (and award marks for) than a small one! Read the question carefully; examiners often want a **whole nucleotide** to be identified in your diagram and to ensure your diagram includes **all 4 complementary bases.** You don't have to remember the number of hydrogen bonds between the bases. Also, remember to draw DNA strands as **antiparallel** (one upside-down versus the other) but you don't have to be able to draw a helix shape!

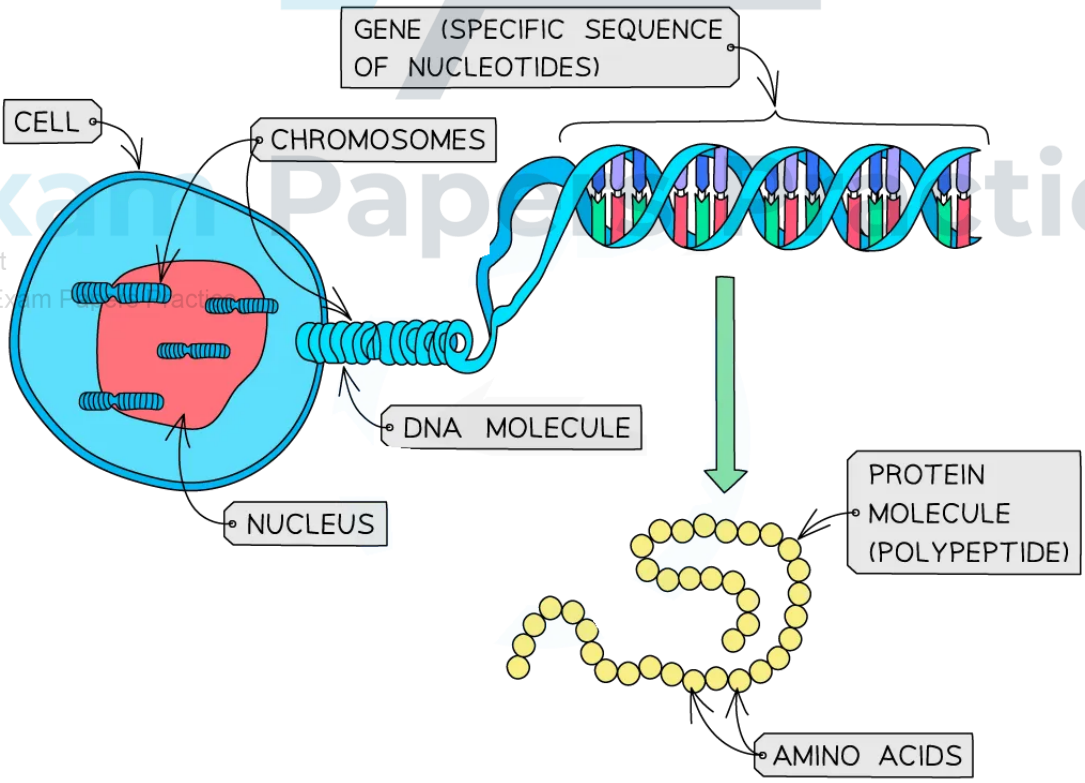
Basis of Genetic Code

The Genetic Code

Genetic code

- DNA molecules carry the genetic code as a **sequence of nitrogenous bases** in the nucleotides
  - These bases are adenine, guanine, cytosine and thymine
- One of the strands of a DNA molecule will **carry the base sequence** that will be read by enzymes
  - This strand is known as the **coding strand**
- The **sequence of bases** that form genes on the coding strand will determine the **order of amino acids** in the proteins that are synthesised
- The code is read as a **triplet of bases**, called a **codon**, with each sequence of three bases coding for **one amino acid**
  - Remember that there are 20 different amino acids that could be coded for
- The sequence of amino acids will **determine the shape and function of the protein** that is synthesised from the code

From gene to protein diagram



The sequence of DNA bases in the genes codes for the production of a specific protein molecule

Conservation of The Genetic Code

The genetic code is universal

- The genetic code is **universal**, meaning that almost every organism uses the **same code** (there are a few rare and minor exceptions)
- The **same triplet codes code for the same amino acids in all living things** (meaning that genetic information is transferable between species)
  - The universal nature of the genetic code is why genetic engineering is possible
- This provides evidence for a **universal common ancestor** from which all living organisms on Earth evolved
- Over time, mutations have led to **changes in some of the base sequences** of organisms
  - These base sequences form the genome of an organism
  - Some base sequences form part of regions that code for proteins, called **coding sequences**, while others are located in regions that do not code for proteins (**non-coding sequences**)
- Many of these coding and non-coding sequences have remained unchanged in all organisms and are known as **conserved sequences**
- Highly conserved sequences are usually found in the genes that code for **proteins involved with** transcription **and** translation, as well as **histone proteins** which help to package DNA tightly into the nucleus
- The similarity in these sequences indicate that living organisms share a **universal ancestry**

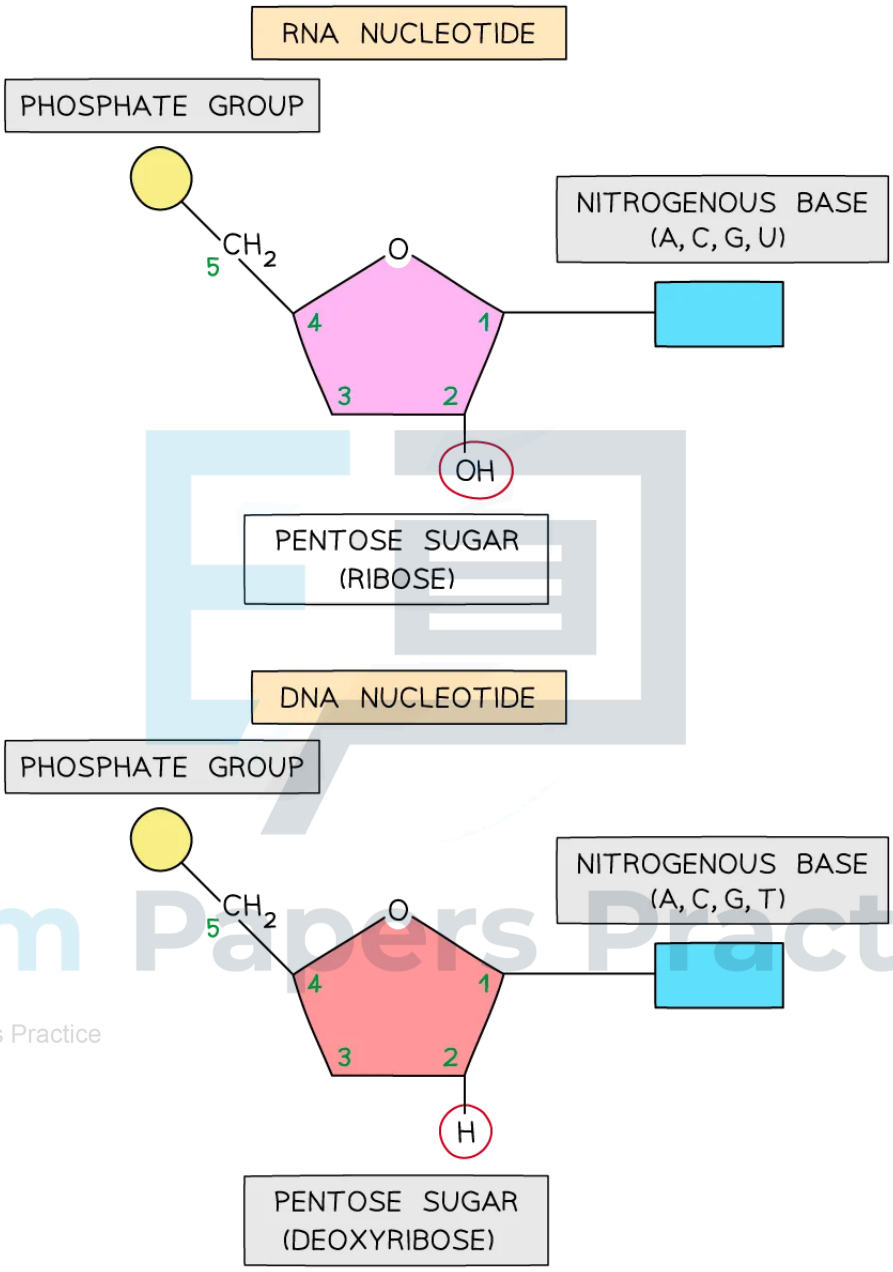
Nucleic Acid Structure & Function

DNA & RNA: Comparison

Differences between DNA and RNA

- **Unlike DNA**, RNA nucleotides **never contain** the nitrogenous base **thymine** (T) – in place of this they contain the nitrogenous base **uracil** (U)
- **Unlike DNA**, RNA nucleotides contain the **pentose** sugar **ribose** (instead of deoxyribose)

Comparing DNA and RNA nucleotides diagram

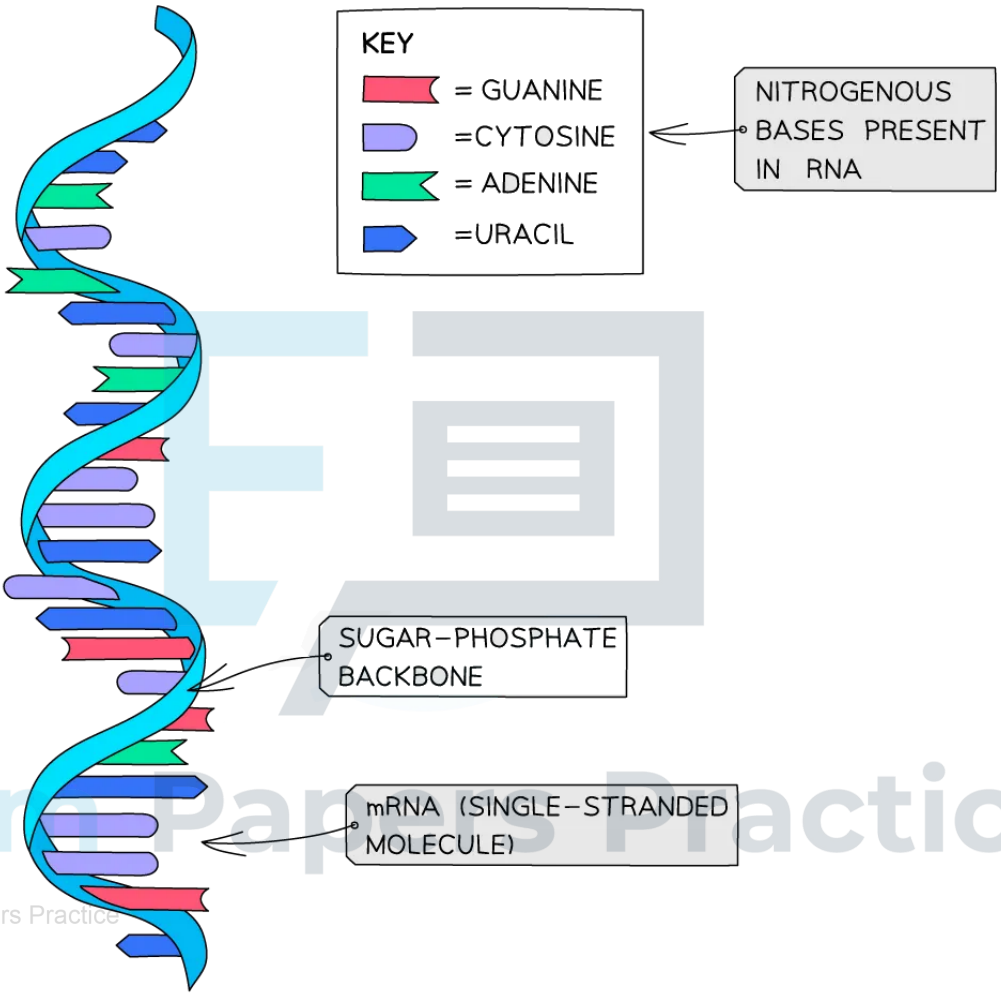


An RNA nucleotide compared with a DNA nucleotide

- **Unlike DNA**, RNA molecules are only made up of **one polynucleotide strand** (they are **single-stranded**)

- **Unlike DNA**, RNA polynucleotide chains are relatively **short compared to DNA**

RNA structure



mRNA as an example of the structure of an RNA molecule

Nucleotide Structure Summary Table

Properties	DNA	RNA
Pentose sugar	Deoxyribose	Ribose

Bases	Adenine (A) Cytosine (C) Guanine (G) **Thymine (T)**	Adenine (A) Cytosine (C) Guanine (G) **Uracil (U)**
Number of strands	Double-stranded (double helix)	Single-stranded

Exam Tip

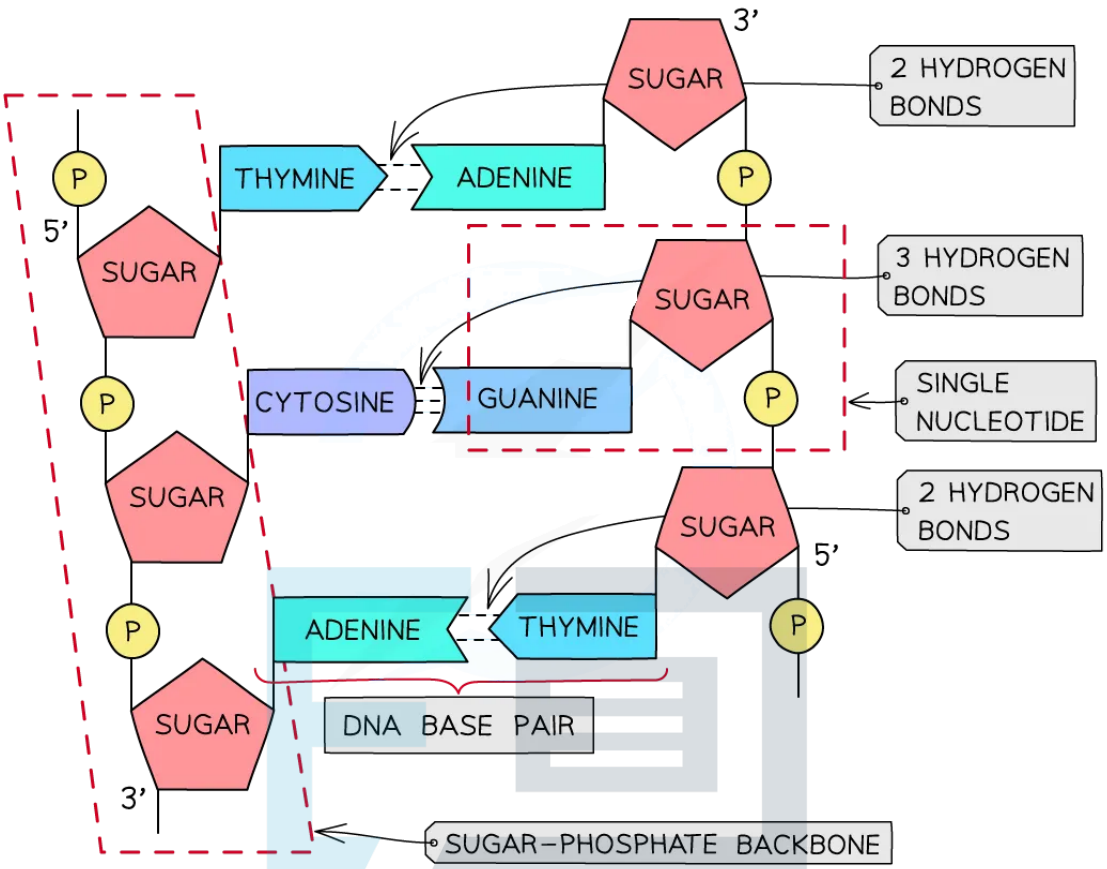
You need to know the difference between DNA and RNA molecules (base composition, number of strands, pentose sugar present). You also need to be able to sketch the difference between ribose and deoxyribose.

Complementary Base Pairing

The role of complementary base pairing

- Complementary base pairing means that the DNA bases on different strands will always pair up in a very specific way:
  - **Adenine** (A) will pair up with **Thymine** (T)
  - **Cytosine** (C) will pair up with **Guanine** (G)
- This is because the hydrogen bonds that hold the two DNA strands together can **only form between these base pairs**:
  - **Two hydrogen bonds** form between A and T
  - **Three hydrogen bonds** form between C and G
- Complementary base pairing means that the **base sequence on one DNA strand** determines the sequence **of the other strand**
  - We say that one strand acts as a template of the other
- This allows DNA to be copied **very precisely during DNA replication** which in turn ensures that the genetic code is accurately copied and expressed in newly formed cells

Complementary base pairs and hydrogen bonding diagram



A section of DNA showing nucleotide bonding and complementary base pair bonding

DNA: Information Storage Molecule

Diversity of DNA base sequences

- Despite the genetic code **only containing four bases** (A, T, C, G), they can combine to form a **very diverse range** of **DNA base sequences** in DNA molecules of **different lengths**
- This means that DNA has an almost **limitless capacity for storing genetic information** in living organisms
- One way in which this storage capacity can be measured is by the **number of genes** contained within the DNA of an organism
- Even the most simplistic forms of life may contain several thousand genes within their DNA

Comparing the Number of Genes between Different Organisms Table

Organism	Human	Dog	Water flea	Bacterium (*E. coli*)	Rice plant
Approximate number of genes	20 000	19 000	31000	4 300	41500

- The storage capacity of DNA can also be measured in the **number of base pairs** contained within the **genome** of an organism
- The **DNA in the nucleus of a human cell** contains about 3.2 gigabases
  - That is about 10^9 DNA base pairs
- These base pairs are contained in DNA with a **length of about 2 meters**, that fits within the nucleus of each human cell
  - Given the fact that a nucleus is microscopic in size, is an indication of how incredibly well packaged this amount of genetic information is
- This gives DNA an **enormous capacity** for storing genetic 'data' **with great economy**

DNA Structure (HL)

Significance of Directionality

Directionality of RNA and DNA

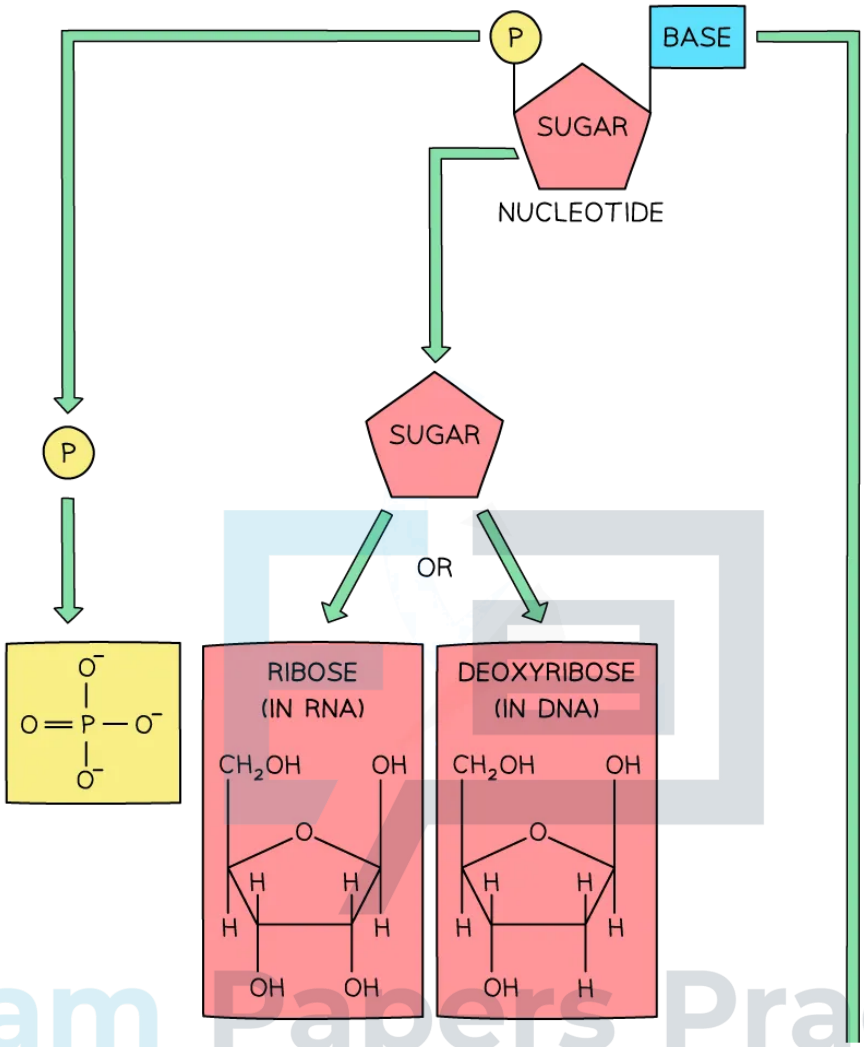
- When nucleotides are linked together to form nucleic acids, such as RNA and DNA, the **phosphate groups** form a bridge between **carbon-3 of one sugar molecule** and **carbon-5 of the next one**
- This means that each polynucleotide strand has a **3' end** where the OH group is located on carbon-3 of the sugar molecule and a **5' end** containing the phosphate group on carbon-5
- In a DNA molecule, one strand runs from **5' to 3'** while the other strand runs from **3' to 5'**
  - This is why the two strands are said to be **antiparallel**
- The **directionality of polynucleotide strands** plays an important role in the processes of:
  - DNA replication
  - Transcription
  - Translation
- During **transcription**, the genetic code on one of the DNA strands (the **coding strand**) is transcribed into a **strand of mRNA**
  - The coding strand is always read in the **3' to 5'** direction by enzymes which will synthesise the mRNA strand in the **5' to 3'** direction
- The mRNA will move into the cytoplasm of the cell, where **ribosomes will translate** the transcribed code in the **5' to 3'** direction
  - The base sequence of the genetic code will determine the **specific order of the amino acids** in the polypeptide chain created during **translation**
- **Directionality in RNA and DNA** are therefore crucially important to ensure that the genetic code is copied, transcribed and translated correctly

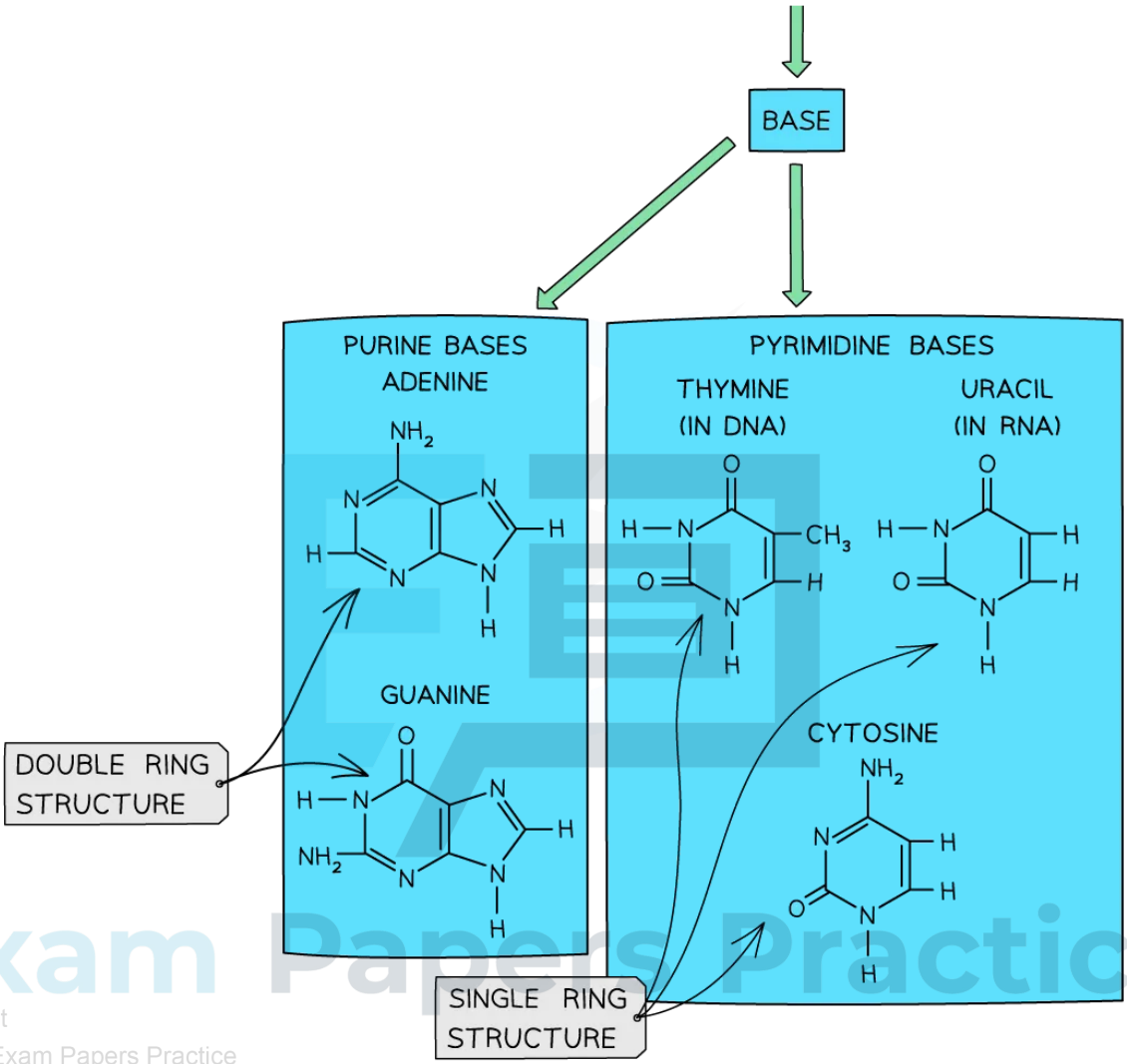
DNA Helix Structure

Purine to pyrimidine bonding in the DNA helix structure

- Francis **Crick** and James **Watson** were two Cambridge scientists who worked together to establish the **double helix** structure of DNA in 1953
- Through trial and error, they managed to build a model of the DNA double helix structure where the different base pairs fit together correctly
- The base pairings **A to T** and **C to G** are **equal in length**, meaning that the DNA helix will have the **same 3D structure** regardless of the base sequence
  - Adenine (A) and guanine (G) are **purine** bases while thymine (T) and cytosine (C) are **pyrimidine** bases
  - Purines are larger in size than pyrimidines due to their two carbon ringed structure
- The stability of the double helix is further increased by the **hydrogen bonds** that form between these **complementary base pairs**

Purines and pyrimidines diagram





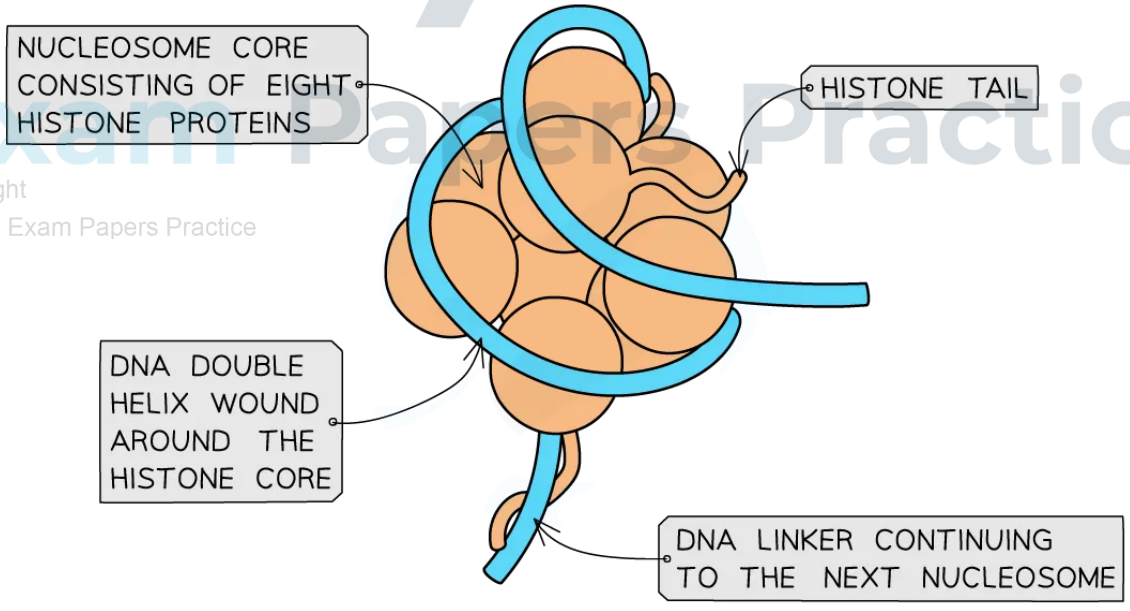
The different sizes of purine and pyrimidine bases mean that they can only pair up in a very specific way. Note that you do not need to know the structural formulae of purines and pyrimidines

Nucleosomes & Molecular Visualisation Software: Skills (HL)

Nucleosomes

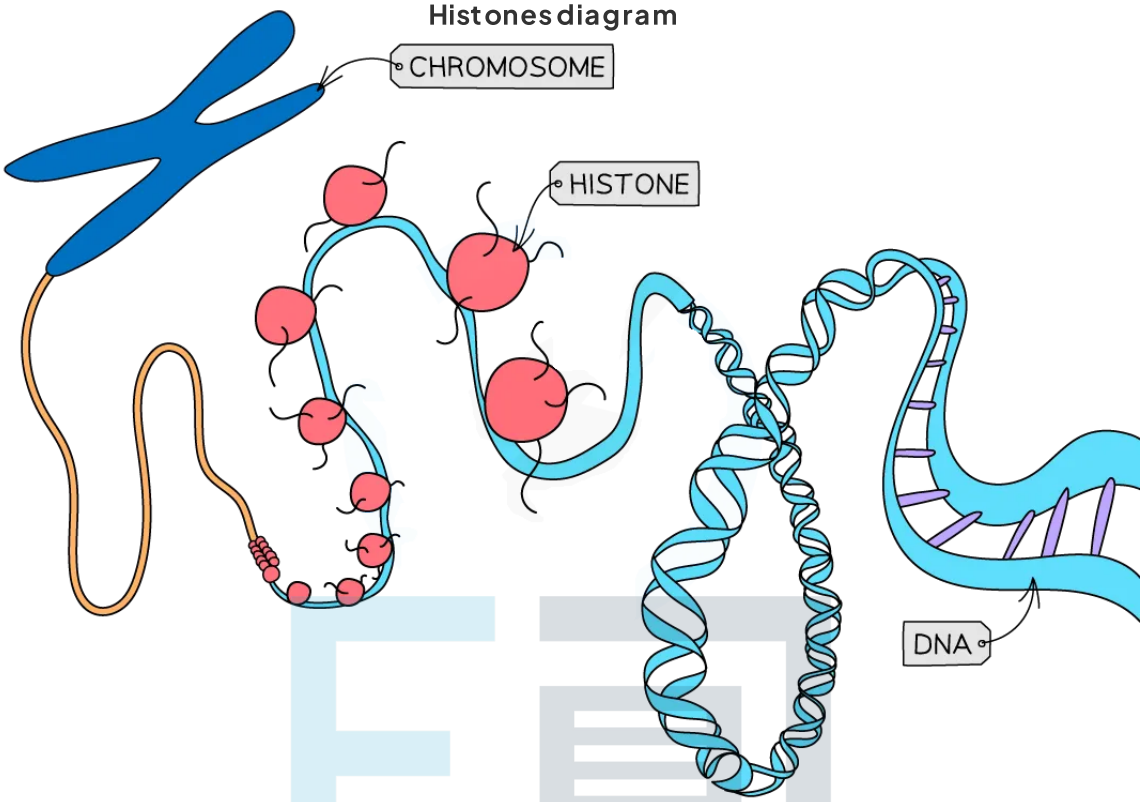
- Unlike most prokaryotic DNA which is referred to as 'naked', **eukaryotic nuclear DNA** is associated with proteins called **histones** (to form **chromatin**)
- Histones package the DNA into structures called **nucleosomes**
  - The nucleosome consists of a strand of DNA coiled around a core of **eight histone proteins** (octamer) to form a bead-like structure
  - DNA takes **two turns** around the histone core and is held in place by an additional histone protein which is attached to linker DNA
  - The DNA molecule continues to be wound around a series of nucleosomes to form what looks like a '**string of beads**'
- Nucleosomes help to **supercoil the DNA**, resulting in a compact structure which saves space within the nucleus
  - Nucleosomes also help to **protect DNA** and **facilitate movement of chromosomes** during cell division
  - An analogy for supercoiling is **twisting an elastic band** repeatedly until it forms additional coils
- Nucleosomes can be **tagged with proteins** to promote or suppress transcription

Nucleosome structure diagram



Structure of a nucleosome

Histones diagram



DNA is wrapped around a series of nucleosomes.
Nucleosomes coil tightly around each other to form the chromosome structure.

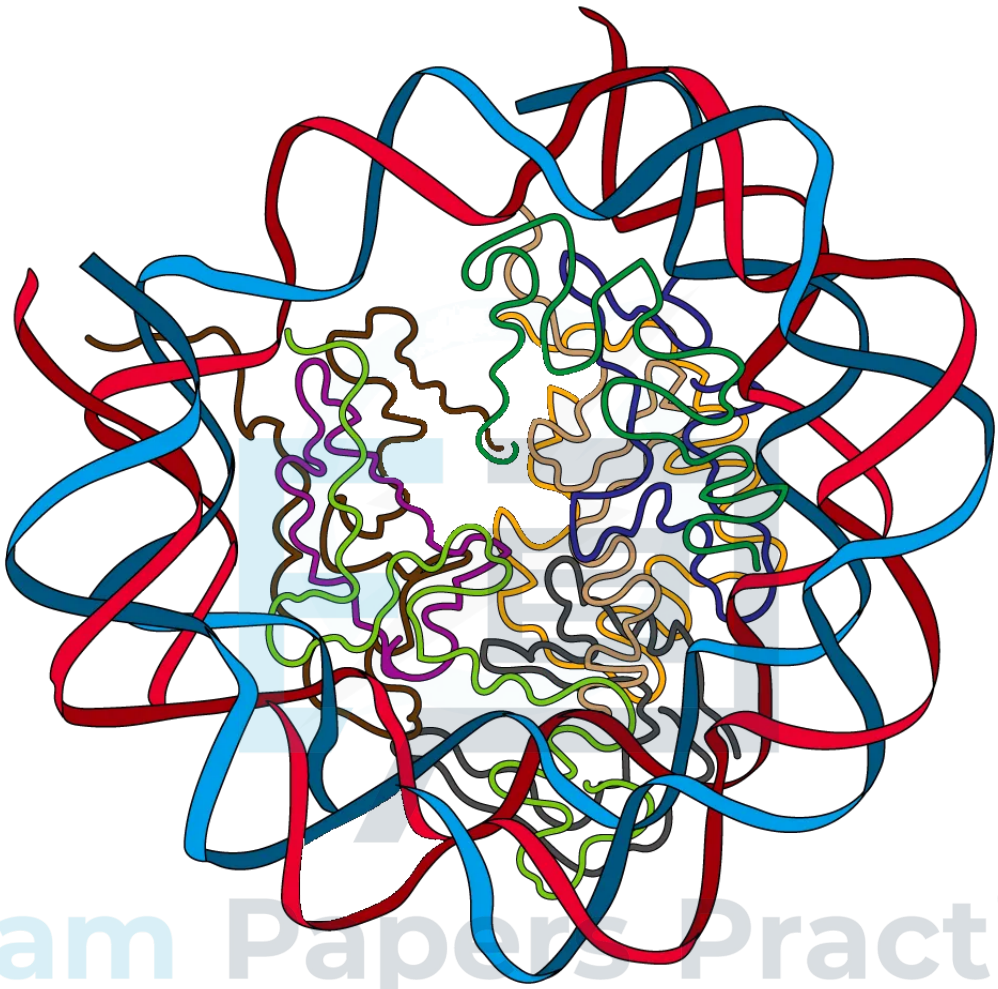
Skills: Molecular Visualisation Software

- **Molecular visualisation software** can be used to help understand molecular structures
  - Macromolecules like **protein**, **DNA, RNA** and **complex carbohydrates** can be **visualised as 3-D structures**
- This allows researchers to analyse macromolecules and/or study interactions between them
  - **Primary sequence** information can be related to structure and function
  - This helps to relate how **structure** might relate to **chemical** or **biological behaviour**
- Macromolecules can be represented in many different ways including **ball and stick atom models** or **simplified ribbon** representations that show the protein backbone
- Most molecular visualisation software is **freely available on the Internet** or can be accessed through many bioinformatics repositories such as the Protein Data Bank (PDB)

Analysing the association between protein and DNA within a nucleosome

- Visit the **Protein Data Bank (PDB) site** and search for: **6T79 structure of human nucleosome** (do not put the search term in quotes)
- Select the "**3D view**" to view the protein structure in **Mol***
  - The 3-D structure of the nucleosome can be viewed
  - The **DNA double helix** can be clearly seen surrounding the **histone proteins**
  - **Rotate** or **zoom** into the image to visualise the different components
  - The DNA can be seen to make **two loops around the histone** octamer core
- Look carefully - the tails of each histone protein can be seen projected from the nucleosome core
  - These can be **chemically modified** to help **regulate gene expression**
- Try **changing different settings** in the viewer or select a different viewer such as JSmol

Human nucleosome diagram



Structure of human nucleosome showing the association between DNA (in 2 loops around the edge) and histones (central region)

The Hershey & Chase Experiment (HL)

The Hershey & Chase Experiment

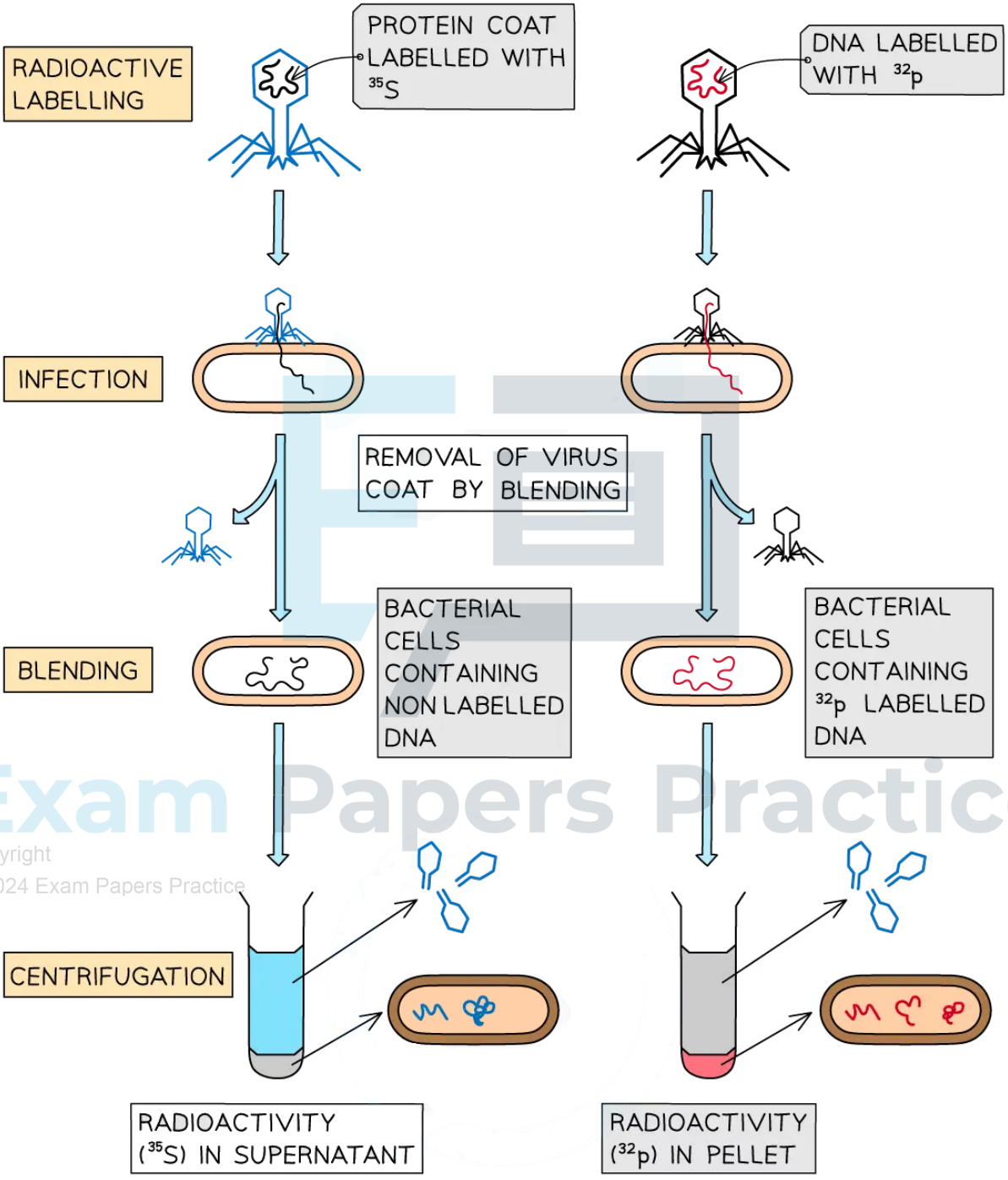
Which Biomolecule is the Heritable Material?

- DNA was identified in 1869 but many scientists assumed that protein was the heritable material
  - Owing to the fact that there are **20 amino acids and only 4 nucleotide bases**
- In the 1950s, Alfred **Hershey** and Martha **Chase** showed that **DNA**, not protein, **is a factor of heredity** responsible for carrying genetic information from one generation to another
- **Viruses** that infect bacteria were used in their experiment as they only consist of DNA encapsulated by a protein coat
- This would allow the biomolecule of heredity (i.e. the one that caused bacterial cells to be used to produce viral progeny) to be easily determined

Analysis of results of the Hershey and Chase experiment

- Hershey and Chase took advantage of the **chemical differences** between DNA and proteins
  - **DNA** contains **phosphorus** but **no sulfur**
  - **Amino acids** (that make up proteins) contain **sulfur** but **no phosphorus**
- **Bacteria** grown in separate media containing either **radioactive sulfur (^{35}S)** or **radioactive phosphorus (^{32}P)** were infected with viruses
- The progeny viruses contained either ^{35}S labelled proteins or ^{32}P labelled DNA
- Unlabelled bacteria were then infected separately with either type of virus
  - Bacteria would be expected to contain the heritable material following infection
- A blender was used to remove attached viruses from the bacterial cells and centrifugation was used to isolate the bacteria
  - Viruses are small so remained in the supernatant in the centrifuge tube
  - Bacteria are larger so formed a pellet
- Only the bacteria infected by **^{32}P labelled viruses (DNA)** were shown to be radioactive
- This suggested that **DNA** (and not protein) was transferred to bacteria and **is the hereditary (genetic) material**

Hershey and Chase experiment diagram



Hershey and Chase's experiment provided unequivocal proof that DNA is the heritable material

NOS: Technological developments can open up new possibilities for experiments – The availability of radioisotopes as research tools made the Hershey-Chase experiment possible

- Radioisotopes were made available to scientists as research tools at the **end of the Second World War**
- This enabled scientists in a **variety of research fields**, such as biochemistry and virology, to do experiments that **were not previously possible**
- Isotopes are particularly useful in studying **chemical changes** that occur during metabolic pathways or life cycles in organisms
- Without the availability of radioisotopes, Hershey and Chase would not have been able to label the different parts of a virus in order to determine that **DNA is the heritable material in organisms**

Chargaff's Data (HL)

Chargaff's Data

NOS: Addressing the problem of induction by the certainty of falsification – How Chargaff's data falsified the tetranucleotide hypothesis of DNA

- **Erwin Chargaff** analysed the **DNA composition of different organisms** during the 1930s and 1940s and made the following discoveries:
  - The number of **purine bases** were **equal** to the number of **pyrimidine bases**
  - The number of **adenine bases** were **equal** to the number of **thymine bases** while the number of **guanine bases** were **equal** to the number of **cytosine bases**
- This means that a **purine base can only pair up with a pyrimidine base** between the sugar-phosphate backbone, since they have different sizes
- This forms the foundation of **complementary base pairing** in DNA

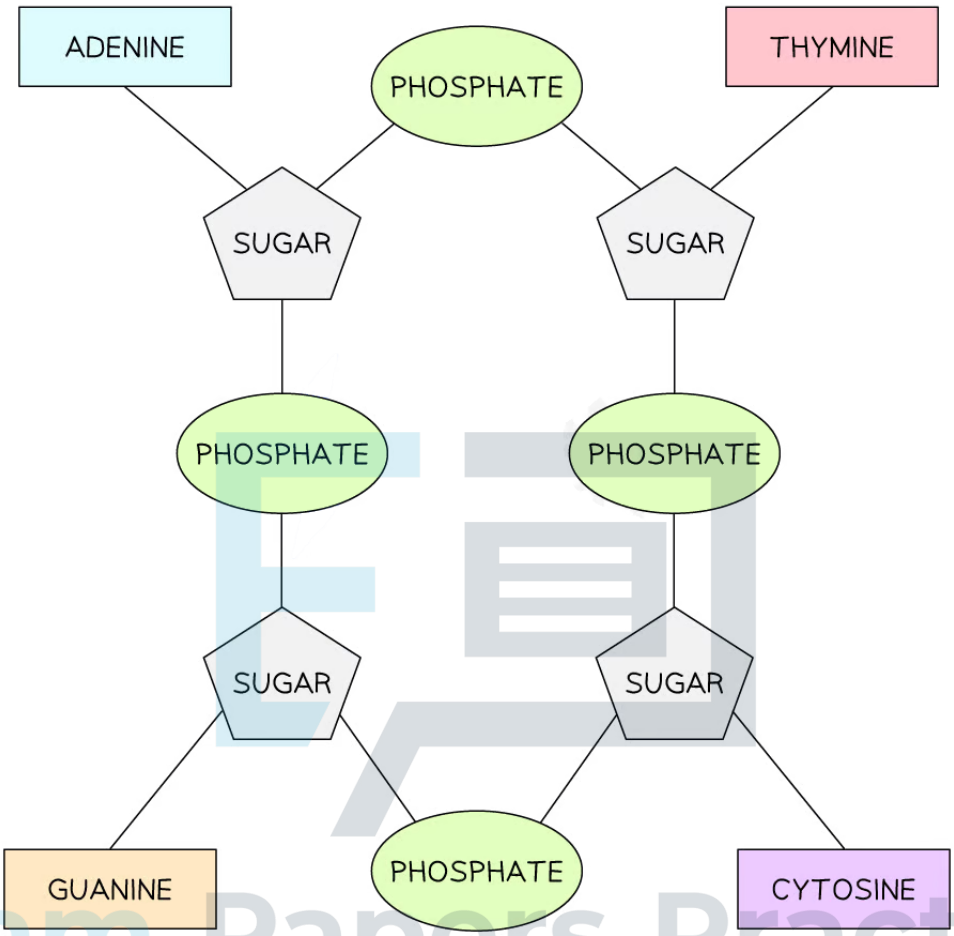
The problem of induction

- The **inductive scientific method** starts with a scientist making observations and collecting raw data
- After data analysis, a **hypothesis is formulated** which is then tested by means of a suitably designed investigation
- This may lead to some **general conclusions** being drawn based on **specific observations**
- Using data gathered in the past to create general **predictions** about what will happen in the future, assumes that the future will be the same as when you gathered your data
- It is therefore **impossible to prove a hypothesis** generated by inductive reasoning as **absolutely true**, since we cannot be sure that the general observations we made in the past will hold true in the future
- This is known as the **problem of induction** and is the main reason why most scientific theories are considered to be tentative
- Even if several investigations supports a hypothesis, it can still be **proven incorrect (falsified) in the future** as new discoveries are made
- For this reason, the philosopher Karl Popper suggested that new scientific knowledge is not gained by inductive steps but rather by the **falsification of existing hypotheses**

Falsification of the tetranucleotide hypothesis

- The biochemist **Phoebus Levene** discovered the **pentose sugars** of DNA and RNA in the early 1900's
- He suggested that the structure of nucleic acid was a **repeating tetramer unit** which he called a nucleotide
  - This was called the **tetranucleotide hypothesis**

Tetranucleotide structure diagram



The tetranucleotide structure of nucleic acid which was suggested by Levene

- At the time of his research, there were **limitations to the analytic techniques** available which made it difficult to determine the relative amounts of nucleotides present in nucleic acids
- The tetranucleotide hypothesis was **falsified by Chargaff's data** in the late 1940s, which showed the organism-specificity of nucleic acids
- When the **structure of DNA** was determined in the 1950s, it further proved that the repeating tetramer unit suggested by Levene would **not be suitable** for carrying genetic information from one generation to the next