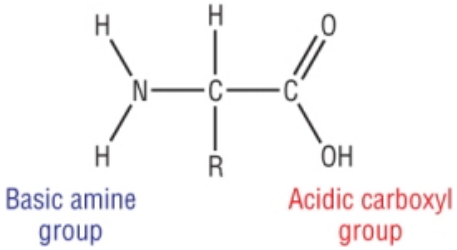
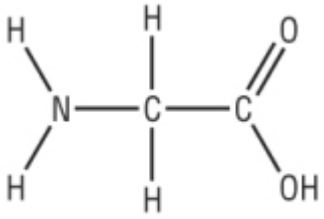
3.13 Amino acids, proteins and DNA

Amino acids

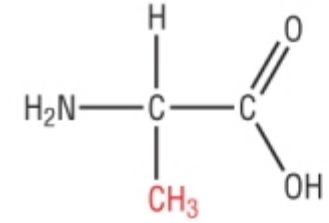
- Have a basic **NH_2**, amine and an acidic **COOH**, carboxylic acid group.
- α **amino acids** have the NH_2 and COOH attached to the **same carbon**:

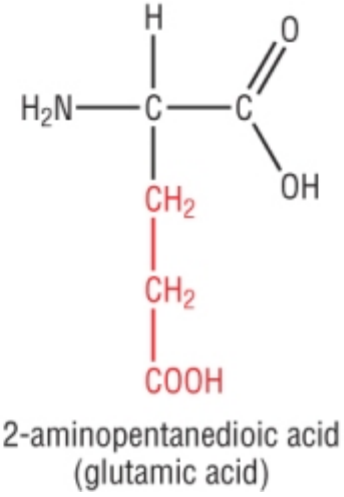


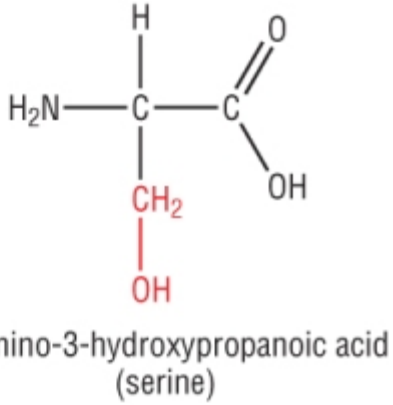
- They are the building blocks for proteins which are held together by peptide links.
- The body has 20 naturally occurring amino acids which join to form proteins, polypeptides, dipeptides, tripeptides and enzymes etc.
- The R is an organic side group and can contain OH, SH, COOH or NH_2 groups.
- Glycine is the simplest amino acid which means it will not contain an R group:



- The names of some amino acids are shown below with different R groups attached.
- All but Glycine exhibit optical isomerism as they have 4 different groups attached to the same carbon.







- Longest chain starting with carboxylic acid as 1, then count the amine / other groups from there:

Naming amino acids

- Give the IUPAC name of the following:

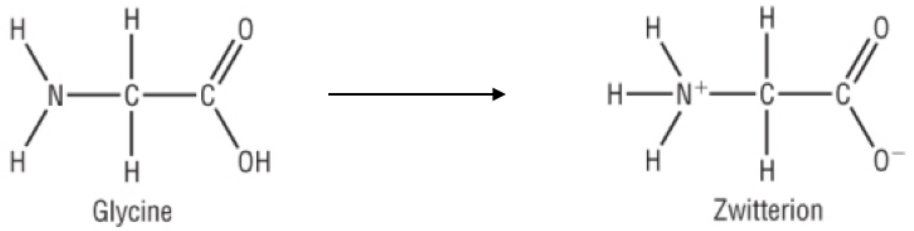
Structure	IUPAC name
$CH_3CH_2CH_2CH_2CH(NH_2)COOH$	
$HOCH_2CH_2CH(NH_2)COOH$	
$HOOCCH_2CH_2CH_2CH(NH_2)COOH$	

- Draw the structure of the following:

Amino acid	Structural formula	Skeletal formula
2-amino-3-methylbutanoic acid (valine)		
2-amino-4-methylpentanoic acid (Leucine)		
2-amino-3-hydroxypropanoic acid (serine)		

Zwitterions:

- The amino acid and carboxylic acid can interact with each other to form a **zwitterion**:



Zwitterion:

Is the pH at which the ion has no net electrical charge

- A proton is transferred from the **COOH** to the **NH_2**
- The **zwitterion** has no overall charge as the **COO^-** cancels out the **NH_3^+**
- This **+ / -** charge increases the attractive forces between amino acids considerably.
- They are often described as having **unusually high melting points**.

Molecule		**Melting point**
Glycine,	NH_2CH_2COOH	262°C
Propanoic acid	CH_3CH_2COOH	-21°C

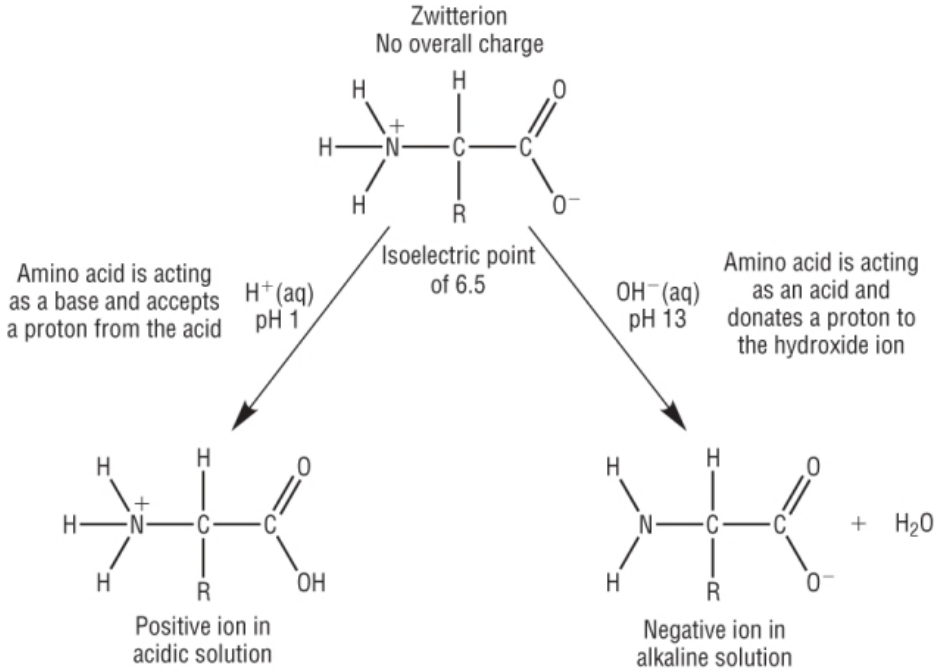
The isoelectric point and R groups:

Amino Acid	**R group**	**Isoelectric Point**
Glycine (Gly)	–H	5.97
Alanine (Ala)	$–CH_3$	6.01
Glutamic acid (Glu)	$–(CH_2)_2COOH$	3.22
Lysine (Lys)	$–(CH_2)_4NH_2$	9.59

- Acidic R groups lower the pH of the isoelectric point.
- Basic R groups increase the pH of the isoelectric point .

The effect of pH on the zwitterion / amino acid structure:

pH < Isoelectric point: **pH > Isoelectric point:**



- In acidic conditions, there is an abundance of H^+ions.
- The amino acid acts as a base and accepts as many H^+ions as possible.

- In Alkaline conditions, there is a deficiency of H^+ions.
- The amino acid acts as an acid and donates as many H^+ions as possible.

Questions:

1) Draw the structures of alanine at the following pH's given its isoelectric point is 6.00:

pH 1.00	pH 6.00	pH 12.00

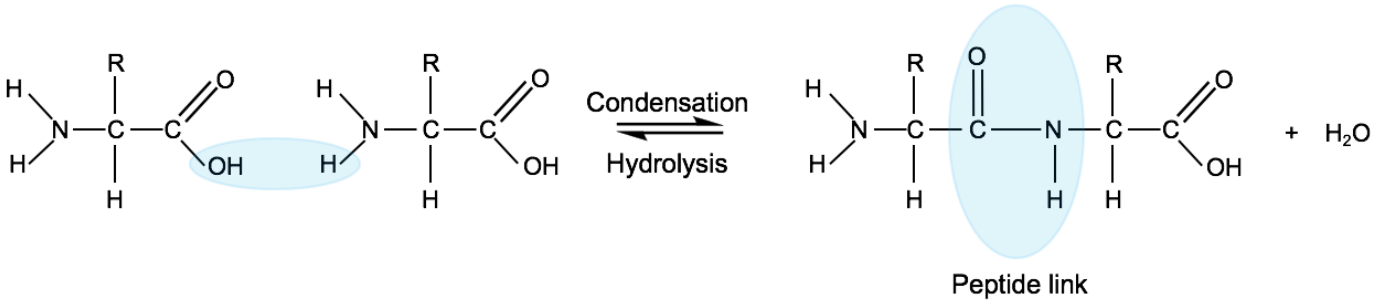
2) Draw the structures of aspartic acid at the following pH's given its isoelectric point is 2.77:

pH 1.00	pH 2.77	pH 13.00

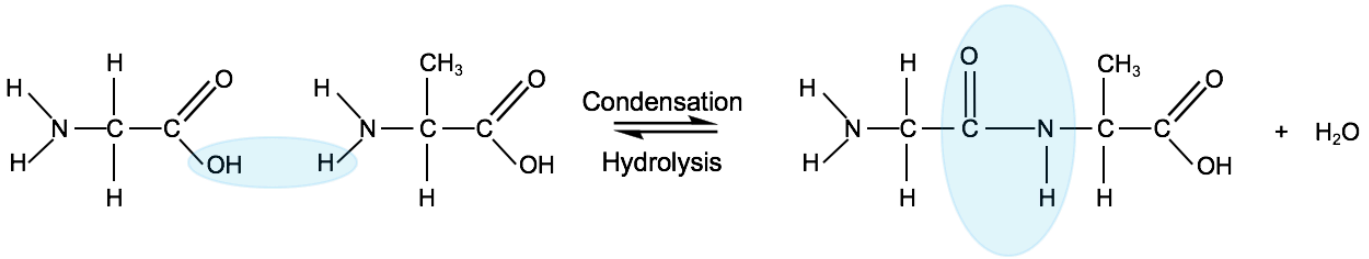
Proteins

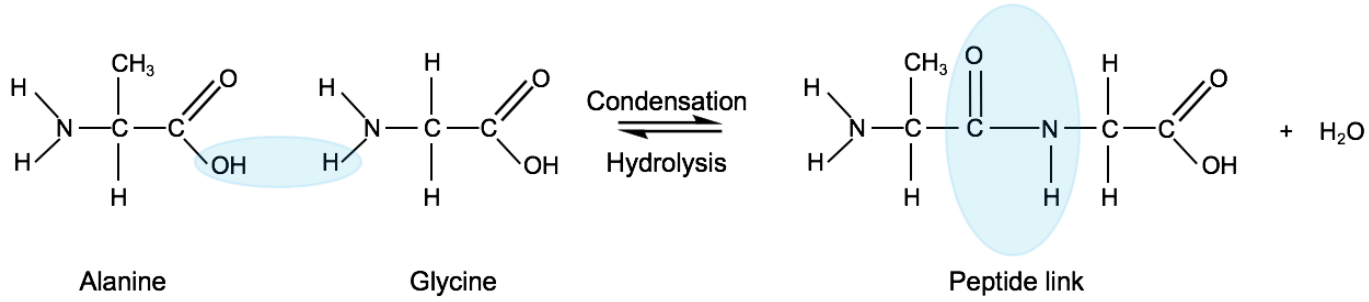
Amino acids and condensation reactions:

- When 2 amino acids join together we call it a dipeptide.
- When 3 amino acids join together we call it a tripeptide.
- When many join together we call it a protein or polypeptide.
- Polypeptides are synthetic.
- Proteins are natural and are usually larger than polypeptides.
- The term '**Peptide link**' is the biochemical name for the **amide link**. (see 3.12 Polymers)
- Whether it is a dipeptide, tripeptide, protein or polypeptide, they all behave the same:
- The forward reaction is called a condensation reaction as water is given off.
- The reverse reaction is called hydrolysis:



- The functional group is an amide group (CONH)
- 2 different amino acids can join to form 2 different dipeptides:

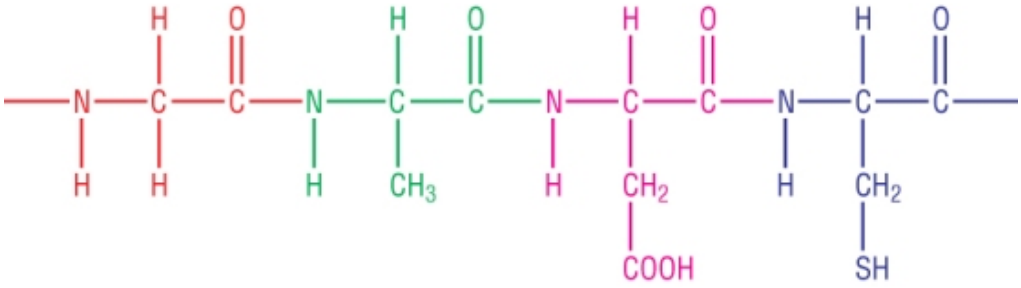




- Dipeptides still have an **NH_2** and a **COOH** group which means they can undergo further condensation reactions.

Proteins (and polypeptides):

- This is a long chain of amino acids joined by peptide linkages:



- Each linkage forms a molecule of water.

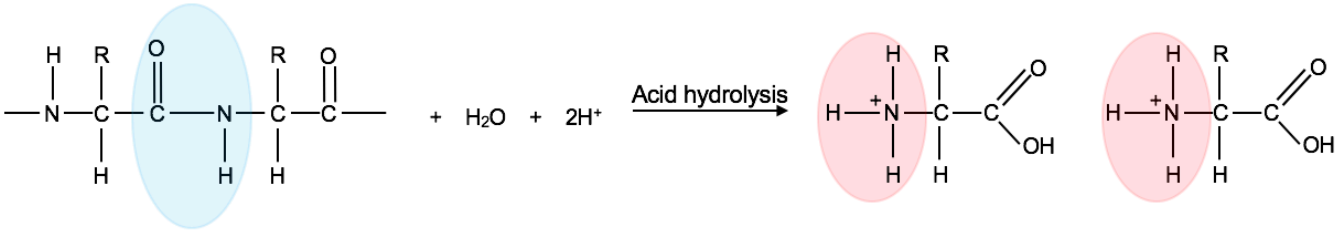
Hydrolysis of polypeptides and proteins:

- Proteins, polypeptides, dipeptides and tripeptides can all be hydrolysed back into their constituent amino acids.
- This is done by using an acid or alkaline catalyst. (review polyamide hydrolysis).

Acid hydrolysis:

Conditions: 6 mol dm^{-3} HCl / reflux 24 hours:

- As this is in acidic conditions, the alkyl ammonium salt is formed:



Peptide link

Abundance of H^+ ions in acidic conditions

- Remember if HCl is used in the equation then the chloride ion would need adding to the balanced equation above.

Alkali hydrolysis:

Conditions: Solution of aq NaOH above 100°C

- As this is in alkaline conditions, the carboxylate salt is formed:

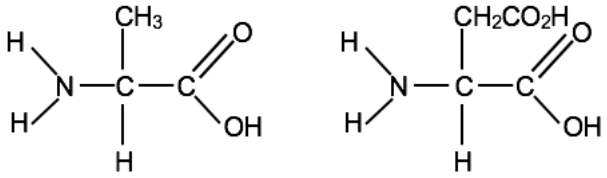
Peptide link

Deficient of H^+ ions in alkaline conditions

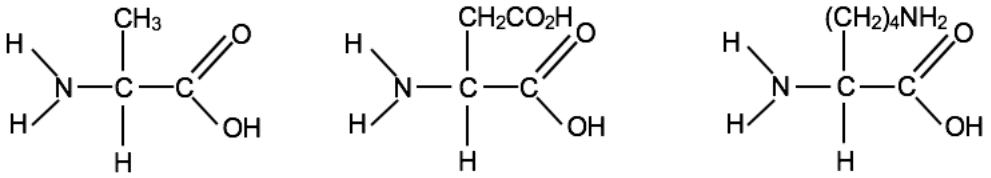
- Remember if NaOH is used in the equation then the sodium ion would need adding to the balanced equation above.

Questions:

1) These 2 amino acids can join in 2 different ways to give 2 different dipeptides. Draw them below:



2) These 3 amino acids can join to form a tripeptide. Draw the tripeptide below in the order they are written below:

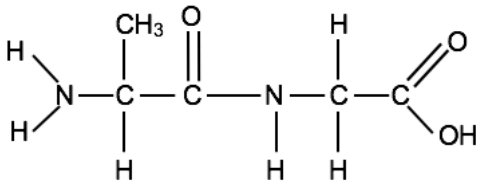


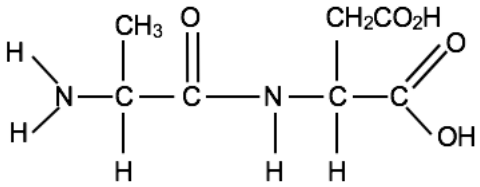
3) Write equations for the reaction of alanine (2-amino propanoic acid) with:

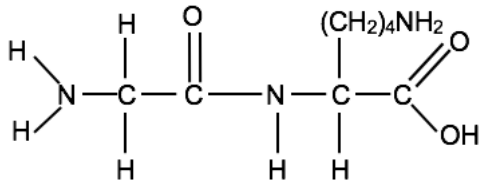
HCl

NaOH

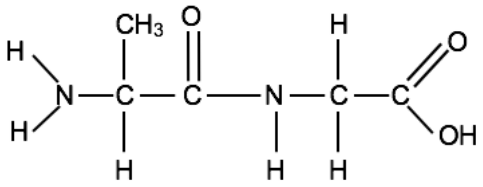
4) Write equations for the acid hydrolysis of the following using HCl:

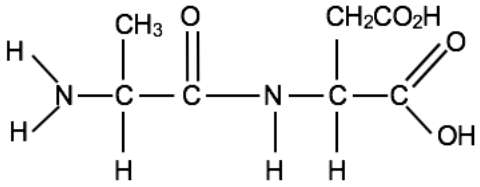


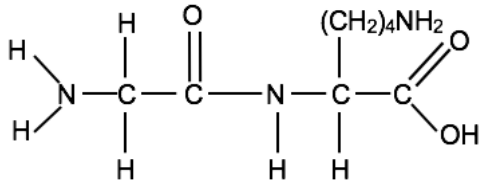




5) Write equations for the alkaline hydrolysis of the following using NaOH:

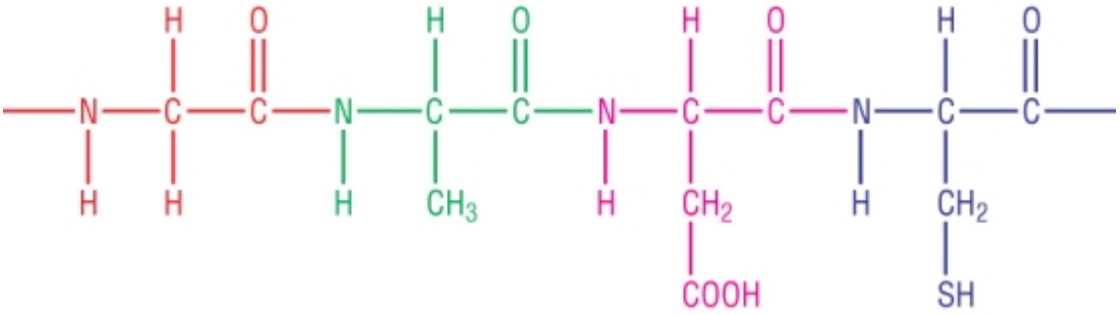




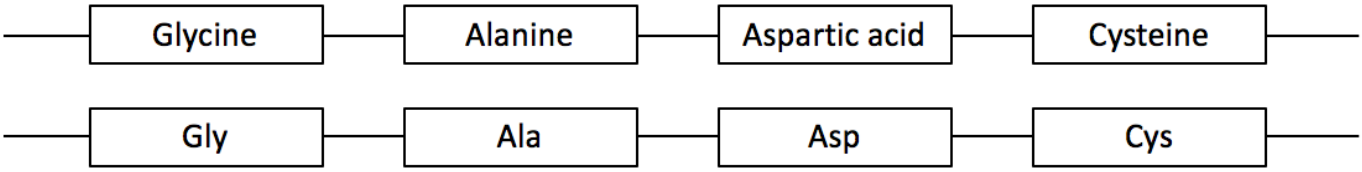


Simple structure of protein:

- This is the sequence of amino acids in a protein chain:



- Each amino acid in the protein chain is referred to as an amino acid residue.
- These can be represented by a 3-letter code:

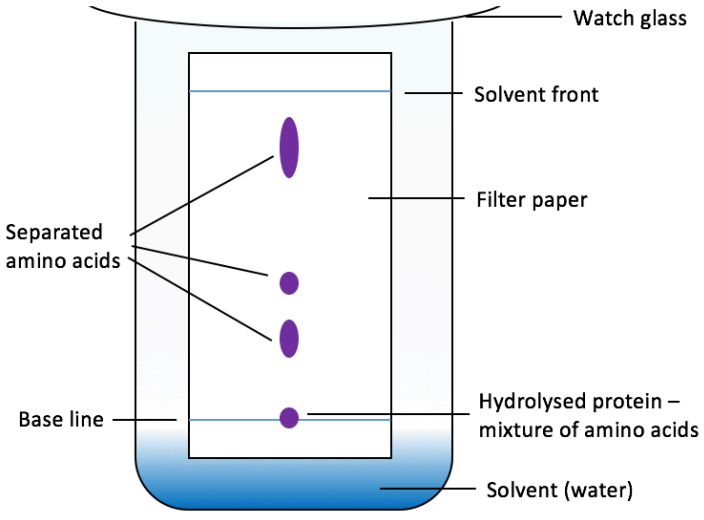


- These can be **hydrolysed** and **separated** by chromatography.
- This allows identification of amino acids in proteins

Separation and identification – Thin layer chromatography

- Amino acids have different R groups which means they will each have different solubilities in the same solvent.
- Chromatography can therefore be used to separate and identify the amino acids:
- As amino acids are clear and colourless, **ninhydrin** a developing agent is used to locate the different amino acids – seen as a violet spot under UV light.

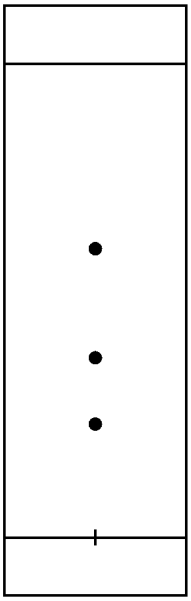
TLC:



$$R_f = \frac{\text{Distance moved by component}}{\text{Distance moved by solvent front}}$$

Example

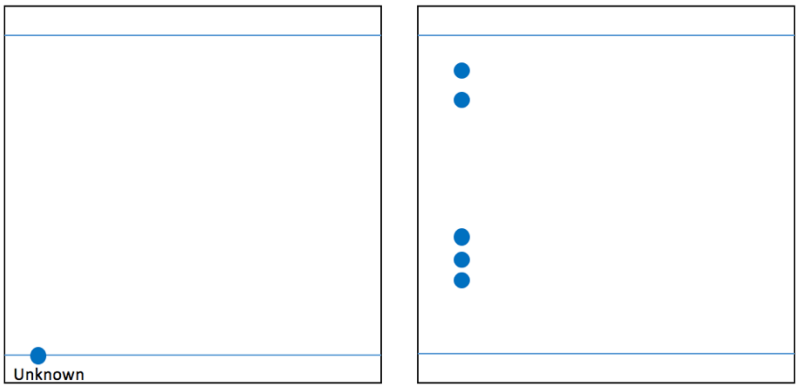
A section of protein is hydrolysed and the amino acids produced are analysed by TLC.
Use the data below to identify the amino acids in the tripeptide.



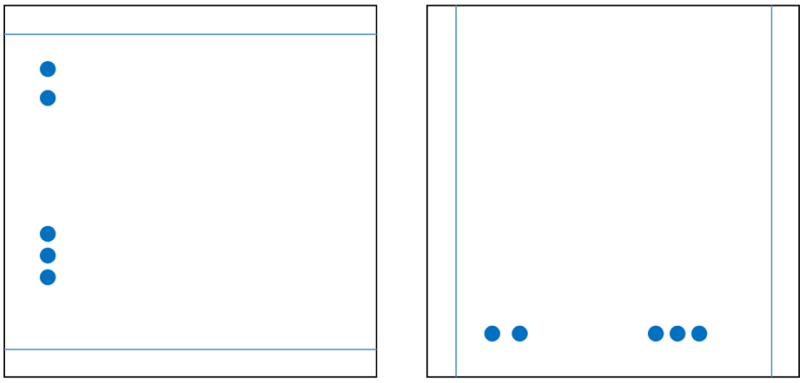
Amino acid	R_f value
Tryptophan	0.66
Valine	0.61
Methionine	0.55
Alanine	0.38
Glutamic acid	0.30
Aspartic acid	0.24
Arginine	0.20

2-Dimensional TLC

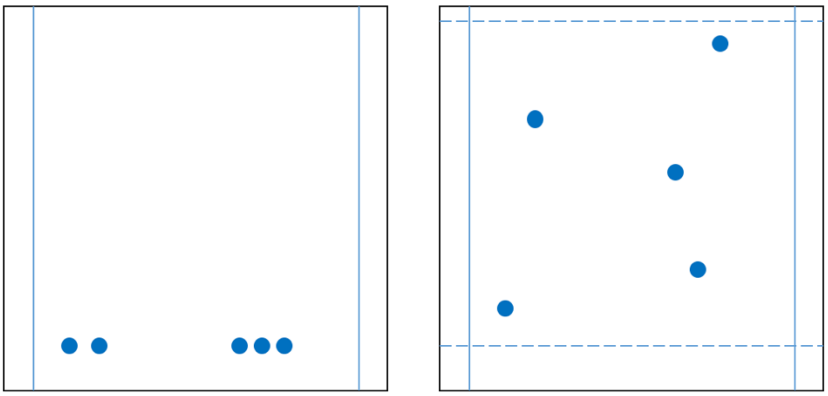
- Amino acids can often have similar R_f values as R groups can be similar
- 2 – dimensional TLC is used to separate components:
- A square TLC plate is used.
- The mixture is spotted in one corner and a chromatogram:
- Separation occurs along one edge of the plate:



- The plate is turned through 90°.



- The chromatogram is run again with a different solvent.



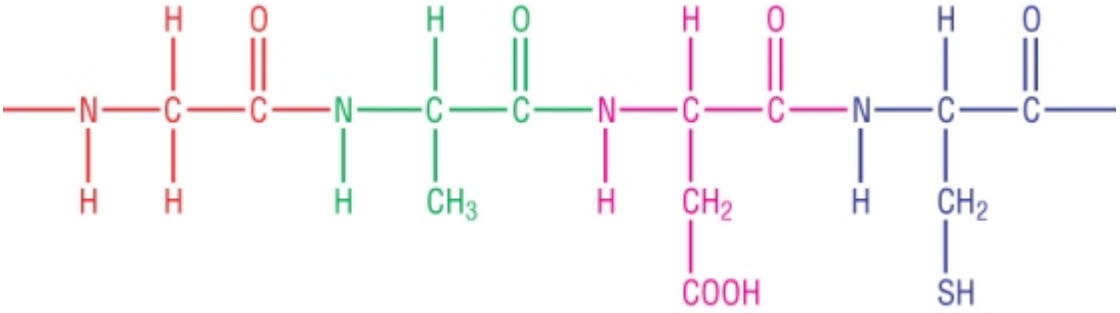
- This gives ***two R_f*** values for each component - one for each solvent.
- The two R_f values are compared to known R_f values for each solvent.
- It gives a greater confidence in the identification of the amino acids.

Structure of protein:

- This long chain of amino acids have complex structures.
- Best thought of in 4 levels:
  - Primary
  - Secondary
  - Tertiary
  - Quartenary (although this structure isn't required)

1) Primary Structure:

- This is the sequence of amino acids in the protein chain:



- Each amino acid in the protein chain is referred to as an amino acid residue.
- These can be represented by a 3-letter code:

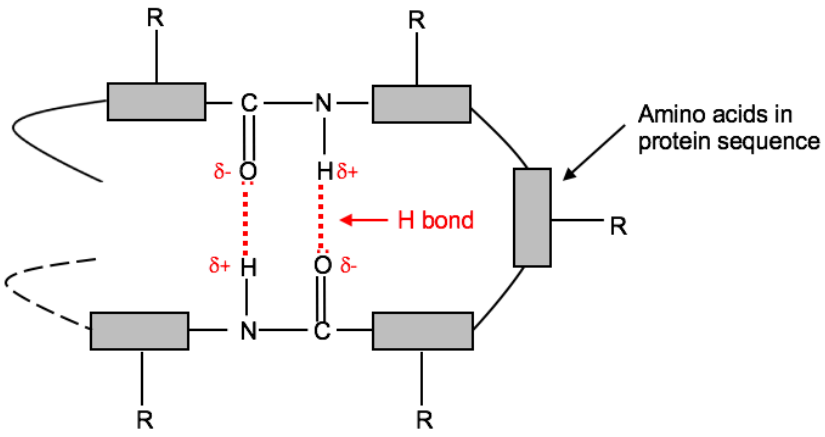
Glycine	Alanine	Aspartic acid	Cysteine
Gly	Ala	Asp	Cys

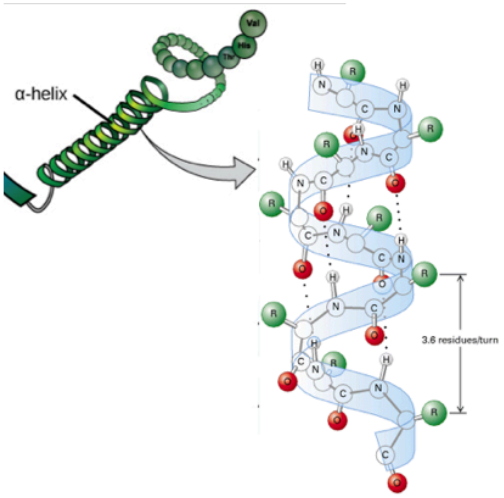
2) Secondary structure:

- The peptide links can form hydrogen bonds with each other in one of 2 ways:

a) α Helix:

- The primary structure forms a helical structure held together by hydrogen bonds.

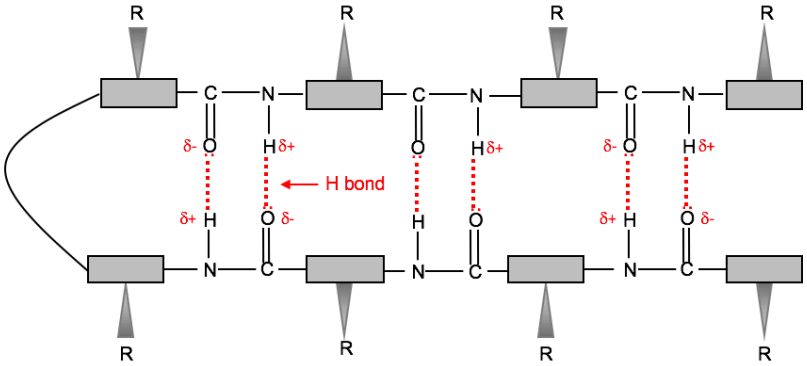


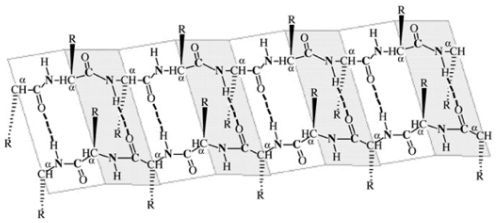


- The R groups point outwards of the helix structure.

b) β Pleated sheets

- The primary structure folds back on itself and held together by hydrogen bonds.
- The given shape looks like a 'pleat' with the R groups alternating up and down along the plane.





- The R groups point outwards of the sheet structure.

3) Tertiary structure:

- This is where the secondary structure folds and coils on itself to give a 3-dimensional structure.
- Additional bonds and intermolecular forces hold the polypeptide or protein in this shape.
- This structure is held together by interactions between different R side groups:

a) Van der Waals' forces of attraction

➢ Between the alkyl R side groups

b) Hydrogen bonds

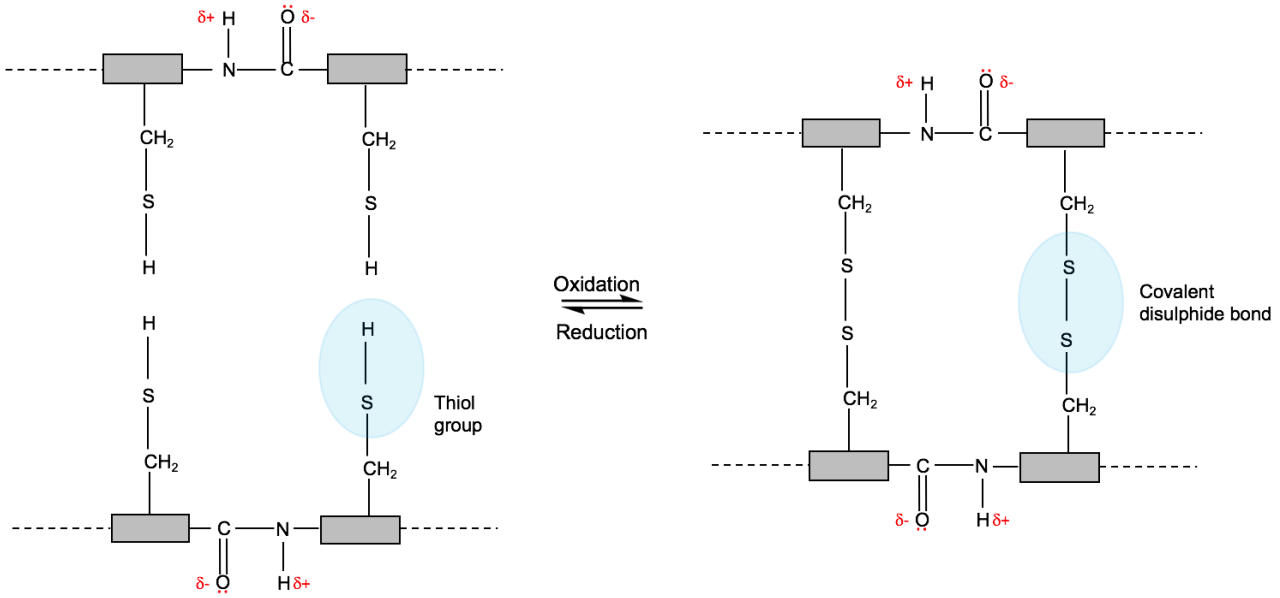
- Between OH / NH / C=O in the R side groups and the main chain.

c) Ionic bonds

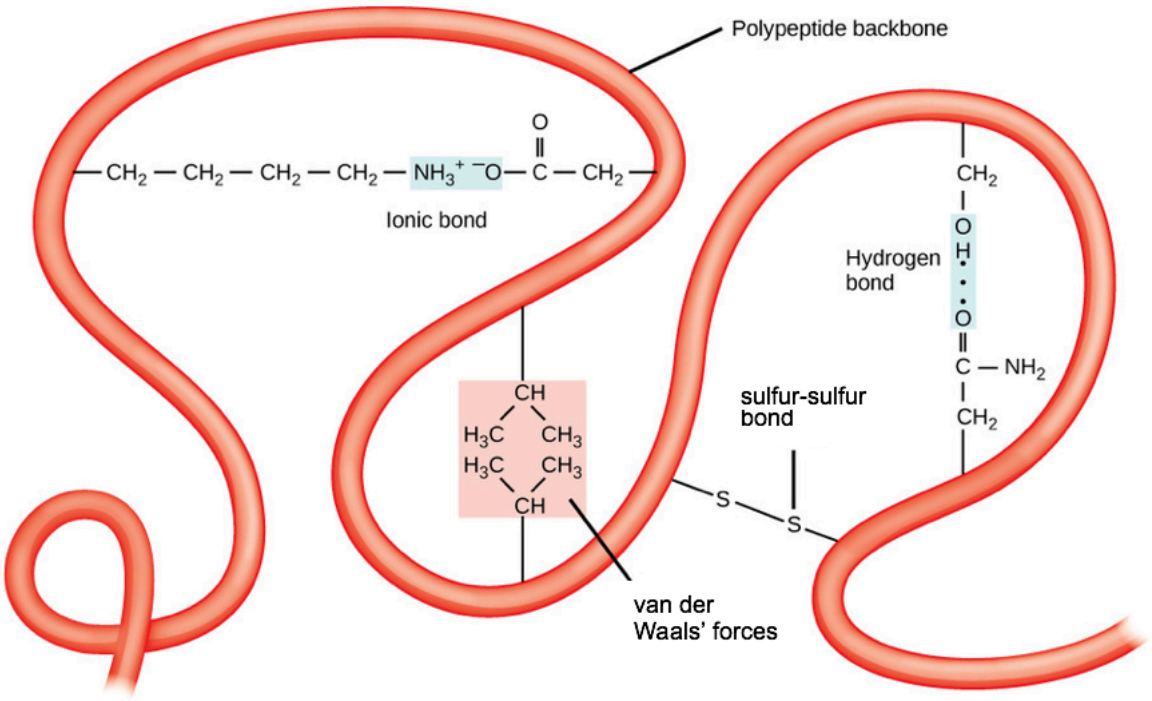
- Between $-COO^-$ and $-NH_3^+$ in the R side groups

d) Disulphide bonds (sulphur – sulphur bonds)

- Suplhur – sulphur covalent bonds are formed between cysteine reisidues



Summary:



- Temperature and pH can affect hydrogen bonding, ionic bonds and disulphide bonds which in turn will change the shape of the protein or polypeptide.

Questions:

1) Name and give an example of the types of bonds that:
   a. Are responsible for the primary protein structure

   b. Are responsible for the secondary structure

   c. Are responsible for the tertiary structure (all 4 types)

Enzymes

- These are proteins that catalyse chemical reactions in living things.
- They often contain non – protein components which are called co – factors.
- These **co – factors** are usually small **organic molecules** or **metal ions**.

Stereospecific nature of enzymes:

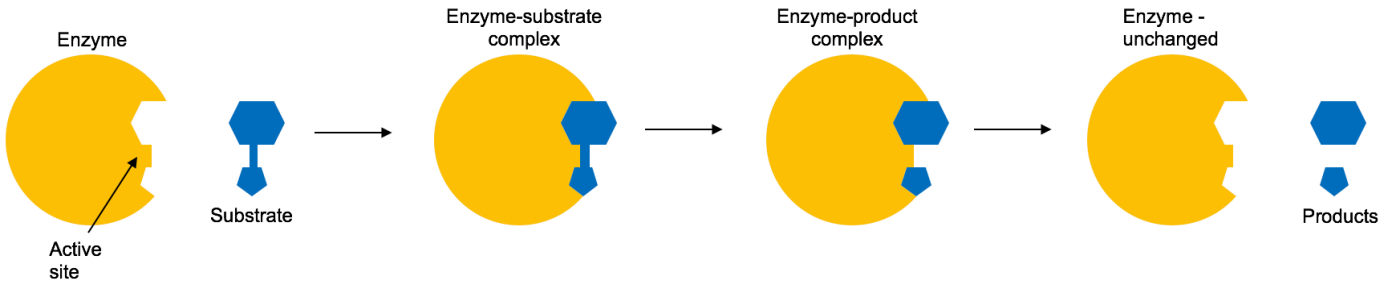
- This means that they only react with specific substances which are called **substrates**.

Active site:

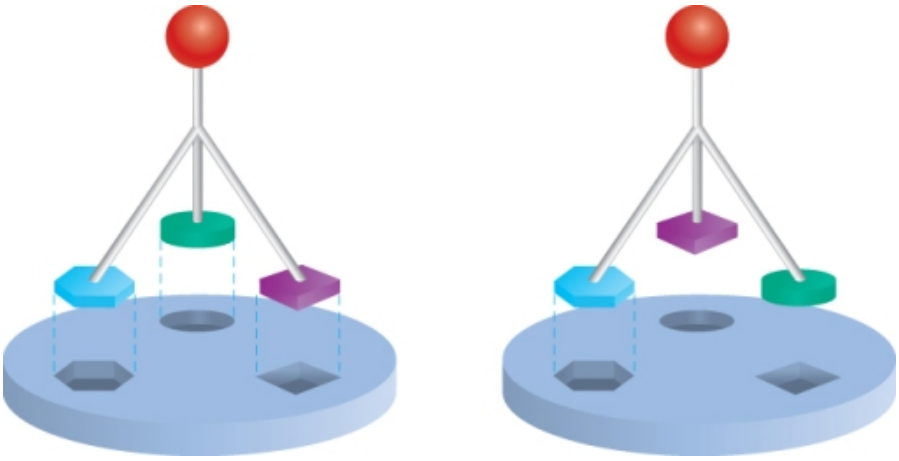
- This is the area within the enzyme where the catalytic activity takes place.
- Enzymes have a definitive shape due to the structure of the protein (earlier).
- The shape contains a hollow which has a specific shape due to its structure (earlier).
- Only molecules that exactly fit this shape will fit into the hollow or active site.

Substrates:

- These are the molecules that exactly fit in the active site to be catalysed:



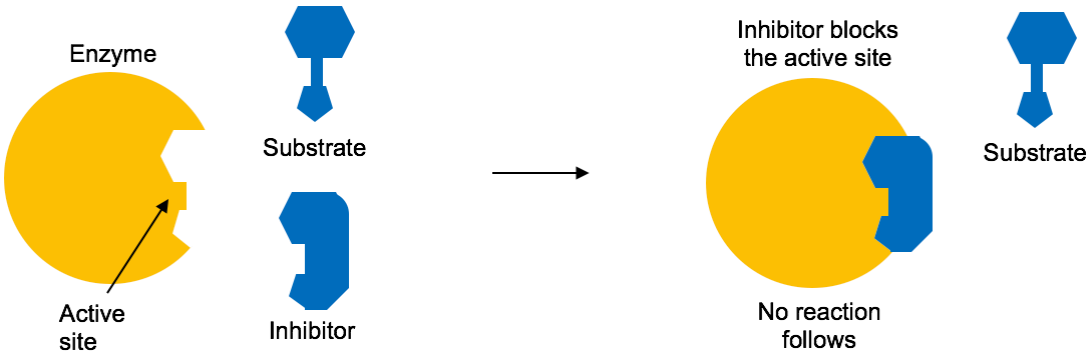
- This is the 'lock and key' mechanism:
- The substrate, like a key, must fit exactly in the lock, the active site.
- Within this hollow there are amino acid residue R groups.
- These interact with the substrate holding it in place.
- The activation energy is lowered / bonds in the substrate are weakened.
- Catalysing the reaction.
- They are so stereospecific that enzymes only catalyse **one enantiomer of optical isomers**:



- This is one way chemists can synthesise just one of the enantiomers in chiral drug synthesis – avoiding undesirable side effects.

Inhibitors:

- These are molecules with a similar shape to the substrate.
- They compete with the substrate for the active site.
- The difference is that they block the active site and no reaction follows.



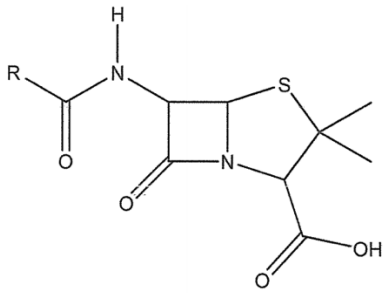
- The level of inhibition depends upon 2 factors:

  1) The relative concentration of the inhibitor : substrate:
     - High concentration of inhibitor : substrate → Little substrate gets to the active site.
     - Low concentration of inhibitor : substrate → More substrate gets to the active site.

  2) How strongly the inhibitor bonds to the active site

Drugs:

- Many reactions in living systems work by enzymes catalysing the reaction.
- Many drugs work using inhibitors to stop an enzyme working.
- Penicillin works in this way:



penicillin

- It blocks the active site in the enzyme that bacteria use to build cell walls.
- The bacteria cells effectively 'burst'
- This was discovered by accident by Alexander Fleming when clearing out some old bacteria plates.
- On an old bacterium plate, he noticed the bacteria didn't grow near some mould.

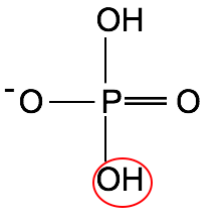
Drug design:

- Historically enzyme inhibitor drugs were produced by trial and error.
- Once the shape of a molecule is found to fit the active site, it is developed further to see if small changes in the shape would make the molecule a more effective inhibitor.
- Modern day techniques use computer modelling.
- This models the shape of the enzyme and inhibitor and looks at how changes in shapes of the inhibitor may increase its effectivity.
- These are then tested in the lab.

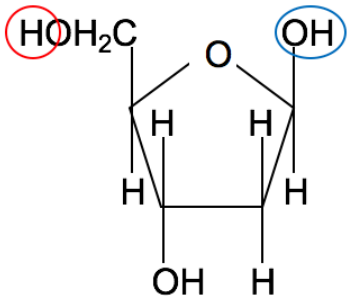
DNA:

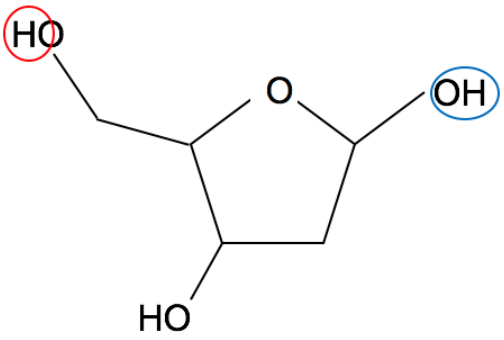
- **D**eoxyribo**n**ucleic **a**cid: Contains all the genetic information of living organisms.
- It is made up of 3 components:

1) Phosphate group:

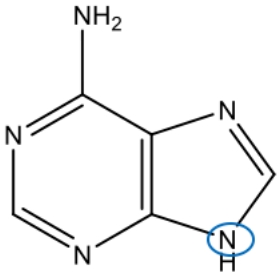


2) Pentose sugar – 2 deoxyribose:

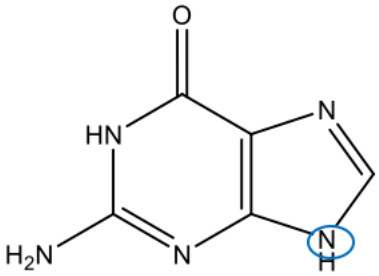




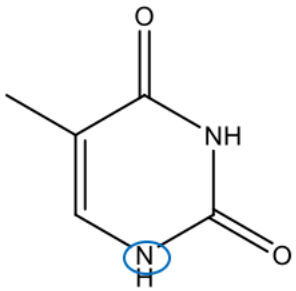
3) Base:



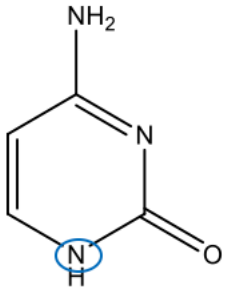
Adenine



Guanine



Thymine

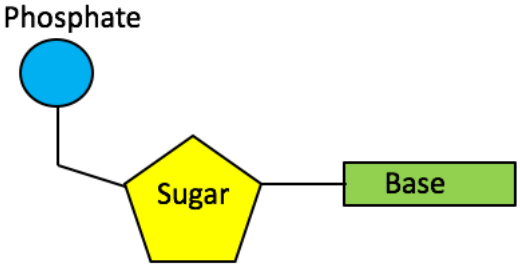


Cytosine

- The circled N are the atoms that form a bond to the 2-deoxyribose molecule.
- The red circles highlight where these molecules join each other at.

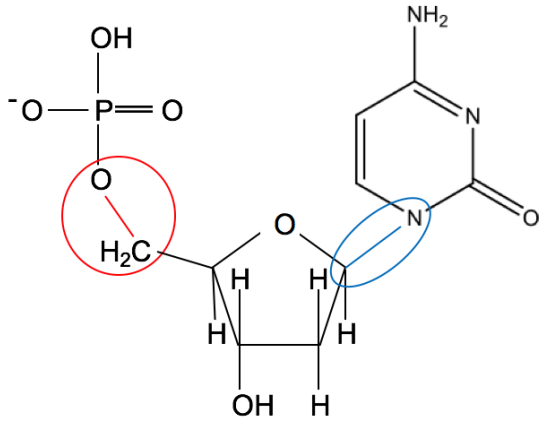
Nucleotides:

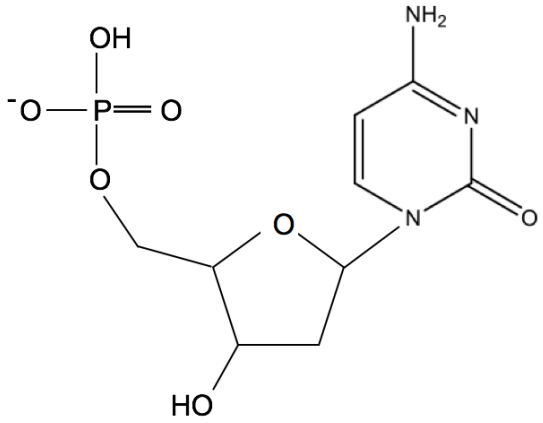
- These are formed from these 3 components joined together.
- A simplified model:



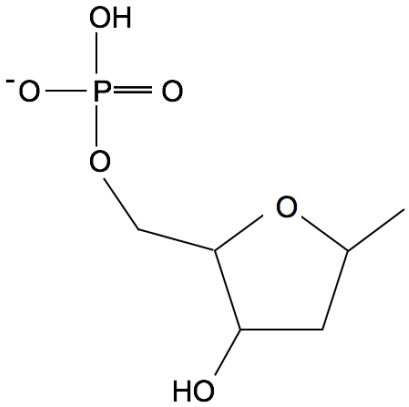
- This simplifies structure allows you to work out the structures of the nucleotides:

1) Cytosine nucleotide:

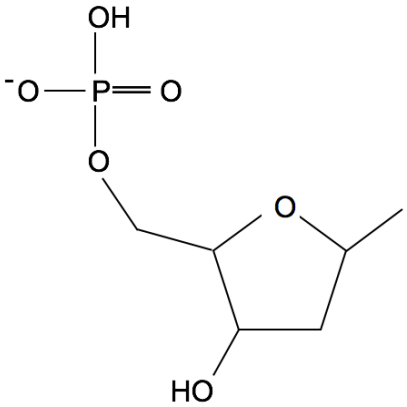




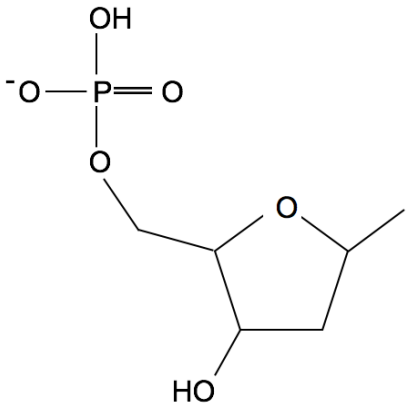
2) Thymine nucleotide:



3) Adenine nucleotide

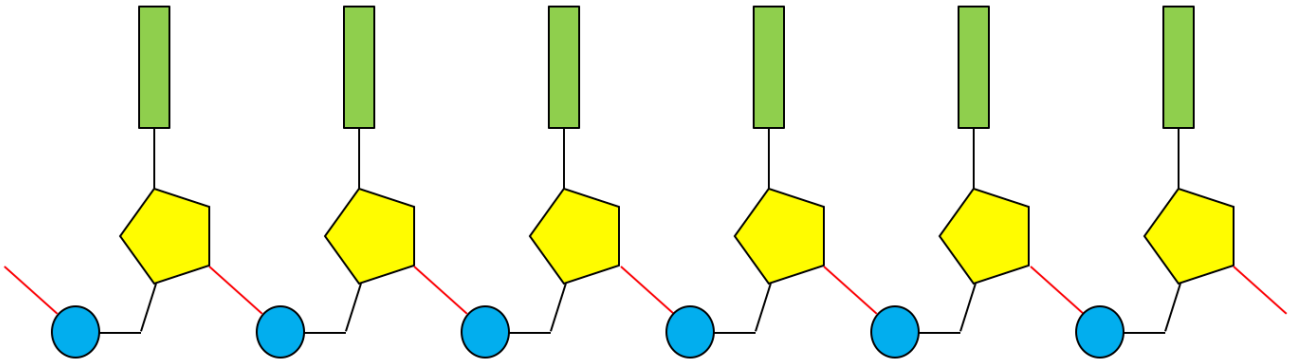


4) Guanine nucleotide:

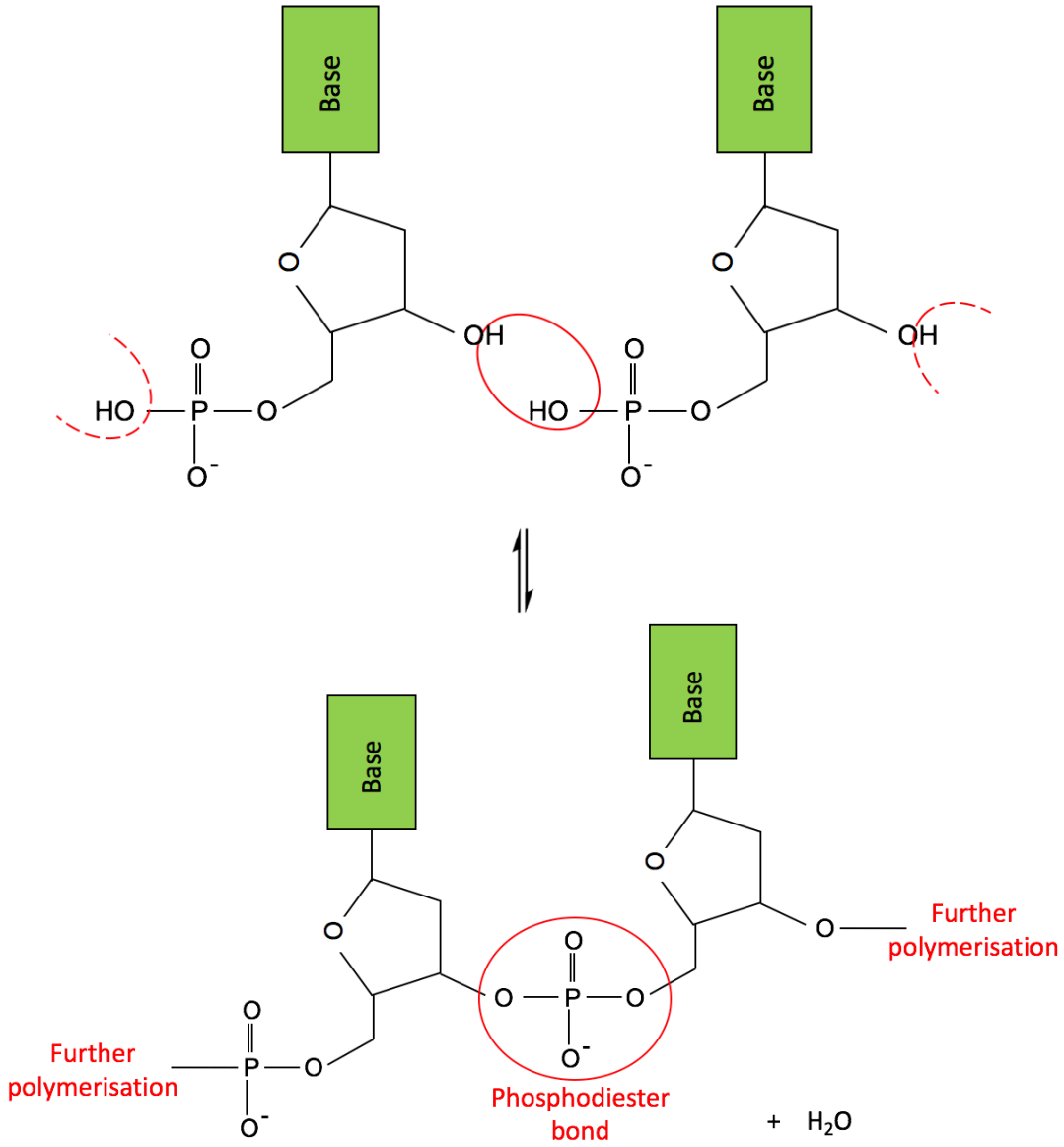


Polynucleotides:

- The nucleotides, above, join together to form a polynucleotide.
- They join by forming a covalent bond between the phosphate on one nucleotide and the sugar of another.



- It is a condensation polymerisation reaction where a water molecule is lost.
- A phosphodiester bond (linkage) is formed
- It makes a **sugar – phosphate** backbone along the chain
- The bases are attached to the sugar on the back bone:



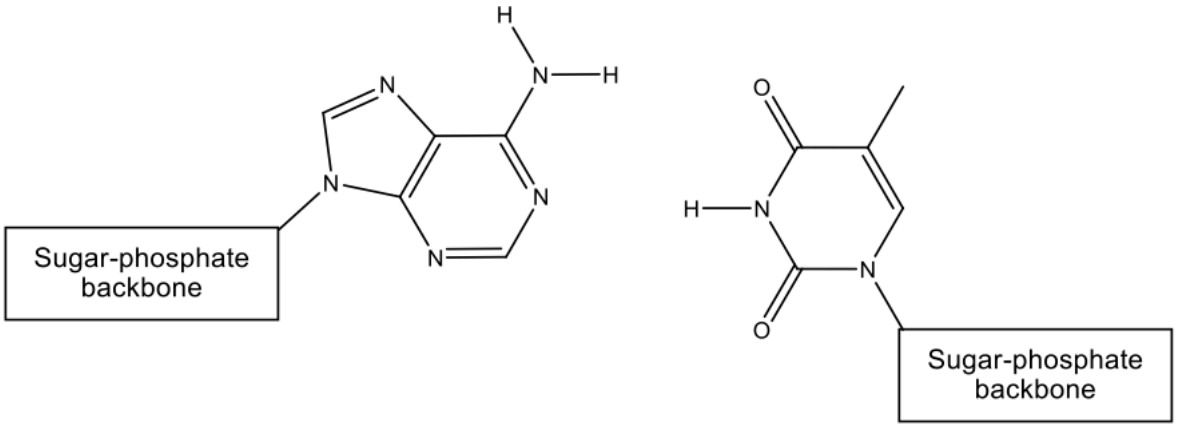
- This polynucleotide chain is one of the 2 strands that make up DNA

DNA structure:

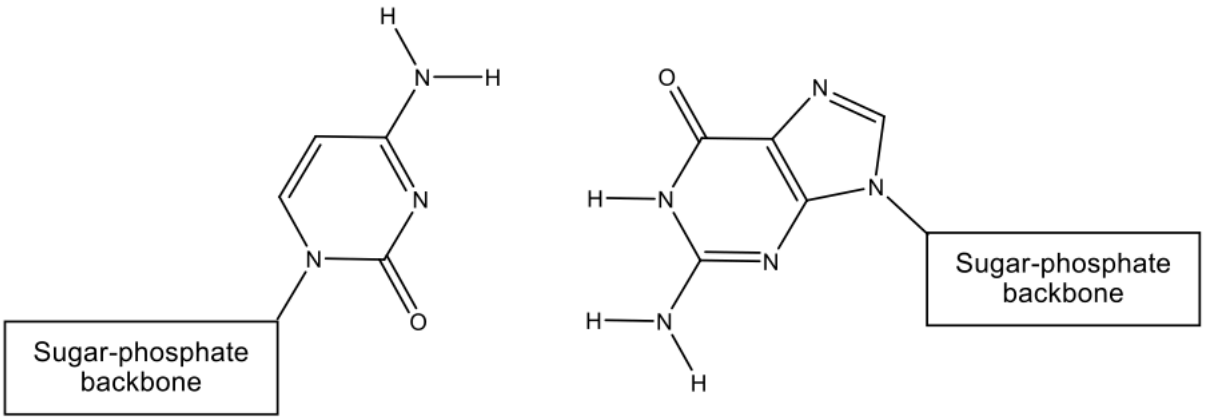
- DNA is formed from 2 polynucleotide strands that form a **double helix**.
- The 2 polynucleotides are held together by **hydrogen bonds** between the bases:
- Each base can only hydrogen bond with one particular base.

- Draw the Hydrogen bonds that exist between these base pairs:

Adenine–Thymine



Cytosine– Guanine



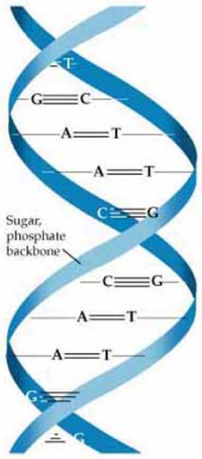
- An easy way to remember this is:

AT — **G**reenhead **C**ollege

Adenine with **Thymine** — **Guanine** with **Cytosine**

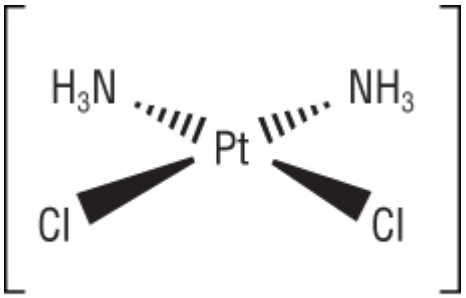
The structure:

- The base pairs always line up in a polynucleotide sequence giving the **complementary strand**.
    - Only **adenine** has the right atoms in the right place to hydrogen bond with **thymine**
        - 2 hydrogen bonds between the pairs.
    - Only **guanine** has the right atoms in the right place to hydrogen bond with **cytosine**
        - 3 hydrogen bonds between the pairs.



- Any other pairing would not have the right atoms the right distance or alignment to form these hydrogen bonds.
- The structure must twist for the atoms in the bases to align → **double helix structure**.

Anti cancer drugs: Cisplatin:

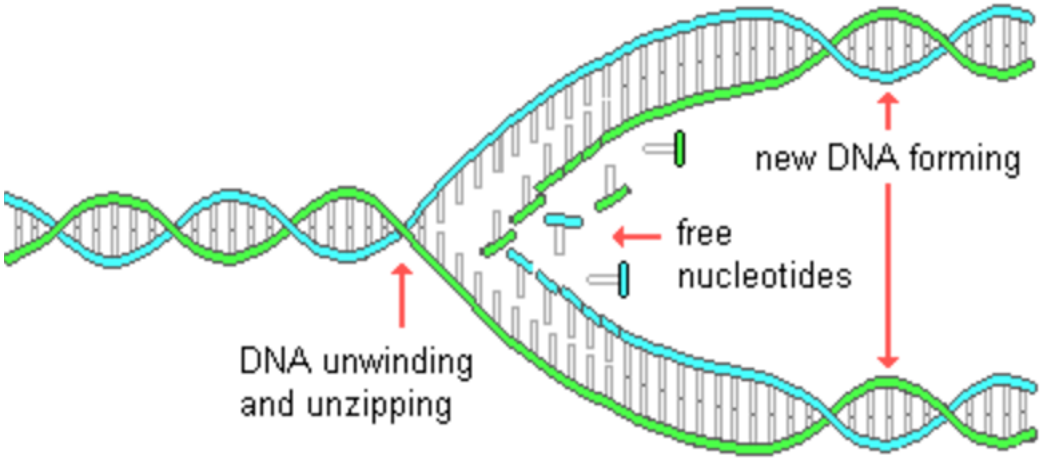


Cis-platin

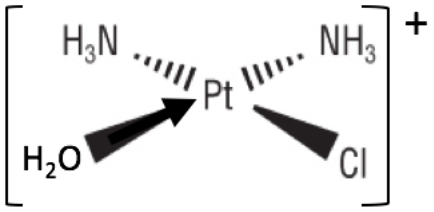
- Cancers are cells in the body that divide and reproduce uncontrollably to form tumours.
- To divide and reproduce, the DNA strands (polynucleotides) must unwind to allow the complimentary base pairs to align forming an identical DNA strand
- Cisplatin is a drug that interferes with this reproductive procedure stopping reproduction.

DNA replication:

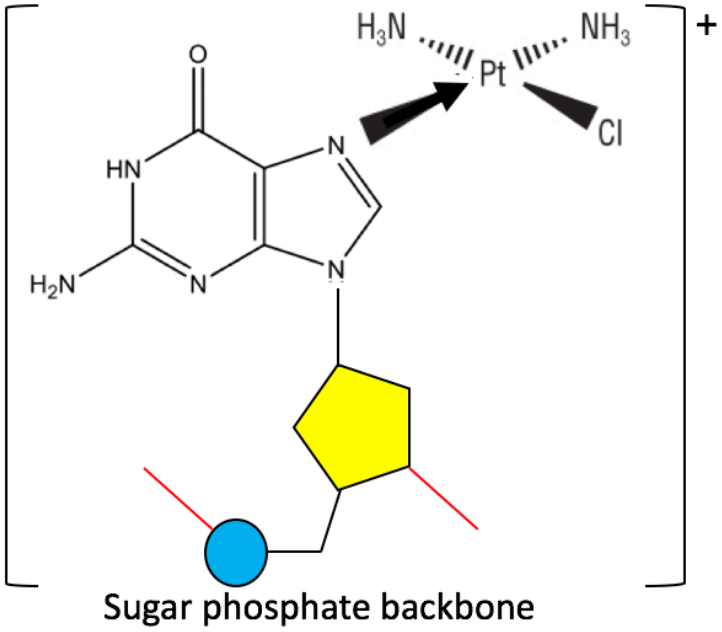
- As the DNA strands (polynucleotides) unwind, free nucleotides align with their base pair on the unwound strand (polynucleotide).
- A new strand (polynucleotide) forms against the original unwound strand producing a new double helix of DNA – Replication.



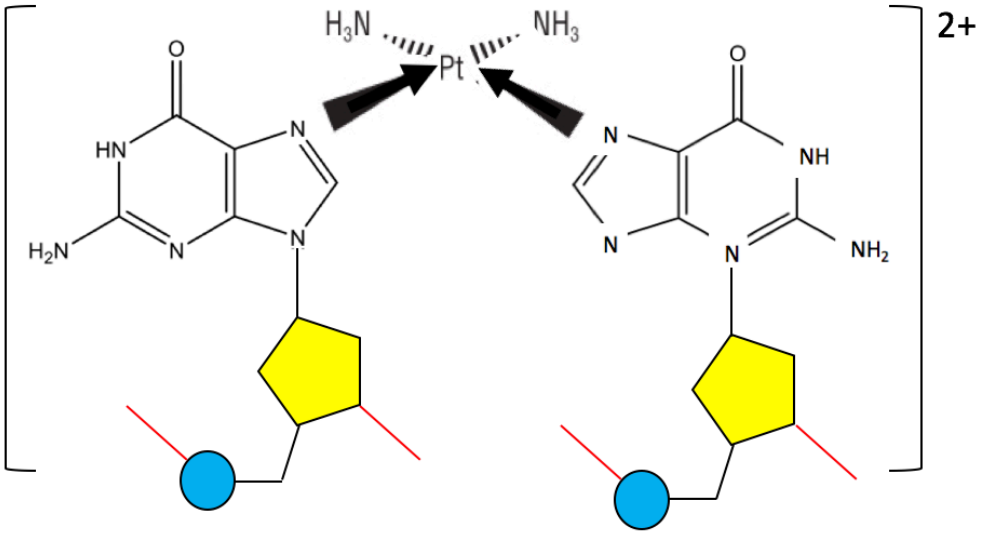
- As Cisplatin enters the cell it becomes hydrolysed as [Cl^-] within the cell is lower:



- An N in the guanine base forms a dative covalent bond with the Pt replacing the water ligand (ligand substitution).



- A second N from a neighbouring guanine molecule can replace a second chloride.
- This second guanine could be on the same strand or the opposite strand (polynucleotide).



- The cisplatin by binding to the DNA blocks replication.
- When bound to the DNA strand(s) it causes a kink.
- This means they can't unwind and be replicated.

Side effects:

- Cisplatin will bind to the DNA in normal cells that replicate quickly as well as cancer cells.
- These cells include: Hair, nails, blood
- This causes side effects:
  - Hair loss
  - Supress immune system
  - Kidney damage

- These side effects can be lessened by:
  - lower doses
  - Direct delivery to tumour

Balance:

- The benefits and side effects of these drugs are always considered before a license is granted.
- Doctors will always discuss the possible side effects of these drugs with the patient.

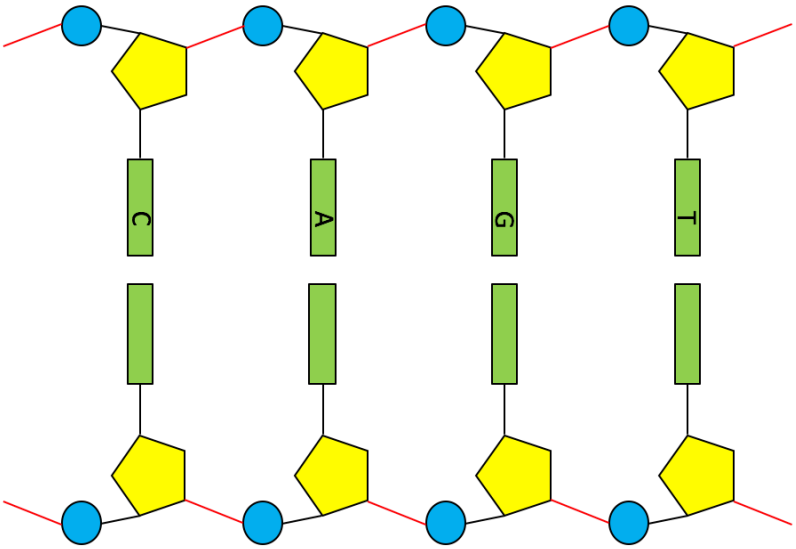
Questions:

1) Name and draw the structures of the 3 components that make up DNA:

2) Use your data sheet to combine these 3 components making a nucleotide using the guanine base:

3) This question is about polynucleotides.
   - What type of polymerisation makes a polynucleotide?
   - What is the name of the bond formed between the monomers?
   - Draw a section of 2 nucleotides with cytosine and thymine:

4) Place a letter in the complimentary bases to show how the 2 polynucleotides would pair.



5) Explain why these bases pair up. In your answer, you should include the number of hydrogen bonds between each pair:

6) Explain how DNA replicates:

7) How does Cisplatin disrupt the replication of cancer cells? In your answer include:
   - The base involved.
   - The type of reaction occurring between the base and Cisplatin
   - How this prevents replication