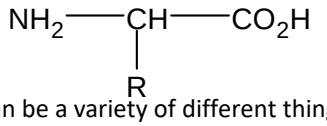
## 6.11 Amino Acids, Proteins and DNA

There are about 20 naturally occurring amino acids.
The human body can synthesise some of these amino acids. The ones the body cannot synthesise are called essential amino acids and these must be present in food. Naturally occurring amino acids almost all have the same general structure – called **α amino acid**

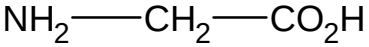
**General structure of an α amino acid**




The R group can be a variety of different things depending on what amino acid it is.

The alpha in '**α**' amino acid means both $NH_2$ and COOH groups are joined to the same second carbon atom.

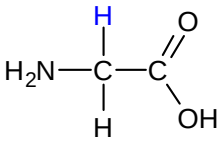
The simplest amino acid is glycine, where the R is an H



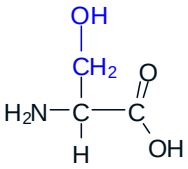

The structures of the naturally occurring amino acids are given on page 3 of this chapter.

**Naming amino acids**

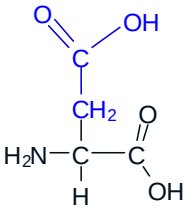
Each amino acid has a common name. The common name is sometimes abbreviated to a unique 3 letter code.
Some of the simpler amino acid can also be named using IUPAC organic nomenclature



Glycine (Gly)
(2-)aminoethanoic acid



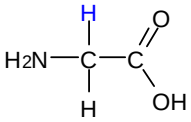
Serine (Ser)
2-amino-3-hydroxypropanoic acid



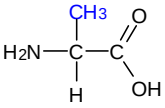
Aspartic acid (Asp)
2-aminobutanedioic acid

The R-groups can give the amino acids different properties.
The R group maybe

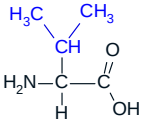
- Non polar and hydrophobic such as glycine, alanine , valine



glycine

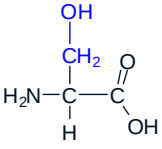


alanine

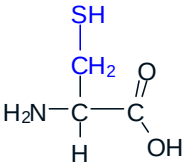


valine

- Polar, neutral and hydrophilic such as serine and cysteine

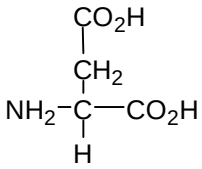


serine

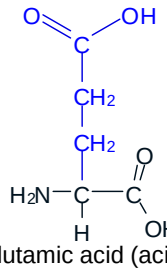


cysteine

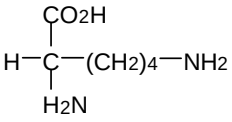
The R group can also have an extra carboxylic acid or an amine group. These are classed as acidic or basic (respectively) amino acids.



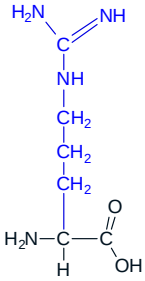
Aspartic acid (acidic)
2-aminobutanedioic acid



Glutamic acid (acidic)
2-aminopentanedioc acid



Lycine (basic)
2,6-diaminohexanoic acid

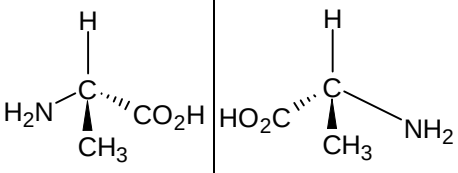


Arginine (basic)

**Optical Activity**

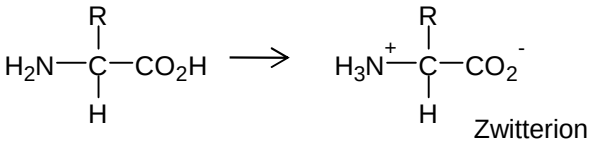
All amino acids, except glycine, are chiral because there are four different groups around the C

They rotate plane polarised light.



**Zwitterions**

The no charge form of an amino acid never occurs.
The amino acid exists as a dipolar zwitterion.



Zwitterion

Amino acids are often **solids**

The **ionic interaction** between zwitterions explains the relatively high melting points of amino acids as opposed to the weaker hydrogen bonding that would occur in the no charge form.

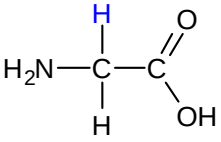
An amino acid exists as a zwitterion at a pH value called the **isoelectric point**

If the side R group of an amino acid contains an acidic of basic group then pH value of the **isoelectric point will be different**

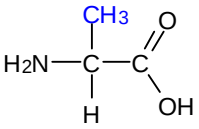
An amine group on the R group may make the isoelectric point be pH > 10
An carboxylic acid group on the R group may make the isoelectric point be pH < 3

## Non- Polar R groups

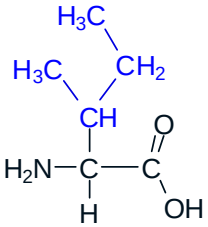
The pH of the isoelectric point is given underneath each one.



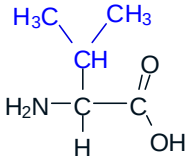
Glycine (gly)
6.0



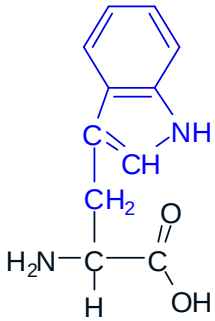
Alanine (Ala)
6.0



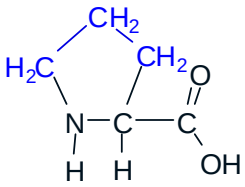
Isoleucine (Ile)
6.0



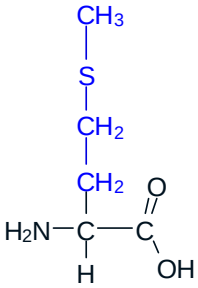
Valine (Val)
6.0



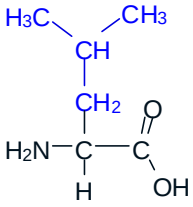
Tryptophan (Trp)
5.9



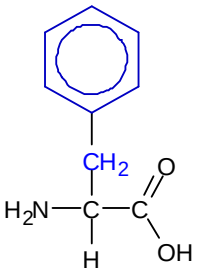
Proline (Pro)
6.3



Methionine (Met)
5.7

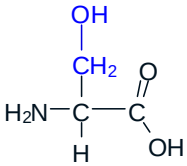


Leucine (Leu)
6.0

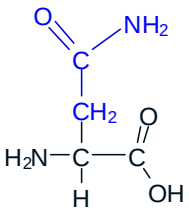


Phenylalanine (phe)
5.5

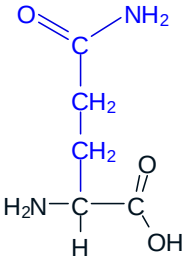
## Polar R groups



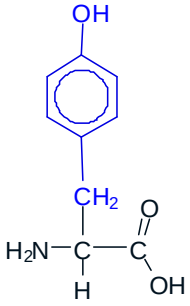
Serine (Ser)
5.7



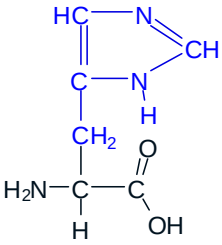
Asparagine (Asn)
5.4



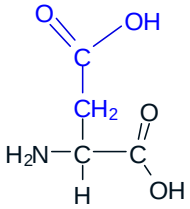
Glutamine (Gln)
5.7



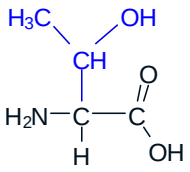
Tyrosine (Tyr)
5.7



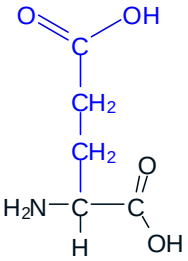
Histidine (His)
7.6



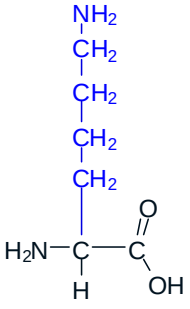
Aspartic acid (Asp)
2.8



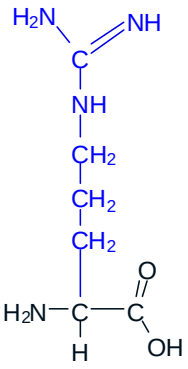
Threonine (Thr)
5.6



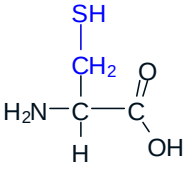
Glutamic acid (Glu)
3.2



Lysine (Lys)
9.7



Arginine (Arg)
10.8

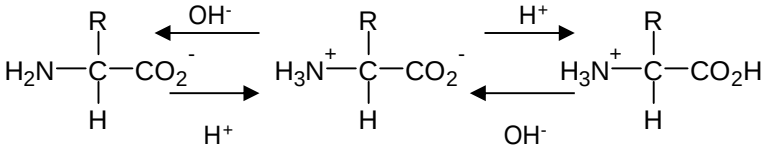


Cysteine (Cys)
5.1

**Acidity and Basicity**

Amino acids show amphoteric behaviour as they can react with both acids and bases.

The amine group is basic and the carboxylic acid group is acidic.




**Species in alkaline solution High pH**

**Species in neutral solution**

**Species in acidic solution Low pH**

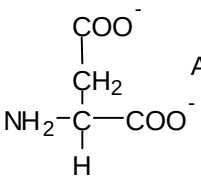
Amino acids act as weak buffers and will only gradually change pH if small amounts of acid or alkali are added to the amino acids.

**Example equations for a zwitter ion of glycine reacting with acid and alkali**

$^{+}NH_3\text{-}CH_2\text{-}CO_2^{-} + HCl \rightarrow Cl^{-}\ NH_3^{+}\text{-}CH_2\text{-}CO_2H$

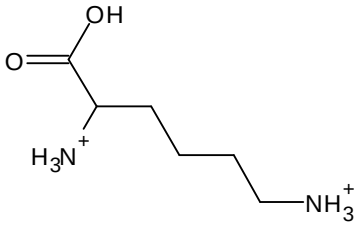
$^{+}NH_3\text{-}CH_2\text{-}CO_2^{-} + NaOH \rightarrow\ NH_2\text{-}CH_2\text{-}CO_2^{-}Na^{+} + H_2O$

The extra carboxylic acid or amine groups on the R group will also react and change form in alkaline and acid conditions



Aspartic acid in high pH

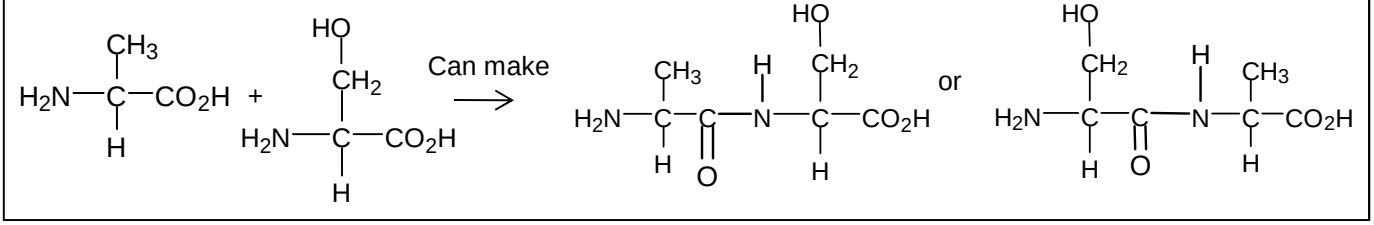
Skeletal formula of lycine in low pH



**Dipeptides**

Dipeptides are simple combination molecules of two amino acids with one amide (peptide) link.

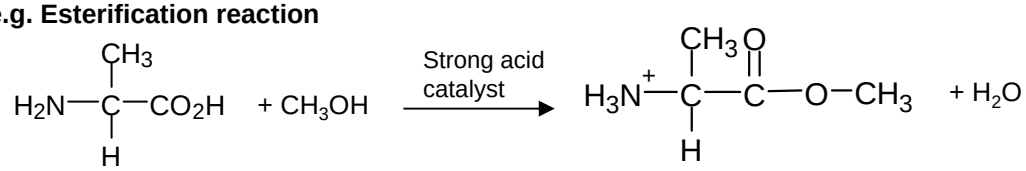
For any two different amino acids there are two possible combinations of the amino acids in the dipeptide.

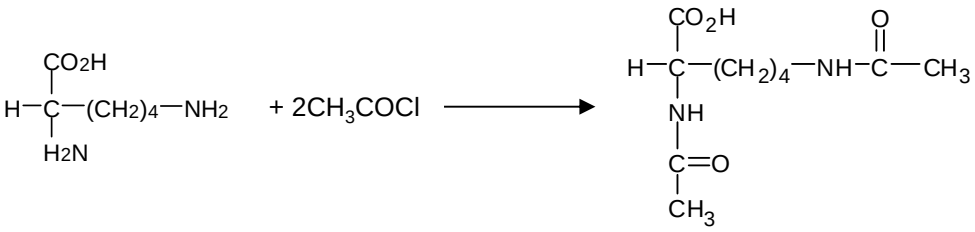



**Other reactions of amino acids**

The carboxylic acid group and amine group in amino acids can undergo the usual reactions of these functional groups met in earlier topics. Sometimes questions refer to these.

**e.g. Esterification reaction**



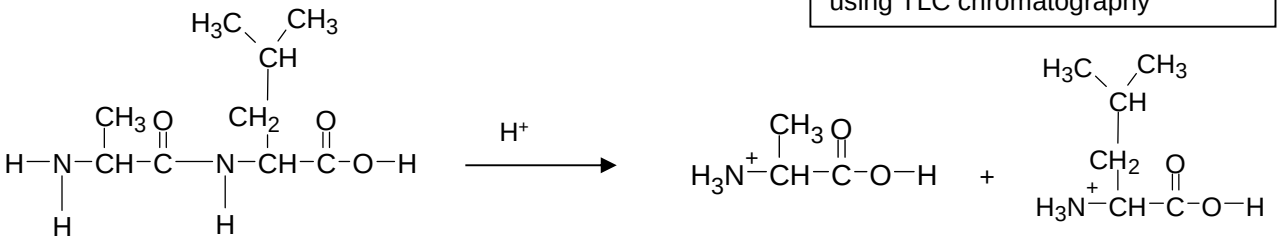



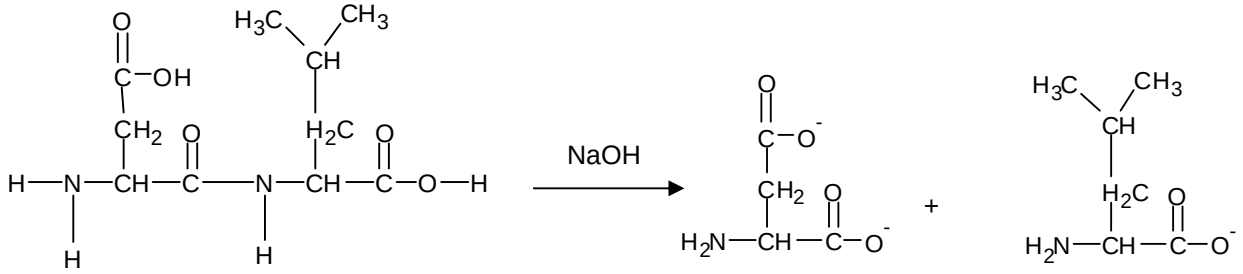
If the R group contains a amine or carboxylic acid then these will do the same reactions as the **α** amine and carboxylic groups

**Hydrolysis of di-peptides/proteins**

If proteins are heated with concentrated acid or alkali they can be hydrolysed and split back into their constituent amino acids.

The composition of the protein molecule may then be deduced by using TLC chromatography



**Chromatography of Amino Acids**

A mixture of amino acids can be separated by chromatography and identified from the amount they have moved.

Method: **Thin-layer chromatography**

a) **Wearing gloves**, draw a **pencil line** 1 cm above the bottom of a TLC plate and mark spots for each sample, equally spaced along line.
b) Use a capillary tube to add a **tiny drop** of each solution to a different spot and allow the plate to air dry.
c) Add solvent to a chamber or large beaker with a lid so that is no more than **1cm in depth**
d) Place the TLC plate into the chamber, **making sure that the level of the solvent is below the pencil line**. Replace the **lid to get a tight** seal.
e) When the level of the solvent **reaches about 1 cm from the top of the plate**, remove the plate and mark the solvent level with a pencil. Allow the plate to **dry in the fume cupboard.**
f) Spray paper with **ninhydrin** and put in oven
Draw around them lightly in pencil.
g) Calculate the Rf values of the observed spots.

Wear plastic gloves to prevent contamination from the hands to the plate

**pencil** line –will not dissolve in the solvent

**tiny** drop – too big a drop will cause different spots to merge

**Depth** of solvent– if the solvent is too deep it will dissolve the sample spots from the plate

**lid**– to prevent evaporation of toxic solvent

Will get more accurate results if the solvent is allowed to rise to near the top of the plate but the Rf value can be calculated if the solvent front does not reach the top of the plate

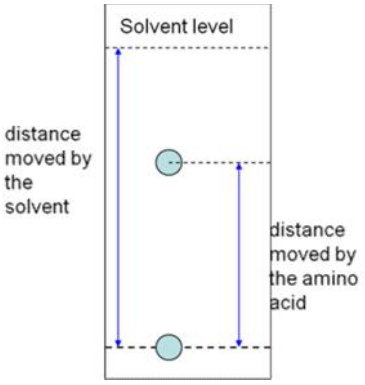
dry in **a fume** cupboard as the solvent is toxic

If ninhydrin is sprayed on an amino acid and then heated for 10 minutes then red to blue spots appear.
This is done because amino acids are transparent and cannot be seen.

$R_f$ value = $\frac{\text{distance moved by amino acid}}{\text{distance moved by the solvent}}$

Measure how far each spot travels relative to the solvent front and calculate the *R*f value.
Each amino acid has its own $R_f$ value
Compare *R*f values to those for known substances.

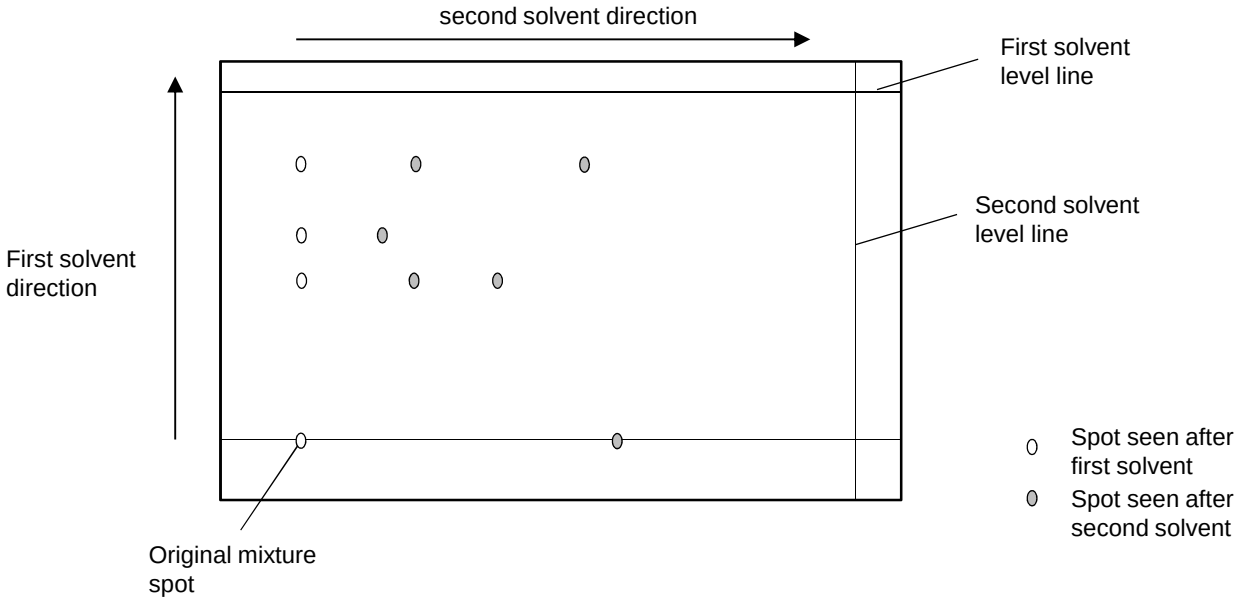
Some substances won't separate because similar compounds have similar *R*f values. So some spots may contain more than one compound



## Two directional chromatography

In order to separate a complex mixture that has components of different solubility in solvents, it may be necessary to do chromatography with two different solvents.

A spot of the mixture on a TLC plate is first separated with one solvent.
Then the TLC plate is rotated 90$^{o}$ and the plate is placed in a second solvent for a second separation to take place

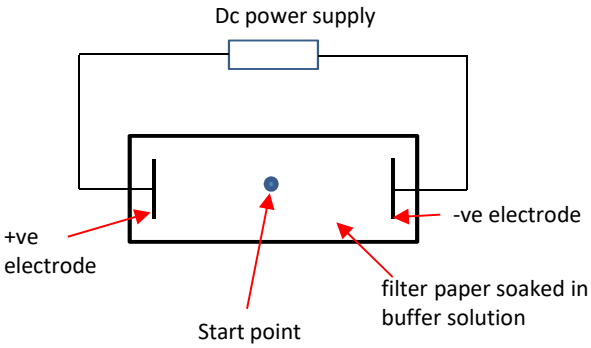



In total this mixture has 6 different components.

This process would be done if components in the mixture have the same Rf value or if some components are not soluble in the first solvent

**Electrophoresis of Amino Acids**

A mixture of amino acids can be separated by electrophoresis



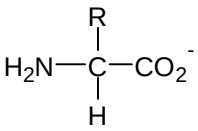

The mixture of amino acids is put at the start point.
A gel or filter paper soaked in a buffer solution is used.
The pH of the buffer solution makes a big difference to the results.

Amino acids are transparent and cannot be seen.
If ninhydrin is sprayed on an amino acid and then heated for 10 minutes then red to blue spots appear.

Different amino acids will move by different amounts depending on their size, charge and the pH of the buffer solution.

## If buffer is alkaline pH>7

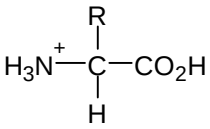
In alkaline conditions the amino acids will be negatively charged and therefore will move towards the positive plate



Amino acids with **smaller sized R** groups will move **further towards** the positive plate.
If the R group contains a $-CO_2H$ group (such as aspartic acid) then as it will have two $CO_2^-$ groups. It will move further than if it was a similar sized neutral R group.

## If buffer is acidic pH<7

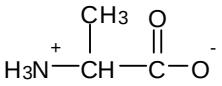
In acidic conditions the opposite will occur. The amino acids will be positively charged and therefore will move towards the negative plate



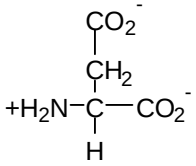
Amino acids with **smaller sized R** groups will move **further towards** the negative plate.
If the R group contains a $-NH_2$ group (such as lysine) then as it will have two $NH_3^+$ groups. It will move further than if it was a similar sized neutral R group.

## If buffer is neutral pH=7

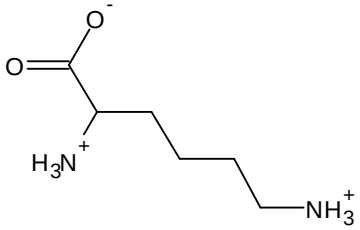
If an amino acid has a non-polar R group, it will exist in zwitterion form and it will not move.



If an amino acid has a R group that contains a $-CO_2H$ group (such as aspartic acid) ) then as it will have two $CO_2^-$ groups. It will move towards the positive plate



If an amino acid has a R group that contains a $-NH_2$ group (such as lysine) then as it will have two $NH_3^+$ groups. It will move towards the negative plate

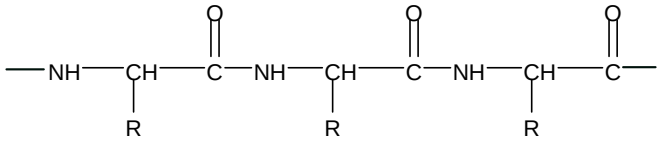


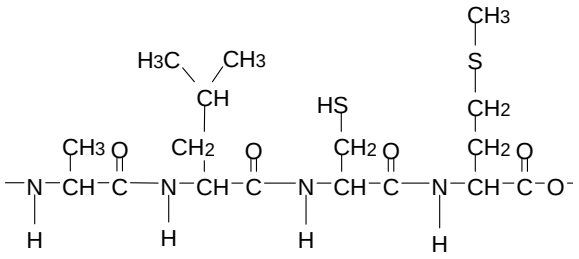
## Proteins

### Primary Structure of Proteins

Proteins are polymers made from combinations of amino acids.
The amino acids are linked by peptide links, which are the amide functional group.

The primary structure of proteins is the sequence of the 20 different naturally occurring amino acids joined together by condensation reactions with peptide links





More commonly the sequence of amino acids in a protein is written using the short codes e.g.
-Ala-Val-Cys-Met-

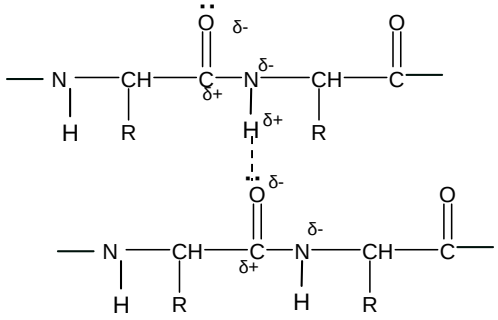
### Secondary Structure of a Protein.

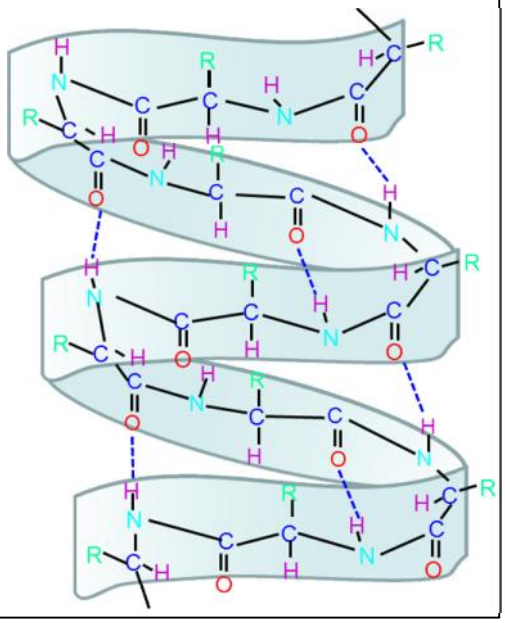
The primary chain can folded in two ways. These are called the secondary structures and are held together by hydrogen bonding.

#### Secondary Structure: α-helix

The 3D arrangement of amino acids with the polypeptide chain in a corkscrew shape is held in place by Hydrogen bonds between the H of $-N^{\delta-}-H^{\delta+}$ group and the –O of $C^{\delta+}=O^{\delta-}$ of the fourth amino acid along the chain

The R-groups on the amino acids are all pointed to the outside of the helix



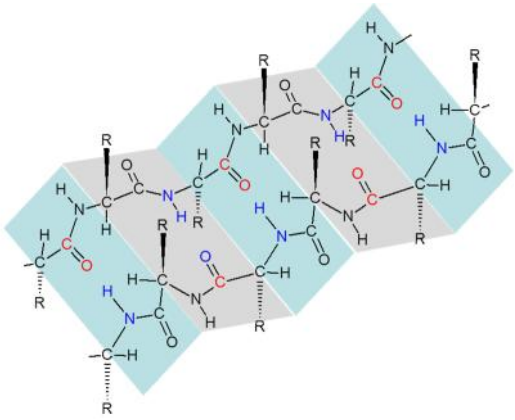


#### Secondary Structure: β-Pleated Sheet Structure of Proteins

The secondary structure can also take the form of a β–pleated sheets

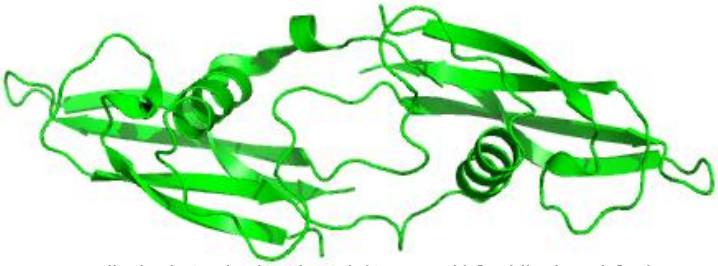
The protein chain folds into parallel strands side by side

The protein chain is held into a the pleated shape by Hydrogen bonds between the H of –N-H group and the –O of C=O of the amino acid much further along the chain in the parallel region .



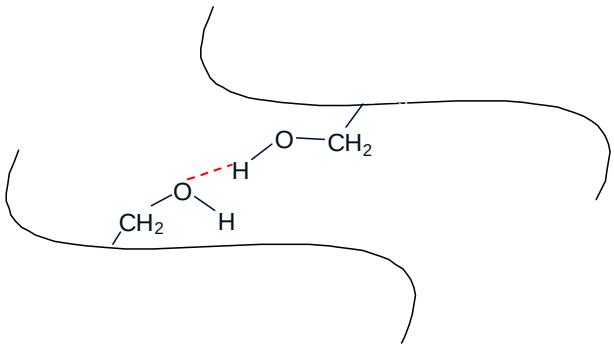
## Tertiary Structures of Proteins

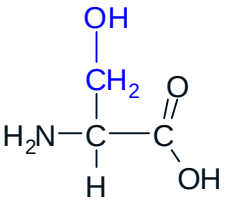
The tertiary structure is the folding of the secondary structure into more complex shapes. It is held in place by interactions between the R- side groups in more distant amino acids . These can be a variety of interactions including hydrogen bonding, sulfur-sulfur bonds and ionic interactions



By Elizabeth Speltz (SpeltzEB) (Own work) [Public domain], via Wikimedia Commons

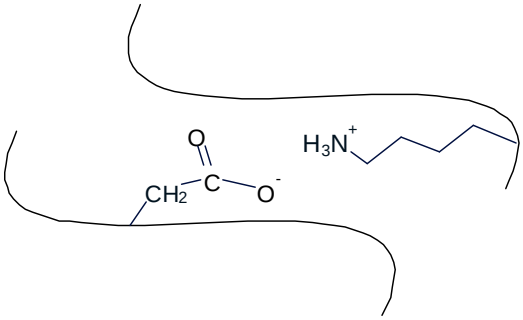
**Hydrogen bonds**



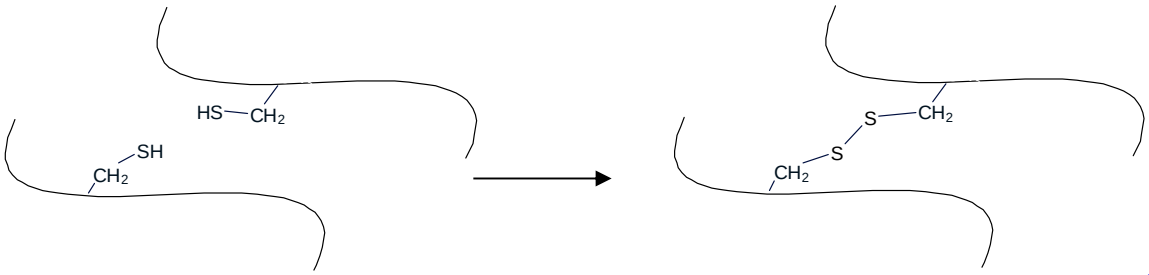

Hydrogen bonds could form between two serine side chains in different parts of the folded chain. (Other amino acids chains can also hydrogen bond)

**ionic interactions**

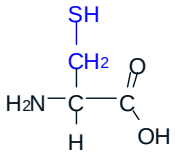



Ionic interactions could form between acidic amino acids such as aspartic acid and basic amino acids such as lysine. There is a transfer of a hydrogen ion from the -COOH to the -$NH_2$ group to form zwitterions just as in simple amino acids.

**Sulfur bridges**




If two cysteine side chains end up near each other due to folding in the protein chain, they can react to form a ***sulfur bridge***, which is a covalent bond.




cysteine (cys)

## Quaternary Structures of Proteins

In some proteins, two or more different protein polypeptides chains are held together into a quaternary structure . Haemoglobin is a protein that contains four different polypeptide chains.

## Fibrous and globular proteins

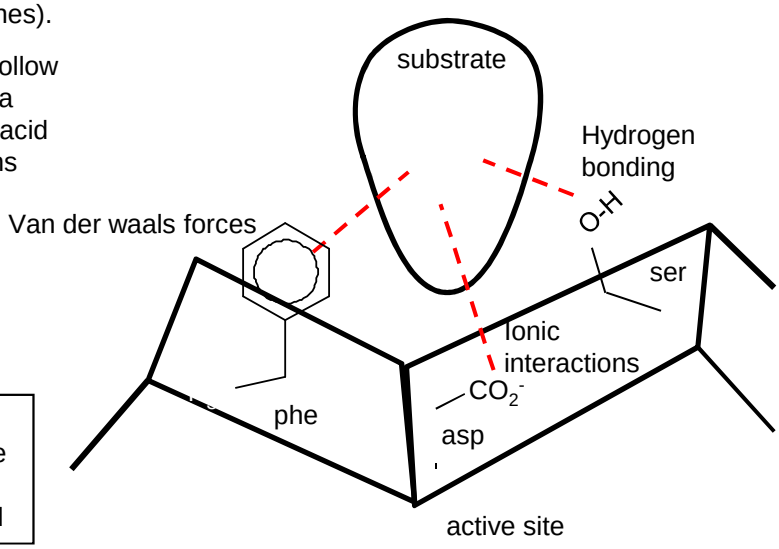
Most proteins in natural organisms can be categorised as Fibrous or Globular. Fibrous proteins have long molecules which has many cross links between the chains. Fibrous proteins form structures such as muscle fibres. Globular proteins are smaller, with chains coiled together to give a round, compact structure.

## Enzymes

Enzymes are proteins (generally globular ones).

The active site of an enzyme is usually a hollow in the globular protein structure into which a substrate molecule can bond to the amino acid side chains through a variety of interactions including

- Hydrogen bonding
- Van der waals forces
- Permanent dipole forces
- Ionic interactions



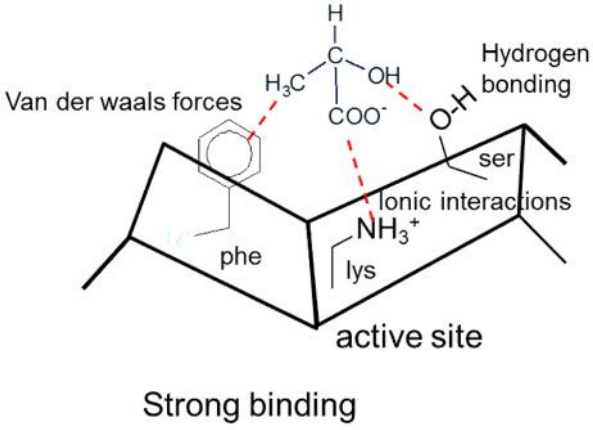

When the enzyme bonds to the active site it is called and enzyme-substrate complex

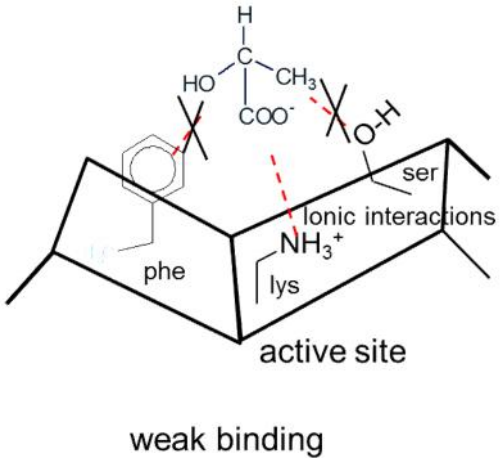
The interactions need to be strong enough to hold the substrate for long enough for the enzyme catalysed reaction to occur but weak enough for the product to be released

Only substrate molecules with the right shape and correct positions of functional groups will fit and bind to the active site- called the **lock and key hypothesis**

**stereospecific active site**

If the substrate is chiral then its likely that only one enantiomer will fit in the enzyme and so only one isomer will be catalysed




Inhibitors play an important role in regulating the activities of enzymes

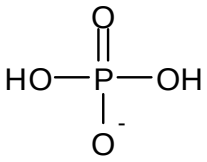
## Drugs as Enzyme Inhibitors

Many drugs act as an enzyme inhibitor by blocking the active site.
The inhibitor will often bind to the active site strongly, so stopping the substrate attaching to the enzyme. These are called competitive inhibitors.
Another type of inhibitors are called non-competitive. These attach elsewhere on the enzyme but in doing so can change the shape of the active site which also stops its effectiveness.
Computers can be used to help design such drugs

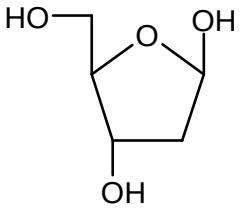
## DNA

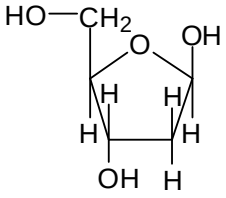
DNA stores genetic information in a cell. It controls inheritance as its structure is copied and passed on from one generation to the next. It encodes the information to produce all the proteins in an organism.

**Key molecules in DNA**



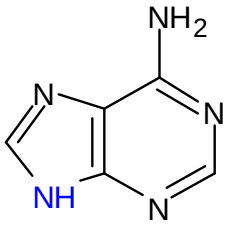
phosphate ion



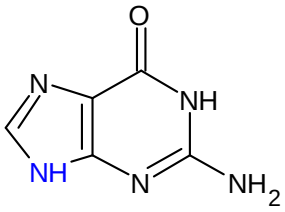


2-deoxyribose (a pentose sugar)

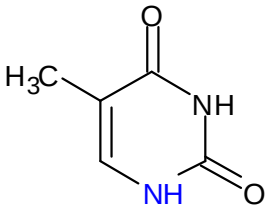
The 4 bases



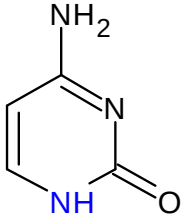
Adenine (A)



Guanine (G)



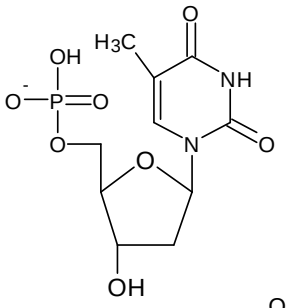
Thymine (T)

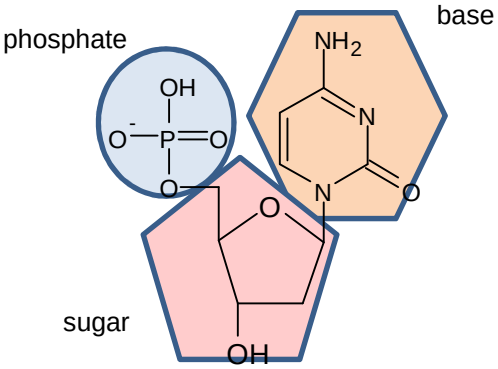


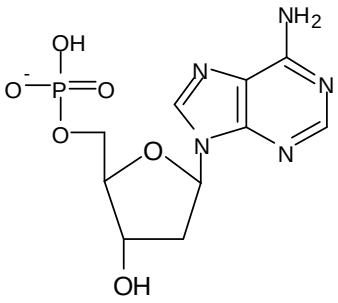
Cytosine (C)

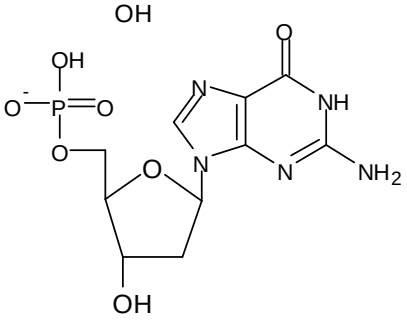
**Nucleotides**

A nucleotide is made up from a phosphate ion bonded to 2-deoxyribose which is in turn bonded to one of the four bases adenine, cytosine, guanine and thymine



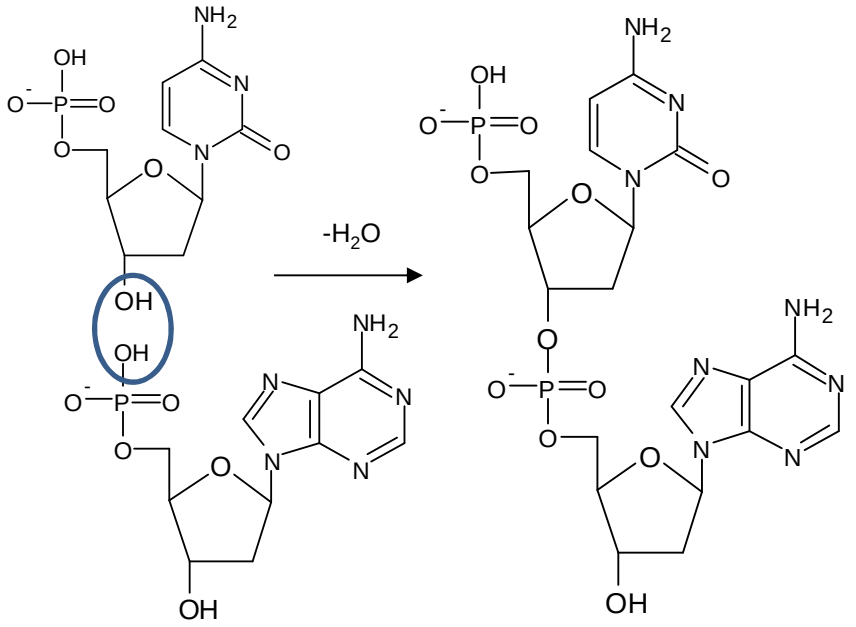




Although the structures will be given in the data sheet you need to learn which atoms on the base joins on to the sugar and how the sugar attaches to the phosphate ions

## Sugar- phosphate chain




A single strand of DNA (deoxyribonucleic acid) is a polymer of nucleotides linked by covalent bonds between the phosphate group of one nucleotide and the 2-deoxyribose of another nucleotide.

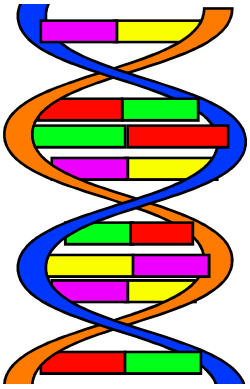
This results in a sugar-phosphate-sugar- phosphate polymer chain with bases attached to the sugars in the chain.

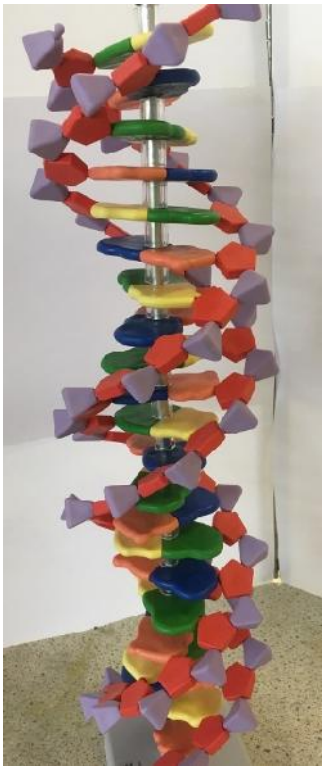
Carefully learn how these join together

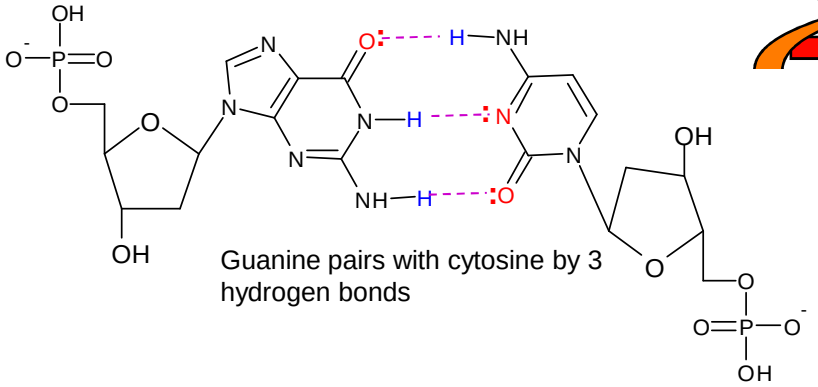
DNA exists as two complementary strands of the sugar phosphate polymer chain arranged in the form of a double helix.

**Complementary** means the two strands must have base sequences that match all A to T and C to G

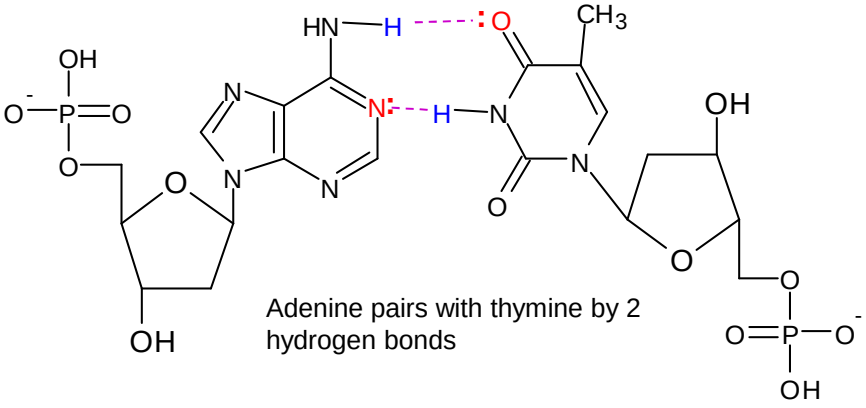
Hydrogen bonding between base pairs leads to the two complementary strands of DNA.







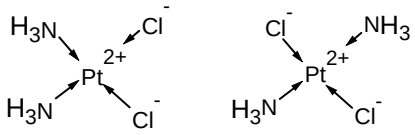

Guanine pairs with cytosine by 3 hydrogen bonds




Adenine pairs with thymine by 2 hydrogen bonds

## Cisplatin

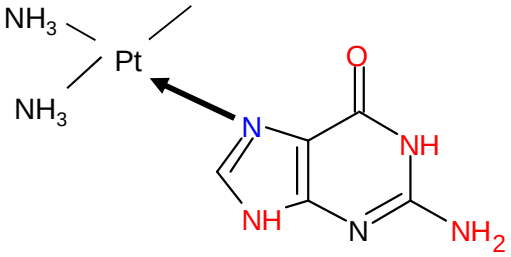
The Pt(II) complex cisplatin is used as an anticancer drug.



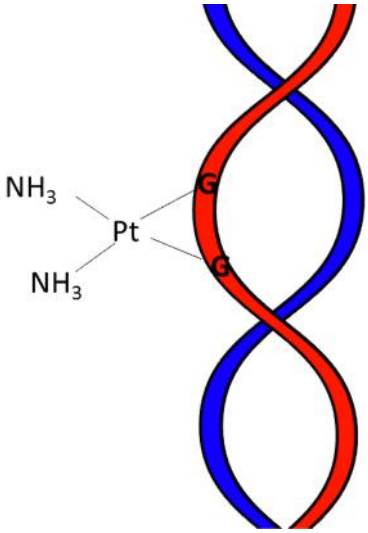

cisplatin transplatin

The cisplatin version only works as two chloride ions are displaced and the molecule joins on to the DNA. In doing this it stops the replication of cancerous cells.

Cisplatin prevents DNA replication in cancer cells by a ligand replacement reaction with DNA in which a dative covalent bond is formed between platinum and a nitrogen atom on guanine




The N and O atoms marked in red can't bond to cis-platin as they are involved in the bonding within the DNA




Cisplatin can also prevent the replication of healthy cells by bonding on to healthy DNA which may lead to unwanted side effects like hair loss. Society needs to assess the balance between the benefits and the adverse effects of drugs, such as the anticancer drug cisplatin.