**3.3.13 Amino acids, proteins & DNA**

**AS Link:**

**3.1.3 – Bonding (ionic bonding & forces between molecules)**

**3.3.1 – Introduction to organic chemistry (nomenclature)**

**A level Link:**

**3.2.5 – Transition metals (cisplatin metal complex)**

**3.3.9 – Carboxylic acids**

**3.3.11 – Amines**

**3.3.12 – Polymers**

**3.3.16 – Chromatography**

**Introduction**

**Amino acids** are the building blocks of **proteins**.

The term amino acid is used for substances that contain a **primary amine** group attached to the carbon adjacent to a **carboxylic acid** group, it is a **bifunctional compound**:

|  |  |
| --- | --- |
| RCH(NH2)COOH |  |

Naming amino acids

**Amines** uses the prefix **amino** and the position it is attached to the carbonyl functional group is sometimes called the **α-position** so a more complete name is for amino acids is **α-aminocarboxylic acid**. Other naming systems use the numerical position of the amine group so they are **2-amino acid** meaning the **amine group** is on the **second carbon** next to the carboxylic acid group.

***Task: Identify the chiral carbon***

Virtually all of the **20 naturally occurring amino acids** exist as **optical isomers**, they are negative enantiomers (D or L anticlockwise).

***Task: Write the systematic name, draw the structural & displayed formula and mark the chiral centre for the following amino acids:***

|  |  |  |  |
| --- | --- | --- | --- |
| **R =** | **Name** | **Systematic** | **Structural formula** |
| H | glycine | 2-aminoethanoic acid  or  α-aminoethanoic acid | HCH(NH2)COOH |
| CH3 | alanine | 2- aminopropanoic acid  or  α-aminopropanoic acid | CH3CH(NH2)COOH |
| CH(CH3)2 | valine | 2-amino-3-methylbutanoic acid  or  α-amino-3-methylbutanoic acid | CH(CH3)2CH(NH2)COOH |
| CH2COOH | aspartic acid | 2-aminobutanedioic acid  or  α-aminobutanedioic acid | CH2COOHCH(NH2)COOH |

There are different categories of amino acids:

|  |  |  |
| --- | --- | --- |
| Non-polar aliphatic R group |  | |
| Aromatic R group |  | |
| Polar uncharged R group |  | Cysteine has a thiol  group (-SH) |
| Positively charged R groups |  | Arginine has an imine group (-C=NH) |
| Negatively charged R groups |  | |

***Starter: 6.4 – Amino acids***

**Acid and base properties**

***Task: What do you know about the properties of the 2 functional groups?***

|  |  |
| --- | --- |
| **Amine group gains a proton**  – acts as a base: | **Carboxylic acid loses a proton**  – acts as an acid: |
|  | |

They are **amphoteric** – they have **both acidic and basic properties**; acting as **acids** due to the carboxylic acid groups and **bases** due to the amine group.

|  |  |
| --- | --- |
| Strongly **acidic** conditions:   * behaves as a **base** * amine group **protonated** * carboxylic acid is undissociated | Strongly **basic** conditions:   * behaves as an **acid** * carboxylic acid is **deprotonated** * amine group unchanged |

In living systems the pH is about 7.3, so amino acids never exists as an uncharged compound. They exist as **dipolar** **ions**, with both a **permanent positively** and **negatively charged group** so they have **no overall charge**.

|  |  |
| --- | --- |
|  | These are called **zwitterions** *(German for hybrid*).  They can exist as **dipolar ions in the solid state** too. |

|  |  |  |
| --- | --- | --- |
| Low pH  Acidic | Neutral  Zwitterion | High pH  Basic |
|  |  |  |
| amine is protonated  carboxylic acid is protonated | amine is protonated  carboxylic acid is deprotonated | amine is deprotonated  carboxylic acid is deprotonated |

Depending on the **pH** of the solution in which the amino acid is dissolved they can be **ionic** or **non-ionic**. It only exists as a **zwitterion** at the point when the pH is such that it has **no overall charge**; this is called the **isoelectric point**. The precise pH of this point varies with the amino acid.

***Task: Draw the structure of the amino acid at the pH shown***

|  |  |  |
| --- | --- | --- |
| pH 12 |  |  |
| pH 2 |  |  |
| pH 4 |  |  |
| pH 13 |  |  |

**A typical amino acid:**

* **white solid** at room temperature
* behaves like an **ionic** salt
* **high melting point** (~290oC) i.e. **evidence of ionic bonding**
* **dissolves** well in **water** but poorly in non-polar solvents

***Starter: 6.4 – Amino acids***

***Application: CGP234 PQ1***

***Fact recall: CGP234 Q1-2***

**Proteins**

|  |  |
| --- | --- |
| Amino acids are **monomer** that link together in a **condensation reaction** to form peptides with **reactive groups at either end**. Long chains of amino acids are called **polypeptides**, which form **proteins**.  Proteins and polypeptides are **broken down into amino acids** by **hydrolysis** reactions. |  |

**Amino acids join** **together** with an **amide bond** or **peptide link** (-CONH-) where the **amine** of one amino acid **joins** with the **carboxylic acid** of another.

|  |  |  |
| --- | --- | --- |
| An amide:  RCONH2 |  | Amide bond: |

Condensation reaction

|  |  |
| --- | --- |
| When amino acids join **water is eliminated** in a **condensation** reaction.  The **dipeptide** formed still has an **amine** and **carboxylic acid group** so it can form other peptide links producing longer chained molecules. **Proteins** (**polypeptides**) are **condensation polymers**. | ***Task: Show a dipeptide forming by condensation*** |
|  |

Hydrolysis

|  |  |
| --- | --- |
| Proteins can be **split into their amino acids** by a **hydrolysis** reaction, **strong acid** or a specific **enzyme** (**protease**) is needed. They are therefore **biodegradable**. |  |

Structure

**Proteins are large** complicated biological molecules; their structure is described in four levels.

* **Primary** – sequence of amino acids to form polypeptide
* **Secondary** – polypeptide chain bonded to itself with **intermolecular** forces to form a protein
* **Tertiary** – protein chain folded in a specific shape – **intramolecular** bonds between chains (disulphide bonds and ionic bonds between ionic R groups)
* **Quaternary** – more than one protein chain bonded together (*not in specification*)

|  |  |  |
| --- | --- | --- |
| **Structure** | **Description** | |
| **Primary** | This the particular **sequence of amino** **acids**, it can be short and then repeated or long with many different amino acids. It is held together by **covalent bonds** between the atoms so it is relatively **stable**; requiring harsh conditions to separate the amino acids. | |
|  |  |

|  |  |  |  |
| --- | --- | --- | --- |
| **Secondary** | Proteins have **complex shapes** the polypeptide can form **intermolecular** forces between polypeptide chains:   * **Van der Waals’ forces** * **Hydrogen bonds** between O on carboxylic acid and H on amine with another section of the chain   - C=O - - - - H-N-  Although **individually weak** the **hydrogen bonds** **provide stability** when there are lots of them. This gives proteins their specific **shape** which is vital to their **function**. However, they can be **easily disrupted** by **gentle heating** or **pH** and this **denatures** the protein **changing its shape**.  There are two distinct shapes: | | |
| **α-helix** | **β-pleated** | |
| This is the most common shape |  | |
| hair, wool | Silk | |
|  | | |
| **Tertiary** | The α-helix and β-pleated **protein chains** are often **coiled** and **folded** into a **three-dimensional shape** called the tertiary structure. This gives the protein its **unique shape** and so determines it **function**. These are held by **intramolecular** bonds:   * **Ionic** (between ionic COO- & NH3+ on R groups) * **Sulphur-sulphur bonds** called **disulphide** bonds or bridges (where sulphur atoms are present in the R groups)   There are 2 types of tertiary structure:  **Fibrous** bundles of polypeptide chains e.g. hair, wool – insoluble in water  **Globular** polypeptides folded into roughly spherical shape e.g. enzymes, haemoglobin – soluble in water | | |
| Disulphide bonding  An **amino acid that is part of a protein** chain is called a **residue**. **Cysteine** **residues** contain a **thiol** group **(-SH**) which can **lose its hydrogen** and form a **covalent** **disulphide bond (-S-S-**) with another thiol group that has also lost its hydrogen. These bonds link different parts of the protein chain and **provide stability** to the tertiary structure. | | |
|  | | | |
| **Quaternary**  *Not in specification* | Sometimes more than one protein chain is bonded to other protein chains to form a large protein structure. This is referred to as the quaternary structure. An example is haemoglobin which consists of four proteins bonded together with plus 4 haem groups containing an iron complex. | |  |

***Application: CGP238 PQ1-3***

***Fact recall: CGP238 Q1-4***

**Finding the structure of proteins**

|  |  |  |
| --- | --- | --- |
| **X-ray crystallography** *Not in specification*  This is one of many techniques are used to determine the secondary and tertiary structure of proteins. |  |  |

Chromatography

To **determine the primary sequence** of a protein it must be broken down into its amino acids, this involves **breaking the amide bonds** in a hydrolysis reaction. The protein is **refluxed in 6 mol dm-3 hydrochloric acid** (conc. H2SO4 or conc. NaOH could also be used). The mixture of amino acids can be identified by **paper** or **thin-layer chromatography** and **compared to known standards.** Another technique for separating and identifying the amino acids is **electrophoresis** which involves applying a current is across the mixture to separate the amino acids.

|  |  |
| --- | --- |
| Amino acids have **different R groups** which will have **different solubilities** in the same solvent. So they can be easily separated and identified using **thin-layer chromatography** (**TLC**). |  |

However, **amino acids are colourless** so a technique needs to be used that allows them to be visible. There are two options:

|  |  |
| --- | --- |
| **ninhydrin**  this turns the amino acids **purple** | **fluorescent dye**  shining **UV light** the amino acids to **glow** |
|  |  |
|  |  |

Identifying amino acids

|  |  |
| --- | --- |
| Once the chromatogram is developed the **distance the solvent travelled** and the **distance** the **amino acids travelled** can be measured. |  |

|  |  |
| --- | --- |
| Then the **retention factor** or **Rf****value** can be **calculated** and used to identify the amino acids by **comparing** to a table of **known Rf values**. Identical amino acids separated under the same conditions (same solid and mobile phase) will have the same Rf value. |  |

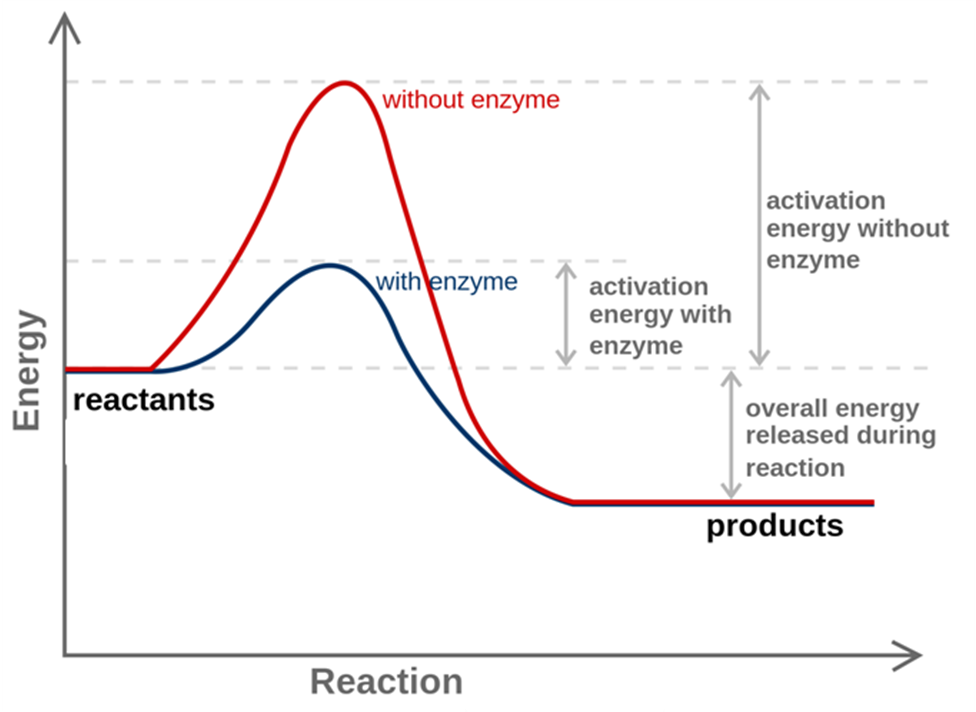
***Fact recall: CGP234 Q3-4***

**Enzymes**

Enzymes **speed up chemical reactions by acting as biological catalysts**; they speed up reactions by 1010 – which means if a reaction is completed in 1 minute with an enzyme it would take 300 years without it!

|  |  |
| --- | --- |
| They catalyse every reaction in living systems but are extremely specific. **Enzymes** are usually **globular proteins**, but some have non-protein components. Their shape has a cleft or crevice called the **active site**, where the reaction takes place. The **molecules** that enzymes **act upon** are called **substrates** and these **fit precisely into the active site** and **held in the right orientation** for the reaction. The **active site** can be so specific that many enzymes only catalyse reactions of **one pair of enantiomers**, they are **stereospecific**. This idea is often called the **lock and key hypothesis**. Now called **induced fit**. |  |

The actual situation is a little more complex. The substrate has to fit into the active site and **temporarily bond with intermolecular forces**; this forms an **enzyme-substrate complex**. These forces then **promote the movement of electrons** within the substrate that **lower the activation energy** for the reaction.



Inhibition of enzymes

Enzymes can be **denatured** by **changes** in **temperature** or **pH**; this **changes their shape** by **disrupting** the **hydrogen bonds** between the polypeptide chains. Living systems maintain an **optimum temperature** and **pH** for the enzymes to function. The optimum temperature is usually around 40oC; a higher temperate will denature the enzyme and a lower one will make the reaction too slow. Most enzymes work at around pH 7, e.g. catalase action in liver with hydrogen peroxide (catalyses the decomposition of hydrogen peroxide to water and oxygen. It is a very important enzyme in protecting the cell from oxidative damage by reactive oxygen species. Some operate at a lower or higher pH, e.g. digestive enzymes such as protease and lipase.

|  |  |
| --- | --- |
| Molecules that have a **similar shape to the substrate can act as enzyme inhibitors**. This **competes** with the substrate for the active site; if they bond with the enzyme they **block the active site** so the substrate can’t fit in it.  The **amount of inhibition depends** on:   * relative **concentrations** of the inhibitor and substrate * how **strongly the inhibitor bonds** to the active site. |  |

Some **drugs** operate in this way, they **inhibit enzyme action**. For example penicillin inhibits the enzymes that control the building of cell walls in bacteria, this causes the cell wall to weaken and eventually burst.

The **active sites** of enzymes are **so specific** it takes **a lot of time and effort** to find a drug molecule that will inhibit the enzyme, especially if the substrate is chiral. **Trial and error** is often used but **computers** are now being **used to speed up the process**. They are able to **model the shape of the enzyme’s active site** and **predict** how well a potential drug molecule will interact with it.

***Fact recall: CGP240 Q1-3***

**DNA**

***DNA model***

DNA is short for **deoxyribonucleic acid**, it’s a **polymer** **found in all cells** that contains the genetic information for that organism.

Nucleotides

The **monomers** that make up DNA are called **nucleotides**, which are made up of three parts. The structures are given in the data sheet.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Phosphate group**  The circled atoms bond with the 2-deoxyribose molecule. | | |  | | |
| **Sugar**  2-deoxyribose  This is a pentose made up of 5 carbons with the hydroxyl group missing on the 2nd carbon (so it’s called deoxy-)  The circled atoms bond with the phosphate. | | |  | | |
| **Base**  4 types  The circled atoms bond with the 2-deoxyribose molecule. |  |  | |  |  |
| adenine (A) | guanine (G) | | cytosine (C) | thymine (T) |
| Purines – two rings | | | Pyrimidines – one ring | |

The **three parts join** together to **form a monomer**, so because there are four different bases there are **four monomers that made up DNA**. The **hydrogen atom** on the nitrogen atoms **circled** are **eliminated** along with an **–OH group** on the **deoxyribose** **sugar** when these groups bond, **water is formed in the condensation reaction**.

Each part can be represents with a shape to make it simpler.

|  |  |
| --- | --- |
| Phosphate |  |
| Sugar |  |
| Base |  |

They are bonded to each other in this arrangement.

|  |  |
| --- | --- |
|  | Simplified version |

***Task: Draw a nucleotide showing each type of base***

|  |  |
| --- | --- |
| Adenine |  |
| Cytosine |  |
| Guanine |  |
| Thymine |  |

Polynucleotides

**Nucleotides** can **polymerise** when a **hydroxyl group** on the **phosphate** on one nucleotide bonds with a **hydroxyl group** on a **sugar** molecule with the **elimination** of a **water** molecule. They form a **polynucleotide chain**, often referred to a **phosphate-sugar backbone**. In this condensation reaction a **covalent phosphodiester bond** is formed. There are still –OH groups on either end of the chain so further links can form.

|  |  |  |
| --- | --- | --- |
|  | This image shows an incorrect bonding between the phosphate and ribose  Oxygens opposite each other form bonds with the ribose molecules | Phosphodiester bond |
|  |

|  |  |  |
| --- | --- | --- |
| Condensation reaction |  |  |
|  |  |  |

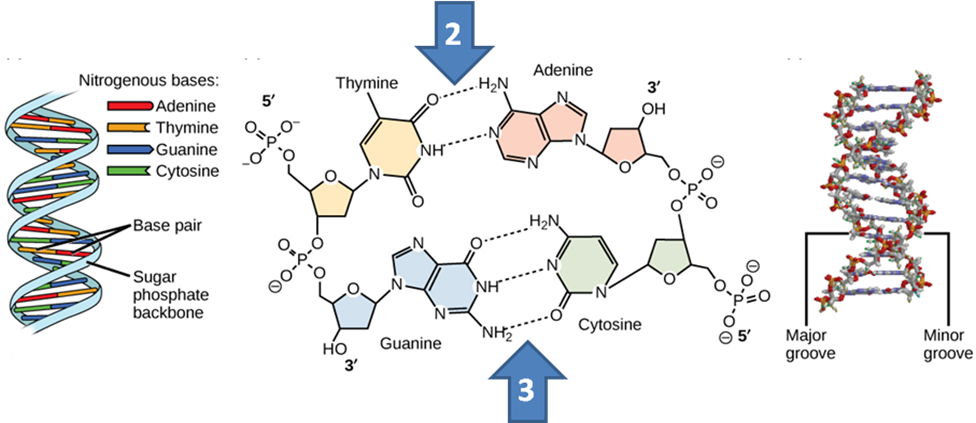
The arrangement of the **bases can be in any order**, this leads to variation in the DNA molecules. The DNA chain can be specified by the order of the bases, e.g. ATCCGTAAG.

DNA structure

DNA exists as a **double helix** held together by **hydrogen bonds**. The **bases pair up** with each other; **adenine and thymine** can form **two hydrogen bonds** with each other and **cytosine and guanine** can form **three hydrogen bonds** with each other. This allows two strands of DNA to bond together but only if their **bases are in a complementary order** – **A matching T** and **C matching G**.

**Two** hydrogen bonds between **adenine** and **thymine**

(Hint: Remember T = Two).



**Three** hydrogen bonds between **guanine** and **cytosine**

(Hint: Remember C = third letter of alphabet).

Only these parings allow hydrogen bonds to form between them due to the bases being in the right position and distance apart and allowing the DNA strands to twist into a double helix shape.

***Task: Draw the complementary base pairs***

|  |  |
| --- | --- |
| adenine : thymine | cytosine : guanine |
|  |  |

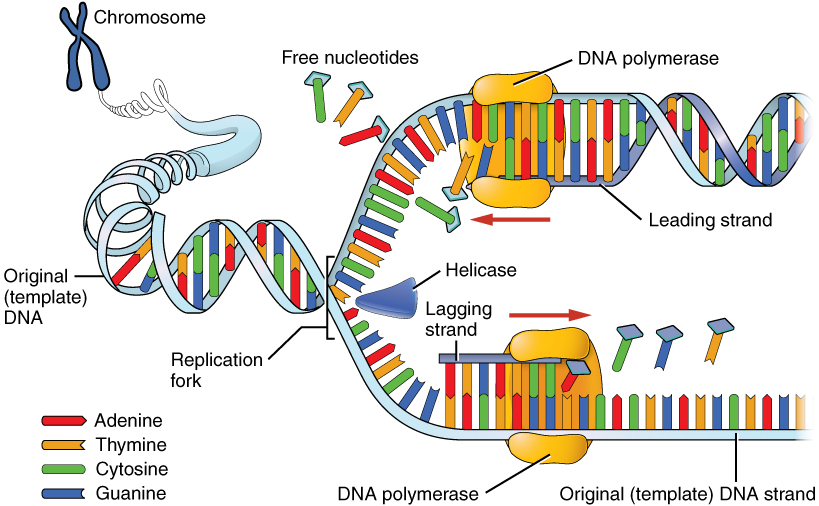
***YouTube clip:***  <https://www.youtube.com/watch?v=o_-6JXLYS-k>

***Sheet: DNA structure***

**DNA replication**

**DNA** is contained in every cell and it provides the instructions for making an organism. So for an organism to function correctly it **must copy itself when cells divide**.

**Hydrogen bonds** are only about **10% the strength of covalent bonds**, so they can be **broken** under conditions that **leave the covalent bonds untouched**. So during **cell division** the two **DNA strands can** part but the **polypeptide chains remains intact** **retaining the sequence of bases**.



In **DNA replication** the **double helix unwinds** and the **two strands separate**. Within the cell there are **free nucleotide molecules** **pair up** with **newly-exposed bases** – T to A and C to G. The **new bases** then **link** through the formation of **phosphodiester bonds** between the phosphates and sugars. A new **complementary strand of DNA** forms resulting in **two identical double helix molecules**.

***Fact recall: CGP244 Q1-3***

Action of anti-cancer drugs

**Cancer** is caused when **cells mutate** and **lose control of their growth and replication** so they **multiply uncontrollably**. For cells to divide they must replicate their DNA. An **anti-cancer drug** called **cis-platin** stops this happening properly and so the **tumour cells** **stop reproducing**. It was discovered in 1965 and is one of the most successful cancer treatments.

**Cisplatin** is a **square planar** complex of **platinum(II) ion** with **two chloride ligands** and **two ammonia ligands**, each forms a dative bond with the ion.

|  |  |  |
| --- | --- | --- |
| **Cisplatin** |  |  |
|  |  | A metal ions can form co-ordinate bonds with ions or molecules that are able to donate a lone pair of electrons. Such groups are called **ligands**, the resulting compound is called a **metal ion complex**. |
| **[Pt(NH3)2Cl2]** |  |  |

It exists as two **geometric isomers**, the other is called transplatin. This form would not work as it doesn’t have two chloride ions adjacent to each other for bonding to the guanine bases. They are on opposite sides.

|  |
| --- |
| **Transplatin** |
|  |

The **cisplatin** works by forming a **co-ordinate** **bonding to nitrogen atoms** on two **guanine bases** (either in the **same strand** or on **opposite** strands) in a section of **DNA**. It bonds **more tightly** that the chloride ions, so can’t be removed. This **distorts the DNA shape** so it can’t **unwind**, **preventing** **replication** of the cell and so it **dies**.

***You tube***: <https://www.youtube.com/watch?v=Wq_up2uQRDo>

|  |  |
| --- | --- |
|  |  |
|  |

The **nitrogen atom** on the **guanine** has a **lone pair of electrons** which are able to form a **co-ordinate bond** with the **platinum**, displacing the chlorine ions. This is called a **ligand substitution** reaction.

Side effects

Like all **drugs** cisplatin has **side effects**. It will also **bind to normal cells** as well as the cancer cells. Since the **cancer cells** are **replicating faster** it has a **greater effect** on these cells. However, **healthy cells that replicate quickly are significantly affected**, so patients undergoing cancer treatment often lose their hair and their i**mmune system is suppressed**. These **side effects can be reduced** by **lowering the dosage** or **targeting the tumour directly** to the cancer cells. Despite the side effects cisplatin is still used because it is effective and the **balance of positive long term effects outweigh negative short term effects**.

***Fact recall: CGP246 Q1-2***

***Sheet: DNA structure drawing questions***

***Sheet: Amino acids, proteins & DNA questions***

***Exam questions: Oxford p214-215 Q1-5***