

Teaching notes: Amino acids, proteins and DNA

These teaching notes relate to section 3.3.13 of our A-level Chemistry specification (7405). The resource aims to provide background material for teachers preparing students for A-level Chemistry. This is an additional resource to cover specification topics teachers may not be familiar with. This information goes beyond the requirements of the specification. It may well be of interest to students, but the specification indicates the limits of the knowledge and understanding of this topic that will be examined.

Important natural polymers in living systems

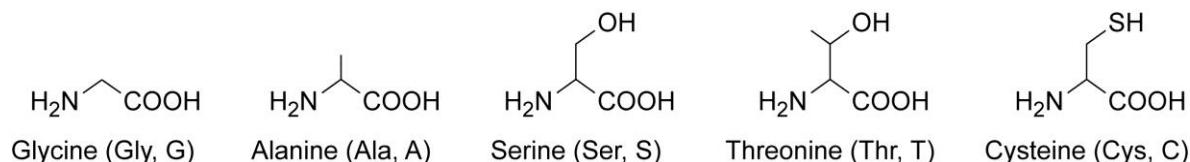
These include proteins, polysaccharides and nucleic acids. Proteins provide structural material in animal cells and can act as catalysts, known as enzymes, for biochemical processes. Polysaccharides provide structural material in plants and also act as an energy store. Nucleic acids provide the coded information for protein synthesis and enable inherited characteristics to be passed from one generation to the next.

Proteins

Amino acids and the peptide bond

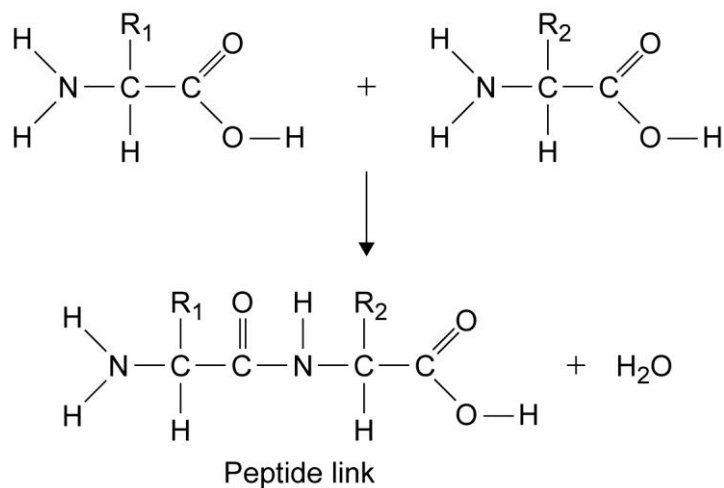
There are about 20 naturally-occurring amino acids in animal cells, each with the general formula $RCH(NH_2)COOH$. Each amino acid has an IUPAC name, a common name and a three-letter abbreviation for the common name. Some examples are shown in **Figure 1** (IUPAC names are not included here).

Figure 1



Peptides are made by linking amino acids. The link between two amino acids is shown in **Figure 2**.

Figure 2

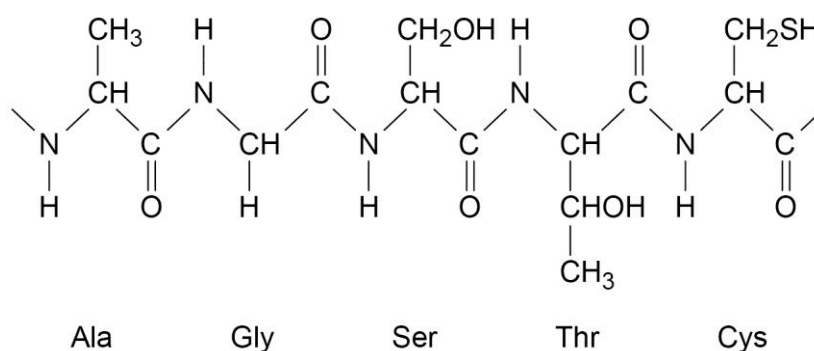


The amide linkage between amino acids is known as a peptide bond. Formation of this bond is difficult to achieve by direct reaction in the laboratory but occurs very rapidly in living cells when catalyzed by an appropriate enzyme called a peptidase.

Primary, secondary, tertiary and quaternary structure of proteins

A protein is made up from one or more long chains of amino acids linked by peptide bonds. The amino acids may be linked in any sequence. For example, Figure 3 shows a section of a protein chain.

Figure 3

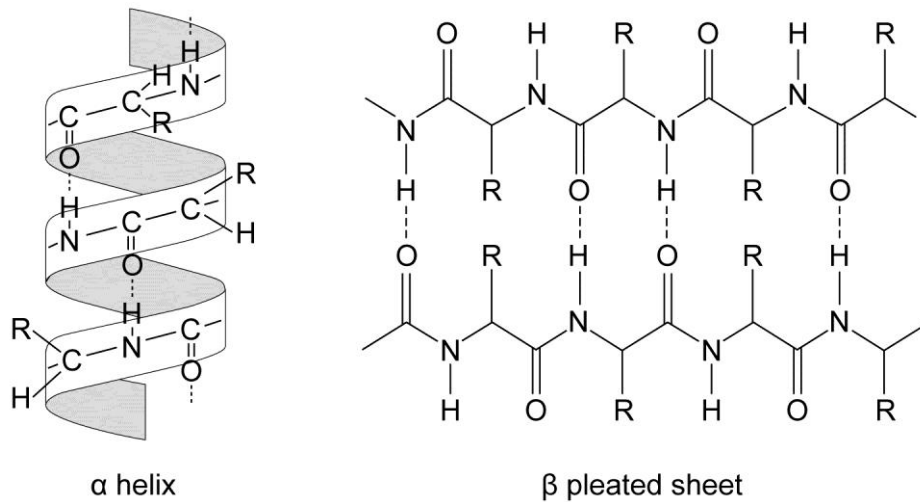


The sequence of amino acids is known as the primary structure of the protein.

Depending on the sequence, the amino acids in the chain may hydrogen bond with other acids forming arrangements of the chain such as an alpha-helix or a beta-sheet. The α term refers to the arrangement of amino acids where all the N-H bonds are on the same side of the protein chain. The β term refers to a section of protein chain in

which the N–H and C=O groups alternate from one side of the protein chain to the other. **Figure 4** shows an alpha helix and a beta pleated sheet.

Figure 4



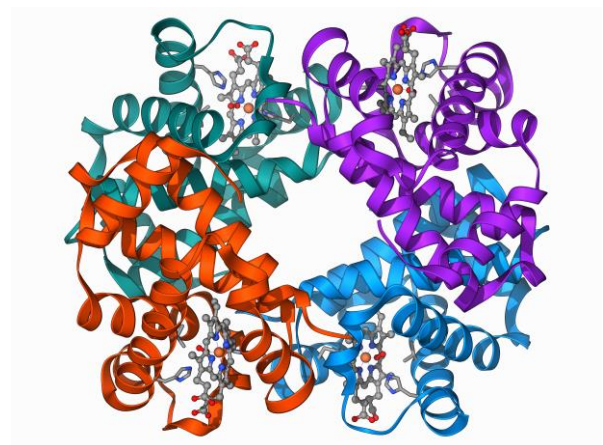
These arrangements are known as the secondary structure within the protein.

A single polypeptide chain, in three dimensions, may arrange itself in helical and sheet forms. A protein may consist of a single polypeptide chain or more than one chain attracted to each other by hydrogen bonds or, in some cases, by S–S bonds formed from the combination of two cysteine side chains.

Figure 5



A single polypeptide chain



Four polypeptide chains (Haemoglobin)

The 3D arrangement of a single polypeptide chain is known as the tertiary structure of the protein. The 3D arrangement of a protein made from more than one single

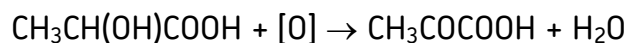
polypeptide chain is known as the quaternary structure. In **Figure 5**, the helices can be seen and also the broad, straightish ribbons that are pleated sheets.

It is the 3D structure that defines the properties of a protein. The 3D structure can be determined using X-ray diffraction.

Enzymes

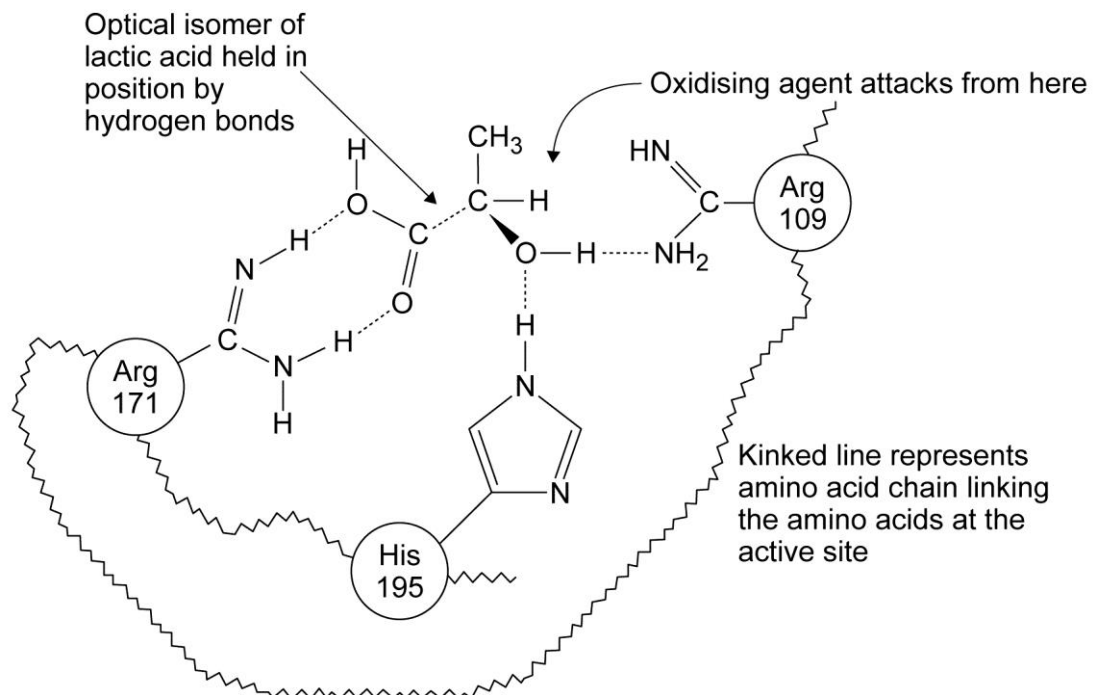
One of the most important functions of proteins is their action as catalysts in organic chemistry. Enzymes typically increase the rate of a reaction by a million times and in some cases by a billion trillion times! They operate by attracting molecules to an active site where reaction occurs with a much lower activation energy. The active site usually involves reactive groups on different amino acids brought close together by the tertiary or quaternary structure of the protein.

For example, the enzyme Lactate Dehydrogenase (LDH) catalyses the reversible oxidation of lactic acid (2-hydroxypropanoic acid) to pyruvic acid (2-oxopropanoic acid)



The lactic acid bonds to the active site of LDH as shown in **Figure 6**.

Figure 6

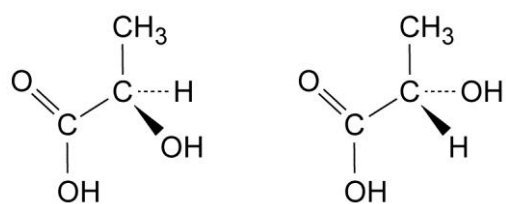


Bonding at active site of Lactate Dehydrogenase (LDH) enzyme

The active site involves the side chains on two separate arginine molecules and a histidine molecule, positioned in the sequence 109, 171 and 195 out of a total of about 330 amino acids in the whole protein.

Lactic acid exists as two optical isomers as shown in **Figure 7**.

Figure 7



Due to its shape and the orientation of the functional groups in space, only the isomer on the left is able to form the necessary hydrogen bonds to the enzyme before the oxidising agent attacks.

Enzymes and drugs

Enzymes can be deactivated by a molecule, similar to the substrate, which binds more strongly to the active site and hence prevents the substrate from bonding and reacting.

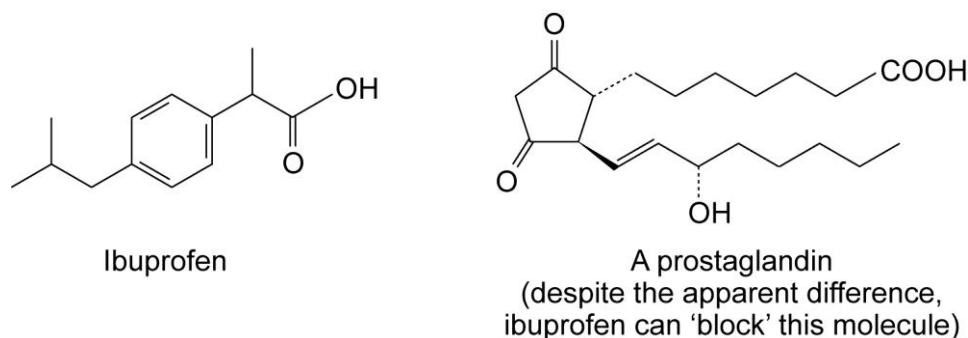
In the case of the LDH, this enzyme is particularly active in the protozoa that cause malaria. Drug companies have tried to find specific molecules that block the active site of LDH. At the moment, an effective drug is the compound oxamic acid (aminooxoethanoic acid), $\text{H}_2\text{NCOCOOH}$. This molecule has a structure related to lactic acid but it bonds more strongly to the active site of the enzyme and so prevents attachment of lactic acid.

In modern research laboratories, if the structure of the enzyme and its active site are known, computer graphics can be used to design a molecule that binds to the active site. Chemists can then attempt to synthesise the 'blocking' molecule. If this molecule inhibits an enzyme that catalyses a process causing illness, the molecule will behave as a drug. Drugs often work best when they are pure optical isomers and in many cases, only one optical isomer is pharmacologically active.

Examples of such drug molecules that inhibit enzymes are ibuprofen, aspirin and methotrexate.

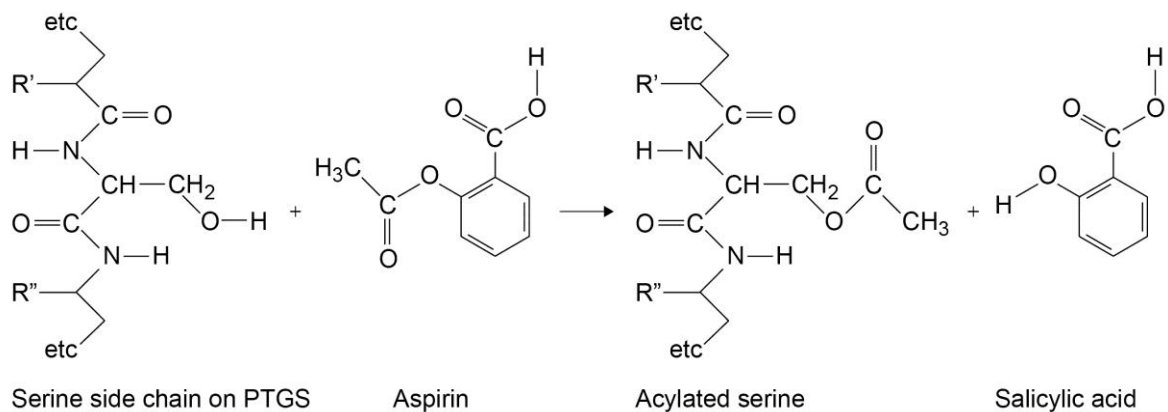
Ibuprofen, a painkiller, works by blocking some enzymes involved in the synthesis of prostaglandins. The structures of these molecules are shown in **Figure 8**. Prostaglandins cause pain. Ibuprofen is most effective if it is prepared as an optically pure isomer. This makes the synthesis of ibuprofen more complicated.

Figure 8



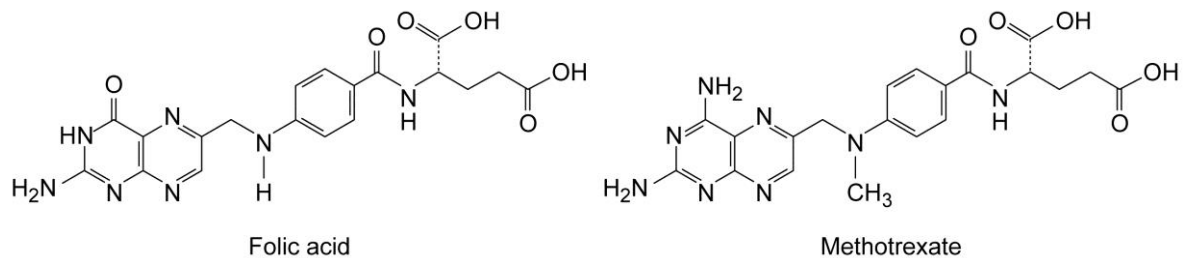
Aspirin, shown in **Figure 9**, works in a similar fashion to ibuprofen by blocking the active site of the enzyme prostaglandin synthase (PTGS). In this case, however, the aspirin reacts to acylate the OH side chain on a serine molecule at the active site. This reaction only involves one part of the active site and is not stereospecific. The same reaction can occur with other serine side chains in other enzymes so the action of aspirin is not as specific as with ibuprofen.

Figure 9



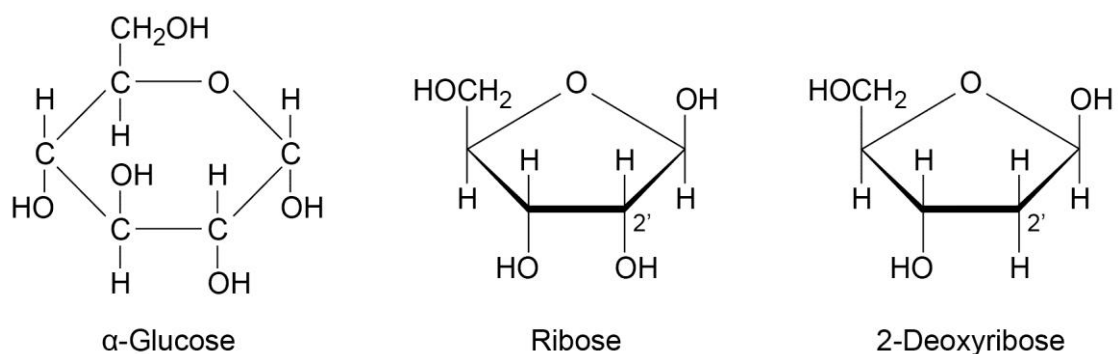
The anticancer drug methotrexate, shown in **Figure 10**, works by blocking the enzyme dihydrofolate reductase. This enzyme uses folic acid as a substrate from which nucleotides are synthesized. Synthesis of nucleotides is a necessary requirement for rapid growth of cancer cells. Methotrexate is similar to folic acid. It binds to the active site but then blocks further reactions and so cancer cells cannot replicate.

Figure 10



Sugar molecules, also known as carbohydrates, are used in living systems as a source of energy. Sugars are found as six-membered rings (hexose sugars) and five-membered rings (pentose sugars). Examples are shown in **Figure 11**.

Figure 11

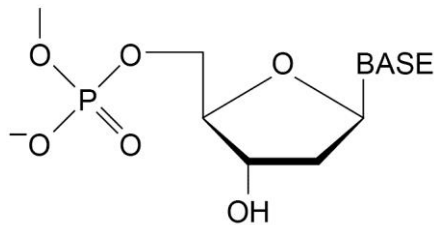


DNA (deoxyribonucleic acid)

All animals start life as a single cell. The cell divides and differentiates into cells with different functions as these cells continue to divide. All the information required to build and operate cells and the living organisms derived from them is carried on a molecule of DNA. The DNA molecule is the same in every cell of the same animal and it is inherited from a combination of the parents' DNA. Each organism and each individual has a different DNA. The only exception is with identical twins.

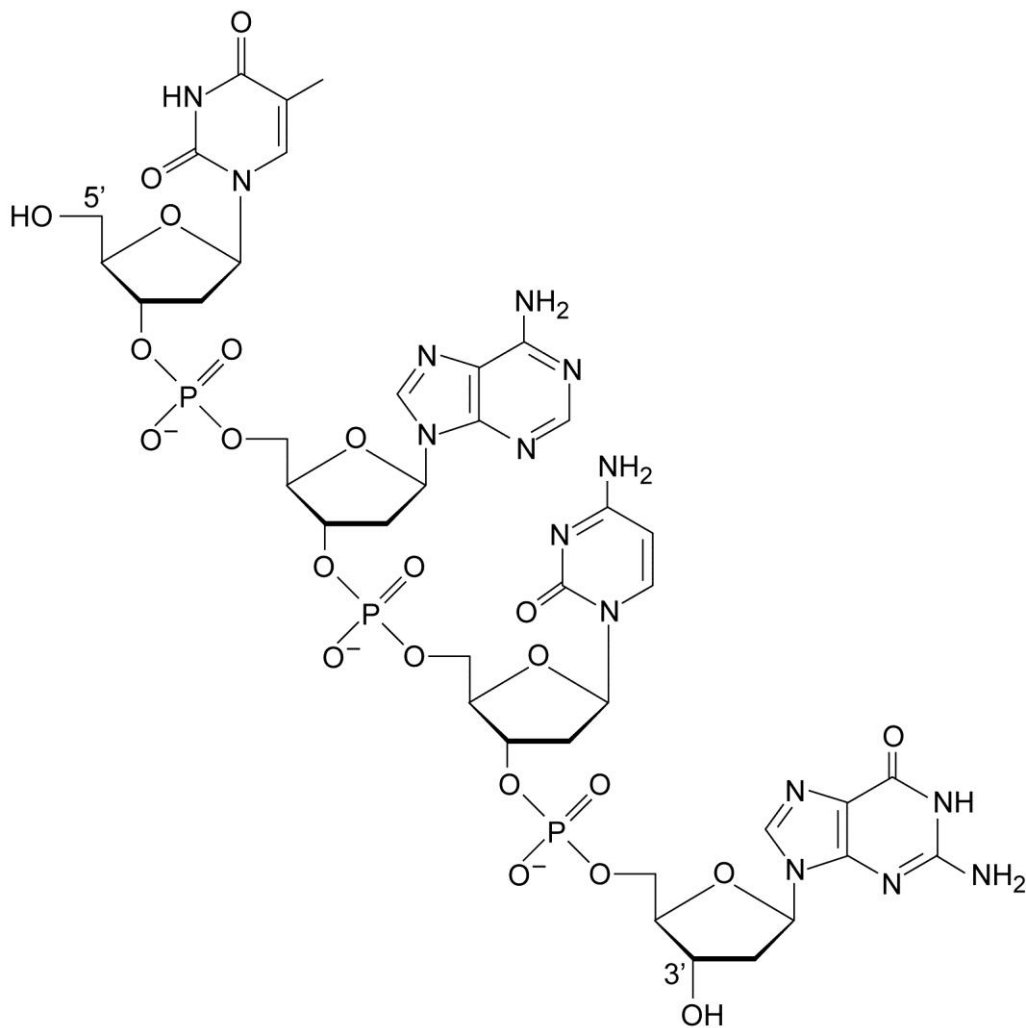
A single strand of DNA is a polymer of nucleotides. A nucleotide is a deoxyribose sugar molecule attached to a phosphate ion (PO_4^{3-}) and one of four possible organic molecules called bases, adenine, thymine, guanine and cytosine. Nucleotides link together to form a strand or chain when a phosphate ion bonds to a deoxyribose molecule. The repeating unit in the chain is shown in **Figure 12**.

Figure 12



Each of these nucleotides links to form a long strand of DNA as shown in **Figure 13**.

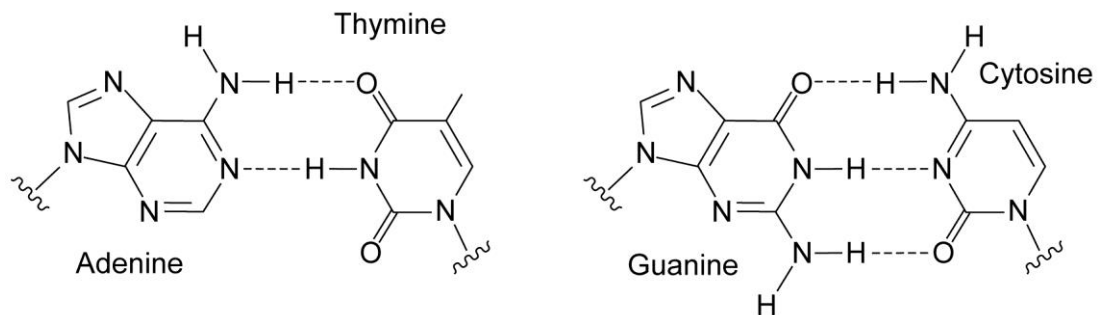
Figure 13



Part of a strand of a DNA molecule

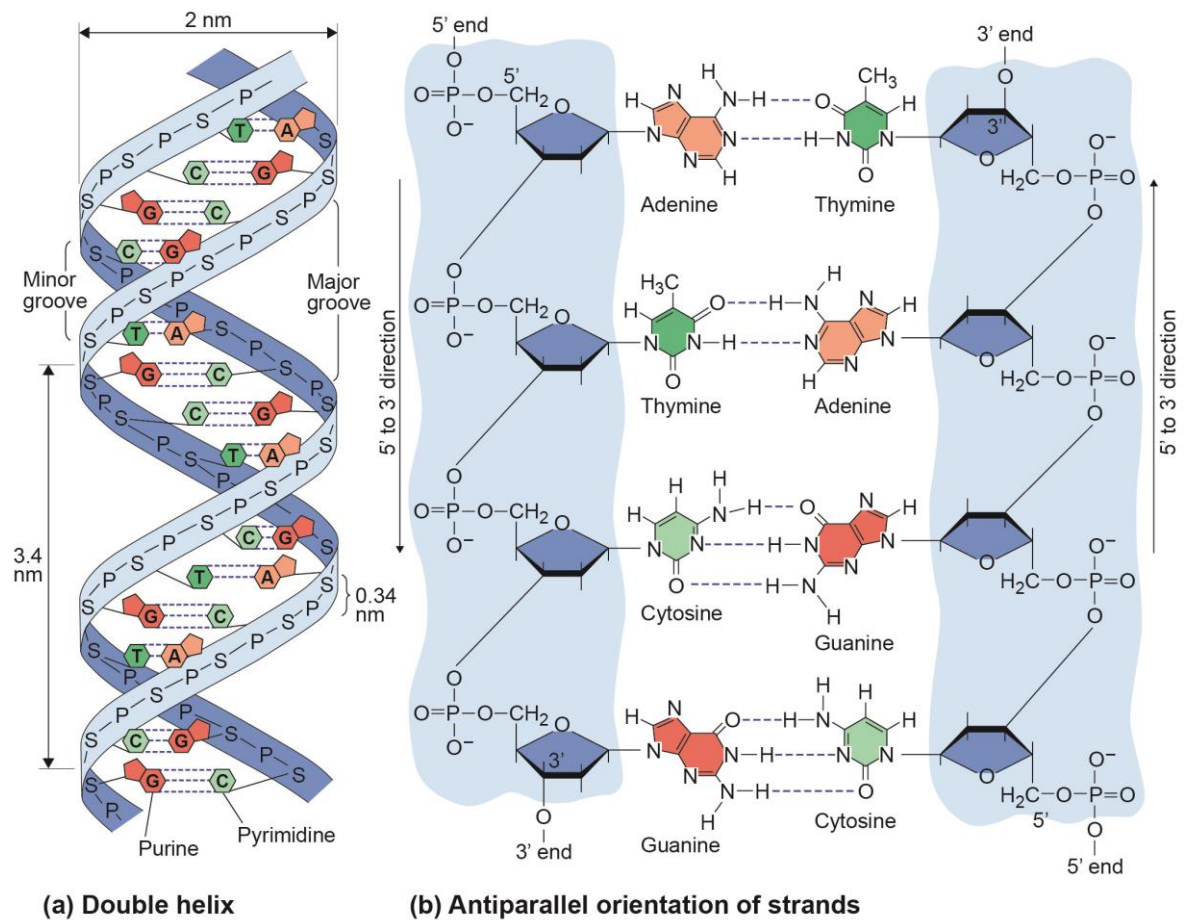
In cells, DNA consists of two complementary strands arranged in double helices. This complementary arrangement is made possible because pairs of bases hydrogen bond with each other as shown in **Figure 14**.

Figure 14



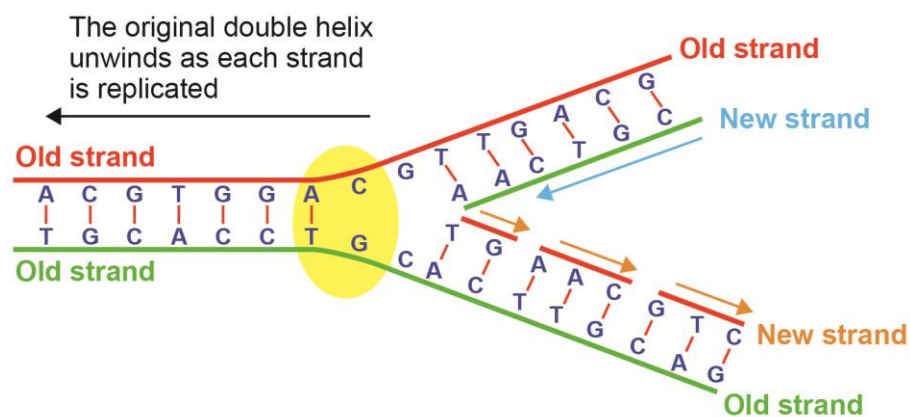
The two complementary strands are organised in a double helix, shown in a three-dimensional representation in **Figure 15**.

Figure 15



When a cell divides, each strand of DNA separates and enzymes bring together nucleotides, creating a complimentary copy so that two new double helices are formed, each identical to the previous one as shown in **Figure 16**.

Figure 16



The sequence of bases on a strand of DNA is not random. The sequence provides information for the sequence of amino acids in the synthesis of proteins. The 'code' used is called the triplet code where three bases together identify a single amino acid. A molecule called messenger RNA copies a section of a DNA molecule and moves with that copy to a section of the cell where amino acids are assembled into proteins. Examples of the triplet code are not unique but amongst others, TCT identifies serine, GGT identifies glycine and CTC identifies leucine etc.

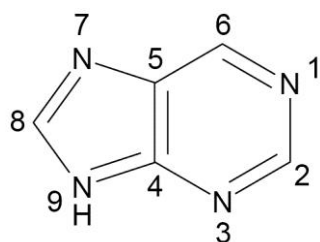
If the DNA sequence is incomplete, damaged or changed in some way (for example by radiation or by the carcinogenic chemicals in tobacco smoke), it can lead to medical problems. For instance, it may cause cells to self-replicate in a fast, uncontrolled manner. This self-replication is a prime cause of cancer.

The anti-cancer drug cisplatin works by undergoing a ligand exchange reaction with one of the nitrogen atoms in a guanine base on DNA.

The drug is usually injected in the form of $(\text{NH}_3)_2\text{PtCl}_2$. This neutral complex, after passing through the cell membrane of a cancer cell reacts with water to form the complex ion $[(\text{NH}_3)_2\text{PtCl}(\text{H}_2\text{O})]^+$.

A guanine molecule is shown in **Figure 17**.

Figure 17



The nitrogen 7 on a guanine unit (even though the guanine component is still bonded to the deoxyribose in a DNA strand) forms a co-ordinate bond to the platinum replacing water. Another guanine on the same strand of DNA or on the complementary strand then bonds to the platinum by replacing the chlorine ligand. If the same strand of DNA is bonded twice to the platinum, the strand becomes kinked and will not replicate. If two DNA strands are bonded to the platinum, this prevents the strands from separating and replication is also prevented in this case.

A problem with this type of chemotherapy treatment is that it also affects normal cells and a patient can have unpleasant side-effects.

DNA profiling

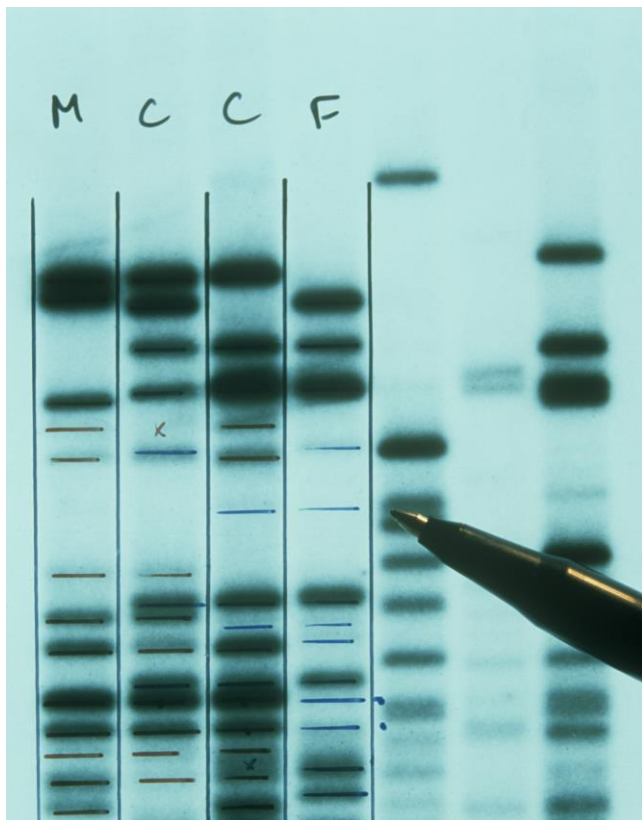
Because each DNA sequence is unique to an individual, the sequence of bases in a sample of DNA is a method for identifying that individual. Similarly, because we inherit long sections of DNA from our parents, the sequence enables family members to be recognised.

The technique of comparing DNA from one source with that from another is called DNA profiling. In this process, a small amount of DNA is required. It is broken down, using enzymes, into sections which are copied, using other enzymes, until many millions of copies of each section are made. The sections (STR or short tandem repeats) are then separated by electrophoresis and compared with those from a known source which should have similar STRs in the same proportions.

This technique is used to identify criminals and to trace ancestry.

Figure 18 shows the results of DNA profiling for two parents and two children. The STR strands have been separated by gel electrophoresis, a type of chromatography.

Figure 18



In Figure 18, the DNA fingerprints of two children are labelled as (C). Some bands are common with their mother and other bands are common with their father.

References

Figure 5: Insulin monomer molecule, Evan Oto, Science Photo Library.

Figure 5: Haemoglobin molecule, Laguna Design, Science Photo Library.

Figure 18: DNA fingerprinting, David Parker, Science Photo Library.

Permission to reproduce all copyright material has been applied for. In some cases, efforts to contact copyright-holders may have been unsuccessful and AQA will be happy to rectify any omissions of acknowledgements. If you have any queries, please contact the Copyright Team, AQA, Stag Hill House, Guildford, GU2 7XJ.

Copyright © 2016 AQA and its licensors. All rights reserved.