Review on the Protein Structure

Proteins are chains of amino acids that fold into three-dimensional shapes. The shape of the protein is very important to its function and the three-dimensional structure is specified by an amino acid sequence. Protein structure has 4 levels of organisation known as primary, secondary, tertiary and quaternary. Proteins are firstly manufactured as a primary sequence composed of a linear sequence of amino acids joined by peptide bonds which continue to fold into secondary, tertiary and finally quaternary structures. Twenty different amino acids are incorporated into proteins, the sequence of amino acids of a protein is termed its primary structure (Loughlin, 2017).

A primary structure is the simplest level of protein structure it is a sequence of amino acids in a polypeptide chain. Each chain has its own set of amino acids assembled in a particular order with a typical basic chemical structure, as shown below in the figure 1. A central carbon atom (the a-carbon) bonded to a hydrogen atom, a basic amino group comprising of a nitrogen atom and two hydrogen atoms (-NH2), a carboxyl group (-COOH) and a specific side chain or R group consisting of varying atoms. The R group gives each amino acid its identity' they can be polar, nonpolar or even unchanged (Loughlin, 2017).

The amino acids of a polypeptide are attached to each other by covalent bonds known as a peptide bond, each bond forms a condensation reaction. During protein synthesis the carboxyl group of the amino acid at the end of the growing polypeptide chain reacts with the amino group of an incoming amino acid, releasing a molecule of water. The resulting bond between amino acids is a peptide bond. Because of the structure of the amino acids, a polypeptide chain has two ends that are chemically distinct from each other. At one end the polypeptide chain has a free amino group called the amino terminus (N-terminus) and the other end that has a free carboxyl group known as the carboxyl terminus (C-terminus). The interactions between amino acids cause a protein to fold; from an amino acid sequence of a polypeptide to a three-dimensional structure of a mature functioning protein (Loughlin, 2017).

The two most important protein secondary structures, the alpha helix (a helices) and the beta sheet (ß sheet) were predicted by Linus Pauling (1951) citied in Loughlin (2017, p. 9) . With the use of X-ray diffraction (Loughlin, 2017, p.12, Box 1.3) Pauling was able to determine the shape of proteins, discovering the spiral structure of proteins; the polypeptide backbone. He recognised that folding of peptide chains among other things such as steric hindrance should maintain the bond angles and planar structure of the peptide bond as well as preventing the atoms from coming to closely together that they repel each other. Both types of the secondary structures a helix and ßsheet are held in shape by hydrogen bonds, which form between

carbonyl (C=O) and amine (N-H) shown below, pulling the polypeptide chain into a helical structure allowing the side chain to stick out and freely interact. The majority of characteristics of proteins are consistent with their secondary structures, they can either be fibrous which are important in forming biological structures or globular, spherical in shape with recognisable regions of a helices and ß sheet structures which are connected non-uniform shapes known as random coils (Loughlin, 2017).

As described in Loughlin (2017) the final three-dimensional structure of a polypeptide is called its tertiary structure as shown in Figure 2, the protein molecule will fold in such a way as to achieve maximum stability or less energy. The tertiary structure is mainly down to the various types of bonding interactions of the different amino acid side chains; hydrogen bonding, disulfide bonds, non-polar hydrophobic interactions, polar hydrophilic interactions and ionic interactions, basically a whole range of covalent bonds. For example R groups with like charges repel one another, while those with opposite charges can form ionic bonds, likewise polar side chains can form hydrogen bonds. Hydrophobic interactions in which amino acids with non-polar hydrophobic R groups collect together on the inside of the protein leaving the hydrophilic amino acids on the surface to interact with surrounding water molecules via hydrogen bonds or ionic interactions. Disulfide bridges are covalent linkages between the sulfur-containing side chains of two cysteines are much stronger than other types of bonds and help to stabilise the folded structure of the protein. They act like a molecular fastening by keeping parts of the peptide chain firmly attached to one another. In order for some proteins to function a little helper known as a cofactor will combine with a polypeptide chain as it fold, this may be important for structural reliability; if they are not present the protein does not fold properly and becomes unstable (Loughlin, 2017).

The quaternary structure of a protein is the fourth level of organisation and refers to the number and organisation of the protein subunits in relation to one another, to form a multimeric protein. Multimeric proteins are heteromeric if their subunits are different; haemoglobin is heteromeric as it is made up of four subunits (two each of two different subunits). If the subunits in a multimeric protein are the same then it is said to be homomeric an example is the enzyme glyceraldehyde 3-phosphate dehydrogenase (GAPDH) which consists of four identical subunits (The Open University, 2017, Activity 1.1). Generally the same type of interactions such as non-covalent interactions and disulfide bonding that contribute to the stability of the tertiary structure also hold the subunits together to give quaternary structure (Loughlin, 2017).

To gain functional stability most proteins have to fold into three-dimensional structures however in a cellular environment newly synthesised proteins are at risk of misfolding so many larger polypeptides require specialised chaperone proteins to help them fold. Their role is to stabilise unfolded proteins and assist in their correct covalent folding or unfolding. Molecular chaperones also help to refold any proteins that have formed incorrect structures by preventing the

polypeptides from combining with other molecules, forming large aggregates. Random aggregates are death for cells; diseases such as Alzheimer's disease, cystic fibrosis and Huntington's disease are caused by unnatural aggregation of proteins (Loughlin, 2017).

Proteins are the most common naturally occurring substance in living organisms, they are the engines and workers in our bodies, each having its own precise function. Polypeptides fold in a manner to make mature proteins, this process depends on the amino acids in the protein and their chemical and structural properties. The function of a protein depends on its shape which is determined by its sequence of amino acids, by folding into a specific three dimensional shape allows interactions between amino acids, enabling proteins to perform their biological function. However when things go wrong as a result of accumulation of proteins formed by misfolded proteins a wide range of degenerative and neurodegenerative diseases can occur (Loughlin, 2017).