PROTEINS

Content Statements:

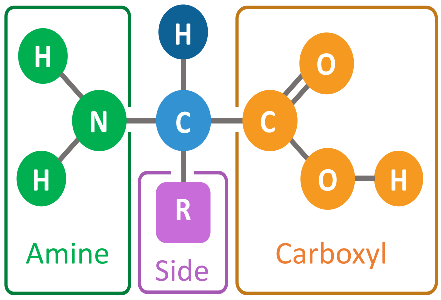
B1.2.1	Generalised structure of an amino acid
B1.2.2	Condensation reactions forming dipeptides and longer chains of amino acids
B1.2.3	Dietary requirements for amino acids
B1.2.4	Infinite variety of possible peptide chains
B1.2.5	Effect of pH and temperature on protein structure

PROTEINS

Proteins are an extremely diverse class of organic compounds that fulfil a wide array of functions within a cell. Proteins are very prevalent within an organism and constitute more than 50% of the total dry mass of a single cell. Proteins function as the 'worker' molecules of a cell – they are encoded by nucleic acids (DNA) and are expressed in accordance with the specific genetic instructions of a particular cell.

AMINO ACIDS

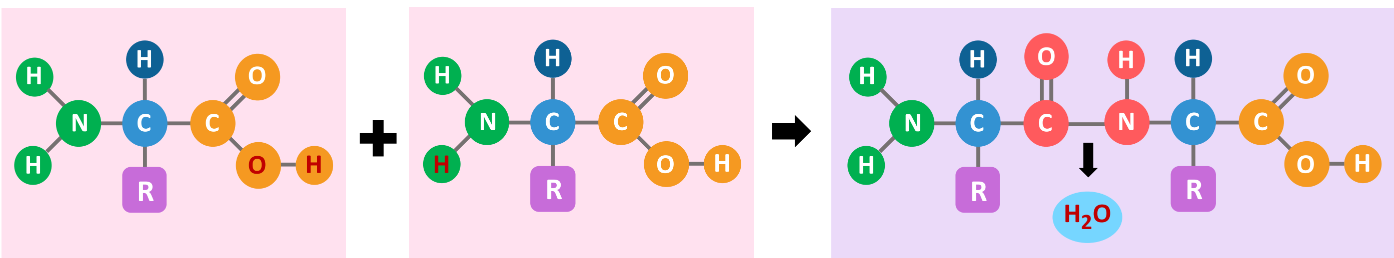
Proteins are comprised of long chains of recurring monomers called **amino acids**. Amino acids all share a common basic structure, with a central carbon atom bound to an amine group, a carboxyl group and a variable side chain. There are 20 different amino acids, each with a distinct side chain (i.e. *R group*). The different chemical properties of these side chains cause a protein to fold differently according to the sequence of amino acids (different order = altered protein structure).



Not all amino acids required for polypeptide synthesis can be produced by the body. Some amino acids can only be acquired through the diet – these are considered **essential** amino acids. In humans, there are nine essential amino acids and a further seven are considered conditionally essential (they can be produced, but at certain times – such as during pregnancy – the rate of production may be less than is required by cells).

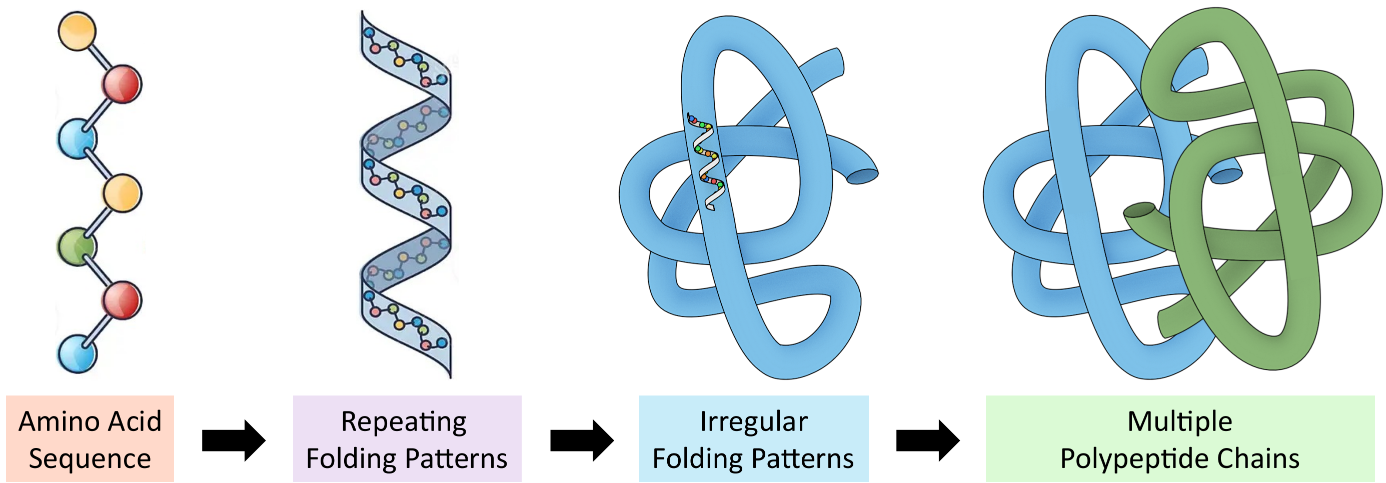
POLYPEPTIDES

Amino acids can be connected via **peptide bonds** to form dipeptides. This involves a condensation reaction (water is released as a by-product). Additional covalent attachments will result in long polypeptide chains. A polypeptide chain will fold into different arrangements according to the order of amino acids in the chain. Some proteins will even be composed of more than one polypeptide chain. Because a polypeptide chain can be of any length, proteins possess infinite variety in both their sequence and their potential structure.



PROTEIN FOLDING

The overall shape and structure of a protein plays an integral role in determining its biological capabilities. The structure of a protein is determined by the order of the amino acids in a polypeptide sequence. The different variable groups have different chemical properties, which will cause a protein chain to fold into different arrangements depending on the relative positions of the variable groups. Hydrogen bonds can also form between the amine and carboxyl groups of different amino acids, creating an additional layer of structural complexity. Some proteins are even composed of multiple (more than one) polypeptide chains.



DENATURATION

Denaturation is a change in protein structure that causes a loss of biological activity. The structural change is usually permanent, however some proteins may be refolded with the assistance of chaperone proteins.

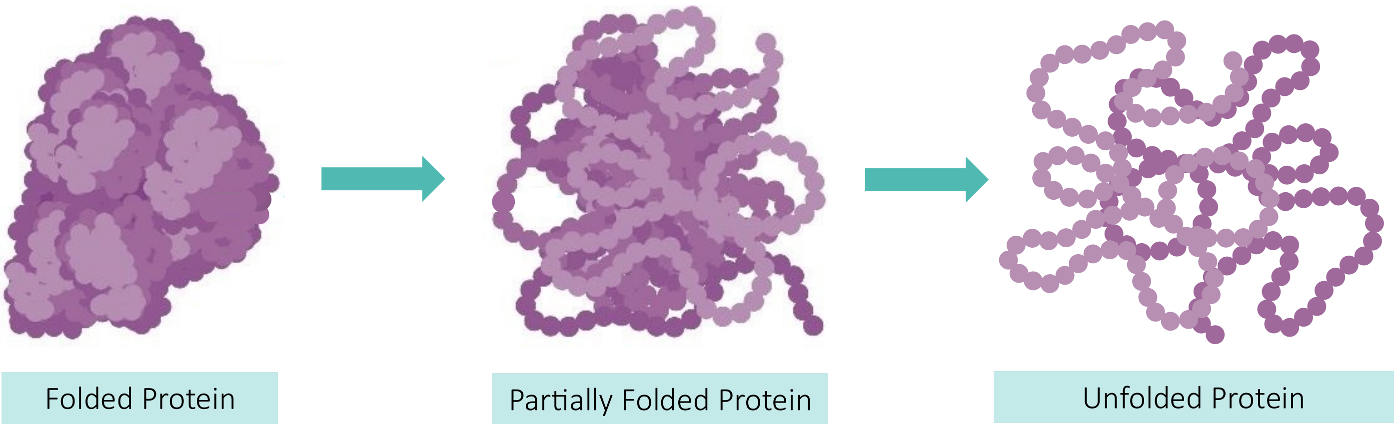
- The denaturation of a protein can most typically be caused by two conditions – temperature and pH.

Temperature:

- High levels of thermal energy may **disrupt the hydrogen bonds** that hold a polypeptide chain together
- As the bonds are broken, the protein will begin to unfold and lose its capacity to function as intended
- Denaturation temperatures may vary, but human proteins function best at body temperature (37°C)

pH:

- Amino acids are zwitterions, neutral molecules with both negative (COO^-) and positive (NH_3^+) regions
- Changing the pH will **alter the charge of a protein**, which in turn will alter protein solubility and shape
- All proteins have an optimal pH which is dependent on the environment in which it typically functions



PROTEIN FUNCTIONS

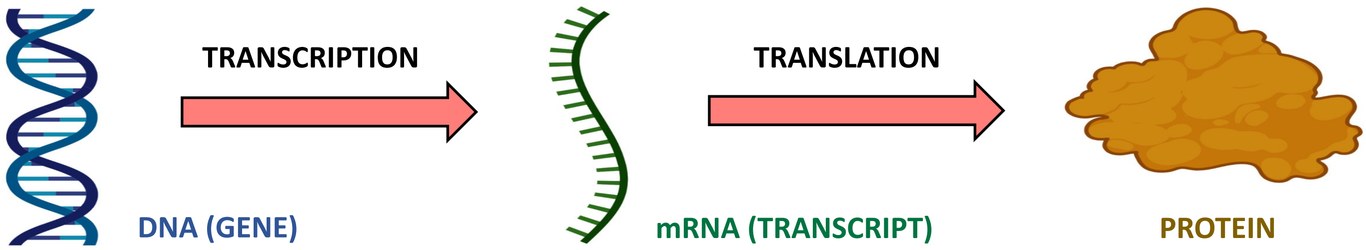
Proteins are a very diverse class of organic compounds that serve a wide variety of different roles in cells. Examples of protein functions include:

- **S**tructure *(e.g. collagen is found in skin, keratin is found in hair)*
- **H**ormones *(e.g. insulin and glucagon regulate blood sugar levels)*
- **I**mmunity *(e.g. immunoglobulins target foreign pathogens)*
- **T**ransport *(e.g. protein channels enable facilitated diffusion)*
- **S**ensation *(e.g. rhodopsin is an eye pigment required for vision)*
- **M**ovement *(e.g. actin and myosin are used in muscle contraction)*
- **E**nzymes *(e.g. Rubisco is responsible for carbon fixation in plants)*

Hint: Use a mnemonic device!

PROTEIN PRODUCTION

The amino acid sequence of a polypeptide is encoded by nucleic acids. A sequence of DNA that codes for a polypeptide is called a **gene**. The gene sequence is converted into an amino acid sequence in two stages – transcription and translation. *Transcription* involves the conversion of a gene sequence (DNA) into an mRNA transcript, which is then transported to the ribosome. The ribosome then synthesises a polypeptide chain from the transcript instructions (*translation*). Some proteins can undergo post-translational modifications – such as phosphorylation, glycosylation or proteolytic cleavage – before adopting their final functional form.



Typically, one gene will code for a one polypeptide – however there are exceptions to this rule. Some genes do not code for proteins at all. For instance, the genes encoding for tRNA and rRNA sequences do not get translated – only mRNA sequences are used by the ribosome to assemble polypeptide chains. Additionally, some mRNA transcripts may be edited (spliced) to create alternative protein variants from a single gene.

PROTEOME

The totality of all proteins expressed within a cell, tissue or organism is called the **proteome**. The proteome of any individual will be unique, and protein expression levels can change with time and conditions. While the human genome is estimated to contain nearly 20,000 protein-producing genes, the human proteome is predicted to contain roughly 100,000 different proteins (some genes can be alternatively spliced and some proteins can be modified after translation). The study of the proteome (**proteomics**) has many applications:

- It is used in disease identification (by identifying proteins present in diseased, but not healthy, people)
- It is used in the subsequent development of medications (by targeting the proteins present in diseases)

Additionally, by understanding how protein activity may be changed by post-translational modifications (such as phosphorylation), metabolic processes may be selectively improved via therapeutic interventions.