B1.2 Proteins

Form and function—Molecules

Standard level and higher level: 2 hours

Additional higher level: 2 hours

Guiding questions

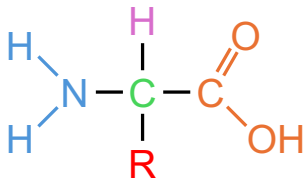
• What is the relationship between amino acid sequence and the diversity in form and function of proteins?

• How are protein molecules affected by their chemical and physical environments?

SL and HL

B1.2.1—Generalized structure of an amino acid

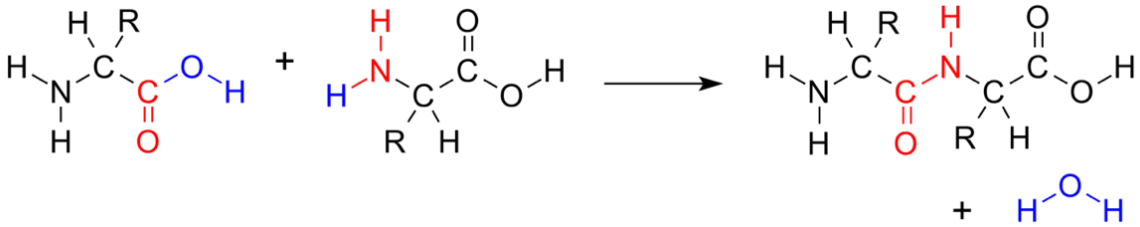
Students should be able to draw a diagram of a generalized amino acid showing the alpha carbon atom with amine group, carboxyl group, R-group and hydrogen attached.



- **Amino acids** are biomolecules that have a generalized structure:
  - **Alpha (α) carbon**: the **central** carbon atom that is bonded to 4 atoms or groups
  - **Amino group (NH_2)**: a **basic** functional group
  - **Carboxyl group (COOH)**: an **acidic** functional group
  - **Hydrogen atom**: covalently bonded to the alpha carbon
  - **R (Residue) group (side chain)**: a **variable** group that gives each amino acid unique properties

B1.2.2—Condensation reactions forming dipeptides and longer chains of amino acids

Students should be able to write the word equation for this reaction and draw a generalized dipeptide after modelling the reaction with molecular models.



- **Condensation reactions** connect amino acids together by forming a **peptide bond** between the amino group of one amino acid and a carboxyl group of the next amino acid
  - A **dipeptide** is formed when **2** amino acids are joined together by **1** peptide bond
  - An **oligopeptide** is formed when **less than 20** amino acids are joined together by peptide bonds
  - A **polypeptide** is formed when **20 or more** amino acids are joined together by peptide bonds
  - A **protein** consists of one or more polypeptide chains

B1.2.3—Dietary requirements for amino acids

Essential amino acids cannot be synthesized and must be obtained from food. Non-essential amino acids can be made from other amino acids. Students are not required to give examples of essential and non-essential amino acids. Vegan diets require attention to ensure essential amino acids are consumed.

- **Essential amino acids** cannot be synthesized by an organism and must be obtained from food
- **Non-essential amino acids** can be synthesized by organisms
  - Essential amino acids can be converted by organisms into non-essential amino acids
- Even though plants synthesize all 20 amino acids, many plant-based food (unlike animal-based) has insufficient amounts of essential amino acids for human needs
  - Vegan diets thus require attention to ensure that essential amino acids are consumed and dietary deficiencies are avoided

B1.2.4—Infinite variety of possible peptide chains

Include the ideas that 20 amino acids are coded for in the genetic code, that peptide chains can have any number of amino acids, from a few to thousands, and that amino acids can be in any order. Students should be familiar with examples of polypeptides.

- 20 amino acids are coded for in the genetic code
- The DNA sequence of a gene determines the amino acid sequence of a polypeptide
- The **proteome** is the set of all polypeptides produced by a cell
- For a polypeptide with ***n*** amino acids, there are $\mathbf{20^n}$ possible sequences (amino acids can be in any order and in any number)

B1.2.5—Effect of pH and temperature on protein structure

Include the term “denaturation”.

- Polypeptides **spontaneously fold** into **3D structures** that are stabilized by **intermolecular forces**
- **Denaturation** is the process by which the folding (3D structure) of a polypeptide is disrupted due to **physical** (e.g. temperature) or **chemical** (e.g. pH) factors
  1. **Temperature**: heat increases the vibrations of atoms, which disrupts intermolecular forces & causes an **often irreversible** conformational change in the 3D structure of the polypeptide
  2. **pH**: deviations away from optimal hydrogen ion concentration disrupts the intermolecular forces & causes a reversible or irreversible conformational change in the 3D structure of the polypeptide
- Denaturation **does not** alter the amino acid sequence of a protein

Additional higher level

B1.2.6—Chemical diversity in the R-groups of amino acids as a basis for the immense diversity in protein form and function

Students are not required to give specific examples of R-groups. However, students should understand that R-groups determine the properties of assembled polypeptides. Students should appreciate that R-groups are hydrophobic or hydrophilic and that hydrophilic R-groups are polar or charged, acidic or basic.

- R-groups determine the properties of assembled polypeptides because they are chemically diverse
  - R-groups can either be **hydrophobic** or **hydrophilic**
    - **Hydrophobic R-groups** associate with each other away from the aqueous cell environment
    - **Hydrophilic R-groups** can be polar or charged
      - **Polar** R-groups can form **polar interactions** & **hydrogen bonds** with each other or with water
      - **Charged** R-groups form **ionic bonds** between charged amino and carboxyl groups
        - Amino groups are **basic**; they can acquire a positive charge by **accepting** an H^+ ion
        - Carboxyl groups are **acidic**; they can acquire a negative charge by **losing** an H^+ ion
  - R-groups containing sulfur can form **disulfide bridges**; covalent bonds between 2 cysteines
- R-groups can be modified (post-translationally) to give rise to amino acids beyond the standard 20

B1.2.7—Impact of primary structure on the conformation of proteins

Students should understand that the sequence of amino acids and the precise position of each amino acid within a structure determines the three-dimensional shape of proteins. Proteins therefore have precise, predictable and repeatable structures, despite their complexity.

- The **primary structure** of proteins is the **linear sequence** (order) of amino acids in a polypeptide chain
- It contains all the necessary information needed to precisely fold the protein into its correct 3D shape
  - Proteins therefore have precise, predictable and repeatable structures, despite their complexity

B1.2.8—Pleating and coiling of secondary structure of proteins

Include hydrogen bonding in regular positions to stabilize alpha helices and beta-pleated sheets.

- The **secondary (2º) structure** of proteins involves folding through **hydrogen bonding**:
  - **Alpha helix**: hydrogen bonding between the N-H and C=O polar bonds of **every 4th amino acid** on the same polypeptide chain
  - **Beta-pleated sheet**: hydrogen bonding between the N-H and C=O polar bonds in **different strands** of the same polypeptide chain
- Since 2º structures do **not** involve R-group interactions, they form spontaneously in most proteins

B1.2.9—Dependence of tertiary structure on hydrogen bonds, ionic bonds, disulfide covalent bonds and hydrophobic interactions

Students are not required to name examples of amino acids that participate in these types of bonding, apart from pairs of cysteines forming disulfide bonds. Students should understand that amine and carboxyl groups in R-groups can become positively or negatively charged by binding or dissociation of hydrogen ions and that they can then participate in ionic bonding.

- The **tertiary structure** is the 3D structure of proteins, which depends on interactions between **R-groups**
  - **Ionic bonds** between charged R-groups
  - **Hydrophobic interactions** between non-polar R-groups
  - **Hydrogen bonding** in R-groups that contain O–H or N–H bonds
  - **Disulfide linkages** that are formed in the endoplasmic reticulum

B1.2.10—Effect of polar and non-polar amino acids on tertiary structure of proteins

In proteins that are soluble in water, hydrophobic amino acids are clustered in the core of globular proteins. Integral proteins have regions with hydrophobic amino acids, helping them to embed in membranes.

- The polarity of R–groups determines how they interact with the surrounding environment
  - Interactions between R–groups & the environment affects the tertiary structure of proteins
- **Non-polar R–groups** associate with each other & away from polar molecules, like water
  - In aqueous environments like the cytosol, they tend to cluster together inwards (away from water)
  - In non–polar environments like the hydrophobic core of the phospholipid bilayer, they face outwards towards the fatty acid chains of membrane lipids
- **Polar R–groups** associate with other polar molecules, like water
  - In aqueous environments like the cytosol, they face outwards towards the polar water molecules
  - In non–polar environments like the hydrophobic core of the phospholipid bilayer, they tend to cluster together inwards (away from the hydrophobic core)
- As these R-groups re-orient themselves depending on the environment, they pull different parts of the polypeptide chain closer together or push them away from the surrounding environment
  - This changes the spatial arrangement of the chain & thus contributes to the 3D tertiary structure

B1.2.11—Quaternary structure of non-conjugated and conjugated proteins

Include insulin and collagen as examples of non-conjugated proteins and haemoglobin as an example of a conjugated protein.

- **Quaternary structure** is the 3D structure of proteins involving **two or more polypeptide chains**
- Proteins can be conjugated or non-conjugated
  - **Conjugated proteins** contain **prosthetic groups** (non-amino acid chemicals)
    - **Hemoglobin** is composed of 4 polypeptides chains & a **heme group** (non-amino acid)
  - **Non-conjugated proteins** do not contain a prosthetic group
    - **Collagen** is composed of 3 helical polypeptide chains that combine to form a triple helix
    - **Insulin** is composed of 2 polypeptide chains linked by disulfide bridges
- Not all conjugated or non-conjugated proteins have a quaternary structure

B1.2.12—Relationship of form and function in globular and fibrous proteins

Students should know the difference in shape between globular and fibrous proteins and understand that their shapes make them suitable for specific functions. Use insulin and collagen to exemplify how form and function are related.

- Proteins can be classified according to their patterns of folding as **fibrous** or **globular**
- **Insulin** is a globular protein
  - Post-translational modifications result in 2 distinct polypeptide chains linked by disulfide bridges
  - The disulfide bridges affect insulin's ability to bind to its receptor
- **Collagen** is a fibrous protein
  - Its amino acid sequence contains a **non-polar glycine** amino acid at **every third position**, giving rise to a **regular primary structure** that allows for helical folding into a fibrous structure
  - Multiple collagen units join together to form **collagen fibrils**, which are very abundant in the extracellular matrix to help structurally support cells and tissues

Property	Fibrous proteins	Globular proteins
Form	Elongated, narrow	Spherical, compact
Function	Structural (support, strength)	Functional (catalyst, transport)
Primary structure	Repetitive	Irregular
Solubility in water	Generally insoluble	Generally soluble
Stability	Less sensitive to heat/pH	More sensitive to heat/pH

NOS: Technology allows imaging of structures that would be impossible to observe with the unaided senses. For example, cryogenic electron microscopy has allowed imaging of single-protein molecules and their interactions with other molecules.

- **Cryogenic electron microscopy (cryo-EM)** enables the study of protein form & function
- Protein samples are frozen at low temperatures & observed under an electron microscope
  - Low temperatures reduce the damage caused by electron radiation
- This allowed imaging of single-protein molecules and their interactions with other molecules

Linking questions

• How do abiotic factors influence the form of molecules?

• What is the relationship between the genome and the proteome of an organism?

References

Ann Clark, Mary, et al. *Biology 2e*. E-book, OpenStax, 2018, https://openstax.org/books/biology-2e/pages/1-introduction. OpenStax.

Billett HH. Hemoglobin and Hematocrit. In: Walker HK, Hall WD, Hurst JW, editors. Clinical Methods: The History, Physical, and Laboratory Examinations. 3rd edition. Boston: Butterworths; 1990. Chapter 151. Available from: https://www.ncbi.nlm.nih.gov/books/NBK259/

Gordon Betts, J., et al. *Anatomy and Physiology 2e*. E-book, OpenStax, 2022, https://openstax.org/books/anatomy-and-physiology-2e/pages/1-introduction. OpenStax.

Sun Z. Foundations for the Study of Structure and Function of Proteins. Basics of Bioinformatics. 2013;303-336. Published 2013 Jul 1. doi:10.1007/978-3-642-38951-1_10

Utiger, Robert D.. "insulin". Encyclopedia Britannica, 3 Feb. 2025, https://www.britannica.com/science/insulin.

Wu M, Cronin K, Crane JS. Biochemistry, Collagen Synthesis. [Updated 2023 Sep 4]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK507709/