



Proteins

Revision notes · IB Biology B1.2: amino acids, peptide bonds, diet, denaturation and, at HL, primary to quaternary structure and globular and fibrous proteins.

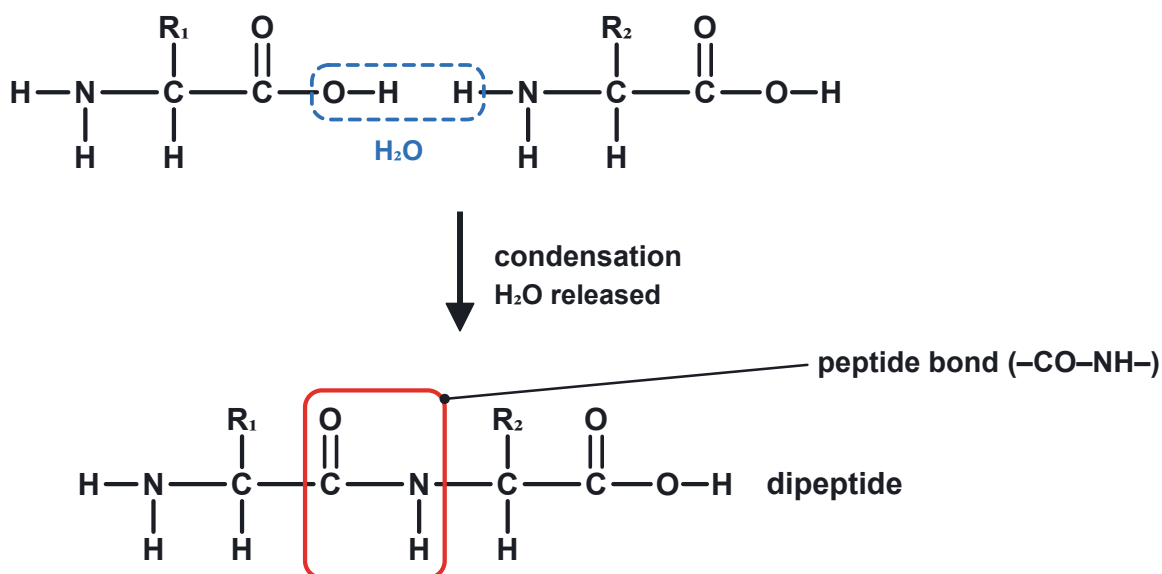
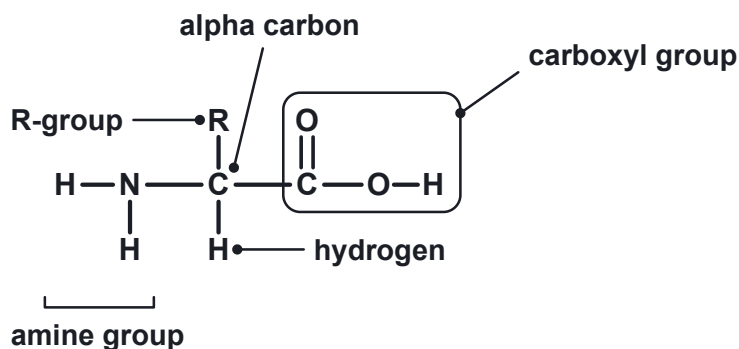
For IB Diploma Biology, Standard Level and Higher Level (first assessment 2025). **HL only** is additional higher level content, for HL students only.

Proteins Spec B1.2

Amino acids and peptide bonds Spec B1.2.1, B1.2.2

Proteins are made of one or more polypeptides, which are chains of amino acids. Every amino acid has the same basic structure, built around a central carbon called the **alpha carbon**.

- an **amine group** (-NH_2)
- a **carboxyl group** (-COOH)
- a hydrogen atom (-H)
- an **R-group**, which differs between the 20 amino acids



[\(opens full size in a new tab\)](#)

Every amino acid has an **amine group**, a **carboxyl group**, H and an **R-group** on the alpha carbon; condensation joins two by a **peptide bond**.

Exam tip: When drawing an amino acid, show every bond to the alpha carbon and write the groups in full: $\text{H}_2\text{N}-$, $-\text{COOH}$, $-\text{H}$ and R . Marks are lost for a missing hydrogen on the alpha carbon.

Forming dipeptides and polypeptides

The carboxyl group of one amino acid reacts with the amine group of the next in a **condensation reaction**. An $-\text{OH}$ and an $-\text{H}$ are removed as water, and a **peptide bond** ($-\text{CO}-\text{NH}-$) forms between the carbon and the nitrogen.



Further condensation reactions add amino acids one at a time to form a polypeptide. In cells this happens on ribosomes during translation. A polypeptide still has a free amine group at one end and a free carboxyl group at the other.

Practical skill: Build two amino acids with molecular models, remove $-\text{OH}$ and $-\text{H}$ to make water, then join the C and N. Drawing the dipeptide afterwards helps you place the peptide bond correctly.

Dietary amino acids and the variety of peptide chains Spec B1.2.3, B1.2.4

Dietary requirements

Essential amino acids	Amino acids that cannot be synthesised and must be obtained from food.
Non-essential amino acids	Amino acids that can be made from other amino acids.

Proteins in food are hydrolysed into amino acids, which are absorbed and used to build the body's own proteins. If any essential amino acid is missing from the diet, proteins that contain it cannot be made.

Some plant proteins contain only small amounts of certain essential amino acids. **Vegan diets** therefore need attention: eating a variety of plant protein sources, such as legumes together with cereals, ensures that every essential amino acid is consumed.

An infinite variety of peptide chains

- **20 amino acids** are coded for in the genetic code.
- A peptide chain can contain any number of amino acids, from a few to thousands.
- The amino acids can be in any order.

For a chain of n amino acids there are 20^n possible sequences. Even short chains have a vast number of possibilities, so the variety of possible peptide chains is **effectively infinite**. Only a small fraction are actually made by living things.

Counting possible peptide sequences

How many different tripeptides could be made from the 20 amino acids? How many sequences are possible for a chain of 10 amino acids?

1. Each position can be any of 20 amino acids, so the number of sequences is 20^n .
2. Tripeptide: $20^3 = 20 \times 20 \times 20 = 8000$.
3. Ten amino acids: $20^{10} \approx 1.0 \times 10^{13}$.

Answer: 8000 tripeptides; about 1.0×10^{13} sequences of 10 amino acids

Examples of polypeptides

Polypeptide	Size	Function
Insulin	Two chains of 21 and 30 amino acids	Hormone that lowers blood glucose
Beta chain of haemoglobin	146 amino acids	Part of the oxygen-carrying protein in red blood cells
Titin	Up to about 34 000 amino acids, the largest known polypeptide	Gives elasticity to muscle

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

A protein's function depends on its precise three-dimensional shape. That shape is held by bonds between parts of the chain, mostly weak ones such as hydrogen bonds and ionic bonds.

Denaturation	A change to the three-dimensional structure (conformation) of a protein so that it can no longer carry out its function. It is usually permanent.
---------------------	---

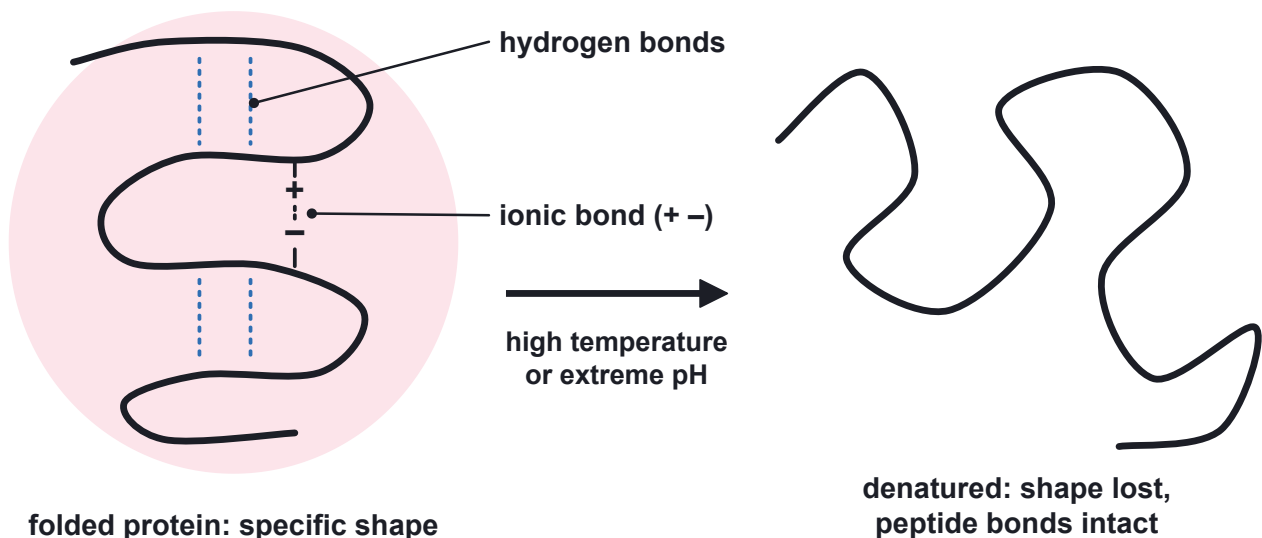
Temperature

Heat makes the atoms in a protein vibrate more. Above a certain temperature, the vibrations break the **hydrogen bonds** and other weak interactions that hold the shape, and the protein unfolds. The temperature at which this happens varies: proteins from organisms living in hot springs stay stable at temperatures that denature most human proteins.

pH

Some R-groups contain amine or carboxyl groups that can gain or lose hydrogen ions. A change in pH changes the **charges on R-groups**, so ionic bonds between them break and new attractions or repulsions form. Each protein is stable over a particular pH range: for example, pepsin works in the acidic stomach while most cytoplasmic proteins need a pH near 7.

- Peptide bonds are not broken, so the amino acid sequence stays the same.
- Soluble proteins often become insoluble when denatured, for example egg white turning solid when heated.
- An enzyme that is denatured loses the shape of its active site, so it can no longer bind its substrate.



[\(opens full size in a new tab\)](#)

Heat and extreme pH break the **hydrogen and ionic bonds** that hold a protein's shape; the **peptide bonds stay intact**.

Practical skill: When testing the effect of temperature or pH on a protein such as an enzyme, change one independent variable, keep substrate concentration, volume and time constant, and repeat each condition at least three times.

Exam tip: Linking question: how do abiotic factors influence the form of molecules? Temperature and pH are abiotic factors that change protein shape, which limits where organisms can live.

R-groups and primary structure Spec B1.2.6, B1.2.7 HL only

Chemical diversity of R-groups

All amino acids share the same backbone, so the differences between them are entirely in their **R-groups**. R-groups vary in size, shape and chemistry. This variety is the basis for the immense diversity in protein form and function.

Types of R-group		
Type	Property	Behaviour in a protein
Hydrophobic (non-polar)	No charge, no hydrogen bonding with water	Tend to be buried away from water or in membranes
Hydrophilic: polar	Uncharged but polar, can form hydrogen bonds	Tend to be on the surface, interacting with water
Hydrophilic: charged, acidic	Carboxyl group loses H^+ and becomes negative	Can form ionic bonds
Hydrophilic: charged, basic	Amine group gains H^+ and becomes positive	Can form ionic bonds

R-groups determine the properties of the assembled polypeptide: how it folds, whether it is soluble, which molecules it binds and how it catalyses reactions.

Primary structure determines conformation

Primary structure	The sequence of amino acids in a polypeptide.
Conformation	The three-dimensional shape of a protein.

The **sequence of amino acids** and the precise position of each amino acid in the chain determine how the polypeptide folds, and so its three-dimensional shape. A polypeptide with a given sequence folds into the same shape every time. Proteins therefore have **precise, predictable and repeatable structures**, despite their complexity.

Changing even one amino acid can change the conformation. In sickle cell anaemia a single substitution in the beta chain of haemoglobin changes how the molecules interact, distorting red blood cells.

Exam tip: Linking question: what is the relationship between genome and proteome? The base sequence of a gene codes for the amino acid sequence, which determines conformation and so function.

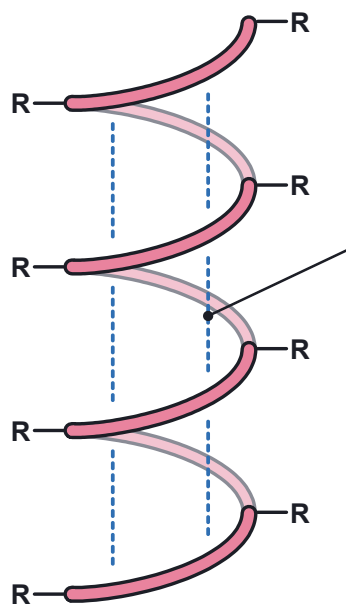
Secondary structure: alpha helices and beta-pleated sheets Spec B1.2.8 HL only

The backbone of a polypeptide contains a repeating $-N-C-C-$ pattern. Each peptide bond has a $C=O$ group, which is slightly negative at the oxygen, and an $N-H$ group, which is slightly positive at the hydrogen. These can form **hydrogen bonds** with each other.

Because peptide bonds occur at regular intervals, hydrogen bonds form **in regular positions**, giving two repeating shapes. R-groups are not involved; they point outwards from these structures.

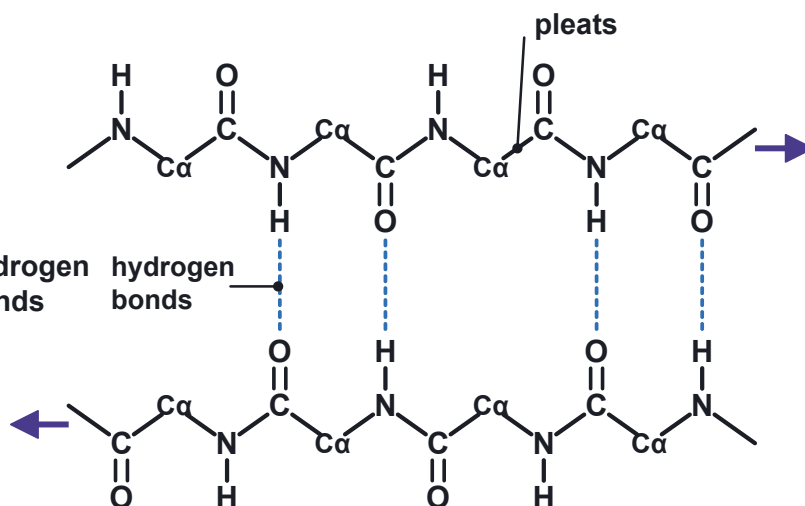
Secondary structures		
Feature	Alpha helix	Beta-pleated sheet
Shape	Chain coils into a spiral	Two or more stretches of chain lie side by side, folded into pleats
Hydrogen bonds	Between $C=O$ of one amino acid and $N-H$ of the amino acid four positions further along the same stretch	Between $C=O$ and $N-H$ on neighbouring stretches of chain
Orientation of bonds	Roughly parallel to the axis of the helix	Roughly at right angles to the stretches

Alpha helix



C=O to N-H four amino acids along; R-groups out

Beta-pleated sheet



Arrows show the direction of each stretch of chain. H and R-groups on each C α are not shown; R-groups point above and below the sheet.

[\(opens full size in a new tab\)](#)

Secondary structure: ==hydrogen bonds between C=O and N-H== of the backbone give the alpha helix and the beta-pleated sheet.

Common mistake: Secondary structure hydrogen bonds are between backbone groups of peptide bonds, not between R-groups. Bonds between R-groups belong to tertiary structure.

Tertiary structure and the effect of polar and non-polar amino acids Spec B1.2.9, B1.2.10

HL only

A polypeptide folds further into a specific three-dimensional shape, its tertiary structure. This is held by interactions between R-groups, which may be far apart in the sequence but end up close together after folding.

Interactions in tertiary structure		
Interaction	Between	Strength
Hydrogen bonds	Polar R-groups	Weak, but many of them
Ionic bonds	A positively charged R-group (-NH_3^+) and a negatively charged R-group (-COO^-)	Stronger than hydrogen bonds; sensitive to pH
Disulfide bonds	Sulfur atoms of two cysteine R-groups (-S-S-)	Covalent, so the strongest
Hydrophobic interactions	Non-polar R-groups clustering together away from water	Weak individually, important overall

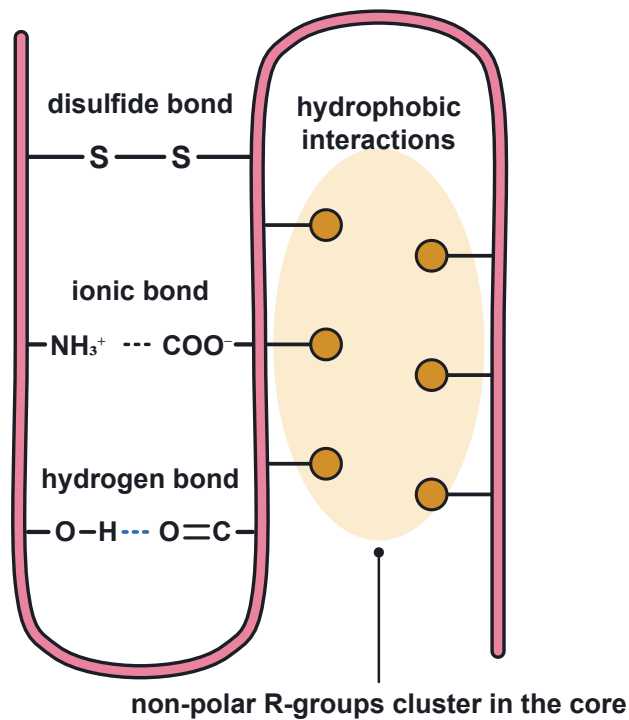
Amine and carboxyl groups in R-groups become charged by binding or dissociation of hydrogen ions: -NH_2 gains H^+ to become -NH_3^+ , and -COOH loses H^+ to become -COO^- . Opposite charges then attract to form ionic bonds. This is why pH affects tertiary structure.

A **disulfide bond** is a covalent bond between the sulfur atoms of a pair of cysteines. It forms between cysteines in the same chain or in different chains.

Polar and non-polar amino acids

- In proteins that are soluble in water, **hydrophobic amino acids are clustered in the core** of the globular protein, away from water. Polar and charged amino acids are on the surface, where they interact with water and keep the protein soluble.
- Integral membrane proteins have regions of hydrophobic amino acids that sit in the hydrophobic core of the phospholipid bilayer, helping them to embed in the membrane. Regions of polar amino acids face the water on either side.

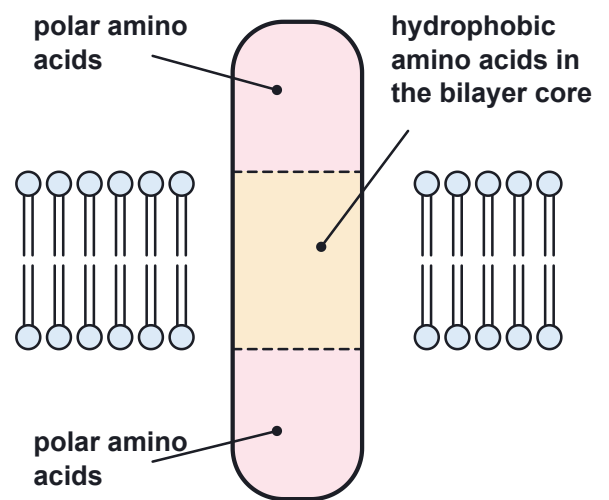
Tertiary structure



[\(opens full size in a new tab\)](#)

Tertiary structure is held by **interactions between R-groups**; in membrane proteins, **hydrophobic amino acids** sit in the bilayer core.

Integral membrane protein



Quaternary structure: non-conjugated and conjugated proteins Spec B1.2.11 HL only

Many proteins consist of a single polypeptide. Others are made of two or more polypeptides linked together; this arrangement is **quaternary structure**. The chains are held together by the same types of interaction as tertiary structure.

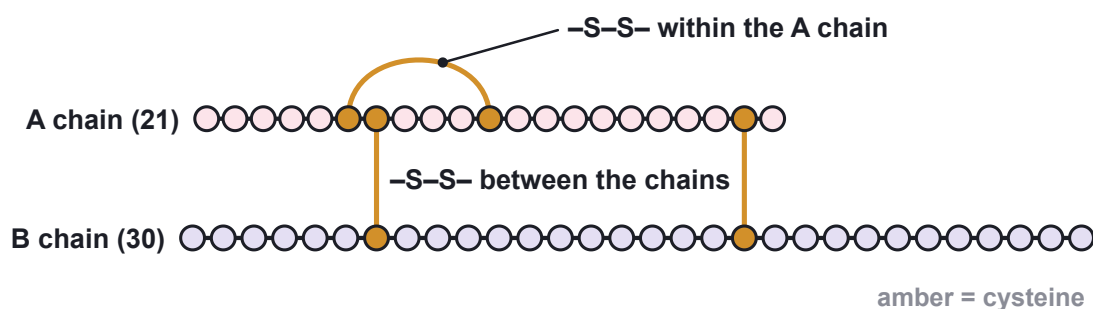
Non-conjugated protein	A protein made only of polypeptides.
Conjugated protein	A protein with one or more non-polypeptide components (prosthetic groups) as well as polypeptides.

Examples of quaternary structure			
Protein	Polypeptides	Non-polypeptide part	Type
Insulin	Two: an A chain (21 amino acids) and a B chain (30 amino acids) linked by two disulfide bonds	None	Non-conjugated
Collagen	Three, wound together into a triple helix	None	Non-conjugated
Haemoglobin	Four: two alpha and two beta chains	Four haem groups, each containing an iron ion that binds one O ₂	Conjugated

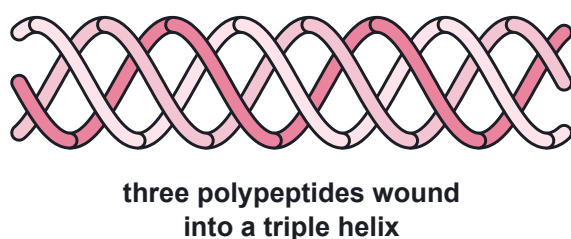
In **haemoglobin** each polypeptide holds one haem group, so one molecule carries up to four oxygen molecules. The haem is not made of amino acids, which is what makes haemoglobin **conjugated**. **Insulin** and **collagen** are made only of polypeptide chains.

Insulin is first made as one long polypeptide, proinsulin. A middle section is then cut out, leaving two chains that are held together by disulfide bonds. The IB guide uses insulin's two linked chains as an example of quaternary structure.

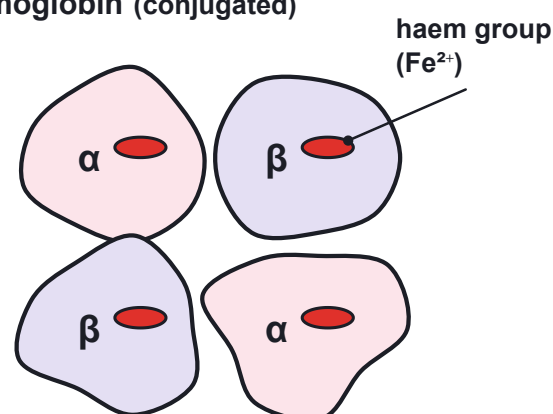
Insulin (non-conjugated)



Collagen (non-conjugated)



Haemoglobin (conjugated)



[\(opens full size in a new tab\)](#)

Quaternary structure: insulin (2 chains) and collagen (3 chains) are **non-conjugated**; haemoglobin has 4 chains plus **haem groups**, so it is **conjugated**.

Exam tip: Nature of science: technology lets us see what the unaided senses cannot. Cryogenic electron microscopy rapidly freezes protein samples and combines many images of single molecules into 3D structures, showing proteins and their interactions with other molecules.

Form and function in globular and fibrous proteins

Spec B1.2.12

HL only

Globular and fibrous proteins		
Feature	Globular	Fibrous
Shape	Rounded and compact, folded into a ball	Long and narrow, in strands
Solubility in water	Usually soluble: hydrophilic R-groups on the surface	Insoluble
Amino acid sequence	Irregular, giving a precise unique shape	Often repetitive
Typical functions	Hormones, enzymes, transport (e.g. haemoglobin), receptors	Structural support and strength
Example	Insulin	Collagen

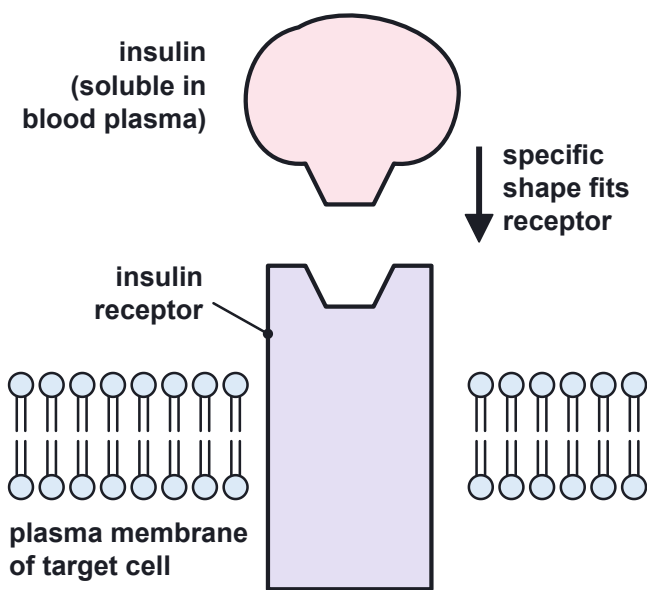
Insulin: a globular protein

- Insulin is a hormone secreted by beta cells of the pancreas when blood glucose is high.
- It is **soluble**, because hydrophilic R-groups are on its surface, so it is carried dissolved in blood plasma.
- Its precise globular shape fits a specific **insulin receptor** on the plasma membrane of target cells, such as muscle, adipose and liver cells. Binding triggers responses that lower blood glucose, for example uptake of glucose by muscle and adipose cells.
- Disulfide bonds between its two chains keep the shape stable.

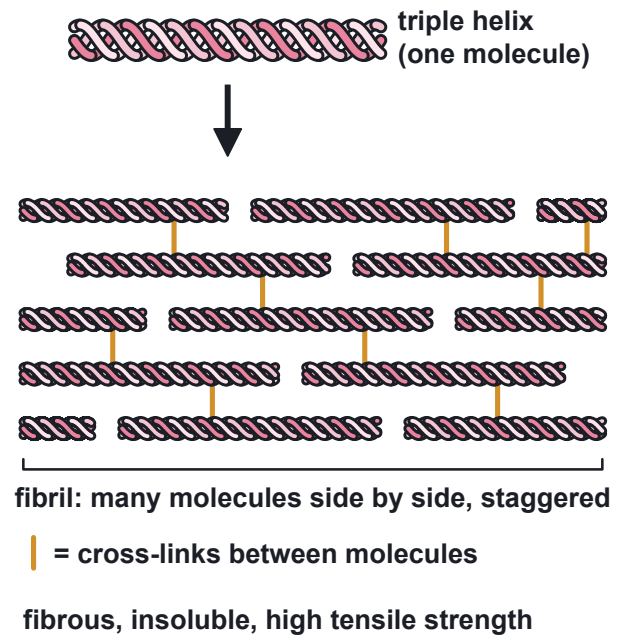
Collagen: a fibrous protein

- Three polypeptides are wound into a **triple helix**, held together by hydrogen bonds between the chains.
- Every third amino acid is glycine, whose very small R-group lets the three chains pack tightly.
- Collagen molecules line up side by side and are cross-linked into fibrils, giving very **high tensile strength** with little stretch.
- It is found in skin, tendons, ligaments, bone and the walls of blood vessels, where it resists pulling forces.

Insulin: globular



Collagen: fibrous



[\(opens full size in a new tab\)](#)

Insulin is **globular** and soluble, with a shape that fits its receptor; collagen is **fibrous**, with cross-linked molecules forming strong fibrils.

Exam tip: In form-and-function answers, pair every structural feature with what it allows: for example, triple helix with cross-links, so high tensile strength, so tendons do not stretch or tear.

Quick check

1. Which four groups are bonded to the alpha carbon of an amino acid?
2. What is an essential amino acid?
3. What is denaturation?
4. **HL only** Which bonds stabilise alpha helices and beta-pleated sheets?
5. **HL only** Why is haemoglobin a conjugated protein?

Quick check answers

Proteins

1. An amine group, a carboxyl group, a hydrogen atom and an R-group.
2. One that cannot be synthesised by the body and must be obtained from food.
3. A usually permanent change in the three-dimensional structure of a protein, so it can no longer carry out its function.
4. Hydrogen bonds in regular positions between C=O and N-H groups of the backbone.
5. It contains non-polypeptide haem groups as well as its four polypeptide chains.