Concepts to remember

Primary structure

The simplest level of protein structure, **primary structure**, is simply the sequence of amino acids in a polypeptide chain. For example, the hormone insulin has two polypeptide chains, A and B, shown in diagram below. (The insulin molecule shown here is cow insulin, although its structure is similar to that of human insulin.) Each chain has its own set of amino acids, assembled in a particular order. For instance, the sequence of the A chain starts with glycine at the N-terminus and ends with asparagine at the C-terminus, and is different from the sequence of the B chain.

[What's up with those S-S bonds?]

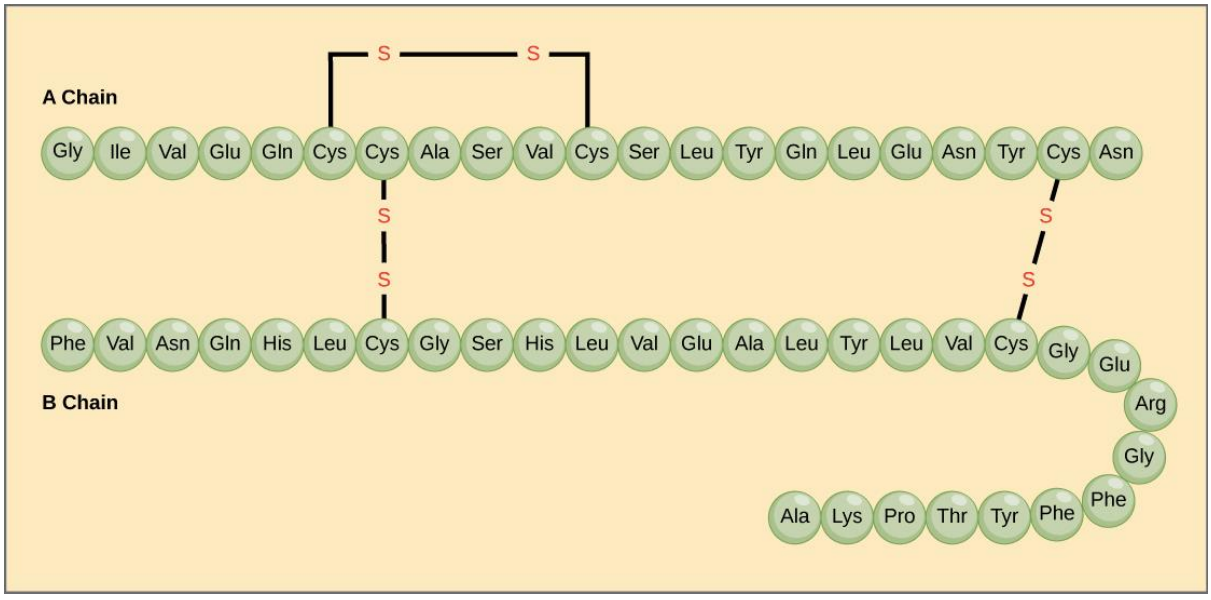


Image of insulin. Insulin consists of an A chain and a B chain. They are connected to one another by disulfide bonds (sulfur-sulfur bonds between cysteines). The A chain also contains an internal disulfide bond. The amino acids that make up each chain of insulin are represented as connected circles, each with the three-letter abbreviation of the amino acid's name.

image credit: OpenStax Biology.

The sequence of a protein is determined by the DNA of the gene that encodes the protein (or that encodes a portion of the protein, for multi-subunit proteins). A change in the gene's DNA sequence may lead to a change in the amino acid sequence of the protein. Even changing just one amino acid in a protein's sequence can affect the protein's overall structure and function.

For instance, a single amino acid change is associated with sickle cell anemia, an inherited disease that affects red blood cells. In sickle cell anemia, one of the polypeptide chains that make up hemoglobin, the protein that carries oxygen in the blood, has a slight sequence change. The glutamic acid that is normally the sixth amino acid of the hemoglobin β chain (one of two types of protein chains that make up hemoglobin) is replaced by a valine. This substitution is shown for a fragment of the β chain in the diagram below.

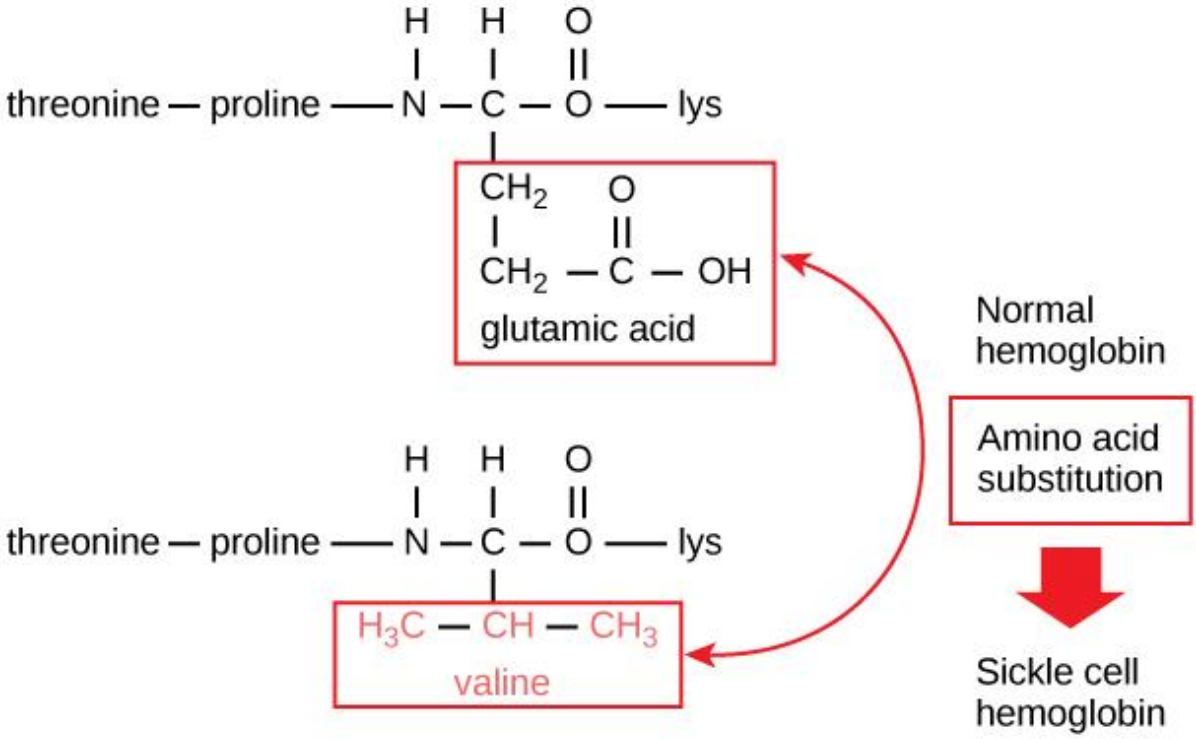
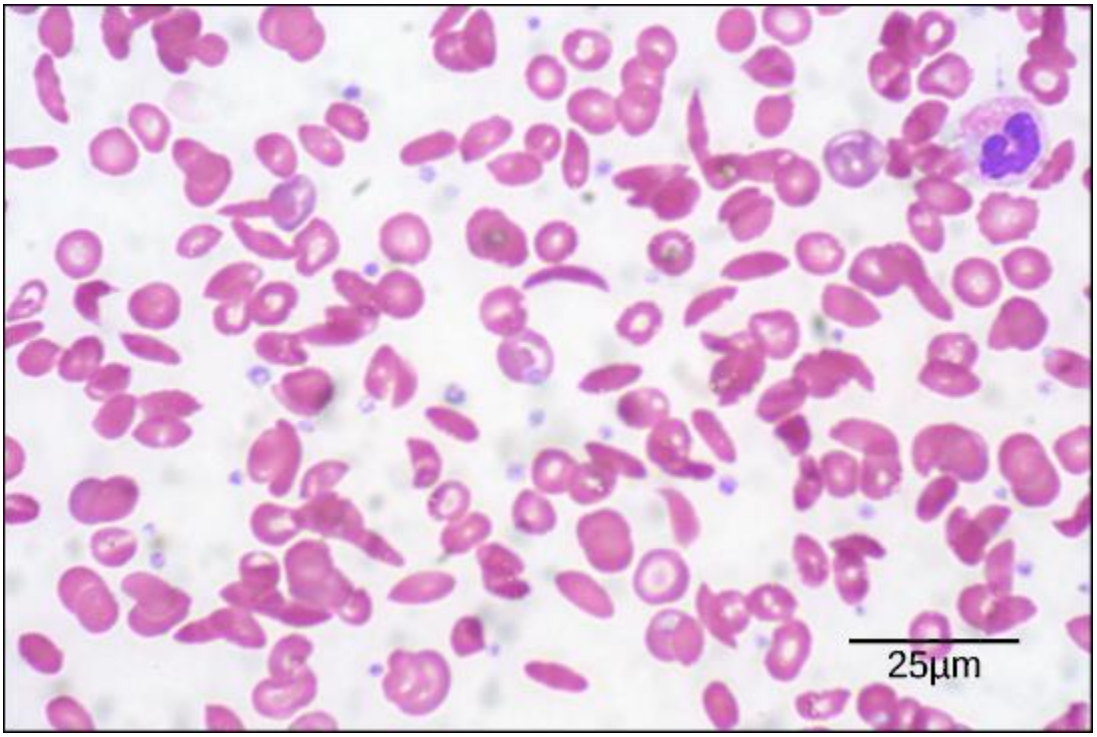


Image of normal and sickle cell mutant hemoglobin chains, showing substitution of valine for glutamic acid in the sickle cell version.

Image modified from OpenStax Biology.

What is most remarkable to consider is that a hemoglobin molecule is made up of two α chains and two β chains, each consisting of about 150 amino acids, for a total of about 600 amino acids in the whole protein. The difference between a normal hemoglobin molecule and a sickle cell molecule is just 2 amino acids out of the approximately 600.

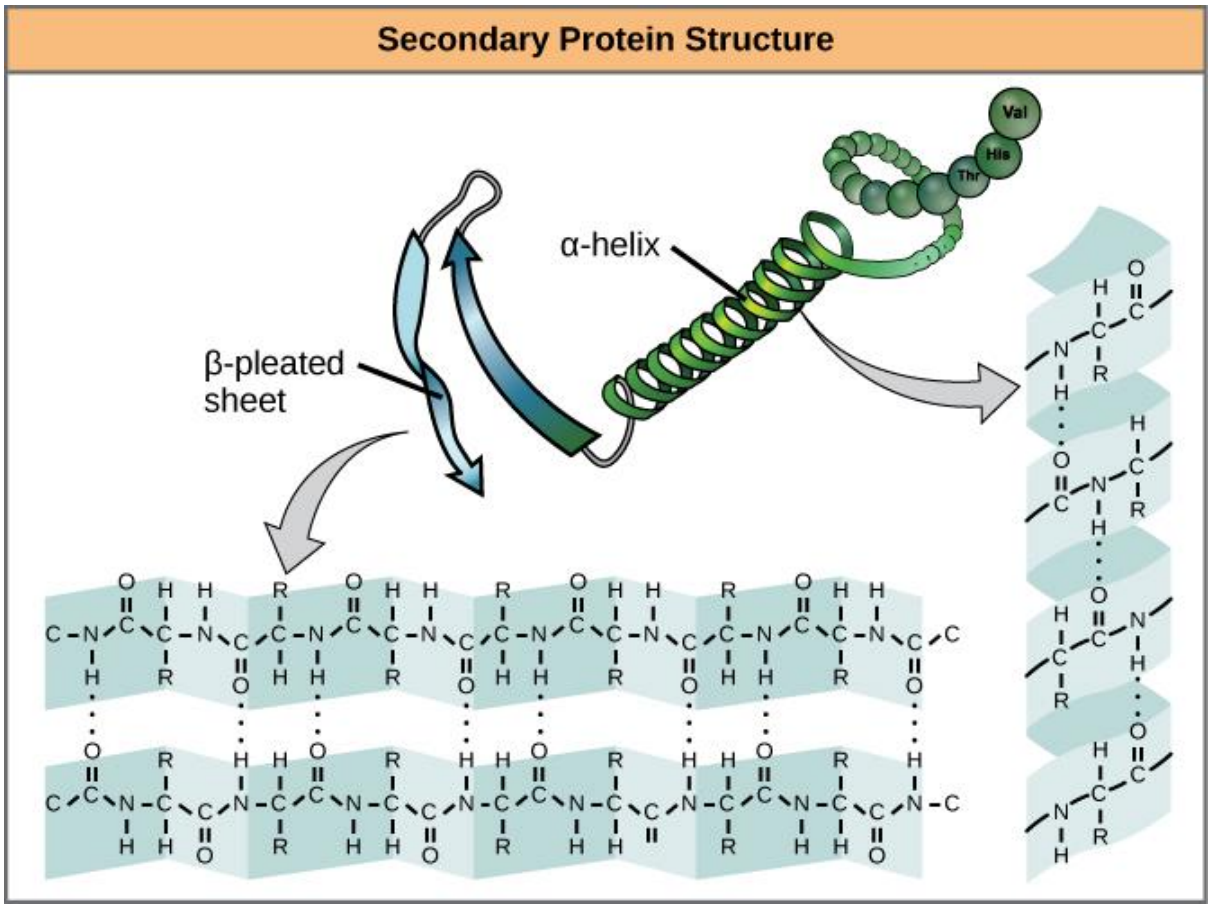
A person whose body makes only sickle cell hemoglobin will suffer symptoms of sickle cell anemia. These occur because the glutamic acid-to-valine amino acid change makes the hemoglobin molecules assemble into long fibers. The fibers distort disc-shaped red blood cells into crescent shapes. Examples of "sickled" cells can be seen mixed with normal, disc-like cells in the blood sample below.



The sickled cells get stuck as they try to pass through blood vessels. The stuck cells impair blood flow and can cause serious health problems for people with sickle cell anemia, including breathlessness, dizziness, headaches, and abdominal pain.

Secondary structure

The next level of protein structure, **secondary structure**, refers to local folded structures that form within a polypeptide due to interactions between atoms of the backbone. (The backbone just refers to the polypeptide chain apart from the R groups – so all we mean here is that secondary structure does not involve R group atoms.) The most common types of secondary structures are the α helix and the β pleated sheet. Both structures are held in shape by hydrogen bonds, which form between the carbonyl O of one amino acid and the amino H of another.



Images showing hydrogen bonding patterns in beta pleated sheets and alpha helices.

In an **α helix**, the carbonyl (C=O) of one amino acid is hydrogen bonded to the amino H (N-H) of an amino acid that is four down the chain. (E.g., the carbonyl of amino acid 1 would form a hydrogen bond to the N-H of amino acid 5.) This pattern of bonding pulls the polypeptide chain into a helical structure that resembles a curled ribbon, with each turn of the helix containing 3.6 amino acids. The R groups of the amino acids stick outward from the α helix, where they are free to interact^3cubed.

In a **β pleated sheet**, two or more segments of a polypeptide chain line up next to each other, forming a sheet-like structure held together by hydrogen bonds. The hydrogen bonds form between carbonyl and amino groups of backbone, while the R groups extend above and below the plane of the sheet^3cubed. The strands of a β pleated sheet may be **parallel**, pointing in the same direction (meaning that their N- and C-termini match up), or **antiparallel**, pointing in opposite directions (meaning that the N-terminus of one strand is positioned next to the C-terminus of the other).

Certain amino acids are more or less likely to be found in α-helices or β pleated sheets. For instance, the amino acid proline is sometimes called a “helix breaker” because its unusual R group (which bonds to the amino group to form a ring) creates a bend in the chain and is not compatible with helix formation^4start superscript, 4, end superscript. Proline is typically found in bends, unstructured regions between secondary structures. Similarly, amino acids such as tryptophan, tyrosine, and phenylalanine, which have large ring structures in their R groups, are often found in β pleated sheets, perhaps because the β pleated sheet structure provides plenty of space for the side chains^4start superscript, 4, end superscript.

Many proteins contain both α helices and β pleated sheets, though some contain just one type of secondary structure (or do not form either type).

Tertiary structure

The overall three-dimensional structure of a polypeptide is called its **tertiary structure**. The tertiary structure is primarily due to interactions between the R groups of the amino acids that make up the protein.

R group interactions that contribute to tertiary structure include hydrogen bonding, ionic bonding, dipole-dipole interactions, and London dispersion forces – basically, the whole gamut of non-covalent bonds. For example, R groups with like charges repel one another, while those with opposite charges can form an ionic bond. Similarly, polar R groups can form hydrogen bonds and other dipole-dipole interactions. Also important to tertiary structure are **hydrophobic interactions**, in which amino acids with nonpolar, hydrophobic R groups cluster together on the inside of the protein, leaving hydrophilic amino acids on the outside to interact with surrounding water molecules.

Finally, there's one special type of covalent bond that can contribute to tertiary structure: the disulfide bond. **Disulfide bonds**, covalent linkages between the sulfur-containing side chains of cysteines, are much stronger than the other types of bonds that contribute to tertiary structure. They act like molecular "safety pins," keeping parts of the polypeptide firmly attached to one another.

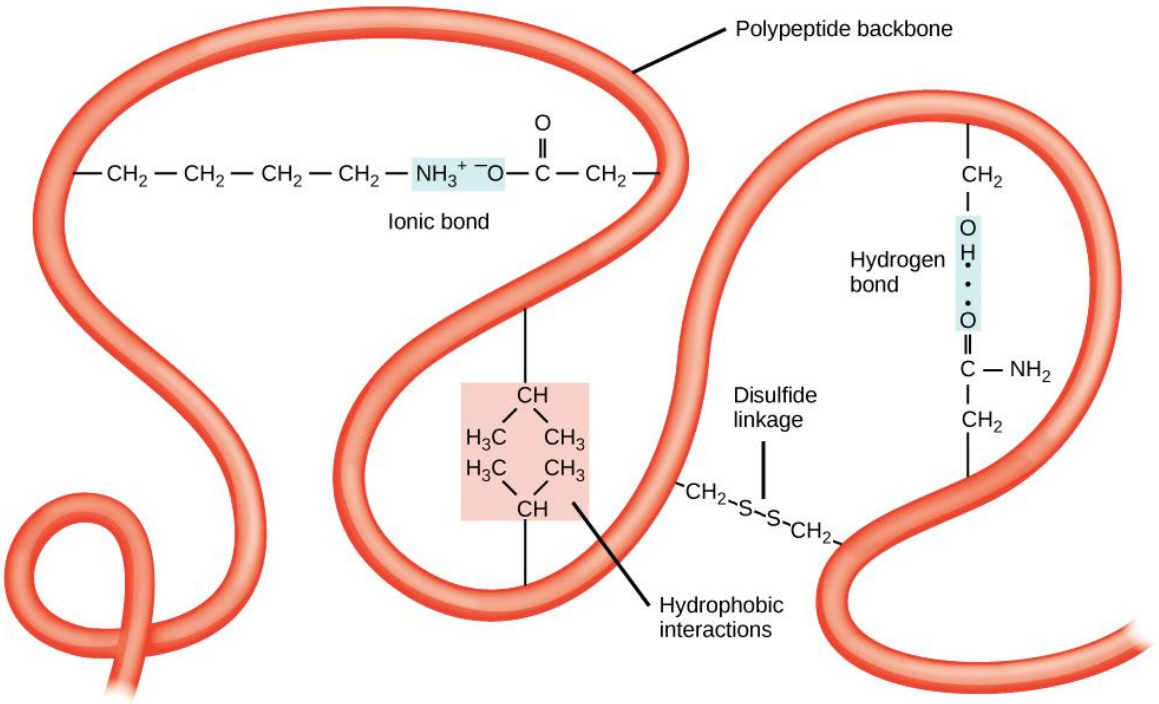


Image of a hypothetical polypeptide chain, depicting different types of side chain interactions that can contribute to tertiary structure. These include hydrophobic interactions, ionic bonds, hydrogen bonds, and disulfide bridge formation.

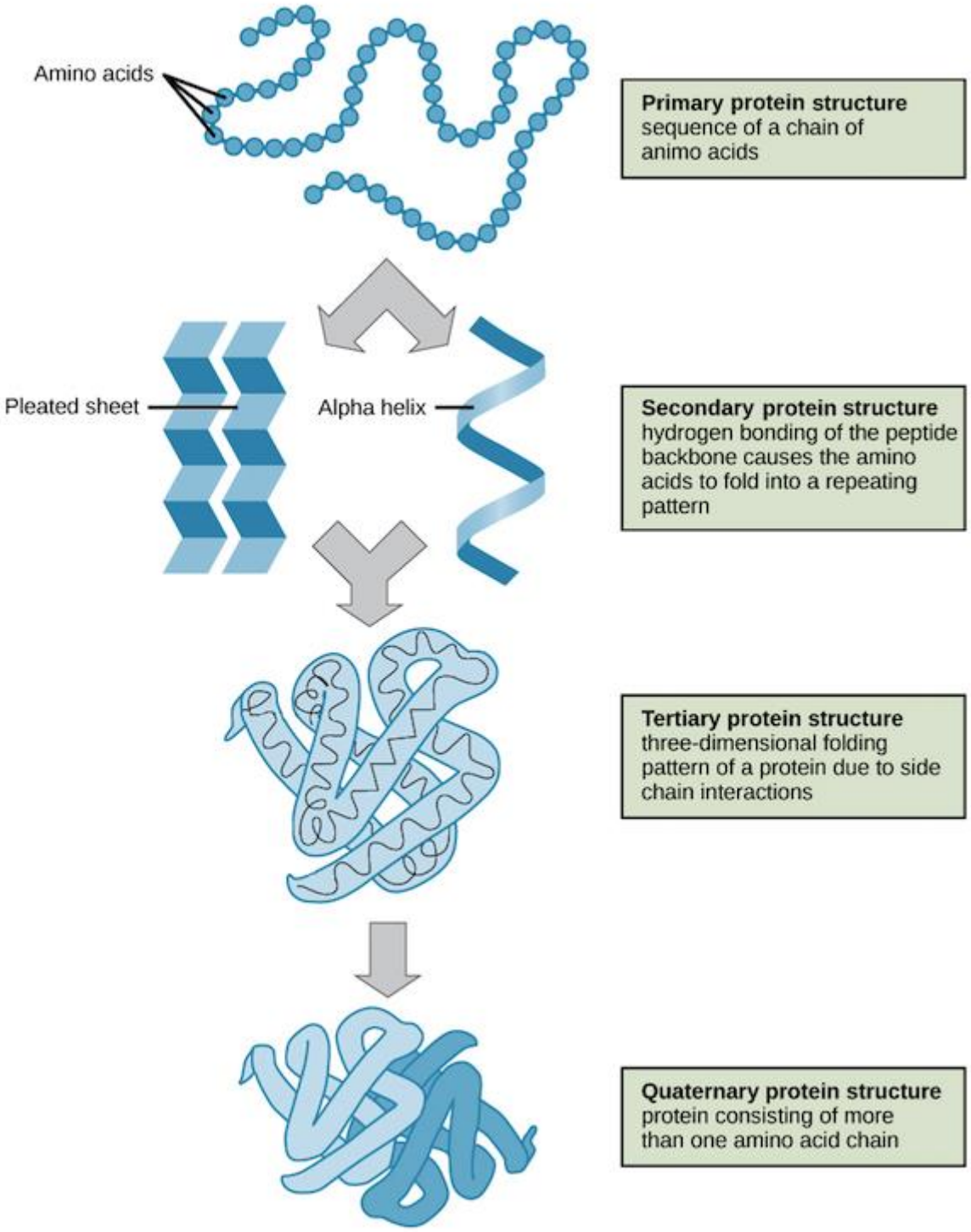
Quaternary structure

Many proteins are made up of a single polypeptide chain and have only three levels of structure (the ones we've just discussed). However, some proteins are made up of multiple polypeptide chains, also known as subunits. When these subunits come together, they give the protein its **quaternary structure**.

We've already encountered one example of a protein with quaternary structure: hemoglobin. As mentioned earlier, hemoglobin carries oxygen in the blood and is made up of four subunits, two each of the α and β types. Another example is DNA polymerase, an enzyme that

synthesizes new strands of DNA and is composed of ten subunits^5start superscript, 5, end superscript.

In general, the same types of interactions that contribute to tertiary structure (mostly weak interactions, such as hydrogen bonding and London dispersion forces) also hold the subunits together to give quaternary structure.



Flowchart depicting the four orders of protein structure