

26.10: PROTEIN STRUCTURE

OBJECTIVES

After completing this section, you should be able to

1. discuss, with reference to a suitable example (either given or of your own choice), the structure of proteins, paying particular attention to distinguishing between the primary, secondary, tertiary and quaternary structure.
2. describe the α -helical secondary structure displayed by many proteins.
3. describe the β -pleated-sheet structure displayed by many proteins.

KEY TERMS

Make certain that you can define, and use in context, the key terms below.

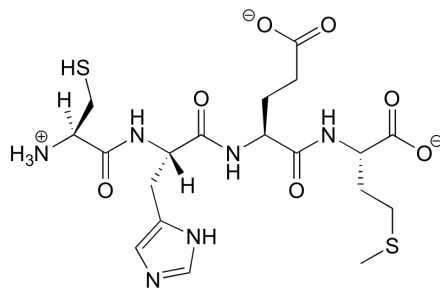
- α helix
- β pleated sheet
- primary structure
- quaternary structure
- secondary structure
- tertiary structure

STUDY NOTES

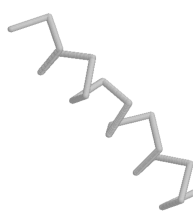
Note that in a diagram of the α -helical structure of a protein, the C-terminal of the protein is at the bottom of the diagram and the N-terminal is at the top. In an α helix, such as the one shown in Figure 26.9.1, the bulky R groups are all found on the outside of the helix, where they have the most room.

THE FOUR LEVELS OF PROTEIN STRUCTURE

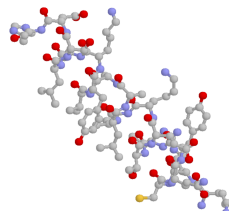
Protein structure can be discussed at four distinct levels. A protein's **primary structure** is two-dimensional - simply the sequence of amino acids in the peptide chain. Below is a Lewis structure of a short segment of a protein with the sequence CHEM (cysteine - histidine - glutamate - methionine)



Secondary structure is three-dimensional, but is a local phenomenon, confined to a relatively short stretch of amino acids. For the most part, there are three important elements of secondary structure: helices, beta-sheets, and loops. In a helix, the main chain of the protein adopts the shape of a clockwise spiral staircase, and the side chains point out laterally.

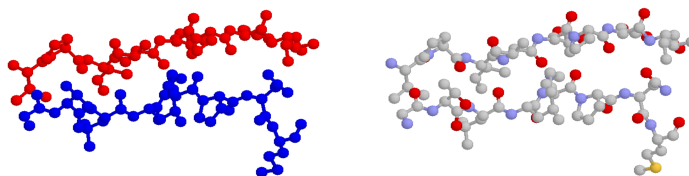


main chain view



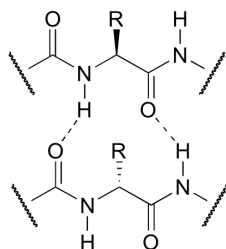
main chain + side chain

In a beta-sheet (or beta-strand) structure, two sections of protein chain are aligned side-by-side in an extended conformation. The figure below shows two different views of the same beta-sheet: in the left-side view, the two regions of protein chain are differentiated by color.



Loops are relatively disordered segments of protein chain, but often assume a very ordered structure when in contact with a second protein or a smaller organic compound.

Both helix and the beta-sheet structures are held together by very specific hydrogen-bonding interactions between the amide nitrogen on one amino acid and the carbonyl oxygen on another. The hydrogen bonding pattern in a section of a beta-strand is shown below.



hydrogen bonding in a β -strand

Secondary structure refers to the shape of a folding protein due exclusively to hydrogen bonding between its backbone amide and carbonyl groups. Secondary structure does not include bonding between the R-groups of amino acids, hydrophobic interactions, or other interactions associated with tertiary structure. The two most commonly encountered secondary structures of a polypeptide chain are α -helices and beta-pleated sheets. These structures are the first major steps in the folding of a polypeptide chain, and they establish important topological motifs that dictate subsequent tertiary structure and the ultimate function of the protein.

A-HELICES

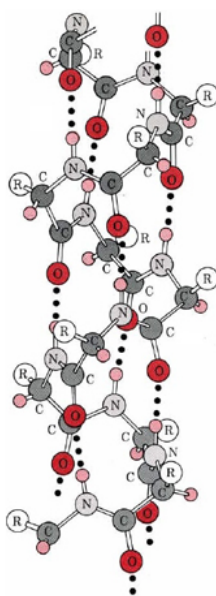


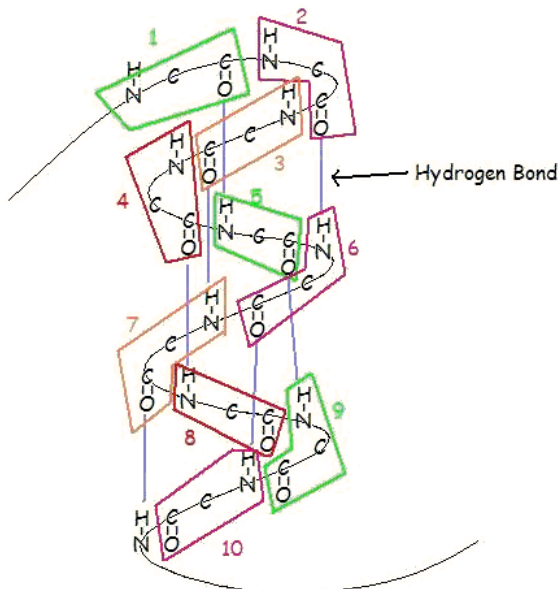
Figure 26.9.1 Ball-and-stick model of the α helix. Hydrogen bonds are shown as dotted bonds. Note that R groups extend almost perpendicular from the axis.

An α -helix is a right-handed coil of amino-acid residues on a polypeptide chain, typically ranging between 4 and 40 residues. This coil is held together by hydrogen bonds between the oxygen of C=O on top coil and the hydrogen of N-H on the bottom coil. Such a hydrogen bond is formed exactly every 4 amino acid residues, and every complete turn of the helix is only 3.6 amino acid residues. This regular pattern gives the α -helix very definite features with regards to the thickness of the coil and the length of each complete turn along the helix axis.

The structural integrity of an α -helix is in part dependent on correct steric configuration. Amino acids whose R-groups are too large (tryptophan, tyrosine) or too small (glycine) destabilize α -helices. Proline also destabilizes α -helices because of its irregular geometry; its R-

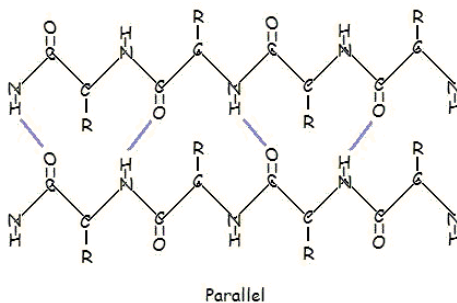
group bonds back to the nitrogen of the amide group, which causes steric hindrance. In addition, the lack of a hydrogen on Proline's nitrogen prevents it from participating in hydrogen bonding.

Another factor affecting α -helix stability is the total dipole moment of the entire helix due to individual dipoles of the C=O groups involved in hydrogen bonding. Stable α -helices typically end with a charged amino acid to neutralize the dipole moment.

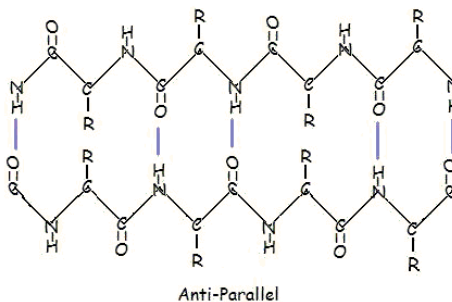


BETA-PLEATED SHEETS

This structure occurs when two (or more, e.g. ψ -loop) segments of a polypeptide chain overlap one another and form a row of hydrogen bonds with each other. This can happen in a parallel arrangement:



Or in anti-parallel arrangement:



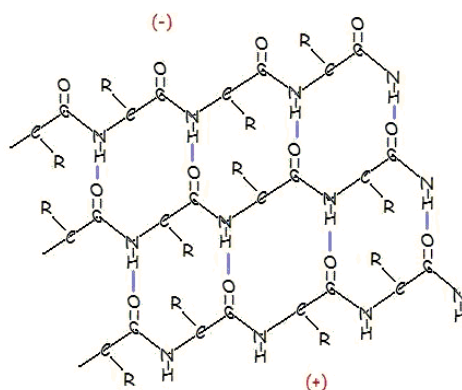
Parallel and anti-parallel arrangement is the direct consequence of the directionality of the polypeptide chain. In anti-parallel arrangement, the C-terminus end of one segment is on the same side as the N-terminus end of the other segment. In parallel arrangement, the C-terminus end and the N-terminus end are on the same sides for both segments. The "pleat" occurs because of the alternating planes of the peptide bonds between amino acids; the aligned amino and carbonyl group of each opposite segment alternate their orientation from facing towards each other to facing opposite directions.

The parallel arrangement is less stable because the geometry of the individual amino acid molecules forces the hydrogen bonds to occur at an angle, making them longer and thus weaker. Contrarily, in the anti-parallel arrangement the hydrogen bonds are aligned directly opposite each other, making for stronger and more stable bonds.

Commonly, an anti-parallel beta-pleated sheet forms when a polypeptide chain sharply reverses direction. This can occur in the presence of two consecutive proline residues, which create an angled kink in the polypeptide chain and bend it back upon itself. This is not necessary for distant segments of a polypeptide chain to form beta-pleated sheets, but for proximal segments it is a definite requirement. For short distances, the two segments of a beta-pleated sheet are separated by $4+2n$ amino acid residues, with 4 being the minimum number of residues.

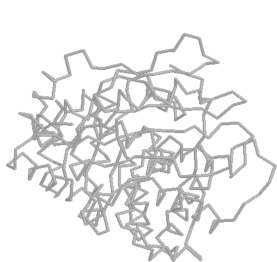
A-PLEATED SHEETS

A similar structure to the beta-pleated sheet is the α -pleated sheet. This structure is energetically less favorable than the beta-pleated sheet, and is fairly uncommon in proteins. An α -pleated sheet is characterized by the alignment of its carbonyl and amino groups; the carbonyl groups are all aligned in one direction, while all the N-H groups are aligned in the opposite direction. The polarization of the amino and carbonyl groups results in a net dipole moment on the α -pleated sheet. The carbonyl side acquires a net negative charge, and the amino side acquires a net positive charge.

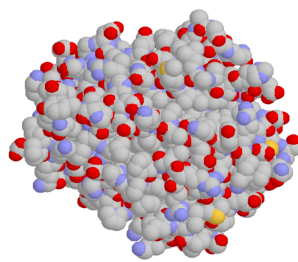


Alpha-Pleated Sheet

A protein's **tertiary structure** is the shape in which the entire protein chain folds together in three-dimensional space, and it is this level of structure that provides protein scientists with the most information about a protein's specific function.

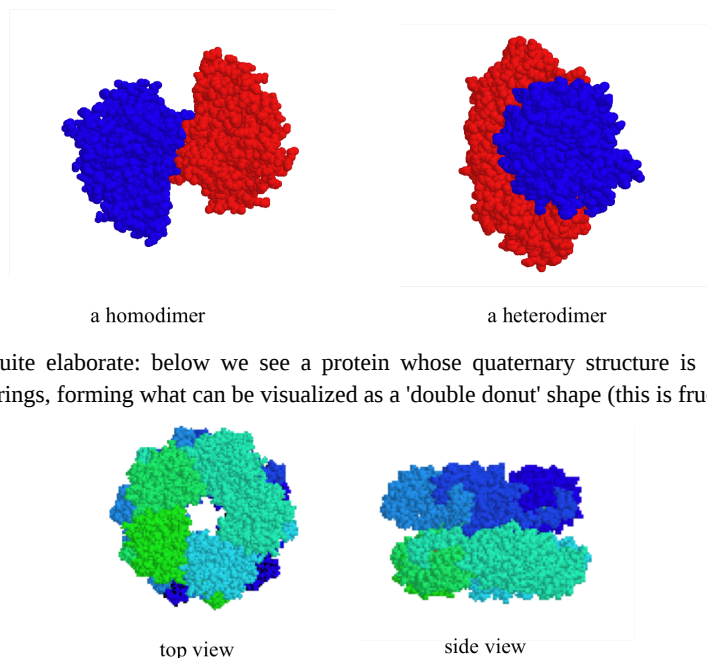


folded protein chain
(main chain view)



folded protein chain
(‘space-filling’ view)

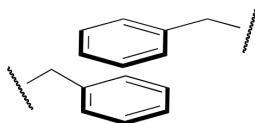
While a protein's secondary and tertiary structure is defined by how the protein chain folds together, **quaternary structure** is defined by how two or more folded protein chains come together to form a 'superstructure'. Many proteins consist of only one protein chain, or **subunit**, and thus have no quaternary structure. Many other proteins consist of two identical subunits (these are called homodimers) or two non-identical subunits (these are called heterodimers).



Quaternary structures can be quite elaborate: below we see a protein whose quaternary structure is defined by ten identical subunits arranged in two five-membered rings, forming what can be visualized as a 'double donut' shape (this is fructose 1,6-bisphosphate aldolase):

THE MOLECULAR FORCES THAT HOLD PROTEINS TOGETHER

The question of exactly how a protein 'finds' its specific folded structure out of the vast number of possible folding patterns is still an active area of research. What is known, however, is that the forces that cause a protein to fold properly and to remain folded are the same basic noncovalent forces that we talked about in chapter 2: ion-ion, ion-dipole, dipole-dipole, hydrogen bonding, and hydrophobic (van der Waals) interactions. One interesting type of hydrophobic interaction is called 'aromatic stacking', and occurs when two or more planar aromatic rings on the side chains of phenylalanine, tryptophan, or tyrosine stack together like plates, thus maximizing surface area contact.



Hydrogen bonding networks are extensive within proteins, with both side chain and main chain atoms participating. Ionic interactions often play a role in protein structure, especially on the protein surface, as negatively charged residues such as aspartate interact with positively-charged groups on lysine or arginine.

One of the most important ideas to understand regarding tertiary structure is that *a protein, when properly folded, is polar on the surface and nonpolar in the interior*. It is the protein's surface that is in contact with water, and therefore the surface must be hydrophilic in order for the whole structure to be soluble. If you examine a three dimensional protein structure you will see many charged side chains (e.g. lysine, arginine, aspartate, glutamate) and hydrogen-bonding side chains (e.g. serine, threonine, glutamine, asparagine) exposed on the surface, in direct contact with water. Inside the protein, out of contact with the surrounding water, there tend to be many more hydrophobic residues such as alanine, valine, phenylalanine, etc. If a protein chain is caused to come unfolded (through exposure to heat, for example, or extremes of pH), it will usually lose its solubility and form solid precipitates, as the hydrophobic residues from the interior come into contact with water. You can see this phenomenon for yourself if you pour a little bit of vinegar (acetic acid) into milk. The solid clumps that form in the milk are proteins that have come unfolded due to the sudden acidification, and precipitated out of solution.

In recent years, scientists have become increasingly interested in the proteins of so-called 'thermophilic' (heat-loving) microorganisms that thrive in hot water environments such as geothermal hot springs. While the proteins in most organisms (including humans) will rapidly unfold and precipitate out of solution when put in hot water, the proteins of thermophilic microbes remain completely stable, sometimes even in water that is just below the boiling point. In fact, these proteins typically only gain full biological activity when in appropriately hot water - at room-temperature they act as if they are 'frozen'. Is the chemical structure of these thermostable proteins somehow unique and exotic? As it turns out, the answer to this question is 'no': the overall three-dimensional structures of thermostable proteins look very much like those of 'normal' proteins. The critical difference seems to be simply that thermostable proteins have more extensive networks of noncovalent interactions, particularly ion-ion interactions on their surface, that provides them with a greater stability to heat. Interestingly, the proteins of 'psychrophilic' (cold-loving) microbes isolated from pockets of water in arctic ice show the opposite characteristic: they have far fewer ion-ion interactions, which gives them greater flexibility in cold temperatures but leads to their rapid unfolding in room temperature water.

This page titled [26.10: Protein Structure](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Steven Farmer, Dietmar Kennepohl, Tim Soderberg, William Reusch, Ed Vitz, & Ed Vitz](#).

- [26.10: Protein Structure](#) by Dietmar Kennepohl, Ed Vitz, Steven Farmer, Tim Soderberg, William Reusch is licensed [CC BY-SA 4.0](#).