Protein Composition and Structure

DNA encodes for a string of one-dimensional amino acids that translate into three-dimensional proteins. Proteins have primary, secondary, and tertiary structures. The primary structure consists of amino acids that are linked by peptide bonds and form linear chains of polypeptides. Secondary structure is the make up of polypeptide chains that fold up into 3D structures. Tertiary structures are the water soluble proteins that condense into compacted structures.

Secondary Structure of a Polypeptide Chain

Ramachandran plot

α-Helices

α-helix

β-PLEATED SHEETS

α-PLEATED SHEETS

Tertiary Structure

Quaternary Structure

Protein structures

• Primary structure (linear polymer of amino acids) (held together with peptide bonds)
• Secondary structure (standard 3-D patterns) (α-helix, β-sheet, held together with H-bonds between backbone atoms)
• Tertiary structure (detailed 3-D conformation) (bonds between side-chain atoms)
• Quaternary structure (combined polymer chains)
Secondary Structure of a Polypeptide Chain

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 alpha-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.

Peptide bonds influence secondary structure
Recall that planar amide bond constrains chain’s bends planes: no rotation around CO-N bonds, but planes rotate around 
\[ \alpha-C-N (\Phi) \] and \[ \alpha-C-C=O \] bonds (\[ \Psi \] )

Ramachandran plot

Shows grouping of \( \phi \psi \) combinations and relates them to structures in real proteins. Repetitive structures (\( \alpha \)-helices, \( \beta \)-sheets) are common

\[ http://commons.wikimedia.org/wiki/Files/alpha_100K.jpg \]

\( \alpha \)-Helices

An alpha-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 alpha-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 alpha-helix is in part dependent on correct steric configuration. Amino acids whose R-groups are too large (tryptophan, tyrosine) or too small (glycine) destabilize alpha-helices. Proline also destabilizes alpha-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 alpha-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 alpha-helices typically end with a charged amino acid to neutralize the dipole moment.

α-helix

- 3.6 amino acids per turn
- 0.54 nm per turn
- side chains pointed out
- H-bonds parallel to axis
- n-4 H-bonds
- dipole moment (neg. at C end)
- no pro, less gly, ser
- limited similar side chain charges

α-helices have a dipole moment; some side chains are preferred
β-PLEATED SHEETS

This structure occurs when two (or more, e.g. psi-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 reverse turn (β-bend):

- R2 (C=O side) is often G,A
- R3 (N-H side) is often D
- Proline is often R2 or R3

**α-PLEATED SHEETS**

A similar structure to the beta-pleated sheet is the alpha-pleated sheet. This structure is energetically less favorable than the beta-pleated sheet, and is fairly uncommon in proteins. An alpha-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 alpha-pleated sheet. The carbonyl side acquires a net negative charge, and the amino side acquires a net positive charge.

![Alpha-Pleated Sheet](https://bio.libretexts.org/Courses/University_of_California_Davis/BIS_105%3A__Biomolecules_and_Metabolism_(Murphy)/Protei)

**Tertiary Structure**

![Tertiary Structure](http://commons.wikimedia.org/wiki/Fi..._Structure.png)

Tertiary structure involves bonds between and among side chains:
• Hydrogen (-O-H…O-)
• Ionic (generally repulsion: -CH2-NH4+:::::::+H4N-CH2-)
• Van der Waal’s (short distance attraction)
• Disulfide (covalent: -CH2-S-S-CH2-)
• Hydrophobic

Tertiary bonds influence the positions of secondary structures.
And the position of a secondary structure in a protein will influence the types of side chains (tertiary structure).
An α-helix on the surface of a protein will have hydrophilic side chains on one side of the helix axis and hydrophobic side chains on the other. An α-helix in the interior of a protein will have primarily hydrophobic side chains. An α-helix exposed to the solution on all sides (unusual) will have hydrophilic side chains on all sides of the helix axis (mostly).

Quaternary Structure

Quaternary structures involve separate polypeptides held together with weak bonds in various symmetries.

Symmetries:
* Homomultimer: a protein with multiple polypeptide chains containing two or more identical components
* Heteromultimer: a protein with multiple polypeptide chains containing two or more different components

Isologous: a protein with multiple polypeptide chains containing two or more identical components
* Heterologous: a protein with multiple polypeptide chains containing two or more different components

Closed: a protein that is not connected
* Open: a protein that is connected

The folding of a protein reduces the free energy (ΔG) of the system.

The folding of a protein involves both protein and solvent.
ΔG = GF - GU
= ΔH - TΔS
= ΔH(protein) + ΔH(solvent) - TΔS(protein) - TΔS(solvent)

ΔG for folding is small (-20 to -60 kJ/mol) and primarily from hydrophobic interactions.

Why so low?
Changes in shape are an important part of protein function and control. For example: a change in shape allows DNA methyltransferase to choose hemi-methylated me CG/GC for bimethylation to me CG/Gme C.