4. Protein Structure

Protein Folding.

Having considered the building blocks of proteins, let’s now consider the protein as a whole. To have a functional protein it must not only have the correct amino acid sequence (primary structure) but it must also have the correct 3D conformation. The process of protein folding is quite complex and not fully understood. Many disease states are thought, however, to be the result of incorrect folding e.g. Alzheimer’s disease, mad cow’s disease, glaucoma.

To understand something of the process we must first consider what holds the protein in its 3D conformation. The main forces which maintain the 3D conformation are the weak inter/intramolecular forces: hydrogen bonding, ionic interactions, Van der Waal’s forces and hydrophobic interactions. Disulphide bonds (strong covalent bonds) are also involved although they contribute less to the overall conformation than might be expected.

Let’s consider these weak forces in more detail.

Hydrogen bonding. This results from the small dipole (uneven electron distribution) that exists in certain side chains and in the peptide bond itself. To have effective H bonding you need an H bond donor and an acceptor. The amide N of the peptide bond is an H bond donor while the carbonyl O (of another peptide bond) can act as an acceptor. The polar non-ionic side chains can also participate in H bonding. Side chains with –OH (Ser, Thr, Tyr), -SH (Cys) and amide N-H (Asn, Gln) all act as H bond donors. Side chains with –C=O (Asn, Gln) can act as acceptors.

Electrostatic or ionic interactions. Side chains with a positive charge (His, Arg or Lys) can interact with those side chains with a negative charge (Asp or Glu) if they are close enough. If side chains with like charges are brought in close proximity they will repel each other and this force also helps maintain 3D conformations. The strength of the interaction is also dependent on the local environment. The presence of high concentrations of salts (high ionic strength), particularly on the protein surface tend to weaken or dampen down the strength of this force.

Van der Waal’s interactions. These interactions between molecules or different parts of molecules can result from a charged or polar group coming in close proximity with a non polar group. The charged group will induce a small dipole on the non-polar group. Sometimes a transient dipole is induced just from fluctuations in the electron distribution in neighbouring atoms. While these interactions are quite weak singly, they make a major contribution in a macromolecule like a protein where many of these interactions occur. Distant parts of the protein can be brought in close proximity and once there van der Waal’s forces will help stabilize the conformation.
**Hydrophobic interactions.** Non-polar side chains will tend to cluster together rather than mix with polar solvents. The old adage “water and oil don’t mix” may be more insightful than your grandmother realized! What drives the non-polar residues to cluster is more to do with entropy; it allows the polar solvent molecules (usually water) more options when the non-polar groups are clustered than when they are scattered throughout the polar solvent. Essentially a non-polar side chain will form a “shell” of ordered water molecules around it. These water molecules have less options for movement than the free flowing molecules contacting only polar groups. To minimize the effect of this constraint i.e. to minimize the loss of entropy the non-polar groups cluster together and bury themselves in the core of a soluble protein away. This removes them from exposure to the polar environment of the protein surface in contact with water. In a hydrophobic environment i.e membrane bound proteins the non-polar side chains arrange themselves on the exterior in contact with the membrane lipids. **It is this interaction which drives soluble proteins to fold.**

If these weak forces are responsible for keeping proteins properly folded you need to ensure the environment of the cell is well regulated, particularly the temperature, pH and ionic strength.

**What 3D structures can proteins assume?**

There are a number of secondary structures a section of polypeptide can form. The double bond character of the peptide bond (called planarity) places constraints on the number of conformations the backbone can assume. Because of its restricted rotation the only bonds free to rotate are the bonds between the alpha carbon and the amide N and the alpha carbon and the carbonyl C.

![A tripeptide at pH 7.](image)

Rotation of the $C_\alpha$-N bond is termed phi ($\phi$) rotation and rotation of the $C_\alpha$-C=O is referred to as psi ($\psi$) rotation.
The other consideration is the polarity of the peptide bond. The amide N can act as an H bond donor and the carbonyl O acts as an acceptor.

As a consequence of the nature of the peptide bond a number of conformations can form. Two basic structures result: the alpha helix and the beta sheet. These are referred to as secondary structures. Sections of a polypeptide chain will form these structures.

**The alpha helix.**

This secondary structure is characterized by the backbone phi and psi rotation. The right handed alpha helix has phi angles of ~ -60° and psi angles of ~ -45°. Left handed alpha helices have phi angles of ~ +60° and psi angles of ~ +45°. The side chains all face outwards, thus steric hindrance from large side chains is not a problem. The helix is maintained by H bonding between peptide N-H and the C=O 4 residues along the chain. All peptide N-H and C=O are H bonded and are all aligned with the N-H pointing in one direction and the C=O
pointing in the other direction. This gives a dipole along the axis of the helix. There are 3.6 residues per turn with a pitch of ~5.4 Å. All alpha helices display these characteristics. Because all the peptide dipoles are satisfied by H-bonding the structure is quite rigid. This is the major secondary structure of globular proteins such as hemoglobin and a key structural element in DNA binding proteins.

Alpha helices can form with all residues except proline. Proline is an unusual amino acid in that it is actually an imino acid, the side chain cyclises back onto the amino group. When proline is incorporated into a peptide chain there is no amide H hence no H bonding can occur from this residue. The side chain constrains the phi rotation preventing the formation of the correct conformation. An internal proline residue will put a bend or kink in an alpha helix.

The beta sheet.

The other commonly found secondary structure is the beta sheet. Again H bonding from peptide N-H and C=O maintains this structure. The strands can form both parallel and antiparallel sheets. Parallel beta sheets form when the peptide strands align in the same direction (N→C) while antiparallel sheets form when one strand is arranged in the N→C direction and the other in the C→N direction. The side chains orientate themselves above and below the plane of the sheet, thus again steric hindrance is not a problem. Antiparallel beta sheets make up silk.
The tertiary structure

The 3D arrangement of the secondary structures within a polypeptide chain is termed the tertiary structure. A polypeptide chain may have a section of alpha helix, followed by a turn (termed a beta turn) then perhaps a stretch of beta sheet etc. Each protein will have its own unique fold ie beta globin always folds the same way.

Proteins often exist in nature as oligomers ie. a number of subunits associate together to form the final functional protein. Each subunit is a single folded polypeptide chain. The arrangement of the monomers or subunits is termed the quaternary structure. Hemoglobin is an association of four globin subunits, each containing a heme group.