Introduction and Protein Structure

Proteins have several layers of structure each of which is important in the process of protein folding. The first most basic level of this structure is the sequence of amino acids themselves.\textsuperscript{1} The sequencing is important because it will determine the types of interactions seen in the protein as it is folding. A novel sequence-based method based on the assumption that protein-protein interactions are more related to amino acids at the surface than those at the core.\textsuperscript{2} This study shows that not only is the amino acids that are in a protein important but also the order in which they are sequenced. The interactions of the amino acids will determine what the secondary and tertiary structure of the protein will be.

The next layer in protein structure is the secondary structure. The secondary structure includes architectural structures that extend in one dimension.\textsuperscript{1} Secondary structure includes α-Helixes (Figure 1) and β-sheets (Figure 2). The α-helices, the most common secondary structure in proteins, the peptide –CO–NH–groups in the backbone form chains held together by NH\textsuperscript{OC} hydrogen bonds.\textsuperscript{3} The α-helices form the backbone of proteins and help to aid in the folding process. The β-sheets form in two distinct ways. They are able to form in both parallel β-pleated sheets and anti parallel β-pleated sheets.\textsuperscript{1} When the α-helix or β-sheet is formed, the excluded volumes generated by the backbone and side chains overlap, leading to an increase in the total volume available to the translational displacement of water molecules.\textsuperscript{4} This is important because it leads to a more thermodynamically stable conformation and leads to less strain on the protein as a whole and thus are aided by the conformation.

Figure 1: (left) typical example to an α-helix, from Wikimedia Commons
The tertiary structure is the next layer in protein structure. This takes the α-Helixes and β-sheets and allows them to fold into a three dimensional structure.\(^1\) Most proteins take on a globular structure once folded. The description of globular protein structures as an ensemble of contiguous closed loops or tightened end fragments reveals fold elements crucial for the formation of stable structures and for navigating the very process of protein folding.\(^5\) The globular proteins generally have a hydrophobic core surrounded by a hydrophilic outer layer. These interactions are important because they lead to the global structure and help create channels and binding sites for enzymes.

The last layer of protein structure is the quaternary structure. The folding transition and the functional transitions between useful states are encoded in the linear sequence of amino acids, and a long-term goal of structural biology is to be able to predict both the structure and function of molecules from the information in the sequence.\(^6\) The Subunit organization is the last level of structure in protein molecules.\(^1\) The organization of the subunits is important because that determines the types of interactions that can form and dictates its use in the body.

### Protein Folding

Proteins are folded and held together by several forms of molecular interactions. The molecular interactions include the thermodynamic stability of the complex, the hydrophobic interactions and the disulfide bonds formed in the proteins. The figure below (figure 3) is an example of protein folding.

![Figure 3: Protein Folding, from Wikimedia Commons](image)

The biggest factor in a protein's ability to fold is the thermodynamics of the structure. The interaction scheme includes the short-range propensity to form extended conformations, residue-dependent long-range contact potentials, and orientation-dependent hydrogen bonds.\(^7\) The thermodynamics are a main stabilizing force within a protein because if it is not in the lowest energy conformation it will continue to move and adjust until it finds its most stable state. The use of energy diagrams and maps are key in finding out when the protein is in the most stable form possible.

The next type of interaction in protein folding is the hydrophobic interactions within the protein. The framework model and the hydrophobic collapse model represent two canonical descriptions of the protein folding process. The first places primary reliance on the short-range interactions of secondary structure and the second assigns greater importance to the long-range interactions of tertiary structure.\(^6\) These hydrophobic interactions have an impact not just on the primary
structure but then lead to changes seen in the secondary and tertiary structure as well. Globular proteins acquire distinct compact native conformations in water as a result of the hydrophobic effect. When a protein has been folded in the correct way it usually exists with the hydrophobic core as a result of being hydrated by waters in the system around it which is important because it creates a charged core to the protein and can lead to the creation of channels within the protein. The hydrophobic interactions are found to affect time correlation functions in the vicinity of the native state even though they have no impact on same time characteristics of the structure fluctuations around the native state. The hydrophobic interactions are shown to have an impact on the protein even after it has found the most stable conformation in how the proteins can interact with each other as well as folding themselves.

Another type of interaction seen when the protein is folding is the disulfide linkages that form in the protein. (See figure 4) The disulfide bond, a sulfur-sulfur chemical bond that results from an oxidative process that links nonadjacent (in most cases) cysteine’s of a protein. These are a major way that proteins get into their folded form. The types of disulfide bonds are cysteine-cysteine linkage is a stable part of their final folded structure and those in which pairs of cysteines alternate between the reduced and oxidized states. The more common is the linkages that cause the protein to fold together and link back on itself compared to the cysteines that are changing oxidation states because the bonds between cysteines once created are fairly stable.

![Disulfide Bonds](https://via.placeholder.com/150)

Figure 4: Disulfide Bonds, shown in the picture in yellow, from Wikimedia Commons

**Misfunctions**

Proteins can miss function for several reasons. When a protein is miss folded it can lead to denaturation of the protein. Denaturation is the loss of protein structure and function. The miss folding does not always lead to complete lack of function but only partial loss of functionality. The miss functioning of proteins can sometimes lead to diseases in the human body.
Alzheimer's Disease

Alzheimer's Disease (AD) is a neurological degenerative disease that affects around 5 million Americans, including nearly half of those who are age 85 or older. The predominant risk factors of AD are age, family history, and heredity. Alzheimer’s disease typically results in memory loss, confusion of time and place, misplacing places, and changes in mood and behavior. AD results in dense plaques in the brain that are comprised of fibrillar β-amyloid proteins with a well-ordered β-sheet secondary structure. These plaques visually look like voids in the brain matter (see figure 5) and are directly connected to the deterioration of thought processes. It has been determined that AD is a protein misfolding disease, where the misfolded protein is directly related to the formation of these plaques in the brain.

![Figure 5: Comparison of healthy brain (left) with brain with Alzheimer's (right)](From Wikimedia Commons)

It is yet to be fully understood what exactly causes this protein misfolding to begin, but several theories point to oxidative stress in the brain to be the initiating factor. This oxidation results in damage to the phospholipids in the brain, which has been found to result in a faster accumulation of amyloid β-proteins.

![Figure 6: Beta-Amyloid Plaque Formation, from Wikimedia Commons](4)

Cystic Fibrosis

Cystic Fibrosis (CF) is a chronic disease that affects 30,000 Americans. The typical affects of CF is a production of thick, sticky mucus that clogs the lungs and leads to life-threatening lung infection, and obstructs the pancreas preventing proper food processing. CF is caused by protein misfolding. This misfolding then results in some change in the protein known as cystic fibrosis transmembrane conductance regulator (CFTR), which can result in this potentially fatal disease. In approximately 70% of CF cases, a deletion of phenylalanine at position 508 in the CFTR is deleted. This deletion of Phe508 seems to be directly connected to the formation of CF. The protein misfolding that results in CF occurs prior to birth, but it is not entirely clear as to why.
Resources

6. Hausratha, A.C. A Kinetic theory of tertiary contact formation coupled to the helix-coil transition in polypeptides. The Journal of Chemical Physics. 2006