# The three-dimensional structure of proteins

When two amino acids are positioned so that the carboxyl group of one is adjacent to the amino group of the other, they can become joined by a dehydration reaction, with the removal of a water molecule. The resulting covalent bond is called a **peptide bond**. Repeated over and over, this process yields a polypeptide, a polymer of many amino acids linked by peptide bonds. The repeating sequence of atoms highlighted in purple in Figure 5.15 is called the *polypeptide backbone*.

Extending from this backbone are the different side chains (R groups) of the amino acids. Polypeptides range in length from a few amino acids to 1,000 or more. Each specific polypeptide has a unique linear sequence of amino acids. Note that one end of the polypeptide chain has a free amino group (the N-terminus of the polypeptide), while the opposite end has a free carboxyl group (the C-terminus). The chemical nature of the molecule is determined by the kind and sequence of the side chains, which determine how a polypeptide folds and thus its final shape and chemical characteristics. The immense variety of polypeptides in nature illustrates an important concept introduced earlier—that cells can make many different polymers by linking a limited set of monomers into diverse sequences.

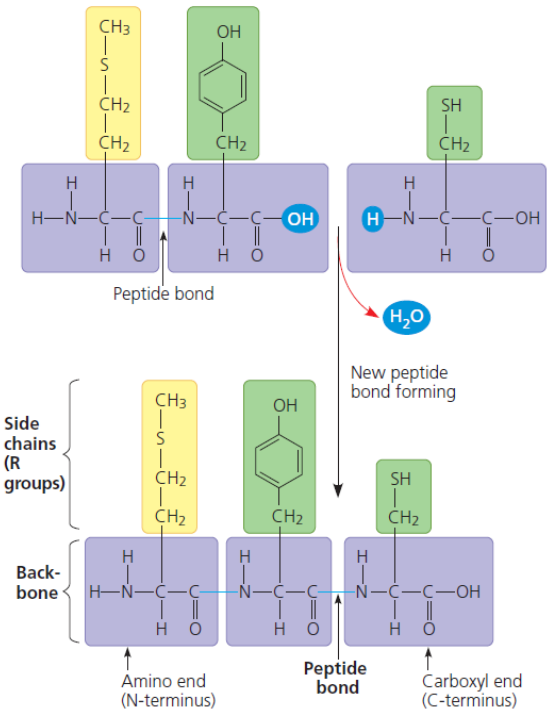



**Figure 5.15 Making a polypeptide chain.** Peptide bonds are formed by dehydration reactions, which link the carboxyl group of one amino acid to the amino group of the next. The peptide bonds are formed one at a time, starting with the amino acid at the amino end (N-terminus). The polypeptide has a repetitive backbone (purple) to which the amino acid side chains (yellow and green) are attached.

# Protein structure and function

The specific activities of proteins result from their intricate three-dimensional architecture, the simplest level of which is the sequence of their amino acids. What can the amino acid sequence of a polypeptide tell us about the three-dimensional structure (commonly referred to simply as the "structure") of the protein and its function? The term *polypeptide* is not synonymous with the term *protein*. To work the proteins must have a specific tri-dimensional **conformation**. The spatial arrangement of atoms in a protein is called its **conformation.**

A functional protein is not *just* a polypeptide chain, but one or more polypeptides precisely twisted, folded, and coiled into a molecule of unique shape, which can be shown in several different types of models. And it is the amino acid sequence of each polypeptide that determines what three-dimensional structure the protein will have under normal cellular conditions.

When a cell synthesizes a polypeptide, the chain may fold spontaneously, assuming the functional structure for that protein. This folding is driven and reinforced by the formation of various bonds between parts of the chain, which in turn depends on the sequence of amino acids. Many proteins are roughly spherical (*globular proteins*), while others are shaped like long fibers (*fibrous proteins*).

A protein's specific structure determines how it works. In almost every case, the function of a protein depends on its ability to recognize and bind to some other molecule. In an especially striking example of the marriage of form and function, Figure 5.17 shows the exact match of shape between an antibody (a protein in the body) and the foreign substance on a flu virus that the antibody binds to and marks for destruction.

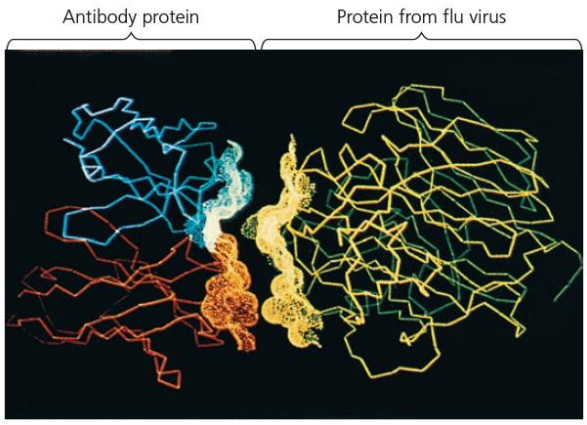



**Figure 5.17 Complementarity of shape between two protein surfaces.** A technique called X-ray crystallography was used to generate a computer model of an antibody protein (blue and orange, left) bound to a flu virus protein (yellow and green, right). This is a wireframe model modified by adding an "electron density map" in the region where the two proteins meet. Computer software was then used to back the images away from each other slightly.

### *Four Levels of Protein Structure*

In spite of their great diversity, proteins share three superimposed levels of structure, known as primary, secondary, and tertiary structure. A fourth level, quaternary structure, arises when a protein consists of two or more polypeptide chains. **Figure 5.18** describes these four levels of protein structure. Be sure to study this figure thoroughly before going on to the next section.

The possible conformations of a protein include any structural state that can be achieved without breaking covalent bonds. A change in conformation could occur, for example, by rotation about single bonds. Of the numerous conformations that are theoretically possible in a protein containing hundreds of single bonds, one or (more commonly) a few generally predominate under biological conditions. The need for multiple stable conformations reflects the changes that must occur in most proteins as they bind to other molecules or catalyse reactions. The conformations existing under a given set of conditions are usually the ones that are thermodynamically the most stable, having the lowest Gibbs free energy (*G*). Proteins in any of their functional, folded conformations are called **native** proteins.
In the context of protein structure, the term **stability** can be defined as the tendency to maintain a native conformation.

# Primary structure

The **primary structure** of a protein is its sequence of amino acids. As an example, let's consider transthyretin, a globular blood protein that transports vitamin A and one of the thyroid hormones throughout the body. Transthyretin is made up of four identical polypeptide chains, each composed of 127 amino acids. Shown here is one of these chains unravelled for a closer look at its primary structure. Each of the 127 positions along the chain is occupied by one of the 20 amino acids, indicated here by its three-letter abbreviation.
The precise primary structure of a protein is determined not by therandom linking of amino acids, but by inherited genetic information.
The primary structure in turn dictates secondary and tertiary structure, due to the chemical nature of the backbone and the side
chains (R groups) of the amino acids along the polypeptide.

## Secondary structure

Most proteins have segments of their polypeptide chains repeatedly coiled or folded in patterns that contribute to the protein's overall shape. These coils and folds, collectively referred to as **secondary structure**, are the result of hydrogen bonds between the repeating constituents of the polypeptide backbone (not the amino acid side chains). Within the backbone, the oxygen atoms have a partial negative charge, and the hydrogen atoms attached to the nitrogens have a partial positive charge; therefore, hydrogen bonds can form between these atoms. Individually, these hydrogen bonds are weak, but because they are repeated many times over a relatively long region of the polypeptide chain, they can support a shape for that part of the protein. One such secondary structure is the $\boldsymbol{\alpha}$ **− helix**, a delicate coil held together by hydrogen bonding between every fourth amino acid, as shown above. Although each transthyretin polypeptide has only one $\boldsymbol{\alpha}$ − helix region (see the Tertiary Structure section), other globular proteins have multiple stretches of $\boldsymbol{\alpha}$ − helix separated by nonhelical regions (see haemoglobin in the Quaternary Structure section). Some fibrous proteins, such as $\boldsymbol{\alpha}$ − keratin, the structural protein of hair, have the $\boldsymbol{\alpha}$ −helix formation over most of their length.

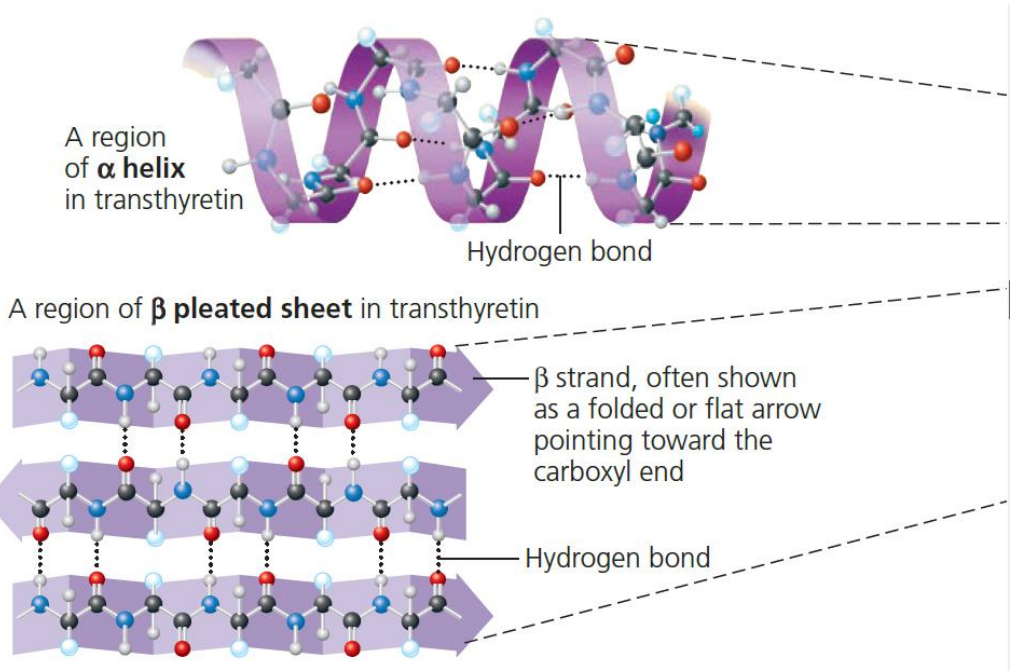
Thus, five different kinds of constraints affect the stability of an $\boldsymbol{\alpha}$ helix:

1. the electrostatic repulsion (or attraction) between successive amino acid residues with charged R groups
2. the bulkiness of adjacent R groups
3. the interactions between R groups spaced three (or four) residues apart
4. the occurrence of Pro and Gly residues
5. the interaction between amino acid residues at the ends of the helical segment and the electric dipole inherent to the $\boldsymbol{\alpha}$ helix.

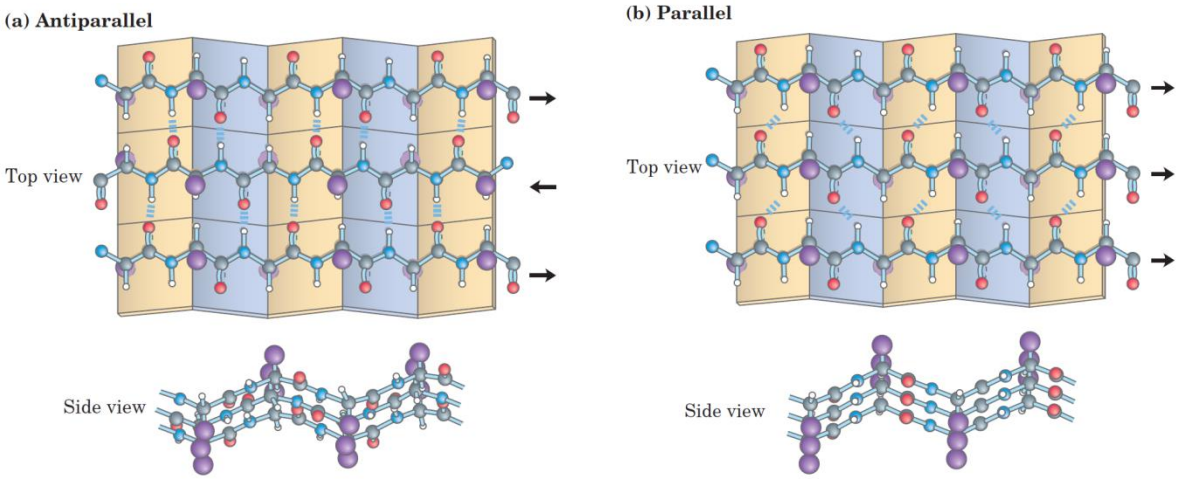
The tendency of a given segment of a polypeptide chain to fold up as an $\boldsymbol{\alpha}$ helix therefore depends on the identity and sequence of amino acid residues within the segment.

The other main type of secondary structure is the $\boldsymbol{\beta}$ **−pleated sheet**. As shown above, in this structure two or more segments of the polypeptide chain lying side by side (called $\boldsymbol{\beta}$ strands) are connected by hydrogen bonds between parts of the two parallel segments of polypeptide backbone.

$\boldsymbol{\beta}$ pleated sheets make up the core of many globular proteins, as is the case for transthyretin (see Tertiary Structure), and dominate some fibrous proteins, including the silk protein of a spider's web.




In this arrangement, called a ***β* sheet,** hydrogen bonds are formed between adjacent segments of polypeptide chain. The individual segments that form a ***β*** sheet are usually nearby on the polypeptide chain, but can also be quite distant from each other in the linear sequence of the polypeptide; they may even be segments in different polypeptide chains. The R groups of adjacent amino acids protrude from the zigzag structure in opposite directions, creating the alternating pattern

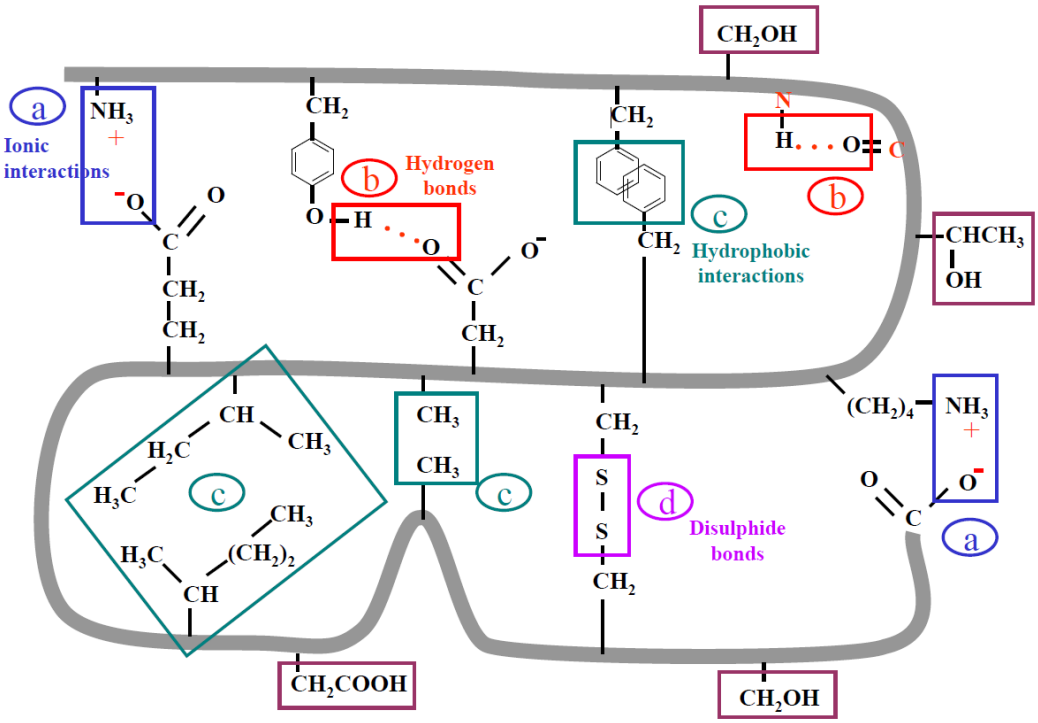



The adjacent polypeptide chains in a ***β*** sheet can be either parallel or antiparallel (having the same or opposite amino-to-carboxyl orientations, respectively). The structures are somewhat similar, although the repeat period is shorter for the parallel conformation (6.5 Å, versus 7 Å for antiparallel) and the hydrogenbonding patterns are different.

# Tertiary structure

Three-dimensional shape is stabilized by interactions between side chains. Superimposed on the patterns of secondary structure is a protein's tertiary structure, shown here in a ribbon model of the transthyretin polypeptide. While secondary structure involves interactions between backbone constituents, **tertiary structure** is the overall shape of a polypeptide resulting from interactions between the side chains (R groups) of the various amino acids.

Knowing the kind of amino acids and their sequence, we can predict how will be the structure of the proteins.




Amino acids that are far apart in the polypeptide sequence and that reside in different types of secondary structure may interact within the completely folded structure of a protein. The location of bends in the polypeptide chain and the direction and angle of these bends are determined by the number and location of specific bend-producing residues, such as Pro, Thr, Ser, and Gly. Interacting segments of polypeptide chains are held in their characteristic tertiary positions by different kinds of weak bonding interactions (and sometimes by covalent bonds such as disulfide cross-links) between the segments.

- **Hydrogen bonds:** A hydrogen bond is a weak type of force that forms a special type of dipole-dipole attraction which occurs when a hydrogen atom bonded to a strongly electronegative atom exists in the vicinity of another electronegative atom with a lone pair of electrons. These bonds are generally stronger than ordinary dipole-dipole and dispersion forces, but weaker than true covalent and ionic bonds.
  More over the hydrogen atom is the only element that have only one electron, this one is attracted by the oxygen, so the only one proton won't be screened and will be a core as high density of positive charge. The sharing of electrons between H and O is therefore unequal; the electrons are more often in the vicinity of the oxygen atom than of the hydrogen. The result of this unequal electron sharing is two electric dipoles in the water molecule, one along each of the H-O bonds; each hydrogen bears a partial positive charge ($\delta^+$) and the

oxygen atom bears a partial negative charge equal to the sum of the two partial positives ($2\delta^-$). The hydrogen bond is used also to stabilized double links of DNA, and so this is the reason why exist a strong association between nucleotides: the specific bonds G −C and A−T provide the highest number of hydrogen bond.

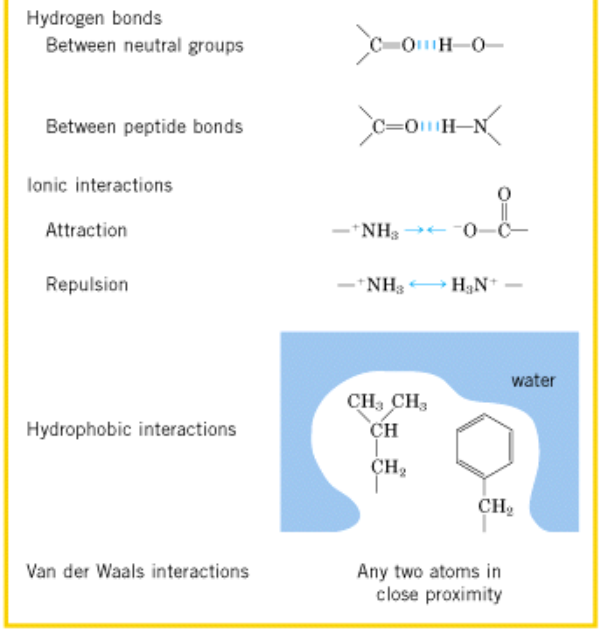
- **Ionic interactions:** In some cases, two atoms are so unequal in their attraction for valence electrons that the more electronegative atom strips an electron completely away from its partner. The two resulting oppositely charged atoms (or molecules) are called **ions**. A positively charged ion is called a **cation**, while a negatively charged ion is called an **anion**. Because of their opposite charges, cations and anions attract each other; this attraction is called an **ionic bond**. Note that the transfer of an electron is not, by itself, the formation of a bond; rather, it allows a bond to form because it results in two ions of opposite charge. Any two ions of opposite charge can form an ionic bond. The ions do not need to have acquired their charge by an electron transfer with each other.

- **Hydrophobic interactions:** As a polypeptide folds into its functional shape, amino acids with hydrophobic (nonpolar) side chains usually end up in clusters at the core of the protein, out of contact with water. Thus, a "*hydrophobic interaction*" is caused by the exclusion of nonpolar substances by water molecules. Once nonpolar amino acid side chains are close together, van der Waals interactions help hold them together. Hydrophobic interaction is formed between aromatic chain or aliphatic chain.

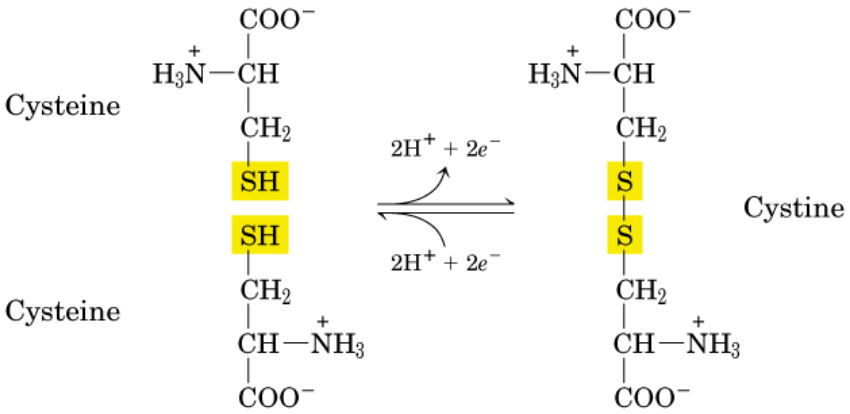
- **Disulfide bonds:** This is the only covalent bond that stabilized the protein structure. Disulfide bridges form where two cysteine monomers, which have sulfhydryl groups (—SH) on their side chains, are brought close together by the folding of the protein. The sulfur of one cysteine bonds to the sulfur of the second, and the disulfide bridge (—S—S—) rivets parts of the protein together.

table 4–4

**Four Types of Noncovalent ("Weak") Interactions among Biomolecules in Aqueous Solvent**



Meanwhile, hydrogen bonds between polar side chains and ionic bonds between positively and negatively charged side chains also help stabilize tertiary structure. These are all weak interactions in the aqueous cellular environment, but their cumulative effect helps give the protein a unique shape.

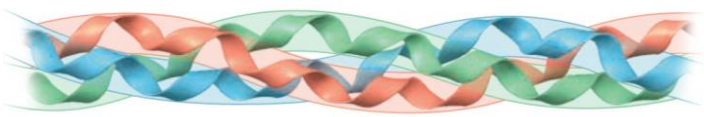



When two Cysteine residuals are next to each other tend to form a covalent bond by an oxidation reaction. Now, there are many oxido reactions in cellular metabolism, so when we have a strong change in oxidative state of the cell, Cysteine reduction can restore a new balance with a consequent change of the protein structure.
We stress this concept because Cysteine have a main role in the regulation of cell's life. Cells, in fact, use chemical elements to communicate with inner organs and the outside cells.

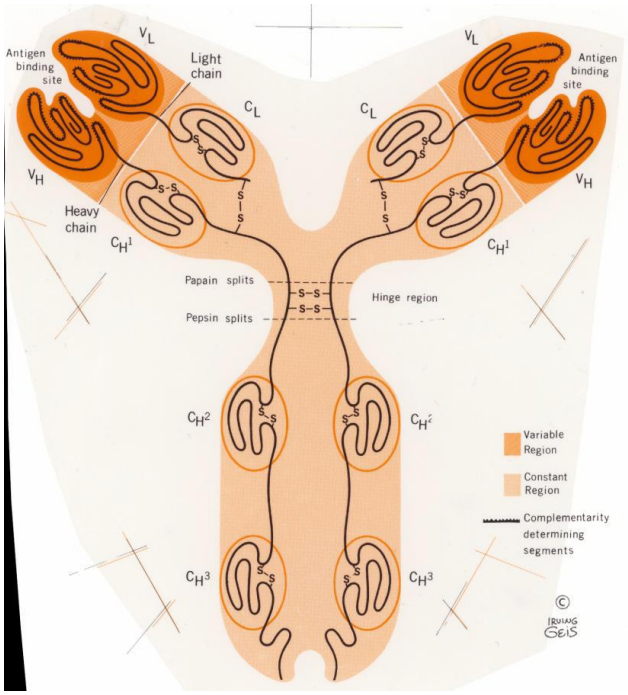
## Quaternary structure

Some proteins consist of two or more polypeptide chains aggregated into one functional macromolecule. **Quaternary structure** is the overall protein structure that results from the aggregation of these polypeptide subunits. For example, shown above is the complete globular transthyretin protein, made up of its four polypeptides. Another example is collagen, which is a fibrous protein that has three identical helical polypeptides intertwined into a larger triple helix, giving the long fibers great strength.

Proteins can have a different distribution of amino acids depending on both the function and localization. A **coiled coil** is a structural motif in proteins in which 2–7 alpha-helices are coiled together like the strands of a rope.



The helical path of the supertwists is left-handed, opposite in sense to the $\alpha$ helix. This class of molecules are constituted by a hydrophilic channel (so the polar side chains protrude inside) that allow the passage of any substances.




# Immunoglobulin

An **antibody (Ab)**, also known as an **immunoglobulin (Ig)** is a large, Y-shaped protein produced mainly by plasma cells that is used by the immune system to neutralize pathogens such as pathogenic bacteria and viruses. The antibody recognizes a unique molecule of the pathogen, called an antigen, via the Fab's variable region. Each tip of the "Y" of an antibody contains a paratope (analogous to a lock) that is specific for one particular epitope (similarly, analogous to a key) on an antigen, allowing these two structures to bind together with precision. Using this binding mechanism, an antibody can *tag* a microbe or an infected cell for attack by other parts of the immune system, or can neutralize its target directly (for example, by inhibiting a part of a microbe that is essential for its invasion and survival). The ability of an antibody to communicate with the other components of the immune system is mediated via its Fc region (located at the base of the "Y"), which contains a conserved glycosylation site involved in these interactions.

In this protein we have a constant part (or scaffold) and a variable part.

A polypeptide chain of a given amino acid sequence can be arranged into a three-dimensional shape determined by the interactions responsible for secondary and tertiary structure. This folding normally occurs as the protein is being synthesized in the crowded environment within a cell, aided by other proteins. However, protein structure also depends on the physical and chemical conditions of the protein's environment. If the pH, salt concentration, temperature, or other aspects of its environment are altered, the weak chemical bonds and interactions within a protein may be destroyed, causing the protein to unravel and lose its native shape, a change called **denaturation.** Because it is misshapen, the denatured protein is biologically inactive.

Most proteins become denatured if they are transferred from an aqueous environment to a nonpolar solvent, such as ether or chloroform; the polypeptide chain refolds so that its hydrophobic regions face outward toward the solvent. Other denaturation agents include chemicals that disrupt the hydrogen bonds, ionic bonds, and disulfide bridges that maintain a protein's shape.

Denaturation can also result from excessive heat, which agitates the polypeptide chain enough to overpower the weak interactions that stabilize the structure. The white of an egg becomes opaque during cooking because the denatured proteins are insoluble and solidify. This also explains why excessively high fevers can be fatal: Proteins in the blood tend to denature at very high body temperatures.

When a protein in a test-tube solution has been denatured by heat or chemicals, it can sometimes return to its functional shape when the denaturing agent is removed. (Sometimes this is not possible: For example, a fried egg will not become liquefied when placed back into the refrigerator!)

We can conclude that the information for building specific shape is intrinsic to the protein's primary structure; this is often the case for small proteins.

If the temperature is increased step by step, and so gradually, the proteins lose their structure, but they potentially can recover again their original structure. In this case denaturation is reversible.

The unfolded state is characterized by higher entropy, the native state, instead, is characterized by lower energy.
There are a lot **disease** which are linked with protein misfolding(the proteins isn't able to acquire the native structure because is mutated ).

Protein-misfolding diseases include conditions where a protein:

1. fails to fold correctly
   (cystic fibrosis, Marfan syndrome, amyotonic lateral sclerosis)
2. is not stable enough to perform its normal function (many forms of cancer)
3. fails to be correctly trafficked
   (familial hypercholesterolemia, α1-antitrypsin deficiency)
4. forms insoluble aggregates that deposit toxically
   (neurodegenerative diseases: Alzheimer's, type II diabetes, Parkinson's and many more)

**Amyloidosis** is a group of diseases in which abnormal protein, known as *amyloid fibrils*, builds up in tissue. There are about 30 different types of amyloidosis, each due to a specific protein misfolding.

**We** have two type of amyloidosis:

I. Systemic amyloidoses → large amounts of fibrils accumulate everywhere.

II. Organ-limited amyloidoses →fibrils accumulate locally in one organ ( e.g., brain )

## Post-translational modifications

Protein synthesis occurs during a process called '**translation**'.

**Post-translational modification** of proteins refers to the **chemical change proteins may undergo after translation**. Such modifications come in a wide variety of types and are mostly catalyzed by enzymes that recognize specific target sequences in specific proteins.

**Post-translational modifications (PTM)** of proteins play an important role in the cellular functions. PTM is the covalent addition of certain functional groups to the proteins. Of all the post-translational modifications, **phosphorylation** and **glycosylation** are the major players in many of the protein functions. Most proteins undergo some modification before undertaking any function assigned to them. A post-translational modification can be a reversible or an irreversible activity. **Proteolytic cleavage** is one of the common modifications where proteins are cleaved to remove some additional amino acid(s) or portion of protein. Examples are zymogens, which are inactive forms of enzymes and are activated by the removal of some portion of the protein.

In chemistry, **phosphorylation** of a molecule is the attachment of a phosphoryl group. Together with its counterpart, dephosphorylation, it is critical for many cellular processes in biology. Phosphorylation is especially important for protein function, as this modification activates (or deactivates) almost half of the enzymes, thereby regulating their function. The attachment of phosphate group depends on both the steric affects and the charge availability, so the post translational modification depends on the nature of protein chain.

**The Post-translational modifications amplify the potential variability of proteins' functions.**

Some of the post-translation modifications are not 'for ever': are used in a way that allow, in a second time, to attach and remove depending on the necessity of protein function.

Post translation modification is one of the tools that the cell must modify the structure and changing the affinity of the protein to legant.

Cells can regulate the more trials **through structural change.**