Doctor 022

Biochemistry

Sheet no. 15

Writer: **Al-Razi Node**

Corrector: **Al-Razi Node**

Doctor: **Dr Mamoun, Dr Diala**

Protein Structure

Basically, to understand proteins and their structure scientists have divided protein structure to four levels of complexity: primary structure, secondary structure, tertiary structure, and quaternary structure.

Primary Structure:

We have mentioned previously that the primary structure is the sequence or order of amino acids, which in turn determines the other levels of structure.

***Remember: Based on that, Artificial Intelligence can be used to predict the protein's structure with good accuracy, it's not 100% accurate but at least it gives information about the structure of proteins.**

Secondary Structure:

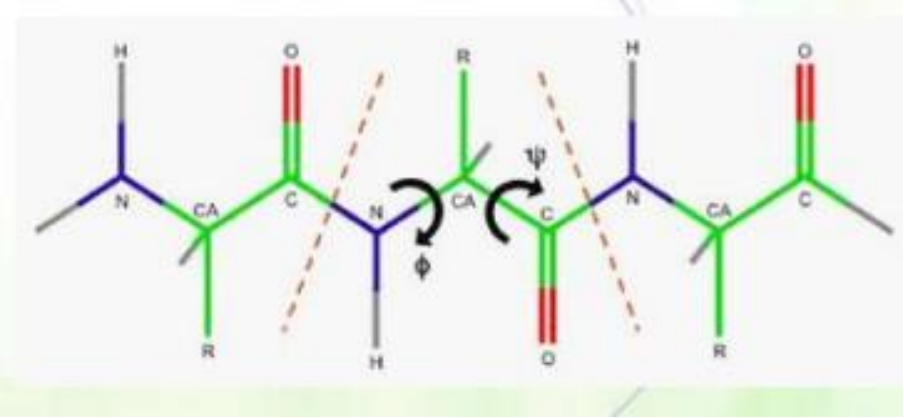
The secondary structure is a localized and organized structure which happens due to specific and distinguishable patterns from the primary sequence. Certain sequence of amino acids in a certain area with defined and organized structure in their local place.

***Remember: the bonds that are around the central carbon can rotate, unlike the peptide bond that is rigid. Therefore, in order for amino acid sequences to form certain patterns, the bonds connected to the alpha carbons of the amino acids have to rotate in the correct orientation.**

*The two bonds within each amino acid residue freely rotate:

- The bond between the a-carbon and the amino nitrogen.
- The bond between the a-carbon and the carboxyl carbon.

*Secondary structure: A hydrogen-bonded, local arrangement of the backbone of a polypeptide chain.



*Polypeptide chains can fold into regular structures such as:

- Regular secondary structures:

*Alpha helix.

*Beta-pleated sheet.

- Nonregular secondary structures:

Turns, Loops, Bends and Coils.

***There are different patterns which are alpha helices and ß-pleated (flat) sheets.**

Let's begin at α-helix:

The Alpha Helix:

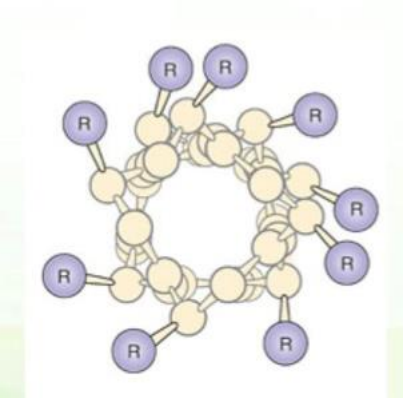
It looks like a helical rod. **(Helix: حلزوني) (exactly like a spring)**

The helix has an average of 3.6 amino acids per turn. **(the 0.6 is a part of the fourth amino acid).**

The pitch of the helix: the linear distance between corresponding points on successive turns. **(From one point to the point directly above it or under it in an exactly straight line),** **It is almost 5.4 Angstrom.**

***Note: 1 Angstrom(Å) = 10^{-10} m = 0.1 nm.**

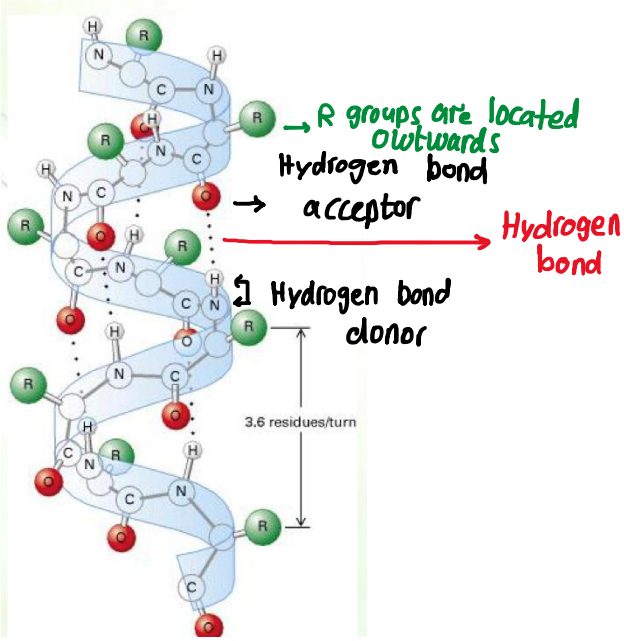
It is very stable because of the linear hydrogen bonding. The trans side chains of the amino acids project outward from the helix, thereby avoiding steric hindrance with the polypeptide backbone and with each other.



***If we look at this cylinder (the helix) from a top view we can notice that the R groups are directed outwards. However, the R groups are not located directly beneath each other to avoid repulsion. So if we look at a certain R group, the other one will be directed sideways otherwise there will be repulsion.**

The other point which is important as well that Alpha helices and all the secondary structures are stabilized by hydrogen bonds between the peptide bonds not the R groups.

As we have mentioned that one of the features of the peptide bonds that they can form hydrogen bonds, so when there is an amino acid and another amino acid directly under it, a hydrogen bond will form between the peptide bonds not the R groups: between the amide group (N-H) of one amino acid which is the donor and the (C=O) of the carbonyl group of the other amino acid which is the acceptor.



➢ Amino acids NOT found in α-helix:

There are some amino acids that shouldn't exist in α-helices, because they affect and disrupt the structure of α-helices, such as:

The plural of helix = helixes or helices

***PROLINE:**

1) No rotation around N-Cα bond.

2) No hydrogen bonding of α-amino group.

***Proline doesn't have a hydrogen bond donor because it has a secondary nitrogen (already bound to 2 carbons), but it has a hydrogen bond accepter.**

*** Proline is rigid, and we need the helix to be smooth during the rotation, so Proline breaks the smoothness of the α-helix (because of its rigidity).**

-Another amino acid that shouldn't exist which is:

***GLYCINE: Too small**

It also disturbs the smoothness because it is too small.

-Other things that affect the structure of α-helix:

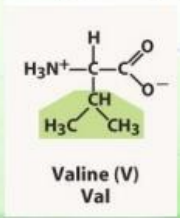
***Close proximity of a pair of charged amino acids with SIMILAR charges.**

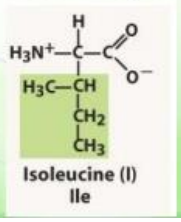
Ex. If I have a positively charged amino acid beside another positively charged amino acid, REPULSION will happen between the R groups.

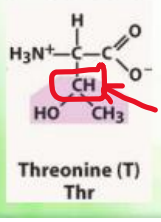
***Amino acids with branches at the β-carbon atom (valine, threonine, and isoleucine).**

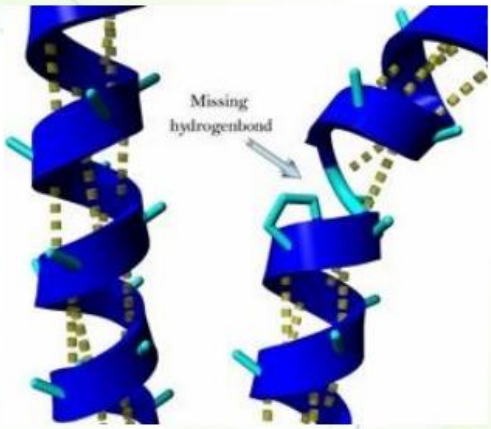
α-Carbon is the central carbon next to β-Carbon. If the β-Carbon was branched, the α-helix is damaged. (The branches on the β-Carbon cause repulsion).

- Glycine: too small
- Proline
  - No rotation around N-Cα bond
  - No hydrogen bonding of α-amino group
- Close proximity of a pair of charged amino acids with similar charges
- Amino acids with branches at the β-carbon atom (valine, threonine, and isoleucine)







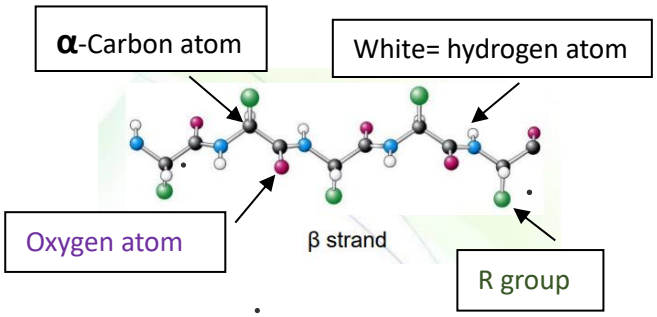


β pleated sheet (β sheet).

Sheets that look like up to down sheets so it shows a Zigzag pattern.

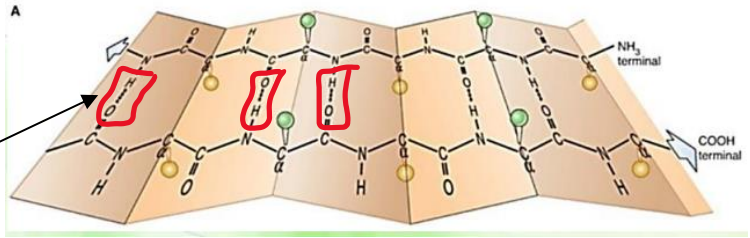
They are composed of two or more straight chains (β strands) that are hydrogen bonded side by side.

so β pleated sheets are composed of several β strands.



***Optimal hydrogen bonding occurs when the sheet is bent (pleated) to form β-pleated sheets**

***Here: Notice that peptide bonds form a zigzag, so we have many β-strands and a Zigzag structure.**



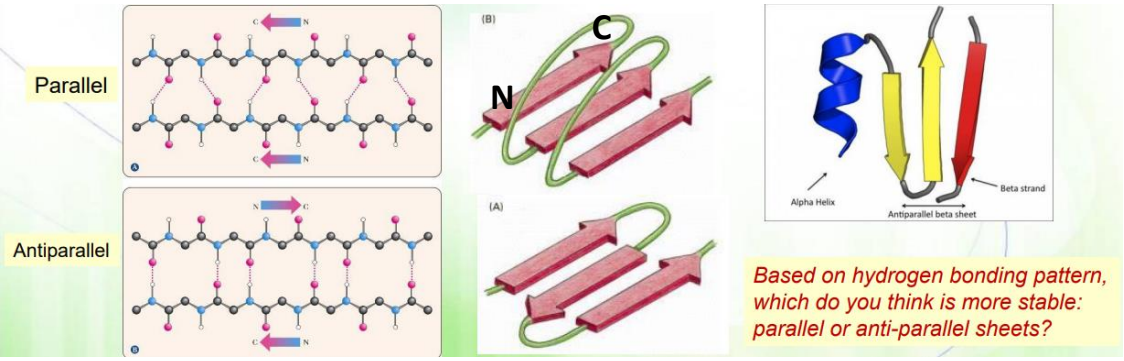
*** β-strands combination is stabilized by hydrogen bonds between the peptide bonds: between amine group (hydrogen bond donor) and the carbonyl group (hydrogen bond accepter).**

More on β-sheets

β sheets can form between many strands in a localized region, typically 4 or 5 but as many as 10 or more.

Such β sheets can be purely antiparallel, purely parallel, or mixed.

We can determine it by going from N to C (as in the figure).



***It can be mixed (parallel + antiparallel).**

As we said, all these β-strands are stabilized by hydrogen bonds between them.

There are two forms of a beta sheet, one is the parallel and the other is the anti-parallel form.

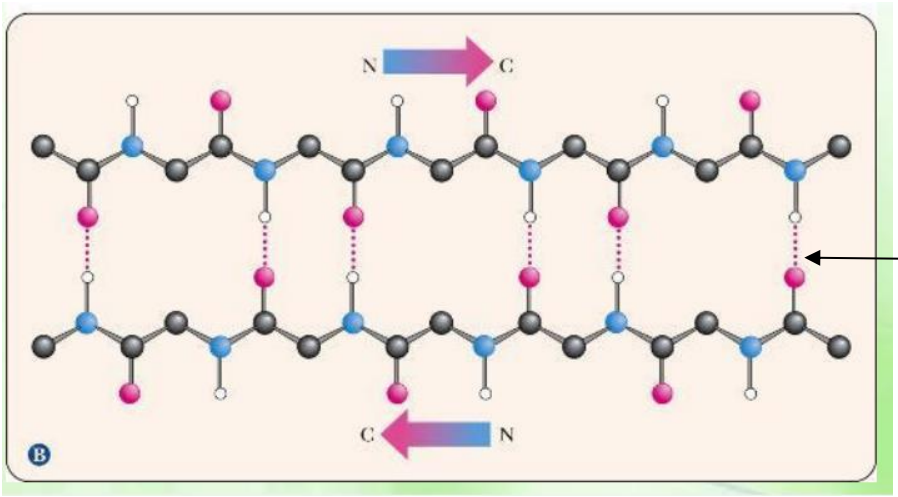
***NOTICE the hydrogen bonding in the parallel and antiparallel forms. Which is more stable?**

ANS: Anti parallel form is more stable than the parallel one due to the following reasons:

- They occupy straight bonds, as they are more stable than bent ones.

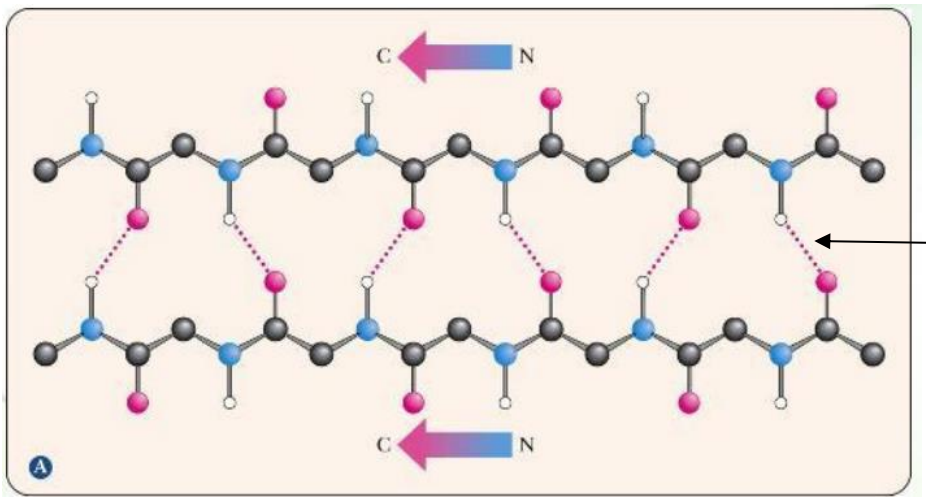
- The hydrogen bond between the groups of the same peptide bonds.

***Remember: Carbonyl and amino groups form hydrogen bonds.**
(Hydrogen bonding between backbone atoms not R group).



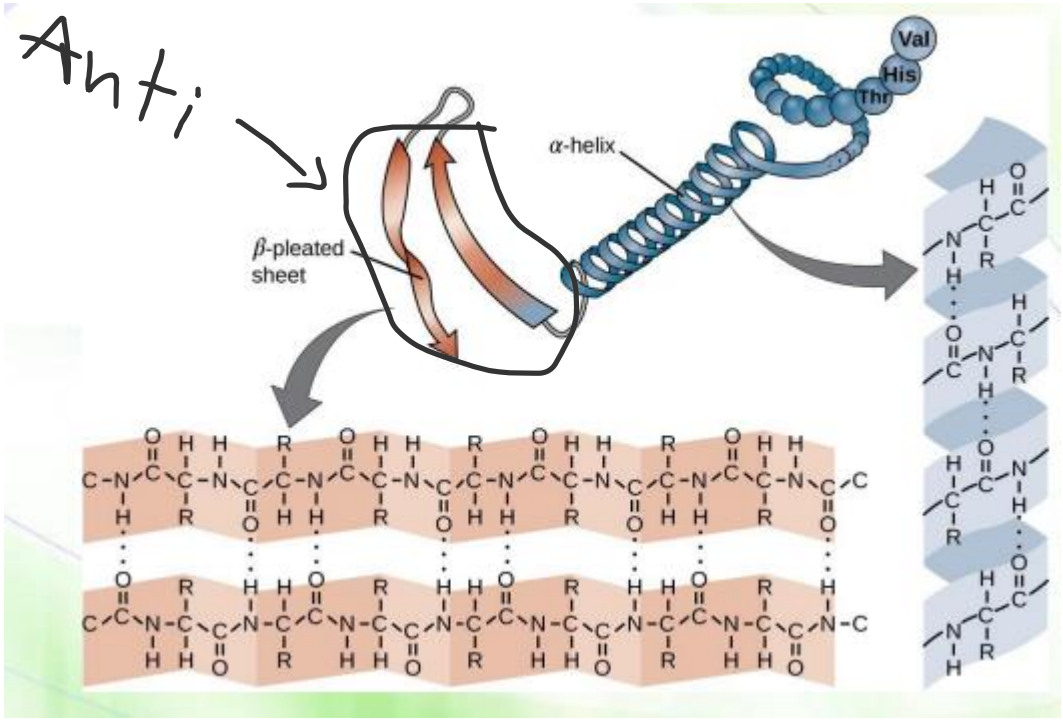
لاحظ كيف بكونوا فوق بعض بالضبط وبشكل عامودي

***Parallel: one hydrogen bond to the right side and another one to the left side, but no hydrogen bonding with the one below, (no hydrogen bonding on the same peptide bonds).**



لاحظ كيف بتكون مائلة الرابطة الهيدروجينية

***NOTE: We can distinguish beta strands by the zigzag form they acquire when looking at them in a form, but in textbooks they are shown or drawn as arrows that point to their direction (from N terminal to the C terminal). Meanwhile alpha helixes are spiral shaped. (see picture below)**



***Amino acids that tend to be present in the beta strand are tyrosine, phenylalanine, and tryptophan (the large anomeric amino acids), valine, threonine, and isoleucine with branched R groups at the β-carbon.**

***Glycine may also be present. It is not harmful for beta sheets but harmful for alpha helixes.**

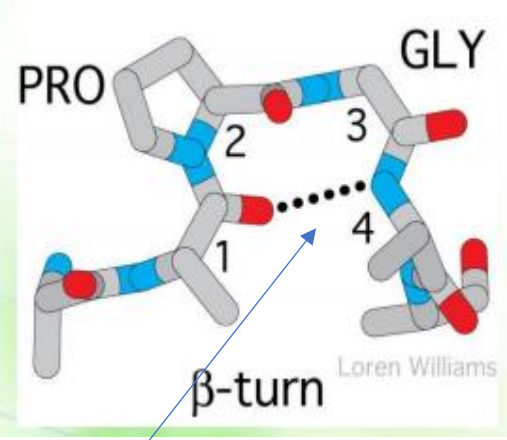
*β-turns:

Turns are compact, U-shaped secondary structures.

They are also known as β turn or hairpin bend.

Now, for the beta sheet structure to form, turns must take place in order for the beta strands to "stack" and make its 3D structure. They tend to be <u>sharp</u>, and are made up of 4 amino acids; the first being any amino acid, the second usually (مش دائما) and preferably would be proline, then glycine ,ending with another random amino acid (check the note below).

*



Note: the composition and the amino acids in the turn is not set in stone and may change from strand to strand but the ones mentioned above are the ones usually present in it.

***(A hydrogen bond forms between the first and the fourth amino acids; forming between the carbonyl group of the first amino acid and the hydrogen bound to the nitrogen of the fourth amino acid.)**

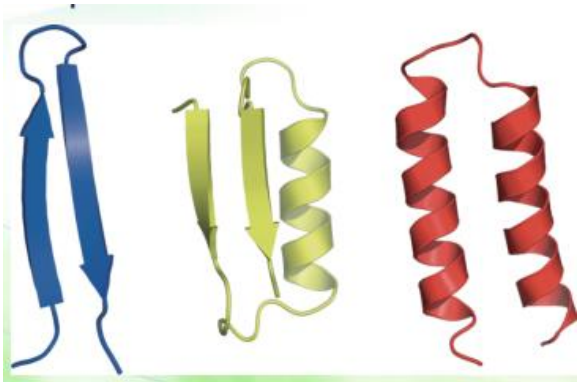
***Note: The importance of proline lies within its rigidity, that causes a kink and breaks the pattern. Where that of the glycine is that it's flexible, small and fits within the turn, not causing any repulsion between amino acids.**

*Loops and coils:

Loops are a diverse class of non-regular secondary structures in proteins with irregular geometry (not a defined structure) and that connect the main secondary structures.

They are found on the surface of the molecule (and contain polar residues) and provide flexibility to proteins. (So, they are flexible)

Amino acids in loops are often not conserved, so they could be anything.



Super-secondary structures:

They are regions in proteins that contain an ordered organization of secondary structures.

(They are collection of multiple secondary structures localized in one region.)

There are at least 2 types:

1- Motifs

2-Domains

Multiple secondary structures

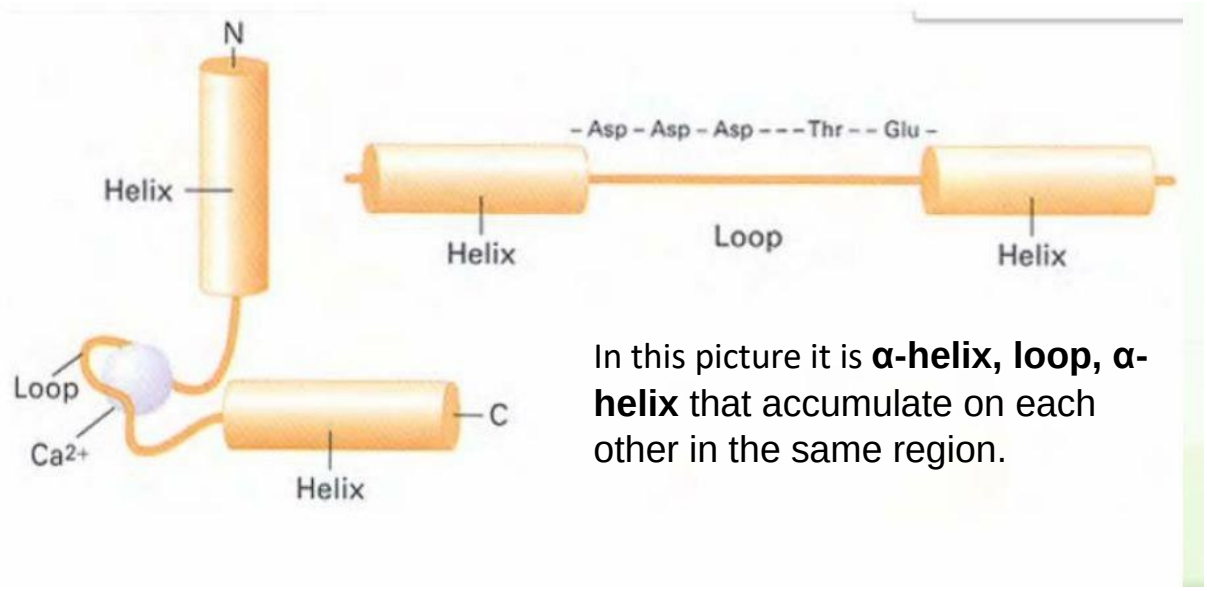
***Let's begin at Motifs:**

A motif (a module) is a repetitive super secondary structure in terms of the primary structure, which can often be repeated and organized into larger motifs, and they can be part of domains.

Example: If we have 6 amino acids that form α-helix followed by a turn, and then another α-helix directly after it. When they get organized together (in the same region), this is what we call a motif.

It usually constitutes a small portion of a protein (typically less than 20 amino acids).

In general, motifs may provide us with information about the folding of proteins (structure), but not the biological function of the protein.

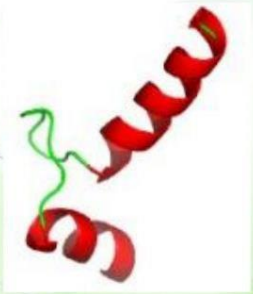


In this picture it is **α-helix, loop, α-helix** that accumulate on each other in the same region.

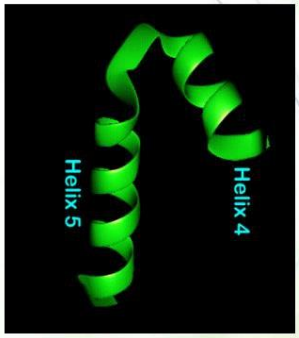
- **Motifs could be simple or complex:**

***Examples of simple motifs:**

Helix-loop-helix is found in many proteins that bind DNA. It is characterized by two α-helices connected by a loop.



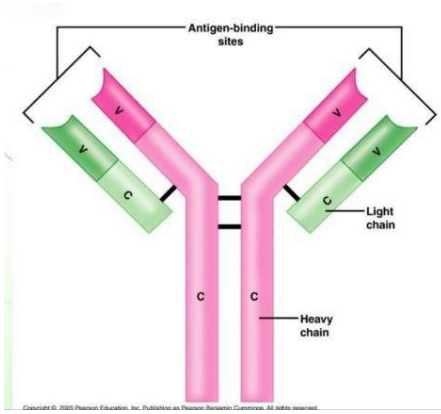
Helix-turn-helix is a structural motif capable of binding DNA. It is composed of two α helices joined by a short strand of amino acids

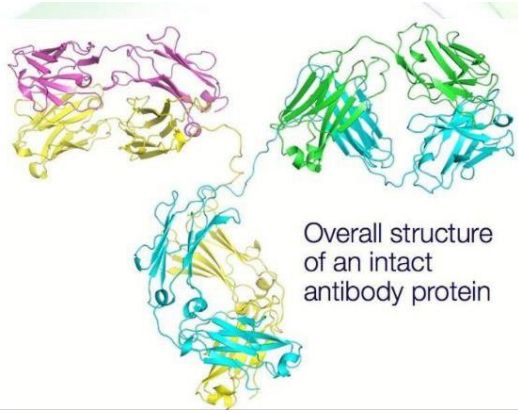


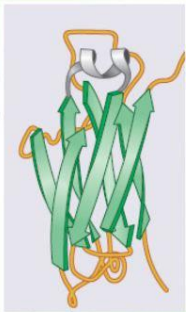
As we said, motifs don't tell us anything about function but usually proteins that bind to the DNA have helix-loop-helix or helix-turn-helix (DNA binding proteins).

***Example of a complex motif:**

The immunoglobulin (antibody) fold or module that enables interaction with molecules of various structures and sizes.







- **The heads of the anti-body (peripheral) are called <u>Antigen-binding sites</u> (the part that interacts with the antigen).**

An antigen is a foreign body. Example: when something enters our body (antigen) like bacteria, viruses, or pollen(لقاح), the B-cells excrete anti-bodies (immunoglobulins).
Now this region (antigen binding site) is a motif and it's a complex motif. Why?? Because it is a β-strand followed by a loop then a β-strand followed by a loop and so on. (Multiple β-strands and multiple loops).

- To summarize:

Secondary structure is well defined structure like: α-helix, β-strands, β-sheets (collection of β-strands), turns, loops, coils. Different shapes but all considered as secondary structure.

Tertiary structure:

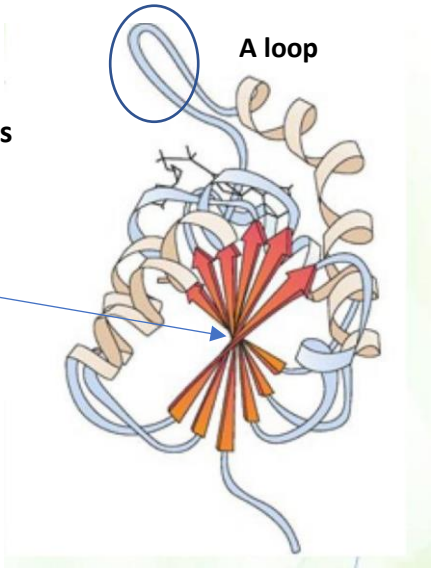
The overall 3-dimensional structure of a protein. (It also shows you the arrangement of R-groups and The overall conformation (folding) of a polypeptide chain.)

The three-dimensional arrangement of all the amino acids residues.

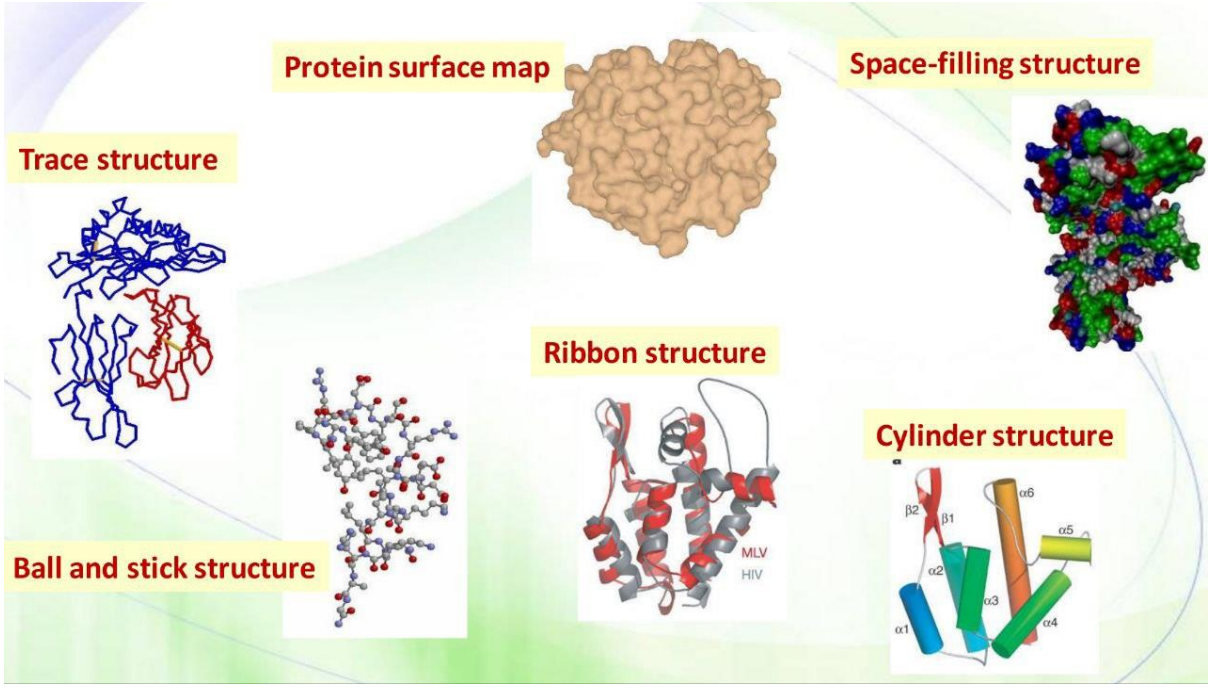
The spatial arrangement of amino acid residues that are far apart in the primary sequence.

We can see at secondary structure all the amino acids in the space.

The arrows showing u β-strands.



Now, how do we (biochemists) illustrate proteins?



- **Trace structure: Represents the backbone only.**

- **Ball and stick structure: Represents the backbone, but it shows the atoms as balls (different atoms have different colors).**

- **Ribbon structure: α-helices are represented as ribbons (شرائط حلزونية) and β-strands are represented as arrows. (Their direction shows the direction of N-terminal to the C-terminal, which is the end of the arrow).**

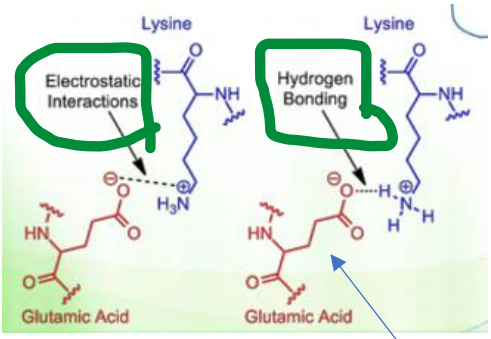
- **Cylinder structure: α-helices are represented as cylinders while β-strands as arrows.**

- **Space filling structure: The atoms are large, so they fill the space.**

- **Protein surface map: This structure shows the surface only from the outside. It doesn't show anything from inside because the importance of this structure is only to show the surface of the protein and where water molecules can interact with the amino acids outside. This is important in designing drugs to see how the drug fits on the surface of the protein.**

*How are tertiary structures stabilized?

By non-covalent interactions between <u>R groups</u> (remember that the secondary structure is stabilized by hydrogen bonds between peptide bonds), there are different types of non-covalent interactions:



The same charged group can form either hydrogen bonding or electrostatic interactions.

-Hydrogen bonds, (it may be between sidechains of R groups and peptide groups).

-Hydrophobic interactions.

-Electrostatic interactions, because these interactions exist between amino acids within the same protein, we call them "<u>salt bridges</u>".

Hydrogen bonds occur not only within and between polypeptide chains but with the surrounding aqueous medium.

Charge-charge interactions (salt bridges between negatively charged amino acid and positively charged amino acid) occur between oppositely charged R-groups of amino acids.

Charge-dipole interactions form between charged R groups with the partial charges of water (hydrogen bond acceptor).

***A smart question: The following two amino acids residues can form salt bridges? (We will have different choices and the right answer is the one that contains negatively charged amino acid with a positively charged amino acid).**

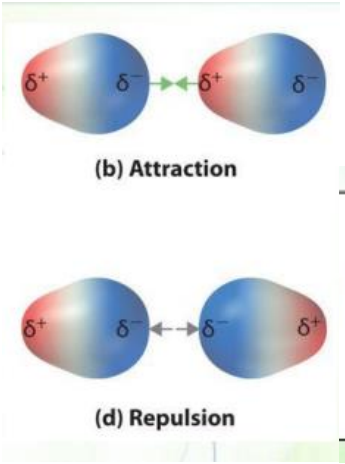
***(Doctor's answer to a student's question: it is possible to have dipole interaction between negative/positive with a dipole for example, hydrogen (partially positive charge) with a negatively charged amino acid)**

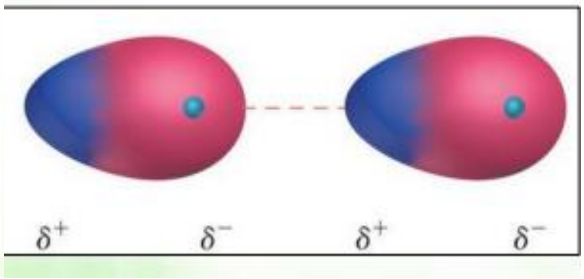
Note: positively charged lysine can form electrostatic interaction (salt bridge) or hydrogen bond with glutamate.

*Van der Waals attraction:

When a movement of electrons occurs, electrons accumulate on one side as a result, we'll have a partial positive charge on one molecule and a partial negative charge on the other. Then they will be attracted by van der Waals.

Van der Waals are very dynamic and transient meaning that they change rapidly.





There are both attractive and repulsive van der Waals forces that control protein folding.

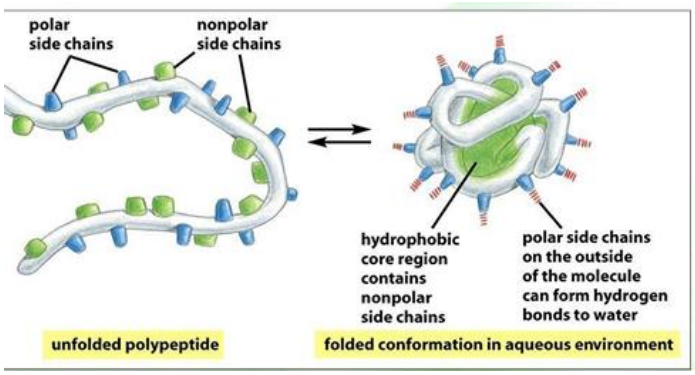
Although van der Waals forces are exteremly weak, they are significant because there are so many of them in large protein molecule.

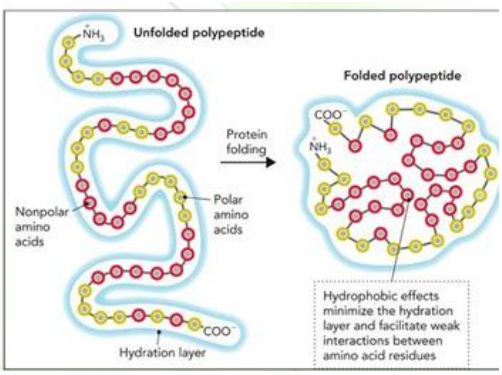
*Hydrophobic interacions:

They are propably the most important type, why?

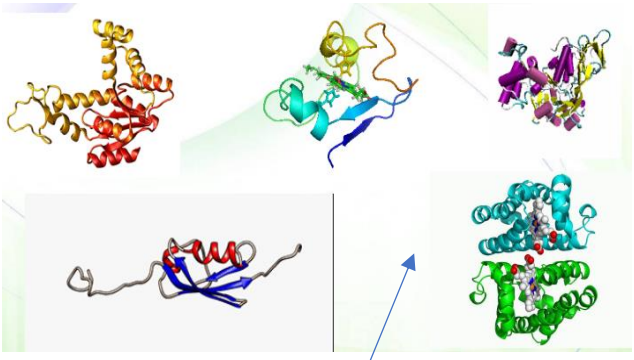
We took in molecular biology that in the process of translation (protein synthesis) the peptides come out of the ribosomes and then, we'll have a stretch of hydrophobic amino acids in the cytosol which is a hydrophillic environment and these hydrophobic amino acids will cluster together so the hydrophobic regions can hide from the water in the cytosol and on the surface we'll have the hydrophillic amino acids because this is more thermodynamically (energetically) stable.

So, A system is more thermodynamically (energetically) stable when hydrophobic groups are clustered together rather than extended into the aqueous surroundings.





But! Can polar amino acids be found inside? YES; because it has a functional role or a structural role (In this case, they form hydrogen bonds with other amino acids or with the polypeptide backbone), and this works on hydrophobic amino acids too, meaning that we can find hydrophobic amino acids on the surface although they are hydrophobic and that's because they have a functional or a structual role and to accomplish their role they must be on the surface.



***The process of protein folding begins when the protein gradually folds into a three dimensional structure, it will form globular structure that would be thermodynamically stable. (You can see a video on youtube of the process.)**

Proteins are NOT static!

Proteins, DNA and RNA are not static in the cell, they are dynamic structures that are stabilized by:

-Non-covalent interactions (as we explained in the previous slides) (they contribute to the stability and the conformation of protein.)

-Disulfide bonds (covalent bonds).

-Metal ions.

Stabilizing factors:

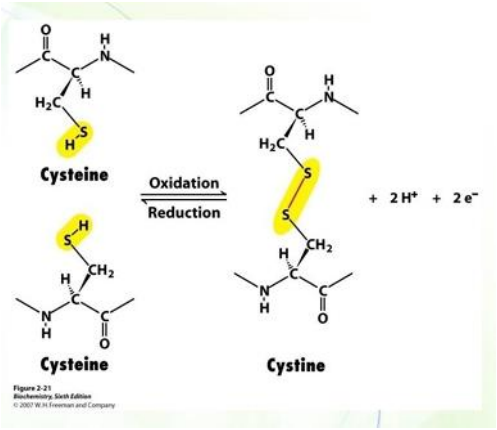
***There are two forces that do not determine the three-dimensional structure of proteins, but stabilize these structures:**

-Disulfide bonds **-metal ions**

Let's begin at Disulfide bonds:

The side chain of cysteine (thiol) contains a reactive sulfhydryl group (-SH), which can oxidize to form a disulfide bond (-S-S-) to a second cysteine.

The crosslinking of two cysteines to form a new amino acid, called cystine.



Cystine helps in stabilizing protein's structure but the conformation of the tertiary structure of the protein is dependent on the non-covalent interactions especially hydrophobic interactions.

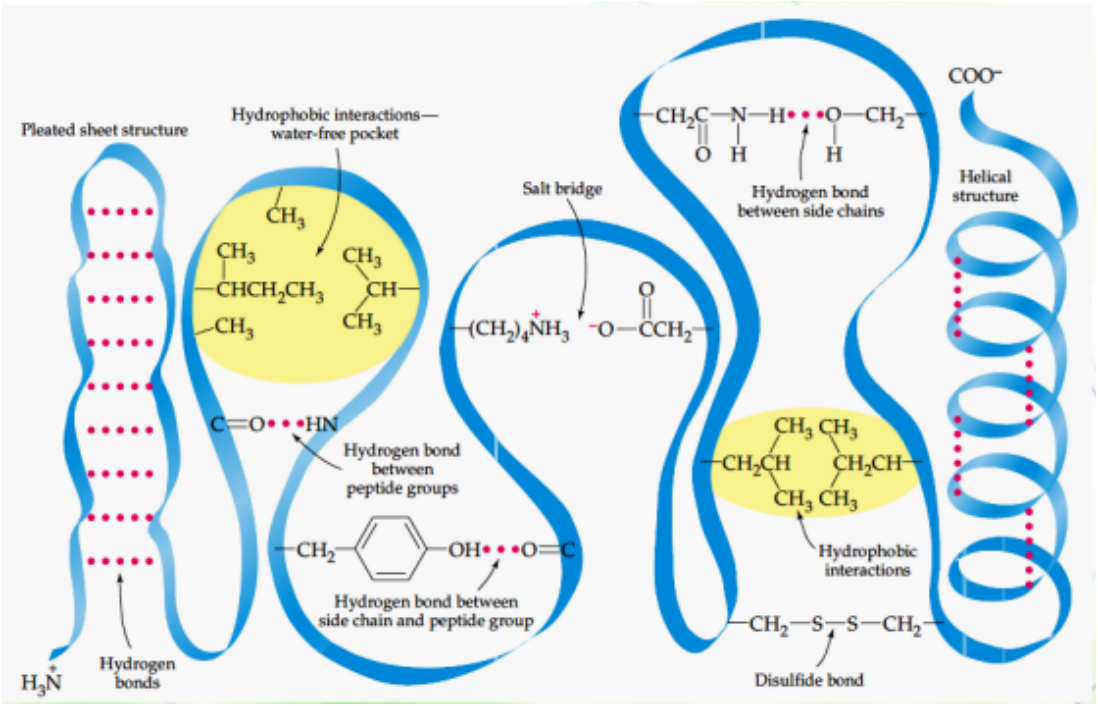
Note: There is a sulfur in methionine's structure but it is in the thioether group so it's already occupied and is therefore hard to oxidize

Metal ions:

Several proteins can be complexed to a single metal ion that can stabilize protein structure by forming:

-Covalent interaction (myoglobin).

-Salt bridges (carbonic anhydrase).



{وأخر دعواهم أن الحمدالله رب العالمين}