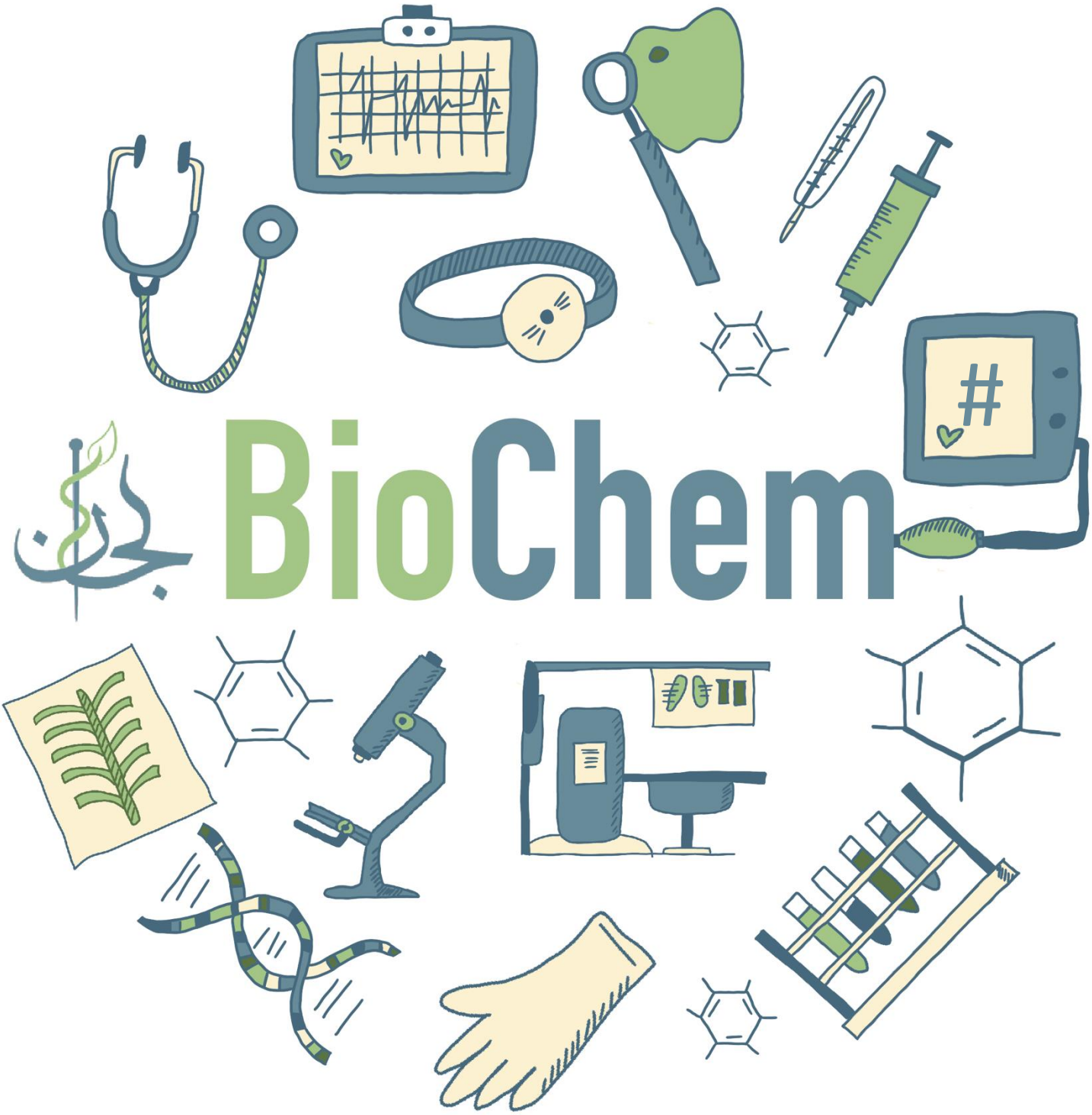


BioChem

Writer: Haitham Alsaifi & Ola Nawafleh

Science: Haitham Alsaifi

Grammar: Ola Nawafleh

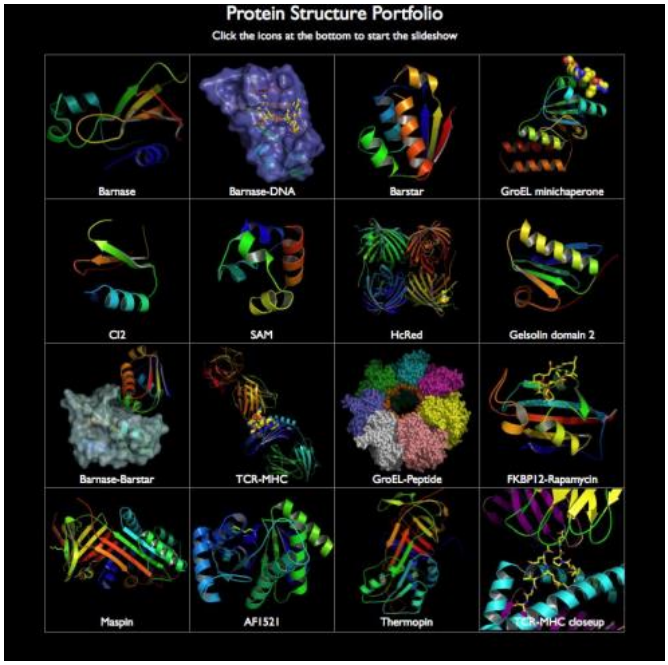
Doctor: Mamoun Ahram

In the previous sheet, we talked about peptides and the formation of a peptide bond between two amino acids. Now, we will talk about polypeptides and how they form a protein with a distinct three-dimensional structure which gives the protein its function.

Overview of proteins

Proteins consist of different amino acids which give the protein its unique three-dimensional structure. The variety of proteins is mainly due to the huge possibilities of sequencing the amino acids, for example, if we have to choose 100 amino acids to make a polypeptide, **what are the possibilities for making different structures of proteins?**

There is a lot of probabilities for that, because for each amino acid there is 20 different possibilities, so for the 100 amino acid there is 20^{100} different structures we can make.



However, the number of ***functional* structures** that can be formed, is about 1 or 2, because there are forces and rules that determine how proteins form their three-dimensional structures. This functional three-dimensional structure is called **native conformation**.

- ✓ Each protein has its own **native conformation**, that determines its function.

Scientists have been working on a project for fifty years, they were trying to predict the three-dimensional structure of proteins according to the amino acids sequence, finally , there were reports **recently** claim that European Scientists have succeeded in creating an artificial intelligence system called "***Alpha fold 2***" that can predict the structure of a protein by only knowing the amino acids sequence. The importance of that is, once you know the three-dimensional structure, *you will be able to*:

1.know **the function** of that protein.

2. know the structure of **abnormal proteins** that cause certain diseases.

3. **Design drugs** that inhibit or activate this protein.

When we talk about protein structure, we simplify it by dividing the complexity of protein structure into **four levels**:

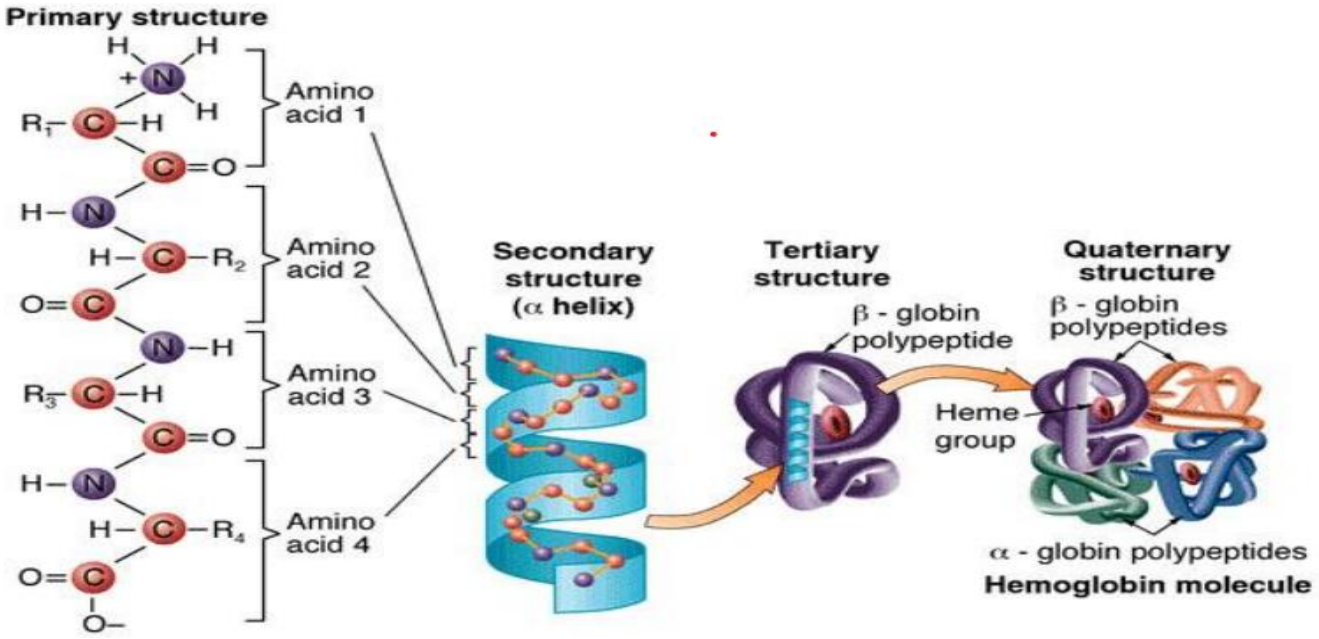
primary structure: the simplest structural level of proteins ,it is basically the sequence of amino acid residues ,like when you say , **Met,Asp,Try,Ser,.....etc.**

Secondary structure: a higher level of protein structure, where you have **localised structures** (it is a part from a polypeptide that is **organised**)

Tertiary structure: the **three-dimensional structure** and/or arrangement of all the amino acid residues of a polypeptide chain.

Quaternary structure: it describes the number and relative positions of the subunits in a **multimeric protein** (a protein that is made from **multiple subunits** (polypeptides)).

☒ **Notice** that in **the primary structure** we have the amino acid sequence starting from the **N-terminus** and finishing by the **C-terminus**, also notice the *zig-zag* shape resulting from the **trans** orientation between bonds. Also, **the secondary structure** where you can see it is a part of the protein having an organized specific structure which is the **α helix** (will be discussed later on). This **α helix** is part of the **the tertiary structure** of one polypeptide. Finally, **the quaternary structure** that consists of multiple polypeptides joined together in a specific three-dimensional way.



Now, we will talk about each structure specifically:

Primary structure

The order in which amino acids are covalently linked together, Example:

N-terminus Leu—Gly—Thr—Val—Arg—Asp—His **C-terminus**

The primary structure is very important, as we said before, knowing the amino acid sequence leads us to know the three-dimensional structure and function of that protein, Also, the primary structure of a protein determines the **other levels** of structure.

Proteins that differ somewhat in primary structure and properties from tissue to tissue, but that retain essentially the same function, are called tissue-specific **isoforms** or **isozymes**.
An example for that is **hemoglobin**, there are different hemoglobin polypeptide chains, the **alpha chain**, **Beta**, **gamma**, **zeta** ,these all considered **isoforms** .

If you look closely at the polypeptide chains, you will see that there are regions that are **highly similar** to each other, but also there are some **differences**, these differences are the reason that give those proteins specific characteristics (structural, functional, regulatory differences).

Furthermore, there are proteins that differs among different species like the protein **GATA2**. if you compare the amino acids that make that protein between humans, mice, and zebrafish you will notice that there are **homologous sequences** among these proteins with some differences indeed.

An example of the importance of the primary structure of a protein is **sickle cell anemia**.

Sickle cell hemoglobin (HbS)

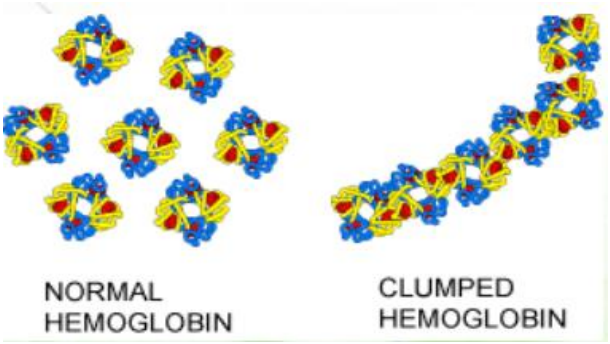
Sickle cell anemia is basically a disease that is cause of a **defective hemoglobin**, which is a protein that is responsible for transporting oxygen in our blood, this indicates how much this protein is **important** to our cells.

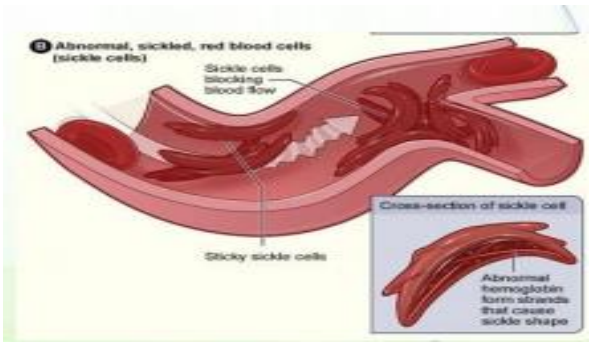
☒ **So, what is the cause of that defect of hemoglobin?**

The cause of that is a change in **the amino acid sequence of hemoglobin**, the change occurs in the **6th position** in **beta** globulin due a genetic (**point**) mutation, instead we have a **glutamate** (a negatively charged amino acid), we have **valine** (aliphatic non-polar amino acid). This is a huge difference, changing a charged amino acid into something that is non-polar, this will disrupt the protein structure**!**

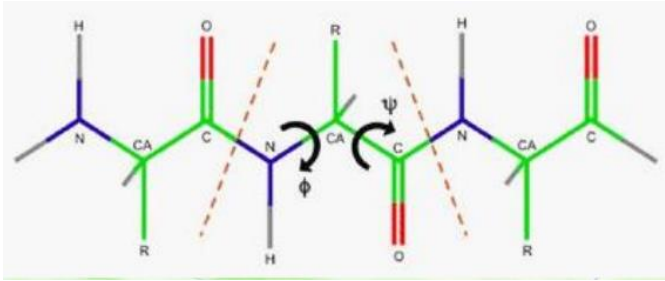
The mutation results in:

1. It changes the protein structure in a way that the proteins are clustered together forming an **array**, instead of having each hemoglobin by **itself**.
2. **Deformation** of the red blood cell, the clumping of hemoglobins will change the structure of the **RBC** and **elongates** it.
3. Clotting in blood vessels and tissues this will result in tissue/organ **failure** because there is **no** oxygen entering the tissues.





Secondary structure



it is organized amino acids into a specific structure. **Recall** that the peptide bond is rigid and doesn't rotate because of the resonance and the alternating double bonds. However, the bonds within the amino acid residue around the central carbon can rotate (the bond between the α-carbon and the **carboxyl carbon** and the bond between the **α-carbon** and the **amino nitrogen**), and this gives the protein its unique three-dimensional structure. That rotation results in hydrogen bonding between the groups that make the peptide bond, and these hydrogen bonds can result in one of the following regular structures:

1.Alpha helix **2.Beta-pleated sheet** **3.Turns** **4.Loops/coils**

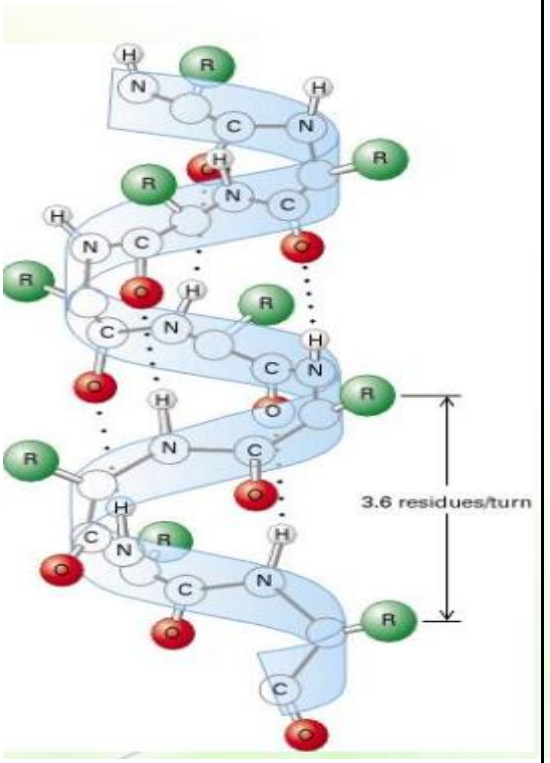
1.The α Helix

The first secondary structure that we want to take about is the α helix :

- **It looks like a helical rod .**
- **The helix has an average of 3.6 amino acids per turn .**
- **The pitch of the helix (the linear distance between corresponding points on successive turns) is 5.4 Å(10^- 10)m .**
- **It is very stable because of the linear hydrogen bonding.**

The R groups are oriented to the outside

3.6 amino acid residue means: 3 amino acid and a part from the 4th one.



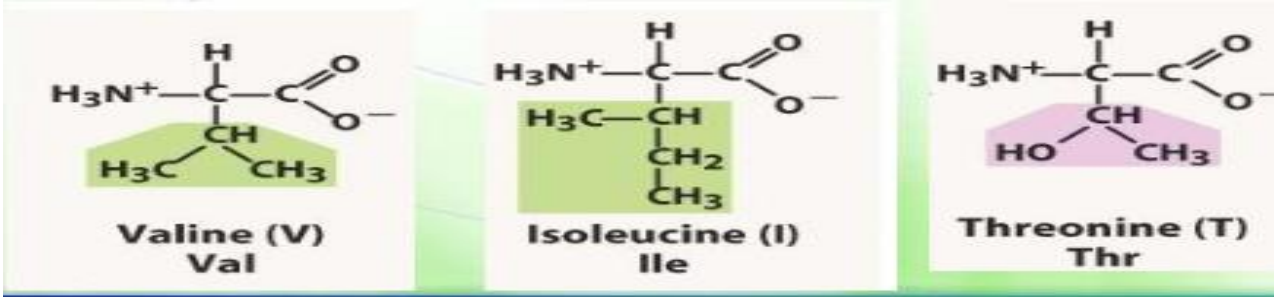
WHAT AMINO ACIDS DO NOT EXIST IN α -HELIX AND WHY ?

There are amino acids that do not prefer to be a part of α helix, like :

- **Glycine**; because it's too small and doesn't have an R group that supports the α helix.
- **Proline**; it breaks the structure because it is a rigid amino acid, it does not allow rotation around N-C bond, and there is **NO** hydrogen bonding of an amino group (because the N is occupied by 3 R groups, so the N can't do hydrogen bonding)

- Amino acids with branches at the β- carbon (**Valine** , **Threonine** and **isoleucine**), because these branches increase the repulsion in the structure. (look at the image below)

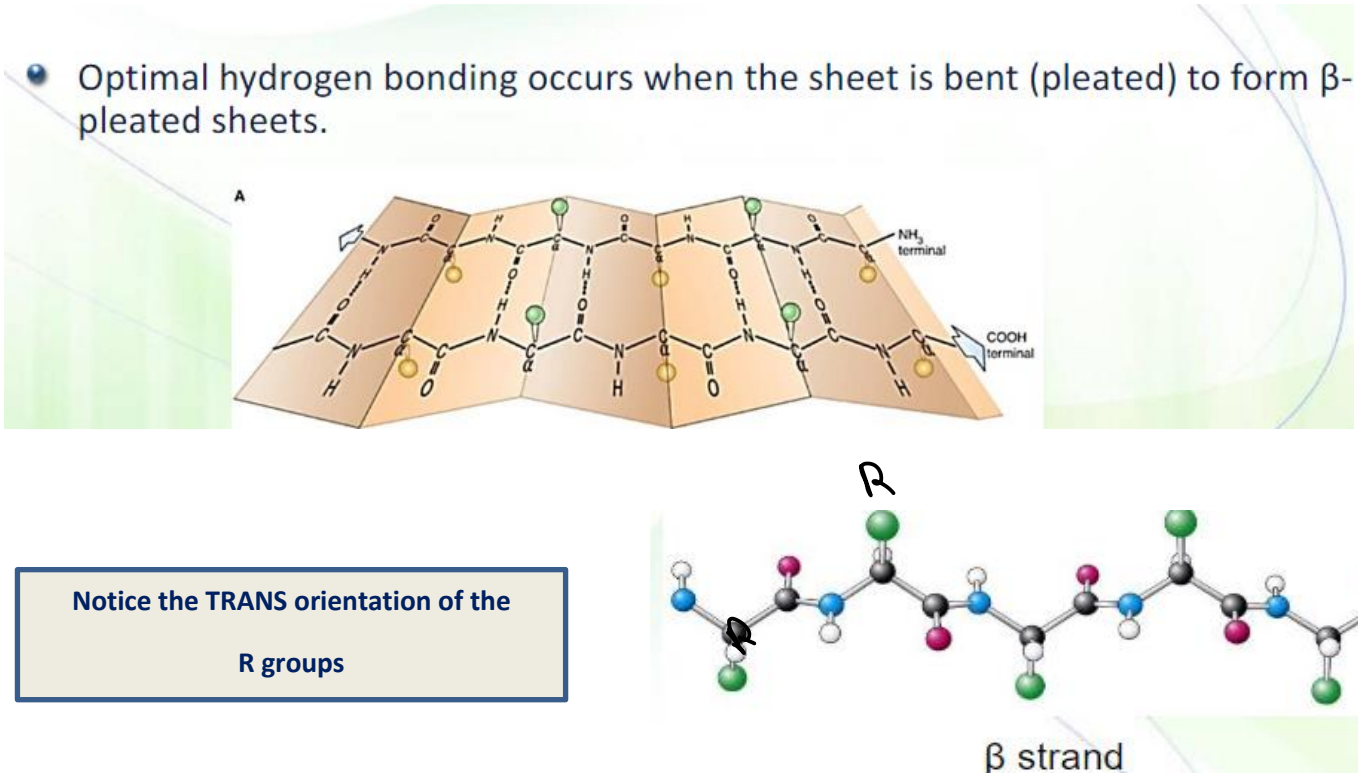
- Close proximity of charged amino acids with **similar charges** (because this causes repulsion). .



2) β-pleated sheet (β sheet)

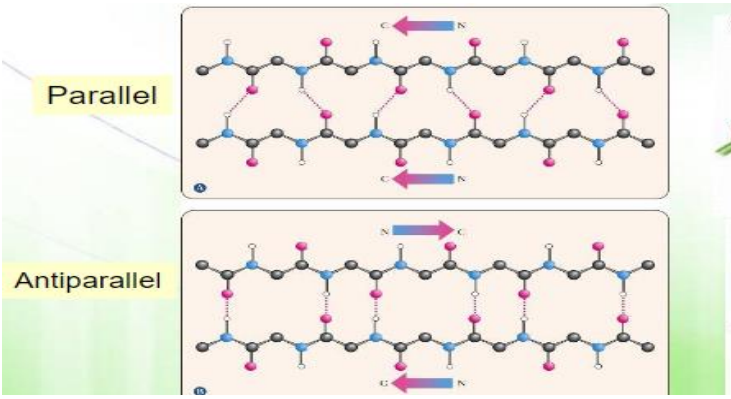
The second type of secondary structure is β-pleated sheet, basically it is a combination of beta strands **(two or more straight chains (β strands) that are hydrogen bonded side by side) .**

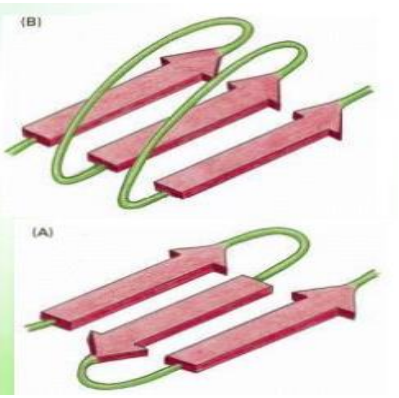
- **Notice** the hydrogen bonding between the groups that form the peptide bond in each strand, which are the **amid group** and **the carbonyl group**.



Notice the TRANS orientation of the R groups

PARALLEL VS ANTIPARALLEL B SHEETS





THEY GO IN THE SAME DIRECTION

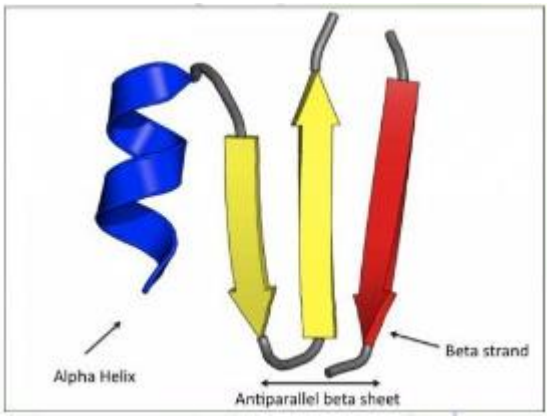
THEY GO IN DIFFERENT DIRECTIONS

Notice that there are different hydrogen bonding pattern between them:

- In **parallel sheets:** one amino acid would form hydrogen bonds with **two** different antiparallel amino acids separated by a third amino acid.
- In **antiparallel sheets**: there are **two** hydrogen bonds between any **two** antiparallel amino acids.

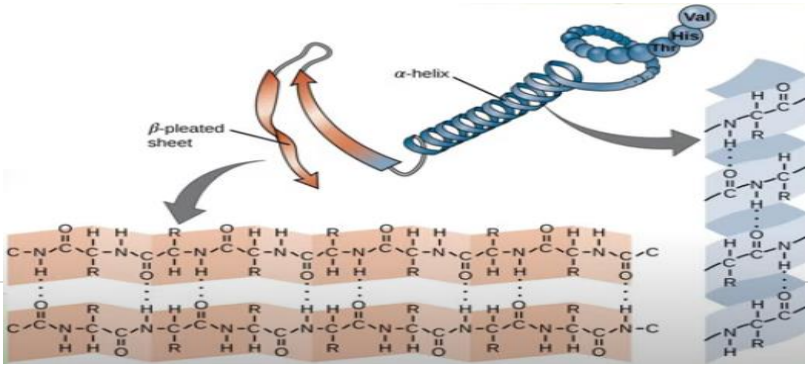
- **Notice that you can have parallel and antiparallel strands in the same beta sheet.**

- β sheets can be formed between many strands, typically 4 or 5 but as many as 10 or more.
- Such β sheets can be purely antiparallel, purely parallel, or mixed.



- We have amino acids that support the β sheets like Valine, threonine and Isoleucine (**the same ones that disrupt α helices are existed with β sheets)** with branched R groups at β-carbon and large aromatic amino acids like (phenylalanine, tryptophan, and tyrosine) tend to be present in β-sheets.
- Proline tends to disrupt β strands (because it is rigid and can't form hydrogen bonds).

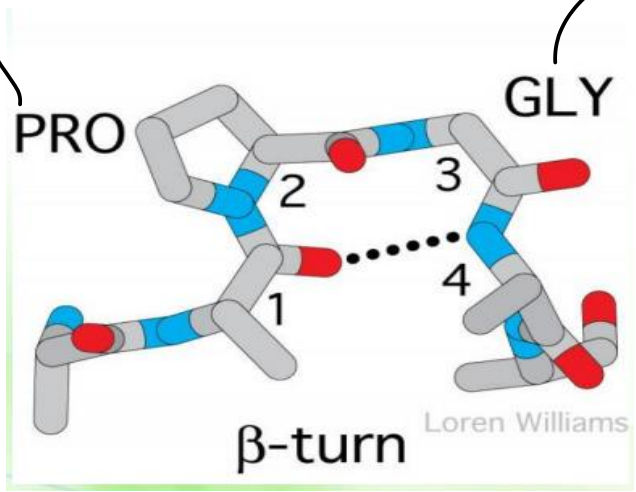
➢ α helices are represented as **helical** rods and β sheets are represented as **thick arrows**.



3) β-turns

Turns are compact, U-shaped secondary structures, they are also known as **β turn** or **hairpin bend**, they are used to form the 3D structure of protein like some proteins that have a globular structure and these molecules help in forming them.

Because its rigid and creates a turn



Because its small and can fit in a small place and allow the peptide to turn smoothly

Hydrogen bonds are formed between groups that make the peptide bond **(this stabilizes the turn)**

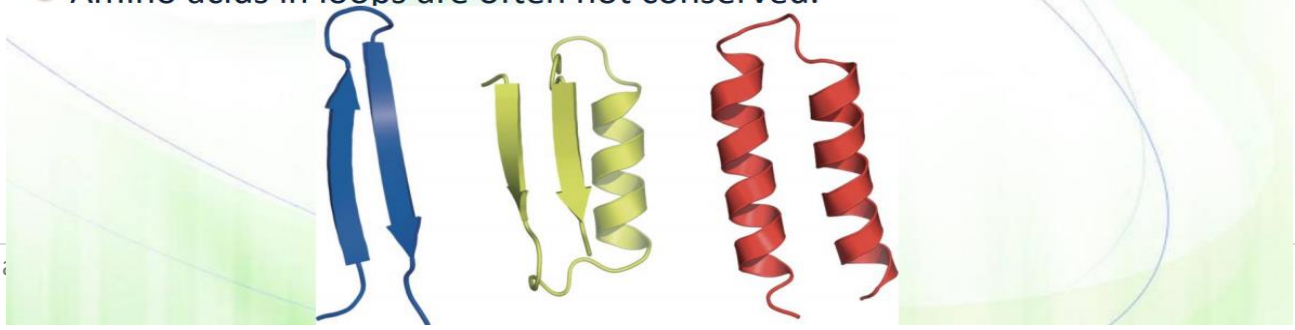
4) loops and coils

Loops are e a diverse class of secondary structures in proteins with irregular geometry and that connect the main secondary structures

The protein will form the same type of loops every single time it produced, but a loop in one protein would have a different structure than a loop in another one.

The amino acids in the loop are not conserved meaning that you can have any types of amino acids making a loop.

- They are found on surface of molecule (and contain polar residues) and provide flexibility to proteins.
- Amino acids in loops are often not conserved.



Super-secondary structures

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

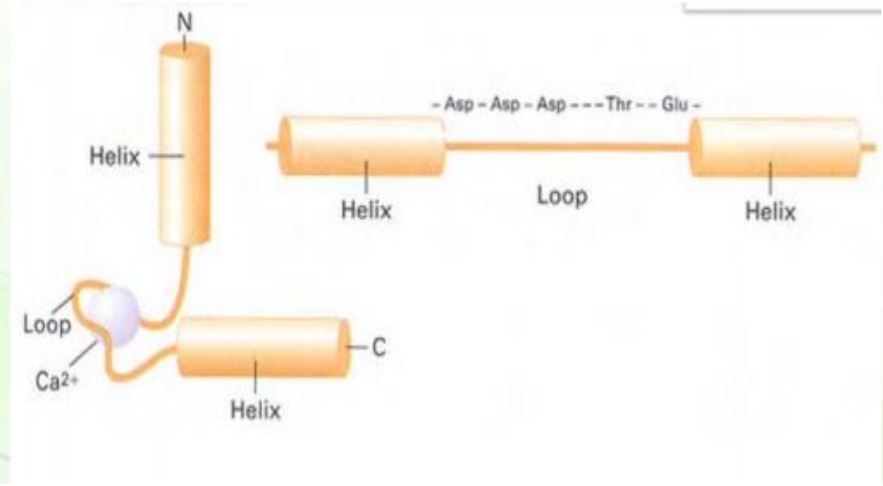
- ✓ ***There are two types of super-secondary structures*:**

1. Motifs
2. Domains (will be discussed in another lectures).

A motif (a module)

A motif is a repetitive super secondary structure, which can often be repeated and organized into larger motifs.It usually constitutes a small portion of a protein (typically less than 20 amino acids). But they can really be large ones like in immunoglobulins.

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

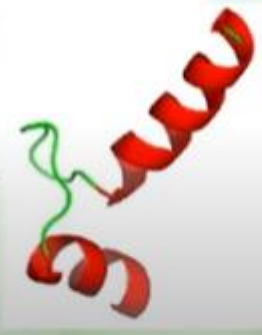


Notice that we have a sequence of amino acids that form **α helix** followed by amino acids that form a **loop** and they are <u>not</u> separated by anything, and this structure is called **α HELIX-LOOP-HELIX MOTIF**.

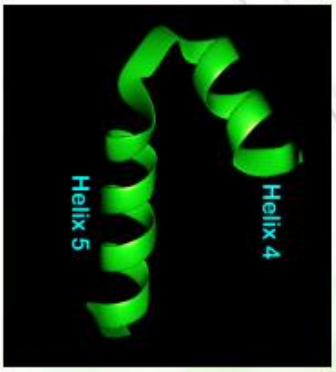
- ITS JUST A STRUCTURE WITH NO FUNCTION

EXAMPLES OF 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

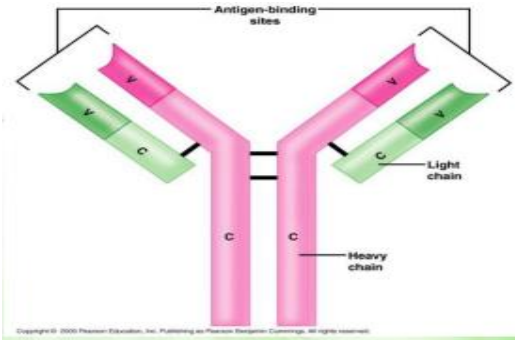


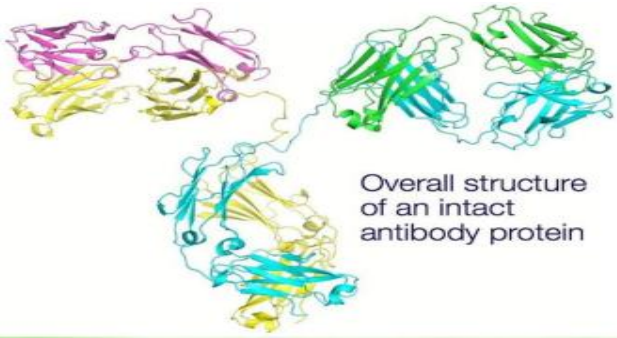
YOU HAVE TO KNOW THAT MOTIF DO NOT SPECIFY A FUNCTION.

Immunoglobulin (a more complex motif):

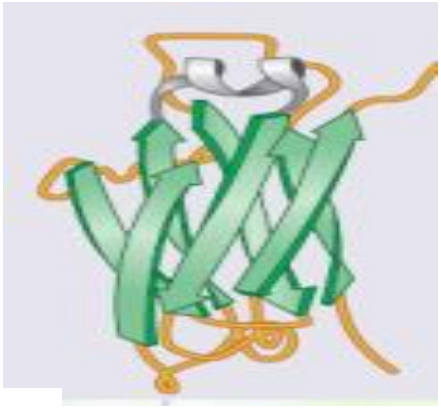
Immunoglobulin is an antibody that recognizes an antigen or foreign bodies, immunoglobulin motif is a part of this protein that recognize the antigen.

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





Notice that these strands connected by loops, so they are repetitive secondary structure (**β strand -loop- β strand**), what is called a motif.



The End

Self-assessment questions

1) what is the direct cause of sickle cell anemia?

A. substitution of single amino acid
B. disruption of disulfide bonds.
C. incorrect folding of hemoglobin
D. all of the above

Answer: A

2) Amino acids found at beta turns must be ……such as…….. .

A. polar such as Arginine
B. small such as glycine
C. hydrophobic such as alanine
D. hydrophilic such as serine

Answer: B

3) A young black man was brought to the emergency room (ER) owing to severe pain throughout his body. He had been exercising vigorously when the pain started. He has had such episodes about twice a year for the past 10 years. An analysis of the blood shows a reduced blood cell count (anemia) and odd-looking red blood cells that were no longer concave and looked like an elongated sausage.

An underlying cause in the change of shape of these cells is which one of the following?

A. Increased ionic interactions between Hb molecules in the oxygenated state
B. Increased ionic interactions between Hb molecules in the deoxygenated state
C. Increased hydrophobic interactions between Hb molecules
D. Increased phosphorylation of Hb molecules in the oxygenated state

Answer: c

4) The three-dimensional arrangement of all atoms is …….. **(tertiary structure)**

5) The order of amino acid residues in the polypeptide chain is……. **(primary structure)**

6) The interaction between subunits in proteins that consist of more than one polypeptide chain……. **(quaternary structure)**

.7) The hydrogen-bonded arrangement of the polypeptide backbone……. **(secondary structure)**