Introduction to Biochemistry and Molecular Biology

Lecture {14}

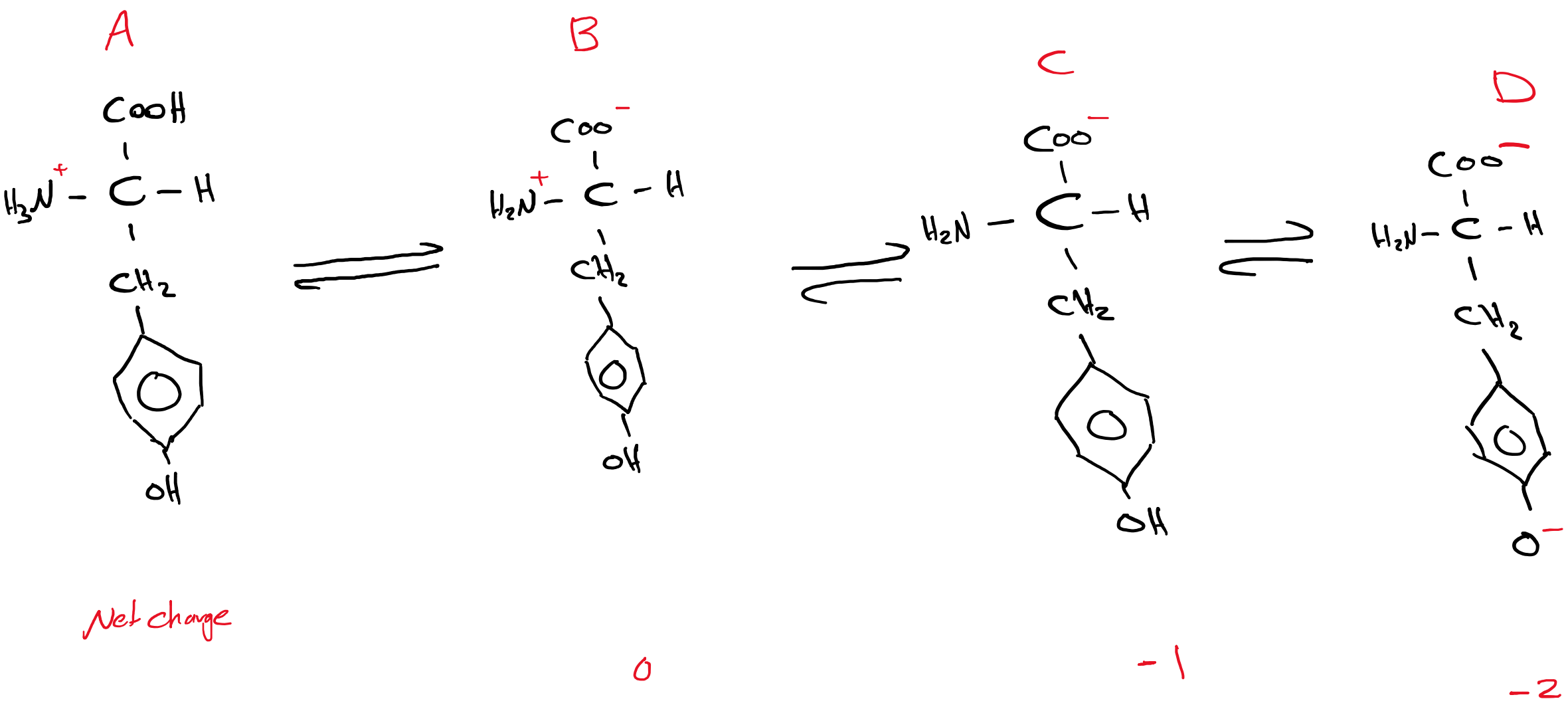
Protein structure

" لو أنَ النّاسَ كُلما استصعَبوا أمرًا تَركوه ، ما قامَ للناسِ دُنيا ولا دين"
عمر بن عبد العزيز

Written by:
Rama Al-oweyrat & Deema Nasrallah

Edited by :
Sadeel Al-hawawsheh

Notes about titration and ionization of Tyrosine



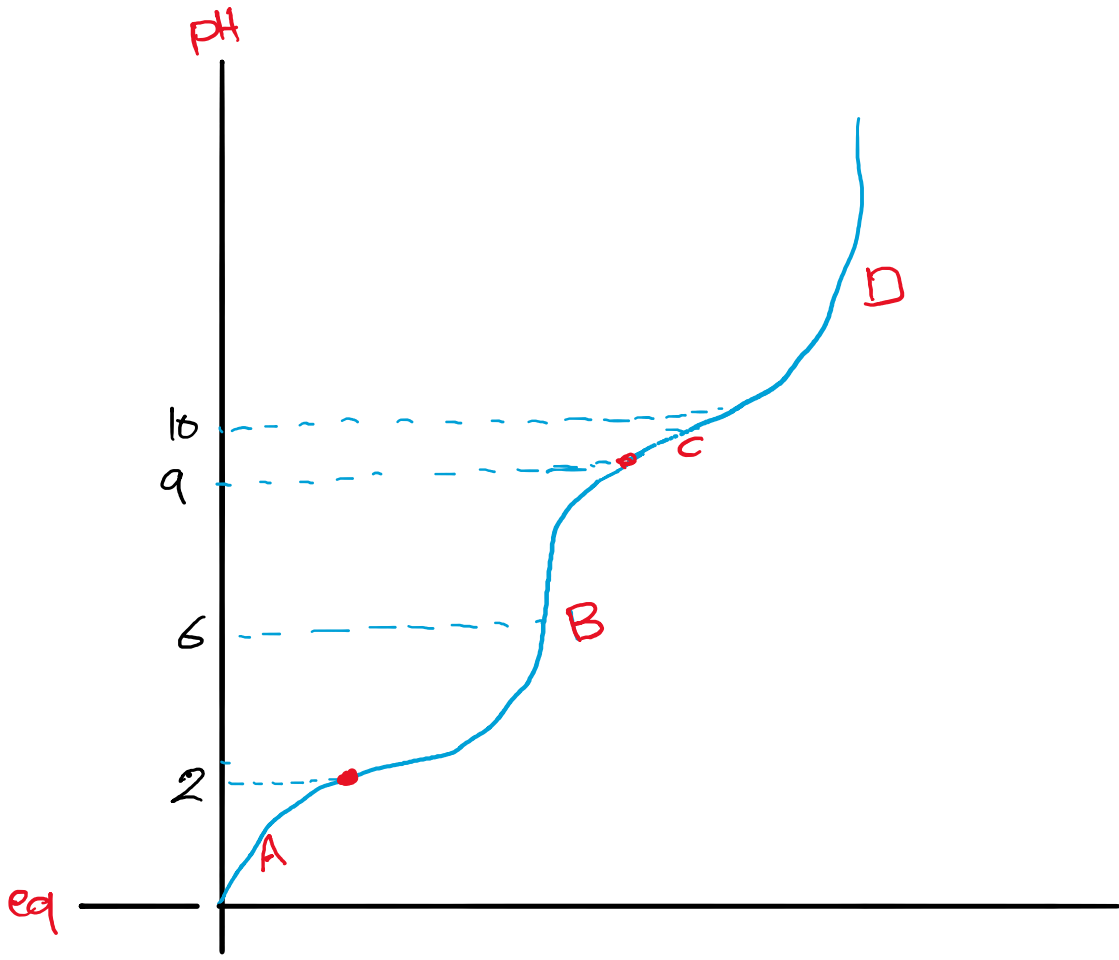
In region A pH is very small and net charge +1
(All groups are protonated COOH , R , NH3+)

In region B net charge =O
(Amino group and R group are protonated but carboxyl group is not protonated

In region C net charge = -1
(R group is protonated but Carboxyl group and amino group are not protonated)

In region D pH is very large net charge = -2
(All groups are not protonated)

Zwitterionic form will be in region B and when we take the average of pKa1 and pKa2 , we will find that it equal to 5.5 not 9.5 as it found in Amino acid File , slide 46



Protein structure

Amino acid→ polypeptide → protein

The difference between polypeptide and protein

Polypeptide→ string of Amion acid that are connected by peptide bond with no specific structure.

* Every time a cell produces a polypeptide, its structure is random (it does not have to be string it can wrap around itself with non specific structures

Protein→ protein has a definite structure (3D structure)
* every time a cell produces protein, each type that will be produced has the same structure.
e.g. : every time a cell produces hemoglobin it has same structure , but when it produces another protein it will have completely different structures
(Simply : comparing different types of proteins , we have different structures)

What about protein structure ? 🤔

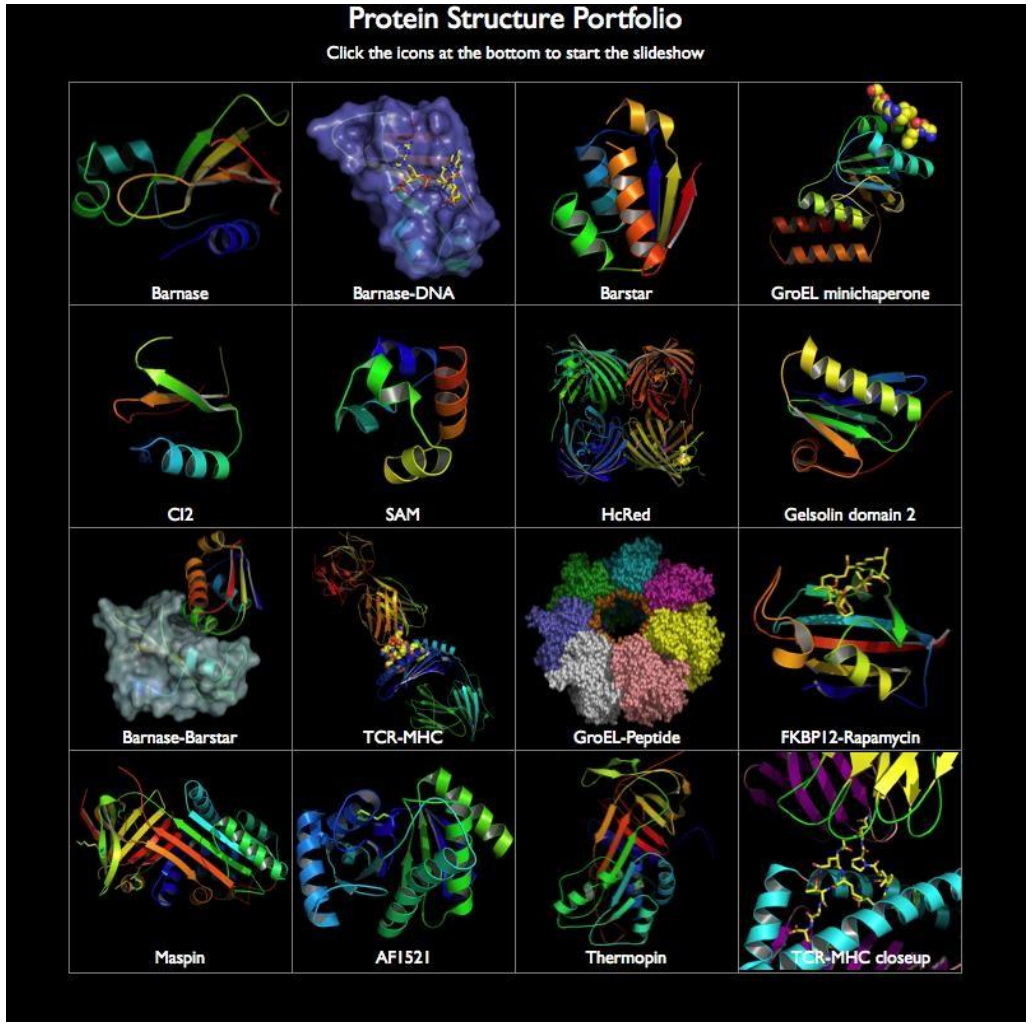
A **protein** consists of a large number of amino acids, so it is believed that there can be a very large number of expected structures for the resulting protein as a result of the amino acids taking a different order each time, but on the contrary when a protein is formed each sequence of Amino acid depends on Many factors, such as chemical and physical factors...etc. to form in certain way that gives the protein a specific shape, so the resulting protein has the most energetically stable structure . E.g hemogloglobin takes up its structure because this is the most energetically stable structure of it

(Native conformation)

Overview of proteins

All proteins have totally different structures (elongated , spherical , triangular , Y-shape)

- Proteins have different structures, and some have repeating inner structures, other do not.
- A protein may have *gazillion* Infinite number possibilities of structures, but a few would be active.
- These active structures are known as native conformations (the 3-dimensioanl structure of a properly folded and functional protein).



Folding protein is important for its stability and function
The function of the protein depends on its structure. If there is a defect in its structure, the function of the protein will also be disrupted.

This is study that conducted in 2021 . since 1960 , scientists were very interested in prediction the protein structure.

Tunyasuvunakool, K., Adler, J., Wu, Z. et al. Highly accurate protein structure prediction for the human proteome. Nature (2021). https://doi.org/10.1038/s41586-021-03828-1

Highly accurate protein structure prediction for the human proteome It is not accurate to a satisfactory level , but it is getting better

Kathryn Tunyasuvunakool, Jonas Adler, [...]Demis Hassabis

Nature (2021) | Cite this article

115k Accesses | **1** Citations | **1308** Altmetric | Metrics

In 2021, two programs used artificial intelligence to predict proteins structures based on the sequence of amino acids. Based on this, the structure of a million proteins was discovered. The number seems huge, but it is very small compared to the total number of proteins in nature . .... Within months, these programs predicted 500 million proteins to a high level of accuracy

AAAS Become a Member

Science Contents News Careers Journals

New public database of AI-predicted protein structures could transform biology

By Robert F. Service | Jul. 22, 2021 , 11:00 AM

Synced

AI TECHNOLOGY & INDUSTRY REVIEW

GLOBAL NEWS INDUSTRY COMPANY RESEARCH COMMUNITY NEWSLETTER CONTACT US

Why did scientists care about knowing the protein structure? 🧐

If I know the structure, I will know the :

1. Function
2. Localization
3. Interactions between proteins
4. I can create drugs that target certain protein
5. I know if one amino acid changed , how the structure and the interactions between proteins would change

Levels of protein structure Four levels of complexity

- Primary structure: the sequence of amino acid residues
- Secondary structure: the localized organization of parts of a polypeptide chain Local region Has a certain pattern
- Tertiary structure: the three-dimensional structure and/or arrangement of all the amino acids residues of a polypeptide chain 3D structure of a polypeptide (in space)
- Some proteins are made of multiple polypeptides crosslinked (connected) with each other. These are known as multimeric proteins.
- Quaternary structure describes the number and relative positions of the subunits in a multimeric protein How polypeptides are together and how they are organized in space and how they make interactions with each others to create the protein

يعني الرابع

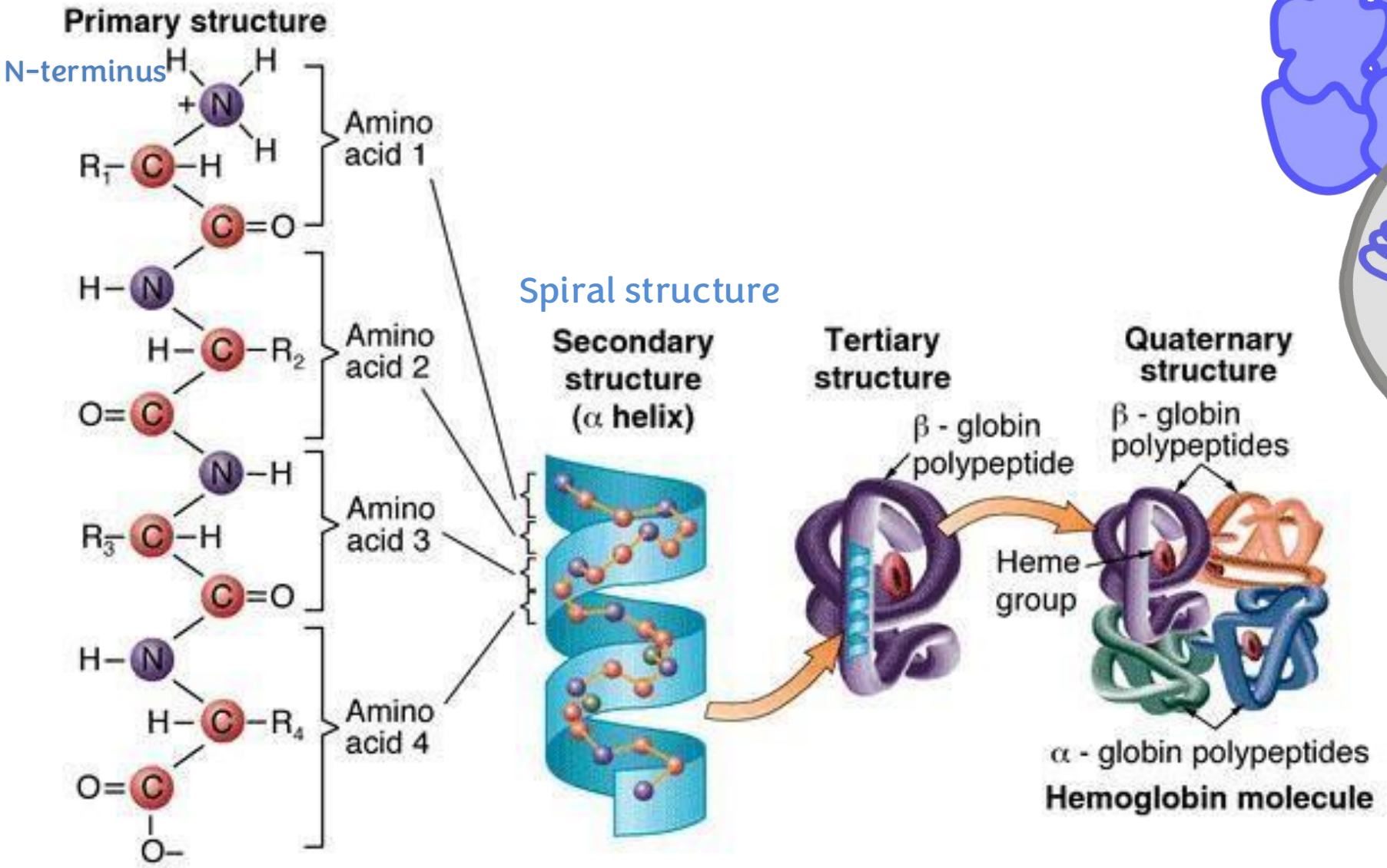
Imagine you are in a helicopter flying above some tall buildings (For example five buildings near each other) **(polypeptide chains-quaternary)**

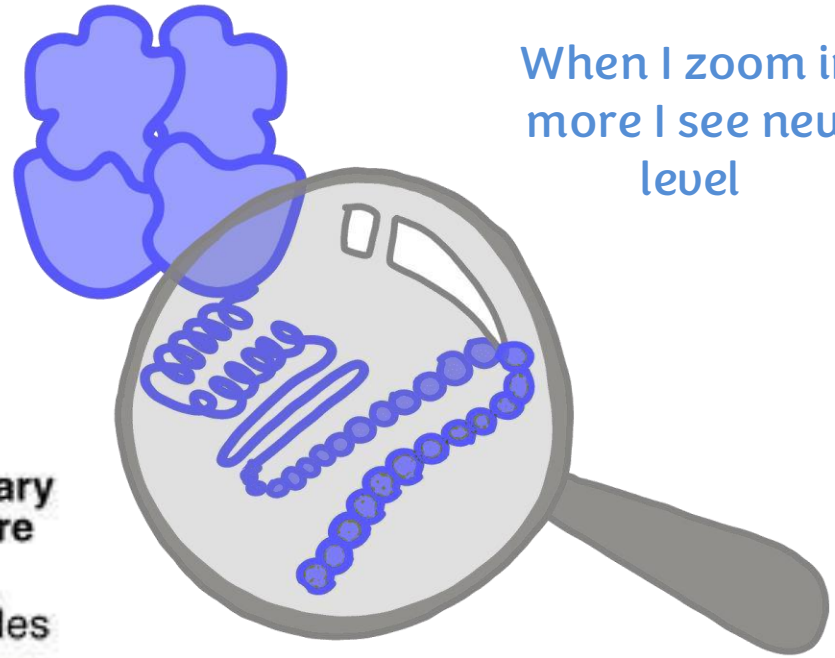
You get off the helicopter in front of one of these buildings and you see its three dimensional structure (its length , width , its doors and windows…) **(The proteins 3D structure-tertiary structure)**

You entered the building and you see that it has many doors with the same pattern & it has groups of windows with the same shape and colour **(secondary structure)**

You get closer and you see the blocks that made up the building **(primary structure)**

Click here





When I zoom in more I see new level

تَبارَكَ الله أحسنَ الخالِقين ، الله سبحانه وتعالي أعطانا أجسام ذكية بتصنع ملايين البروتينات من اكثر من حمض اميني وبتسلسل محدد وبتمر بأربع مستويات معقدة بدون أي خربطة

Primary structure

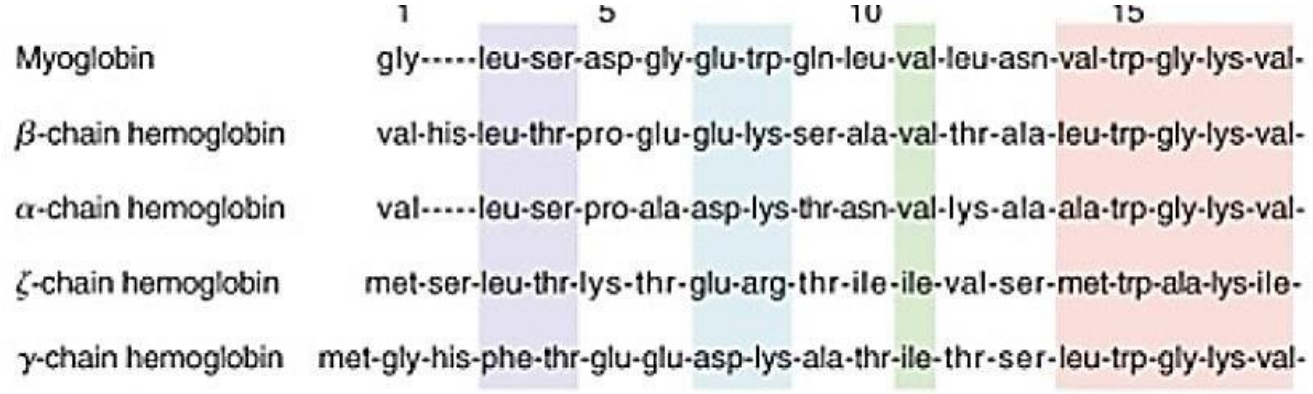
What is primary structure?

The primary structure determines the structures of all levels

- The order in which the amino acids are covalently linked together.
  - Example: Leu—Gly—Thr—Val—Arg—Asp—His
- 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.

Starting from N-terminus and ending with C-terminus

We have similarities in the primary structure in different hemoglobin molecules



Someone asked the doctor how we would know if the AI gives accurate information the doctor said that we test the AI with information we already knew

Artificial intelligence programs took all the information from million proteins (that we have already discovered) and the AI made predictions about the structures of new proteins based on their sequence

We have similarities in certain enzyme (protein) in different organisms

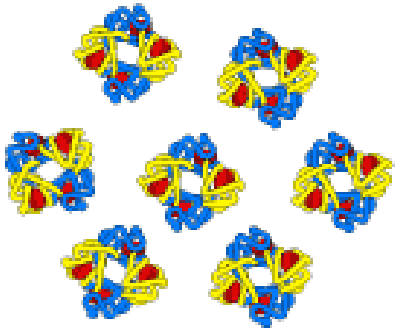
It is very important enzyme that did not change during millions of years and we can use the mouse sequence to understand human diseases (about similar gene sequence differ in 5 gens only)

Why is it the primary structure important?

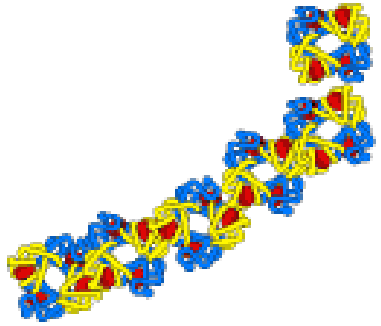
Sickle hemoglobin (HbS) as an example

- A single amino acid substitution can give rise to a malfunctioning protein, as is the case with sickle-cell anemia.
- It is caused by a change of amino acids in the 6th position of β globin (Glu to Val).

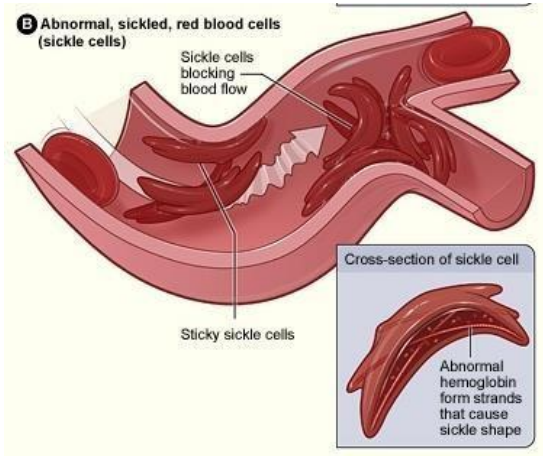
The mutation results in: 1) arrays of aggregates of hemoglobin molecules, 2) deformation of the red blood cell, and 3) clotting in blood vessels and tissues.

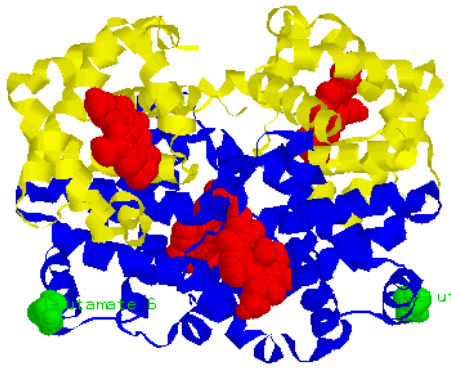


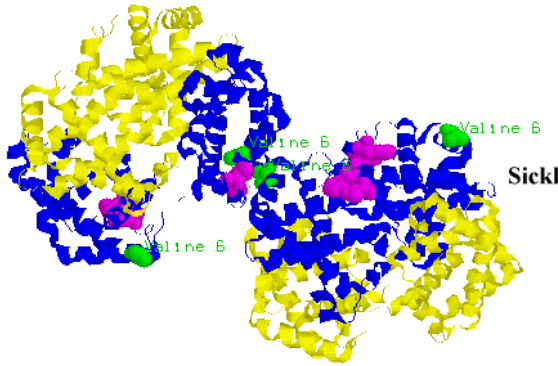
NORMAL HEMOGLOBIN



CLUMPED HEMOGLOBIN







Sickle cell anemia is a disease that occurs in the hemoglobin which is the protein that carries oxygen in our bodies.

What happens in sickle cell anemia?

A change occurs in only one amino acid (due to a change in one letter in the DNA) in the hemoglobin protein (the protein that carries O2 in our body) which leads to a change in the hemoglobin structure creating clusters (aggregates) causing a change in the shape (morphology) of a normal RBC to a sickle shape. As a result, instead of having concave rounded smooth RBC's , they will accumulate on top of each other taking the shape of sickle and clotting blood vessels. We take from this, that one amino acid can affect the secondary, tertiary and quaternary structure of proteins.

In the sickle cell disease, **glutamic acid** (negatively charged amino acid) is replaced with **valine** (non-polar branched aliphatic amino acid), so we replaced a charged amino acid to a non-polar one. However, if we replaced glutamic acid with **aspartate** (also negatively charged) or **lysine** it is highly probable that it **won't** greatly affect the hemoglobin structure.

And these are the individual differences between us in certain proteins. For example, in a certain enzyme or protein I have valine and you have leucine or isoleucine, the function of the enzyme changes slightly but still you are normal and I am normal (individual variations we will talk about in molecular biology).

Secondary Structure

What is a secondary structure? How is it formed?

- Two bonds within each amino acid residue can freely rotate.
  - the bond between the α-carbon and the amino nitrogen
  - the bond between the α-carbon and the carboxyl carbon

But the peptide bond cannot.

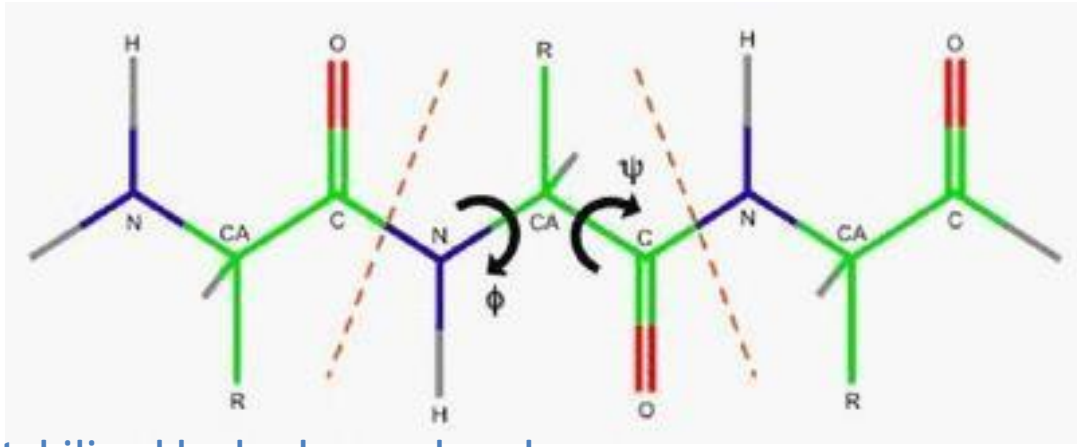
- A secondary structure is hydrogen-bonded, locally arranged structures of the backbone of a polypeptide chain. organized (have a certain pattern)
- Examples:
  - Alpha helix
  - Beta-pleated sheet

  } Regular secondary structures

  - Turns
  - Loops
  - Bends
  - Coils

  } Nonregular secondary structures

Hydrogen bonds between near amino acid in space



Overall, the secondary structures are stabilized by hydrogen bonds between the components of the peptide bond (carbonyl and amine, hydrogen bond accepter and hydrogen bond donor).

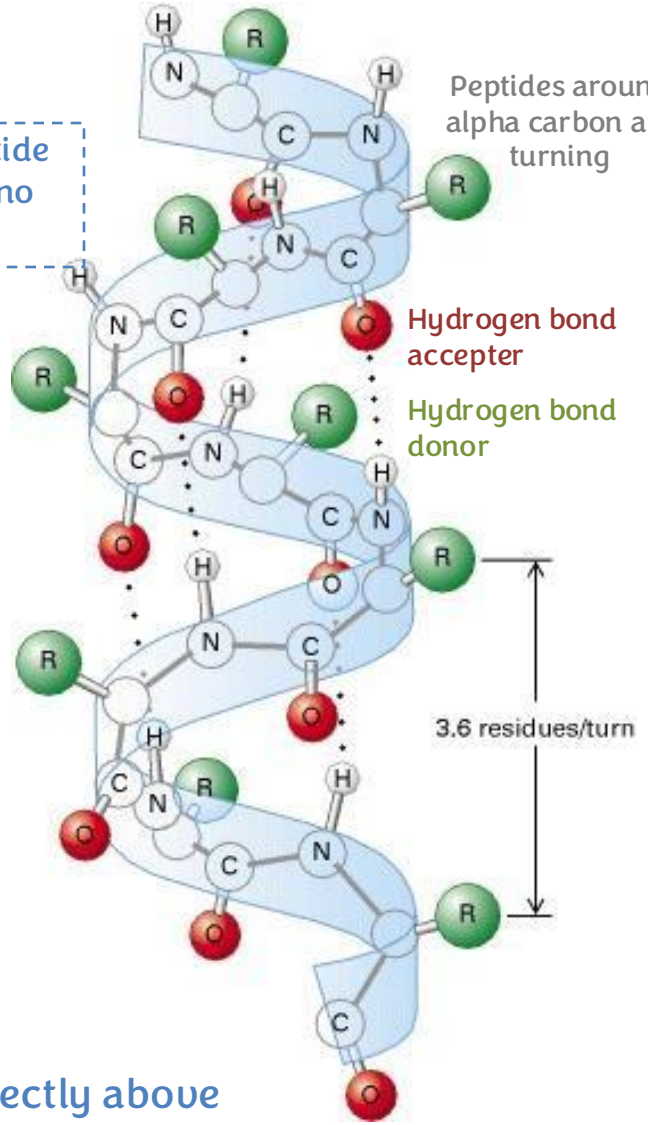
The α helix

Alpha helix is a pattern seen in different proteins

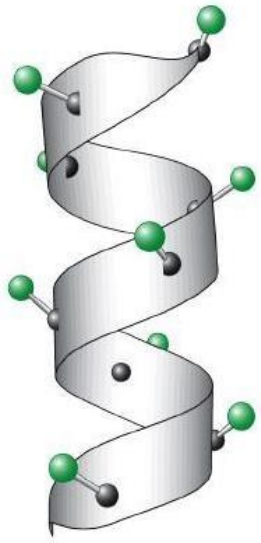
The hydrogen bond doesn't occur between two amino acids that are next to each other, rather between two amino acids on top of each other.

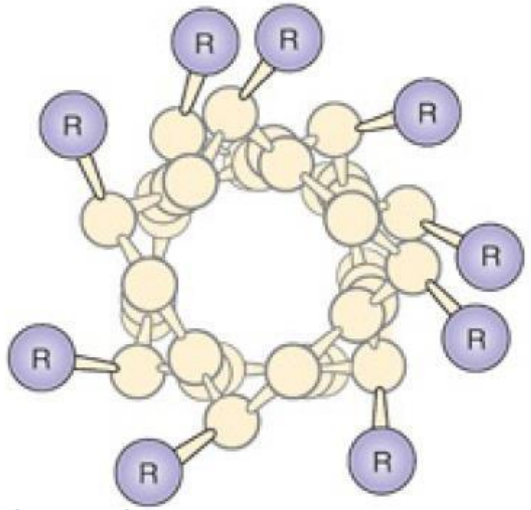
- It looks like a helical rod.
- The helix has an average of 3.6 amino acids per turn.
- It is very stable because of the linear hydrogen bonds.
- The side chains of the amino acids project outward from the helix in order to avoid molecular congestion (known as steric hindrance) with the polypeptide backbone and with each other.

Between the peptide bonds of the amino acids



Peptides have a backbone and side chains.
Backbone: nitrogen, alpha carbon, carbon of the carbonyl.
Side chains: R groups (that extend from the alpha carbon).

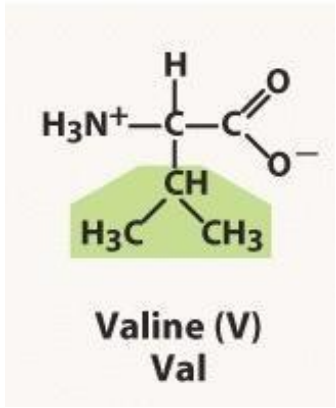


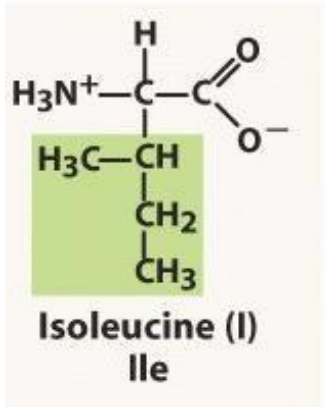


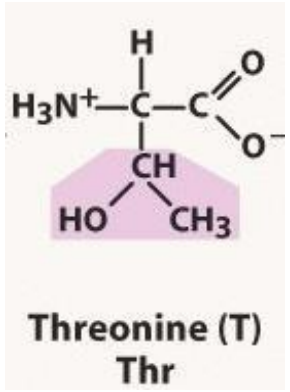
Notice how the R groups are not directly above each other to avoid repulsion.

Amino acids NOT found in α-helix

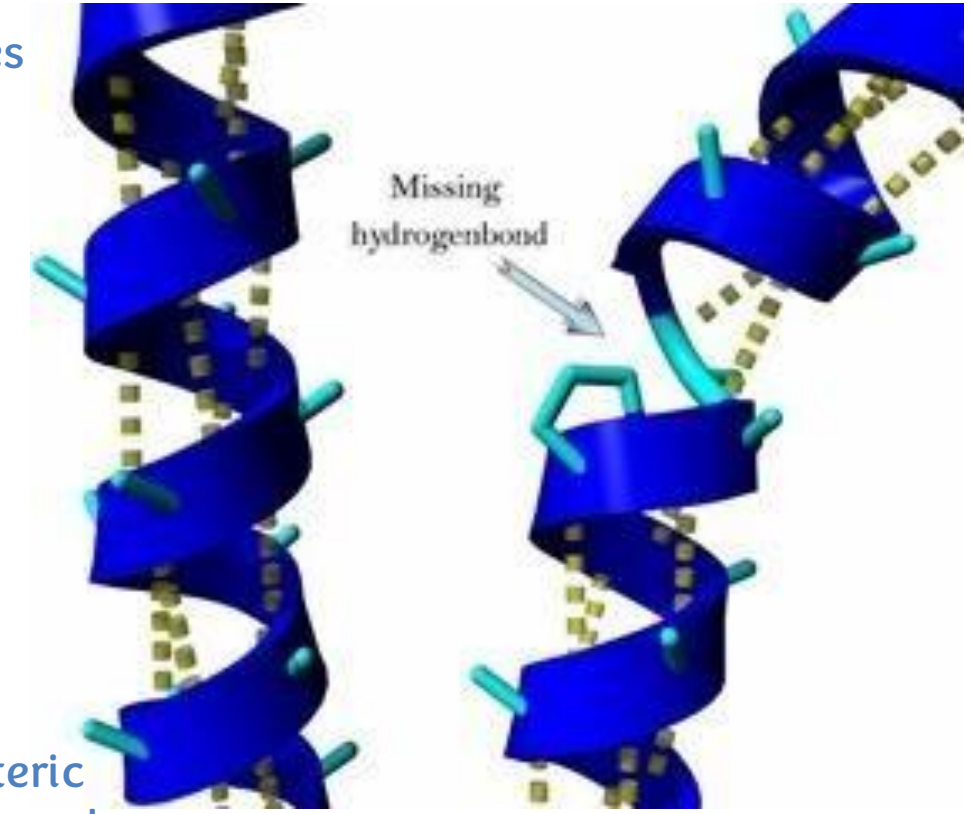
- Glycine: too small Just an H, so it destabilizes alpha helixes
- Proline
  - No rotation around N-C bond
  - No hydrogen bonding of α-amino group
- Close proximity of a pair of charged amino acids with similar charges Although R groups are not near each others , they are big in size
  To avoid repulsion
- Amino acids with branches at the β-carbon atom (valine, threonine, and isoleucine)







To avoid steric hindrance and repulsion



R-groups also give some stabilization for the alpha helix (through van der waal interactions) which is not found in glycine.

Proline is usually found at the end of the alpha helix because it breaks the smooth pattern of it, for two reasons:

1. It is **rigid**, the bond between the alpha carbon and nitrogen cannot rotate due to it's cyclic structure.
2. Nitrogen when it makes a peptide bond it is tertiary N, doesn't have hydrogen so **it is not a hydrogen bond donor**. (We know alpha helix is stabilized by hydrogen bonds between the components of the peptide bond).

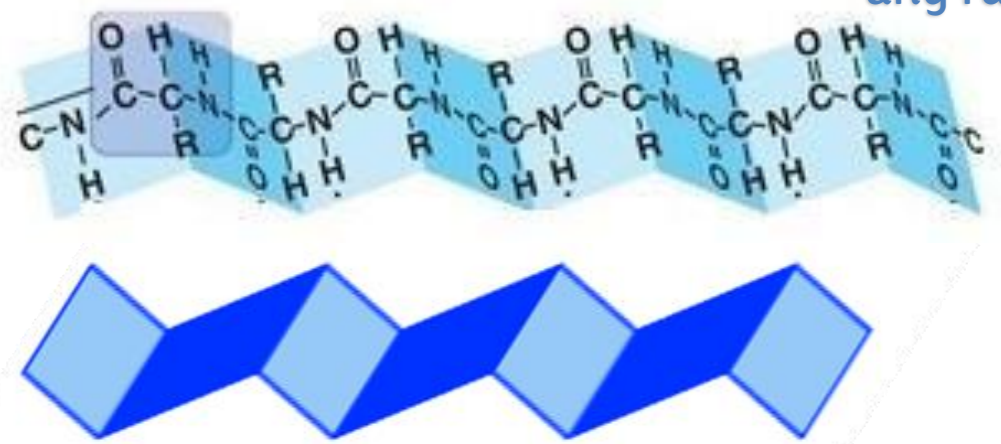
Pleated = folded

β-pleated sheet (β sheet)

No proline in beta strands

Beta strand
Alone it does not form any function

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

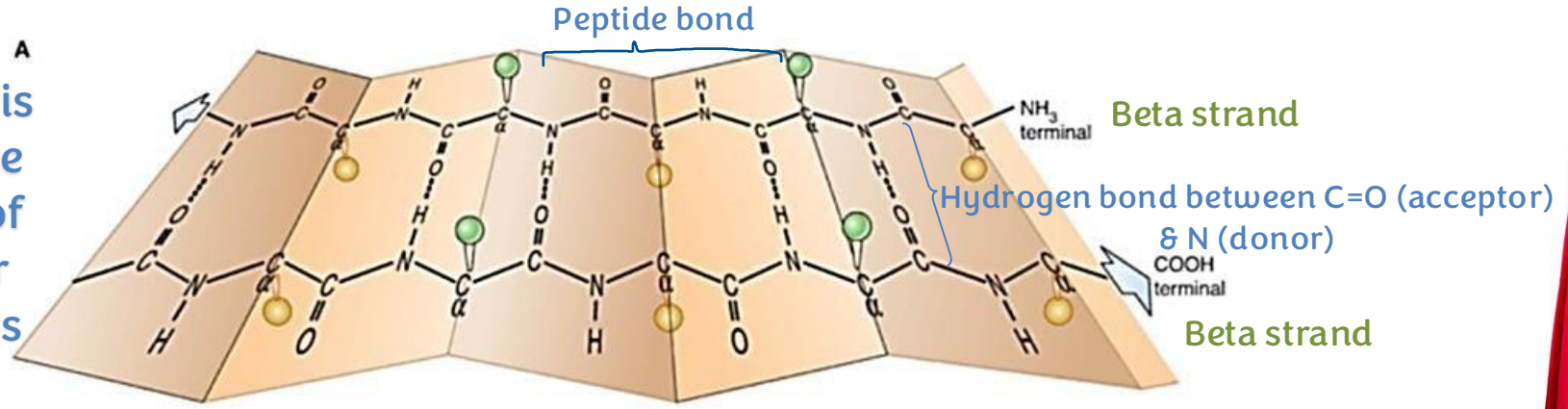


β strand

Stabilization & formation of beta pleated sheet which consists of more than one beta strand

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

A beta-pleated sheet is composed of multiple beta strands on top of or beside each other with hydrogen bonds between the components of the peptide bond.



A pleated skirt

More on β-sheets

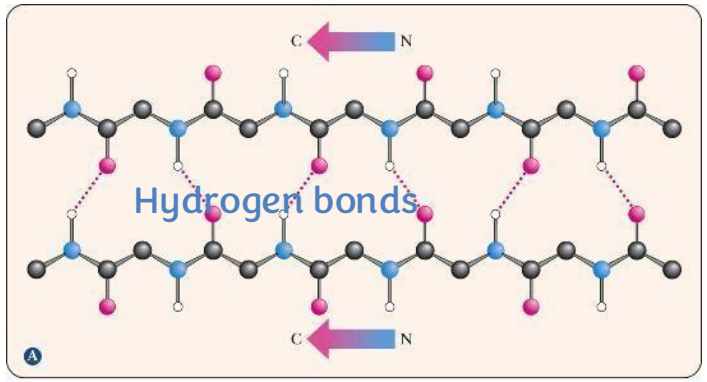
Might be parallel , anti parallel or mixed

- β sheets can form 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.
- Proline tends to disrupt β strands. Why?

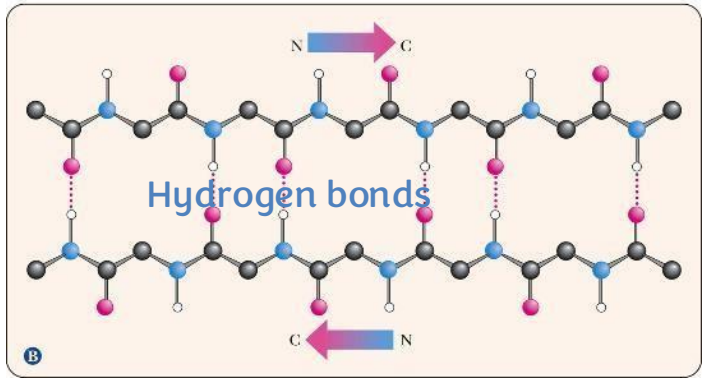
Because it is rigid (not flexible)

Strands goes to the same direction (from N-terminus to C-terminus

Parallel

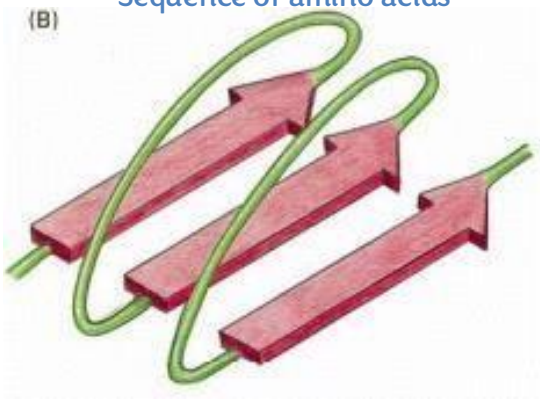


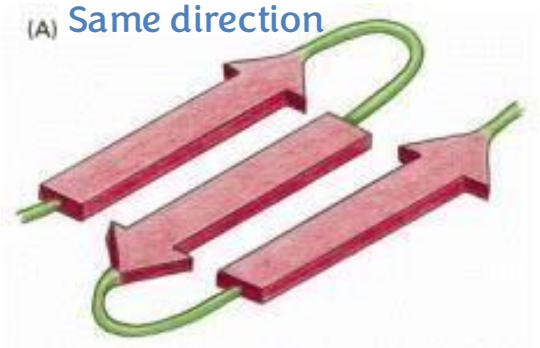
Antiparallel



- Carbon
- Nitrogen
- Oxygen
- Hydrogen

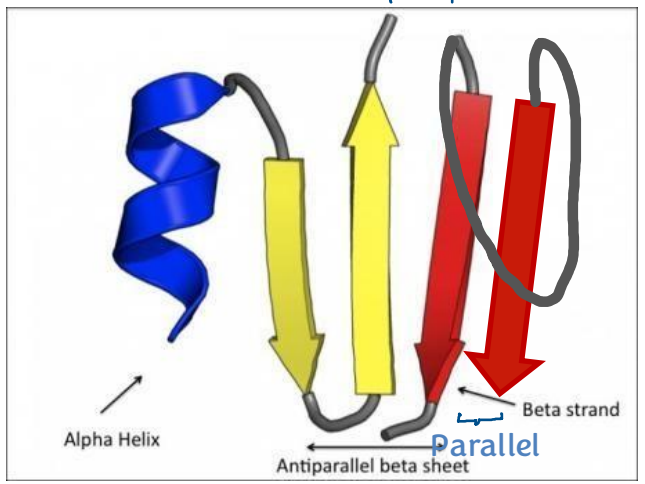
Sequence of amino acids





Opposite direction

Anti-parallel

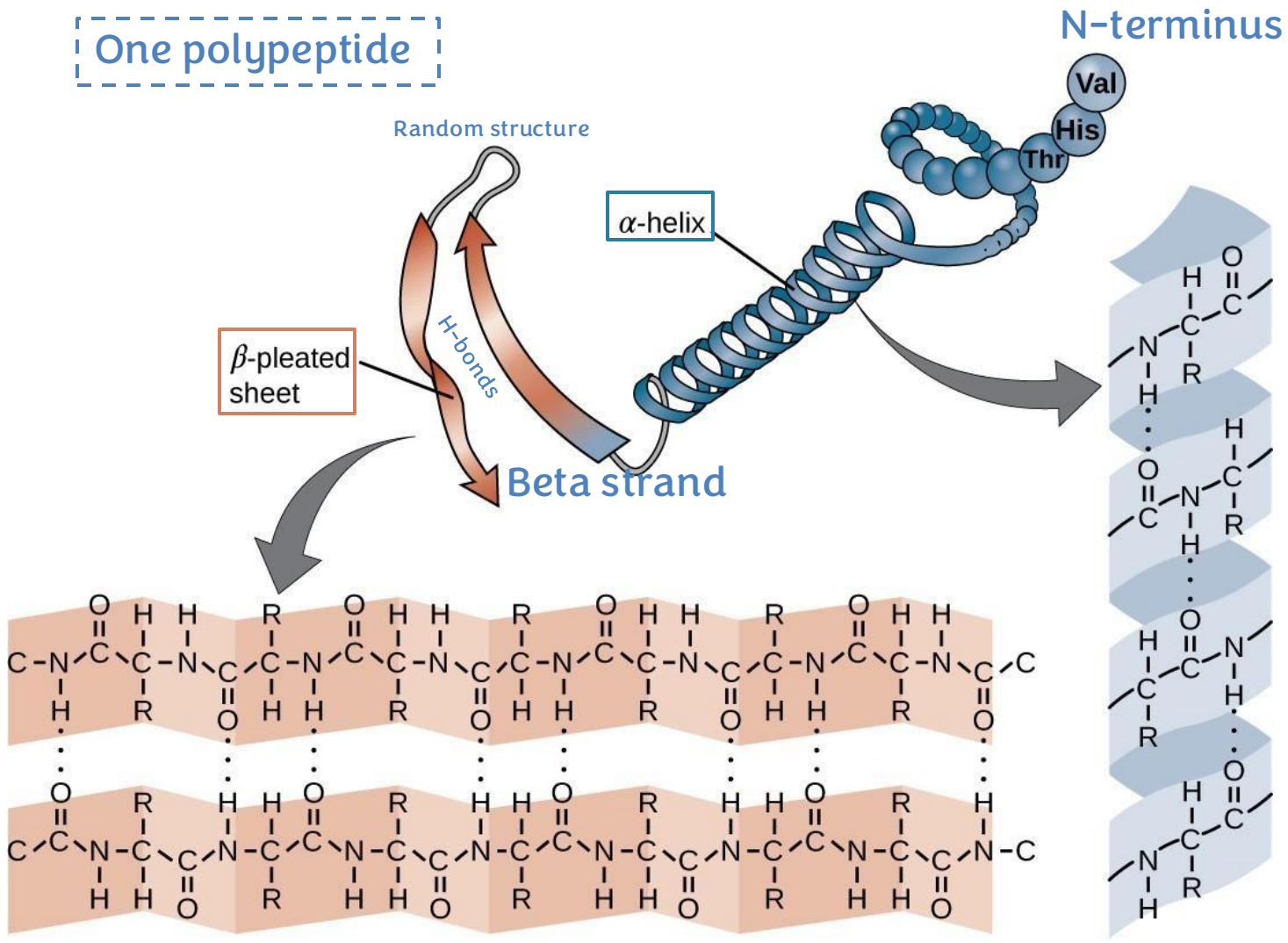


Non regular pattern (mixed parallel with anti parallel within the same sheet

Based on hydrogen bonding pattern, which do you think is more stable: parallel or anti-parallel sheets?

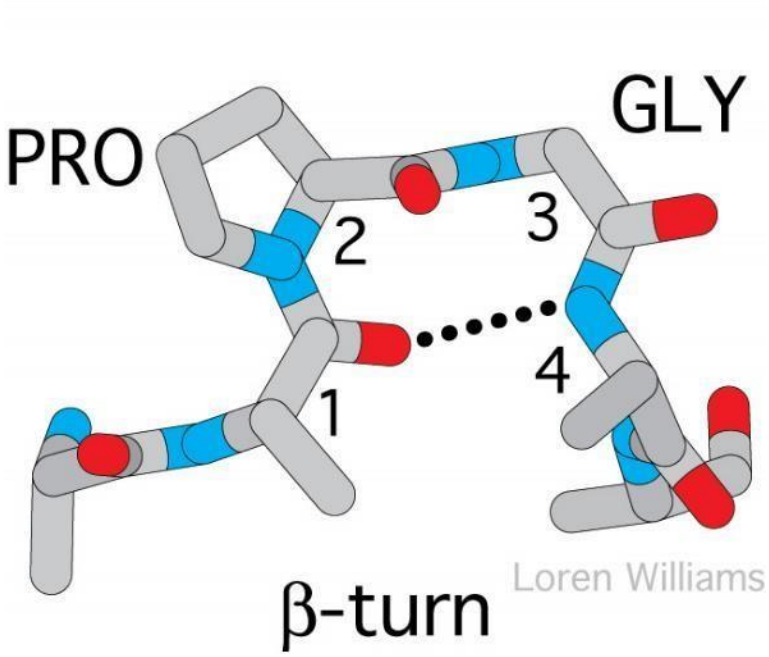
Anti-parallel sheets because the hydrogen bonds are perpendicular

How are they illustrated/drawn?



β-turns

- Turns are compact, U-shaped secondary structures.
- They are also known as β turns or hairpin bend.
- What are they used for? How are they stabilized?
- Glycine and proline are commonly present in turns. Why?

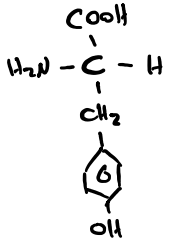
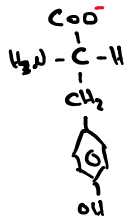


Beta-turns are composed of 4 amino acid residues linked by peptide bond.

In beta-turns usually proline amino acid number 2 because it is **rigid** so it breaks the smoothness, glycine amino acid number 3 because it is **small**, it doesn't cause repulsion and the peptide can continue its turn.

For any feedback, scan the code or click on it.

Corrections from previous versions:

Versions	Slide # and Place of Error	Before Correction	After Correction
V1 → V2	Slide 2 ,3	 Carboxyl group and R group are protonated but R group is not protonated	 R group is protonated but Carboxyl group and amino group are not protonated)
V2 → V3			
V3 → V4			

Additional Resources:

1. Campbell Textbook: sec 4.1+4.2+4.3
2. You Tube : **video explain protein structure**

رسالة من الفريق العلمي:

نسأل الله تعالى أن يرد كيدهم في نحورهم، اللهم أنزل عليهم بأسك الذي لا يرد عن القوم المجرمين،اللهم زلزل أقدامهم ونكس أعلامهم وأذهب ريحهم ولا تجعل لهم قوة ولا ترفع لهم هامة هم ومن يعاونهم، اللهم أنزل مقتك وغضبك عليهم وعلى من والاهم.