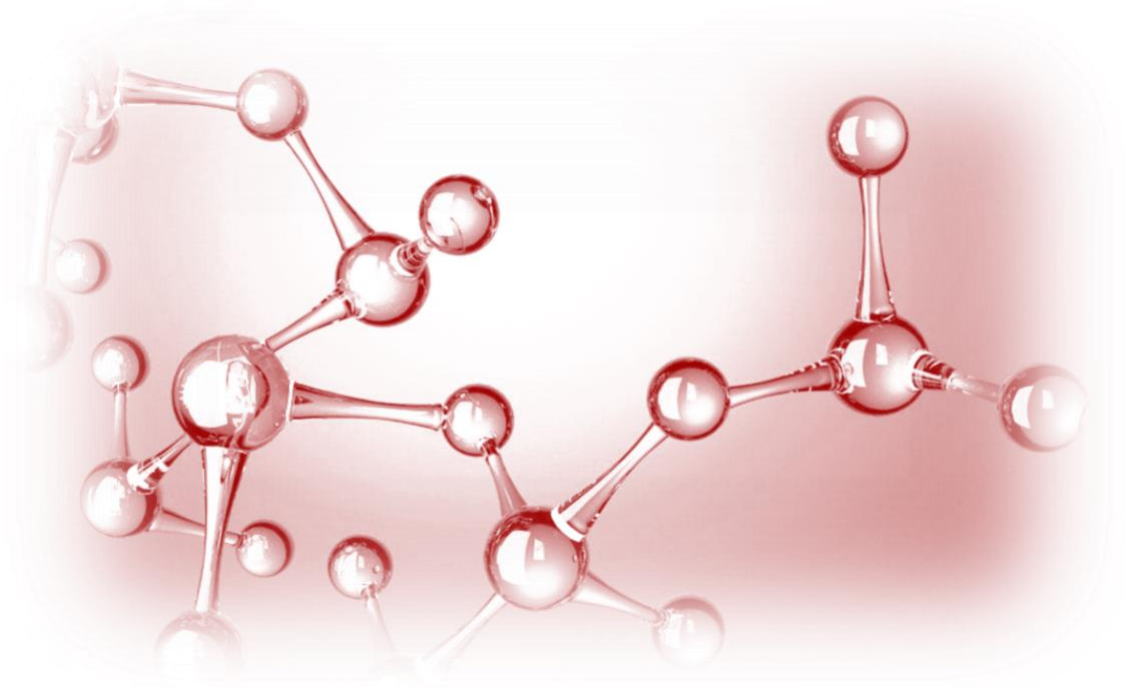
Med Learn

Dr.Ahmad Al-Qawasmi

Biochemistry

Sheet 9

Protein Structure

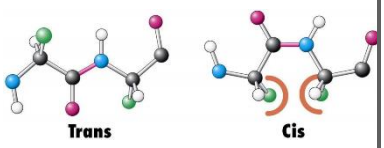


❖ Peptide bond

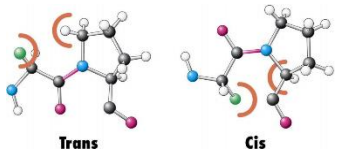
- **Peptide bond:** A covalent bond between 2 amino acids ➔ formed between the **carboxyl group** of an amino acid **with** the **Amino group** of the second one in the backbone of peptides (proteins)
  - ➢ This bond is an **amide bond** → but it is called peptide bond proteins
  - ➢ It is formed by a **condensation (dehydration)** reaction → losing H_2O
  - ➢ R group doesn't participate in the formation of peptide bond

- Peptides & polypeptides
  - ➢ Peptides → consist of **2 to several** residues of amino acids
  - ➢ Polypeptides → consist of **many** residues of amino acids (large number, usually more than 100)
  - ➢ Proteins → consist of **1 or more polypeptide** chains → so proteins are polymers of amino acids

- **Cis & Trans configurations of peptides:**
- According to the orientation of the side chains, peptides can be:
  - ➢ **Cis** → R groups of amino acids have the same orientation → that produces high steric energy (repulsion) leading to steric hindrance
  - ➢ **Trans** → R groups of amino acids have opposite orientation → that produces less steric energy (repulsion) preventing to steric hindrance



  - ➢ So **trans is preferred on cis** (because trans have less energy) → most amino acids are mainly in trans orientation **except proline**
    - In **proline** → Cis & trans conformations have **equivalent energies** so no one of them is preferred on the other
    - Proline present is cis configuration more than other amino acids



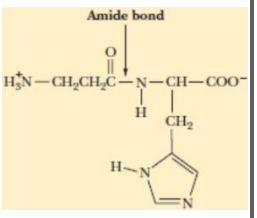
- **Resonance structure:**
  - Peptide bond shifts (alternate) between the **single & double bond forms**, because
    - ✓ The shifting of lone pairs of electrons between the N (of amine) and O (of carboxyl)
    - ✓ So, the double bond shifts → it can be formed between C with O or N
    - ✓ This electron shifting causes the **appearance of positive charge on N & negative on O**
  - So peptide bond is **stronger, rigid & planar** → due to the resonance stabilization (partial double bond) ➔ and this rigidity **prevents rotation around the peptide bond**

- **Hydrogen bonding:**
  - Amino acids can form H-bonds with each others (or with water) by the contribution of H atom of the Amine group of one amino acid (H-bond donor) with O from the other one (acceptor)
  - All amino acids can be H-bond donor or acceptors except **proline which can be acceptor but not donor** → because it **amine group is secondary** (lost the H that can be donated)

- **Notes:**
  - ➢ Peptide (polypeptide) chain is read from **N-terminus** (amino end) **to C-terminus** (carboxyl end)
  - ➢ The order of amino acids in a peptide chain affects its properties
  - ➢ While reading a peptide chain → there will be a repeated sequence of functional groups
    - ✓ **Amide → α-C with R group → carbonyl**

❖ Small peptides with physiological functions

1) Carnosine (β-Alanyl-L-Histidine)



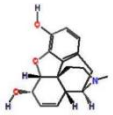
- It is a naturally occuring dipeptide of β-Alanine (unusual amino acid) with L-histidine
  - β-Alanine ➔ usually amino acids are α (because the R group is attached to α-carbon) but here the R group is attached to β-carbon ➔ this β-Alanine is synthesized naturally in the body
- It is highly concentrated in the **muscles & brain tissues**
- **Functions:**
  - **Protection of cell from ROS** (radical oxygen species) & peroxide
  - **Muscle contraction**

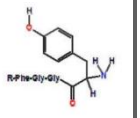
2) Glutathione (γ-glutamyl-L-cysteinylglycine)

- It is a **tripeptide** that consists of → **γ-glutamate, cysteine & Glycine**
  - γ-glutamate ➔ unusual amino acid → with the side chain on γ-carbon
- It functions as **anti-oxidant**
  - Scavenger oxidizing agents by reacting with them → is will be oxidized (lose electrons) from its **cysteine residue** (on thiol group) ➔ then it will react with another glutathione molecule forming **disulfide bridges** → until being recycled (enzymatically) to be reused again

3) Enkephalins

- They are **pentapeptides** in the **brain**→ they are 2 types differ in the amino acid in the C-terminus only
  - Leucine enkephalin → ends with leucine
  - Methionine enkephalin → ends with methionine
  - The amino acids are ➔ **Tyr-Gly-Gly-Phe-(Leu or Met)**
- They are naturally occurring **analgesics (pain relievers)**
  - The relieve pain by binding to specific receptors cause a pain killing response
  - The aromatic side chains of **Tyr & Phe** play important role in their activity
- **Opiates (such as morphine)** have a **similar structure** of enkephalins → so they can bind to the same receptors of pain killing (relieving)
  - The common amino acid (structure) between them is **tyrosine**
  - Morphine is used in hospitals for some cases (such as cancer & surgical operations)


Morphine


Enkephalins

4) Oxytocin & Vasopressin

- **Their structure:**
  - Similarities:
    - ✓ Both of them consist of **9 Amino acids**, with a **cyclic** structure due to the **disulfide bridge** between cysteine residues (Amino acids number **1 & 6**)
    - ✓ Both of them have **amide group** at the C-terminus instead of carboxyl group
  - Differences:
    - ✓ They have the same amino acids → except for residues **3 & 8**
    - ✓ Oxytocin → **Ile & Leu**
    - ✓ Vasopressin → **Phe & Arg**

- Both are secreted from the **hypothalamus** and stored in the posterior **pituitary gland**

- **Functions:**
  - **<u>Oxytocin</u>**
    - ✓ **Induces labor in pregnant women** by controlling contraction of uterine muscle and **stimulates the flow of milk** in a nursing mother
    - ✓ During pregnancy, the number of receptors for oxytocin in the uterine wall increases
    - ✓ As the cervix stretches, sending nerve impulses to the hypothalamus as a positive feedback to release more oxytocin by the posterior pituitary gland
  - **<u>Vasopressin = ADH hormone</u>**
    - ✓ Vasopressin **controls of blood pressure** by regulating smooth muscle contraction
    - ✓ Vasopressin is released by the action of the hypothalamus on the posterior pituitary
    - ✓ Vasopressin **stimulates water reabsorption by the kidney** (an antidiuretic effect) resulting in **water retention and blood pressure increase**
    - ✓ It increases at night

5) Gramicidin S and Tyrocidine A

- **Cyclic structure** formed by peptide bonds
- Two cyclic **decapeptides** produced by the **bacterium Bacillus brevis** → we can't produce them
- Both are **antibiotics**
- Both contain **D-amino acids and L-amino acids**
- Both contain the **amino acid ornithine (Orn)**, a metabolic intermediate
  - **Ornithine:** Naturally occuring amino acid, involved in urea cycle → not used in protein synthesis

6) Aspartame (L-Aspartyl-L-phenylalanine)

- It is a **dipeptide** → with a methanol group on the C-terminus (**methyl ester**)
- It is used as an **artificial sweetener** → 200 times sweet than sugars
- If one of the 2 amino acids or both are replaced by D-isomer → It will be **bitter** rather than sweet
- Used in diet soft drinks → with a controversial safety
- It shouldn't be given for people with PKA

- **PKA (Phenylketonuria)** → A hereditary disorder with metabolic defect (lack Phe hydroxylase)
  - Cause the accumulation of Phenylpyruvate → causing **mental retardation**
  - Must limit the sources of Phe such as aspartame
  - They can use Alatame instead of aspartame → Phe is replaced by Ala

---

❖ Protein Structure

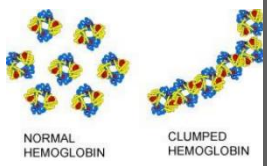
- Proteins are macromolecules composed of a large number of amino acids connected by peptide bond
- Proteins have different structures, with a huge variety (probabilities) of amino acids order
  - So, a protein may has gazillion possibilities of structures but only a few would be active
  - Active structures are called **native conformations** → properly folded (3D structure) & functional

- The difference between a Protein & a polypeptide:
  - **Polypeptide:** A sequence of amino acids linked by peptide bonds
  - **Protein:** Consists of <u>1 or more polypeptide</u> chains → having a certain **3D structure & Function**

- **Levels of proteins structure:**
  - **Primary, Secondary, Tertiary & Quaternary**

1) Primary structure:

- It is the first level of protein structure → represents the **sequence of amino acids from N to C terminus**
  - Such as this sequence Leu – Gly – Thr – Val…
- The primary structure determines the other levels of proteins structure → so any change in the amino acids sequence will **affect the final conformation** of the protein → producing a **malfunctioning protein**
- Such as **Sickle Cell hemoglobin (HbS) & Cystic fibrosis**:
- **Sickle Cell Anemia**
  - A Hereditary disease caused by a mutation changes the amino acid in the **6th position of β-globin from Glu to Val**, results in:
    - **Clumped** Hemoglobin forming arrays of **aggregates**
    - Causing the **deformation of RBCs** → having a **sickle** shape
      - ✓ Normal RBC is a biconcave disc with high flexibility allowing it to move inside small vessels
      - ✓ Sickle shape makes RBCs much less flexible decreasing the efficiency of transporting O_2



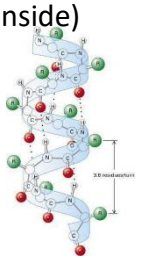
- **Cystic fibrosis**
  - A Hereditary disease caused by a mutation is **CFTR gene** (linked to fluoride ion transport)
    - It causes the exocrine glands excretion to be thicker & harder (mucus) → forming a suitable environment for the growth of bacteria in the respiratory & digestive system

2) Secondary structure:

- **Hydrogen-bonded localized organization of parts** of a polypeptide chain → forming the **basic shape of these part**
  - These shapes can be either **α-helix, β-sheets, Turns or loops**
  - They are formed due to the variation in the **orientation** of the backbone
    - ✓ Peptide bond **can't rotate**
    - ✓ Bonds between α-carbon with Amine group (Phi Φ) and between α-carbon with Carboxyl group (Psi Ψ) can **freely rotate**

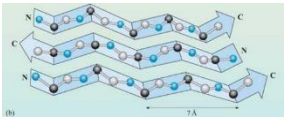
α-helix

- It looks like a helical rod (spring) → with **3.6 amino acid residues per turn** (in average)
- The **pitch of the helix = 5.4 Å** (it is the linear distance between the corresponding points on a turn)
- R groups are **pointed outward** to reduce the steric hindrance → So it is hollow (empty from inside)
- It is very stable → due to the **linear H-bonds**
- Some amino acids can't contribute (found) in α-helix:
  - **Glycine** → because it is very small
  - **Proline** → because:
    - ✓ **No Rotation** on the bond between N & α-C → due to rigidity
    - ✓ **No H-bonding** on the α-amino group → because it is secondary amine (no H to donate)
  - **Similarly charged** amino acids can't present in proximity → to prevent repulsion
  - Amino acids with branches on β-C such as **Valine, Isoleucine & threonine**

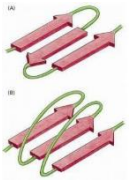


- **Note:**
  - **Integral** membrane proteins are consist mainly of **α-helix**
  - Some proteins have **amphipathic helices** **(such as channels)**
    - ✓ Hydrophobic R groups face the **hydrophobic core of the lipid bilayer (outward)**
    - ✓ Hydrophilic A.A (polar) facing inside allowing **hydrophilic molecules** (ions, water) to pass

β-sheets (β pleated sheets)



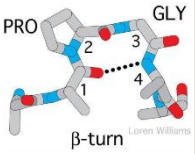
- They are composed of **two or more straight chains** (β-strands) that are **hydrogen bonded side by side**
- R groups have **trans orientation** → having a zigzag shape → giving more space than α-helix
- β-strands forming a β-sheet can be either:
  - **Parallel** to each other (C-terminus faces C-terminus of the other strand and N faces N)
  - **Antiparallel** to each other (C-terminus faces N-terminus of the other strand)



- A protein can contain parallel, anti-parallel or mixed β-sheets
- A.A. having a **branch on β-carbon** (Val, Ile, Thr) & **large aromatic amino acids** (Phe, Trp, Tyr) tend to present in β-sheets → because they have enough space to protrude upward and downward
- **Proline** tends to **disrupt β-sheets** → can't form H-bonds on amine group

β-Turns

- Turns are **compact, U-shaped** secondary structures, also known as β turn or hairpin bend
  - They are important for the 3D structure of proteins (especially globular proteins)
- **Glycine and proline** are commonly present in turns
  - Proline create a kink (sharp turn) because it is rigid
  - The second residue is Proline & the 3rd one is glycine



- Turns are usually **short** and link between Anti-parallel β-sheets
- Loops are usually **long** and link between Parallel β-sheets

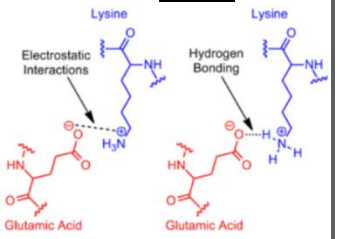
Super-Secondary structures:

- They are regions in proteins that contain an ordered organization of secondary structures
  - So, they are structures between the secondary and the tertiary levels
- **Example: Motifs**
  - Repetitive super-secondary structure, which can often be repeated and organized into larger motifs → motifs consist of a **small portion of a protein** (less than 20 A.A.)
  - They are **structural regions** → indicates the folding of the protein but **do not indicate the biological function of the protein**
  - Example of motifs:
    - Helix-loop-Helix is found in many proteins that bind DNA → two α-helices connected by a loop
    - Helix-turn-helix is capable of binding DNA → two α-helices joined by a short strand of amino acids
    - **Immunoglobulins fold** enables interaction with molecules of various structures & sizes
      - ✓ Immunoglobulins are antibodies (recognize foreign bodies)
      - ✓ They contain repeated structural units (motif)

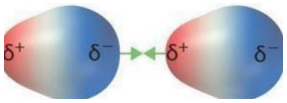
- **Domain:** A domain is a compactly folded region of polypeptide found in proteins with **similar function** and/or **structure** → Domains with similar conformations are associated with the particular function
  - A structural domain may consist of **100 – 200 residues** in various combinations of α helices, β sheets, turns, and random coils
  - **They fold independently of the rest of the protein**
  - Domains may also be defined in functional terms enzymatic activity, binding ability (e.g., a DNA binding domain)

3) Tertiary structure:

- The overall conformation of a polypeptide chain → The **3D arrangement** of all the amino acids residues
  - Tertiary structure represents the **spatial arrangement** of amino acids in 1 polypeptide chain
  - Also, in this stage → modifications occur (such glycosylation)
- This structure is determined by the **interactions between side chains (R-groups)**
- **Shape determining forces:**
  - **Non-Covalent Interactions** determine the 3D structure of the proteins
    - **Hydrogen bonds**
- Occur between Amino acids within or between polypeptide chains or it can be formed with water



    - **Charge-Charge interactions (Salt bridges)**
- Electrostatic interactions occur between **positively charged** R-groups
- **Charge-dipole** interactions form between **water** (**partially** charged) with **Charged R** groups of A.A.
  - The same groups can form either H-bonds or electrostatic interaction
    - **Van Der Waals forces**
- There are both attractive and repulsive van der Waals forces that control protein folding
- Although van der Waals forces are extremely weak (very transient) ➔ but they are **significant** (important) because there are **so many of them in large protein** molecules → having an accumulative effect

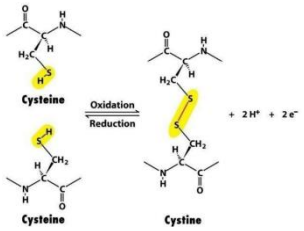


  - They are considered transient → because they depend on the position of electrons around the nucleus which is constantly changing
    - **Hydrophobic interaction**
- **Hydrophobic groups are clustered together rather than extended into the aqueous surroundings** → making the protein more thermodynamically (energetically) **stable** when
  - So non-polar groups are away from the aqueous solution & polar are exposed (faces) it
- the most important force that determines a protein structure is basically hydrophobic interactions

- **Note:**
  - **Polar amino acids can be in the interior** of the protein playing important role in its function
    - ✓ Also, they will form H-bonds with each other & with the back bone
  - Non-polar amino acids can also be on the surface of the protein

- **Stabilizing factors:**
  - They are forces that stabilize the structure of the protein but don't determine its 3D structure

1) Disulfide bonds



- They are formed between 2 Sulfur atoms (in thiol group) of 2 Cysteine residues
  - **Oxidation** (loss electrons) of thiol group → **forming disulfide bond** with another Cysteine
  - **Reduction** → **breaking** disulfide bonds
- The formation of disulfide bonds between 2 Cysteine residues → forming **Cystine** amino acid

2) Metal ions

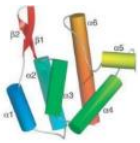
- They can stabilize protein structure by forming:
  - **Covalent interactions**
    - Such as the metal ion (iron) in the heme of **myoglobin** → it is covalently linked to **His**
  - **Salt bridges**
    - Such as Zinc in the **Carbonic anhydrase** → non-covalently bonded to **3 His** residues

- **There are many ways (models) in which we represent (illustrate) proteins:**

A) Ribbon structure

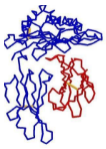
  - α-Helix is represented as a ribbon (helical rod)
  - β-Strands are represented as thick arrows (The direction of the arrow ➔ from N to C terminus)

B) Cylinder structure



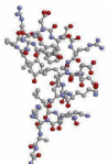
  - α-Helix is represented as A cylinder
  - β-Strands are represented as thick arrows

C) Trace structure



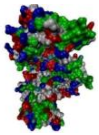
  - We only draw (represent) the backbone

D) Ball & stick structure



  - We draw atoms (as **small balls**), **backbone** & their orientation (angles) representing the exact 3D structure

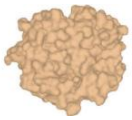
E) Space filling structure



  - Like Ball & stick but more complex → balls are larger & the **backbone isn't seen**

F) Protein surface structure

G) Draws the surface of the protein only (without the interior)



  - Used to design drugs and study their interactions with other proteins

4) Quaternary structure:

- It is the spatial **arrangement of subunits** & the nature of **their interactions**
  - Subunit = polypeptide chain
- It presents only in proteins that consist of **more than one polypeptide** chains (Oligomeric proteins)
  - Monomer → 1 polypeptides chain (1 subunit) ➔ No Quaternary structure
  - Dimer (2 subunits), Trimer (3 subunits) … ➔ have Quaternary structure
  - If subunits forming oligomer are (similar → Homo) / (Different → Hetero)

- These subunits can be connected together by **disulfide bonds or non-covalent interactions**
- **Examples:**
  - **Immunoglobulins**: Consist of 4 subunits of 2 light & 2 heavy chains (hetero-tetramer)
  - **Hemoglobin**: Consist of 4 subunits each 2 of them are identical (hetero-tetramer)
    - Both of them have Quaternary structure

❖ Protein Denaturation & Renaturation:

Denaturation

- It is disrupting the nature of the protein (native conformation), by:

Breaking **non-covalent interaction** → then reducing **disulfide bridges** → disrupting the **3D structure**

- Denatured protein **loses its properties** (activity, solubility …)
  - Denaturation **is mostly irreversible** (it can be reversible according to the structure of the protein and interactions involved)
- **Denaturing agents:**
  - **Heat** → disrupt van der waals (non-covalent) interactions
  - **Extreme pH** → will change the ionization (protonation) state of the A.A groups
  - **Detergents**, such as:
    - **Triton X-100** [nonionic, uncharged] → disrupts hydrophobic forces
    - **SDS** (Sodium dodecyl sulfate) [anionic, charged] disrupts hydrophobic & electrostatic interactions
  - **Urea & Guanidine hydrochloride** disrupt Hydrogen bonds & hydrophobic interactions
  - **Reducing agents,** such as β-mercaptoethanol (βME) & Dithiothreitol (DTT)
    - **Disrupt disulfide bonds**

- The stability of the protein structure is determined by (factor):
  - **The amino acid sequence** (mainly the internal residues)
  - **The peptide bond (rigid → can't rotate)**
  - **The proper angels between A.A,** by:
    - Weak non-covalent interactions between the side chains (mainly)
    - Non-protein molecules, such as heme & zink

Renaturation

- **Returning (re-acquiring)** the Native conformation of the proteins
  - It occurs quickly and spontaneously (when denaturation agent is removed)
  - Disulfide bods are reformed correctly
- Renaturation is not always possible
  - Denaturation of egg proteins (by frying or boiling) is not able to renature
- Renaturation can refold proteins incorrectly (not always correctly refolded)

• If a protein (mainly small proteins) is unfolded, it can refold to its correct structure placing the disulfide (S-S) bonds in the right orientation

• Misfolded & partial folded (not correctly folded) proteins → have their internal hydrophobic regions exposed and interact with other hydrophobic regions on other molecules, and form aggregates
  - These aggregates can be small (soluble dimers or trimers) OR insoluble fibrillar structures
  - Both of them are toxic to cells

• To refold misfolded proteins, we use **chaperons**
• Chaperons (barrel shape proteins) bind to polypeptide chains & help them refold correctly (to the most energetically favorable & stable folding pathway (structure))
• They also prevent hydrophobic regions from associating to other proteins → preventing the formation of aggregates
• So, they contribute in the **quality** of proteins
• Chaperons require energy

- So, if Chaperons are defective (non-functional) that will lead to increase misfolded proteins causing **many diseases,** such as:

1) Prion disease

- Pathological condition result from the **misfolding of prion protein (PrP^c)** in the brain producing PrP^{sc}
  - PrP^c has a lot of **α-helical** conformation, but PrP^{sc} has **more β strands** forming aggregates
- This disease can be acquired by:
  - **Infection**
    - Meaning that it can be caused by a transmissible agent by eating the meat of an affected cow (with misfolded proteins) which bind normal proteins causing them to misfold
  - **Inheritance**
  - **Spontaneous**
- Examples on this disease ➔ **Creutzfeldt-Jacob disease** (in humans), and **mad cow disease** (in cows), and **scrapie** (in sheep)
  - Creutzfeldt-Jacob disease → makes the brain tissue to be spongy

2) Alzheimer's disease

- It is not a transmissible disease but the presence of a family history of the disease increasing the chance to have the disease
- It is caused by the **accumulation of Tau protein or Amyloid Aβ peptides** → forming aggregates → that will damage the neurons and the supportive cells
  - Normally APP (Amyloid precursor protein) is a transmembrane protein is cleaved by the certain type of secretase enzyme (α-secretase) → producing peptides having polar and non-polar parts
  - If this APP is mutated → it will be cut by **other secretases** (β or γ) → causing the non-polar to be cleaved alone without a polar part with it → causing the accumulation of it → **Alzheimer**

❖ Complex Protein Structures

- **Holo- & Apo- proteins**
- **Apoproteins:** Proteins **composed only of amino acids** → without non-protein molecules
- **Holoproteins:** Proteins **conjugated to (linked) to non-proteins molecules** (such as metals & heme)
  - Such as **Glycoproteins** (such as immunoglobulins) → proteins are covalently linked (conjugated) to carbohydrates (sugars)
    - **N-linked** → The sugar is linked to amide nitrogen of the R-group of Asparagine
    - **O-linked** → The sugar is linked to the hydroxyl group of serine, Threonine and occasionally hydroxylysine
  - They can also be linked to lipids forming → **lipoproteins**
  - Proteins can be phosphorylated proteins → **phosphoproteins**
    - Phosphorylation can activate or inhibit metabolic pathways
  - **Hemoproteins** → linked to heme group