1

Bioinformatics II:
Protein Structure & Function

Dr Manaf A Guma

University Of Anbar- college of applied sciences-Heet.

Department of chemistry

1

What is a protein?

• Proteins is a biological molecules that our body is built of.

• It consists of sequence of amino acids connected as a polymer '' multi mere'' by peptide bonds

• It is folded in a specific 3D shape to do a specific function in the living cells.

• Proteins have different structures and functions based on the places where they exist in.

Primary Structure of Protein

2

Protein: amino acid →peptide → polypeptide →protein
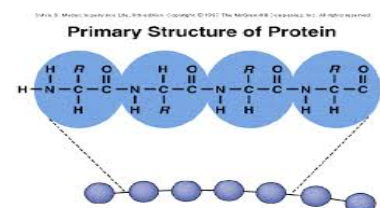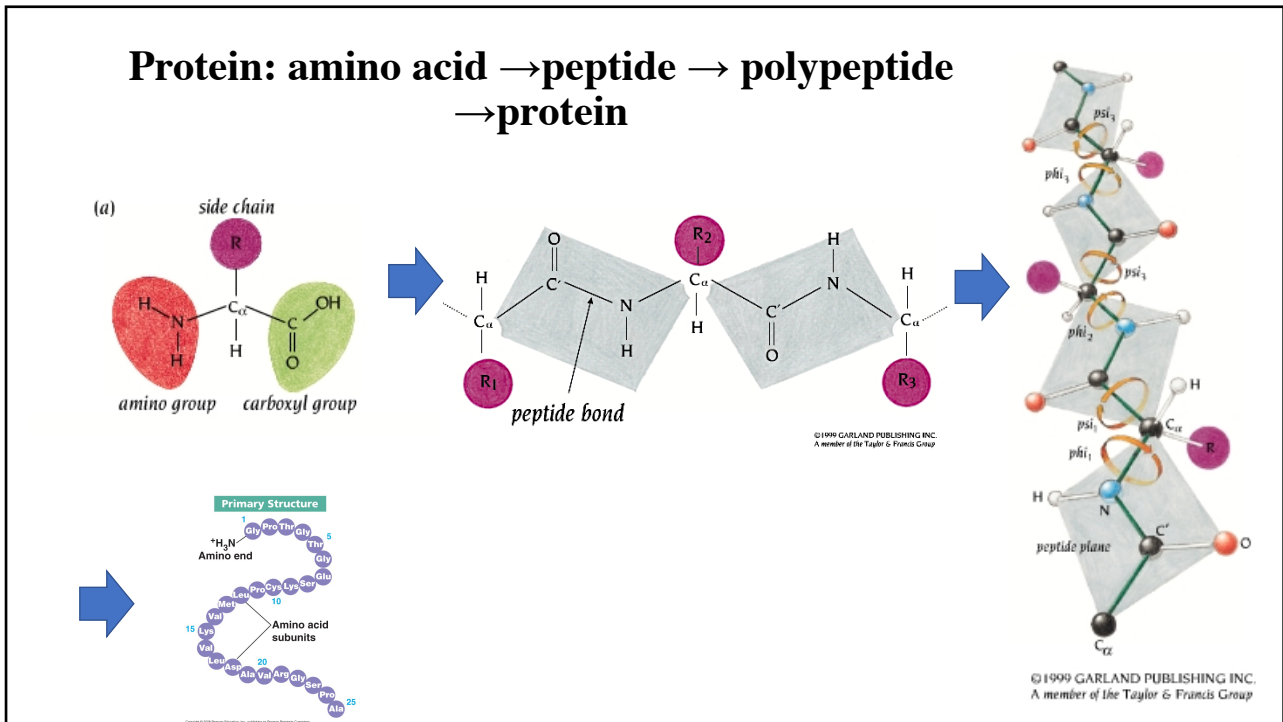
3

CLASSIFICATION OF PROTEINS

Classification Based on Structure
- *Fibrous Proteins*
- *Globular Proteins*
- *Intermediate Proteins*

Classification Based on Composition
- *Simple Proteins*
- *Conjugated Proteins*

Classification Based on Functions
- *Structural Proteins, Enzymes, Hormones*
- *Pigments, Transport Proteins, Contractile Proteins*
- *Storage Proteins, Toxins*

www.easybiologyclass.com

4

**How can we study both: Protein Structure & Function of the protein?**

• *Protein structure* - primarily determined by sequence.

• *Protein function* - primarily determined by structure.

• What can that (protein structure and function) be useful in bioinformatics?

• Most of amino acid sequences for specifics regions of a protein are similar to any other protein.

• So, by identifying protein sequences we can predict the structure and the function of a protein.

5

What are the different structures of protein?

1. Primary structures (simple form)

2. Secondary structures (alpha helix and beta sheets) (3D)

3. Tartary structures

4. Quaternary structures.

They differ by the types of bonds that connect each others.

LEVELS OF PROTEIN STRUCTURE
Primary Structure
Secondary Structure
β-Sheet
α-Helix
Tertiary Structure
Quaternary Structure
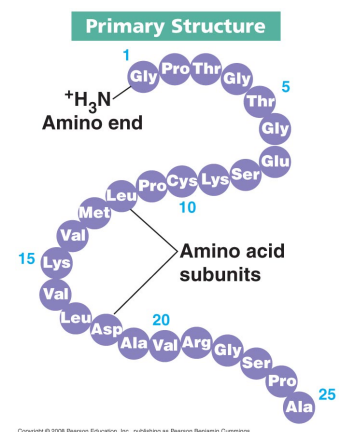Lubrizol Life Science
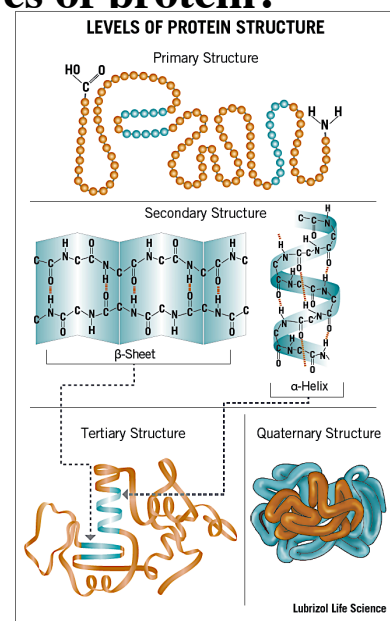
6

What are the basic Levels of Protein Structure?
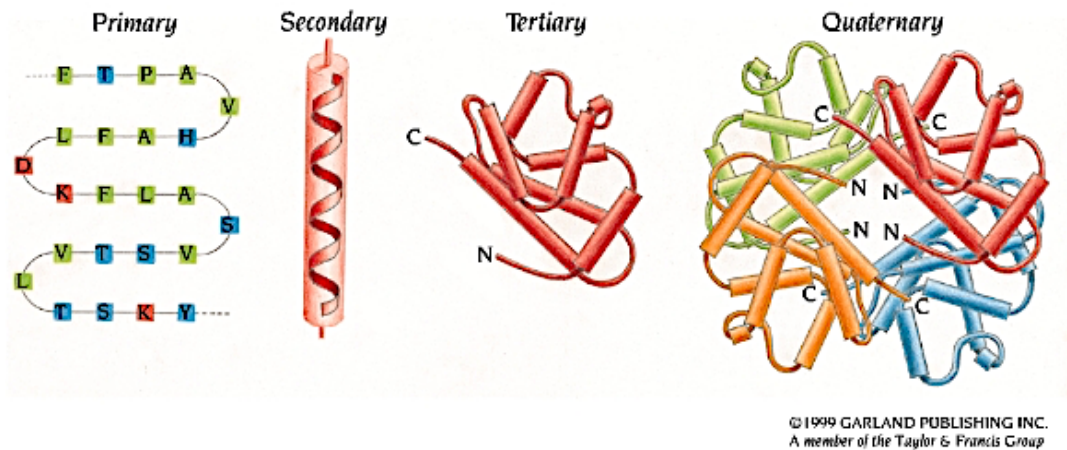


7

What is the Primary structure of protein?

- The Primary structures of protein are:

- Linear sequence of amino acids

- This linear sequence is referred to as a polypeptide chain. The amino acids in the **primary structure** are held together by covalent bonds.
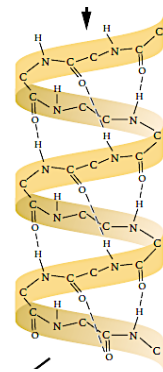

8

What is the secondary structure of protein?

- The Secondary structures of protein are:
- Mostly consists of 2 types which are the α helix and the β pleated sheet.
- Other 2nd str:
- Both **structures** are held in shape by hydrogen bonds, which form between the carbonyl O of one amino acid and the amino H of another
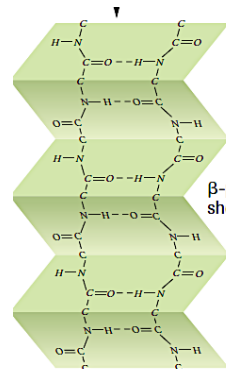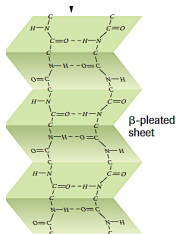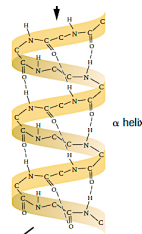
β-pleated sheet

α helix

9

## What are basic differences between a-helix and b-sheet?

β-pleated sheet

α helix

Alpha Helix	Beta Sheet
1. It is rod like structure, coiled peptide chain arranged in spiral structure.	1. It is sheet like structure, composed of 2 or more peptide chains.
2. All bindings are intra-chain (on the inside).	2. Inter-chain between separate polypeptide chains. The intra-chain in a single poly peptide chain folding back on itself.
3. The spiral of a-helix prevents the chain from being fully extended.	3. The chains are mostly extended and flat, Parallel and anti-parallel.
4. All peptide bonds participate in H bonds between the C=O and N-H.	4. The H bonds between N backbone and carbonyl group of **adjacent** chain.
5. 3.6 per turn.	5. number of a.a are Not fixed.

10

5

How many types can b-sheet form?

Parellel form
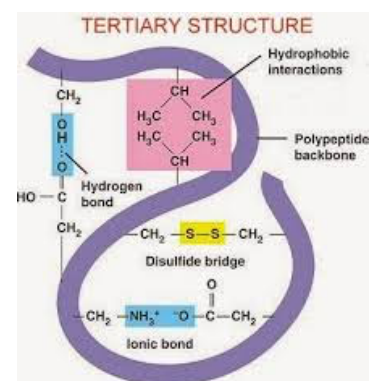
Anti-parellel
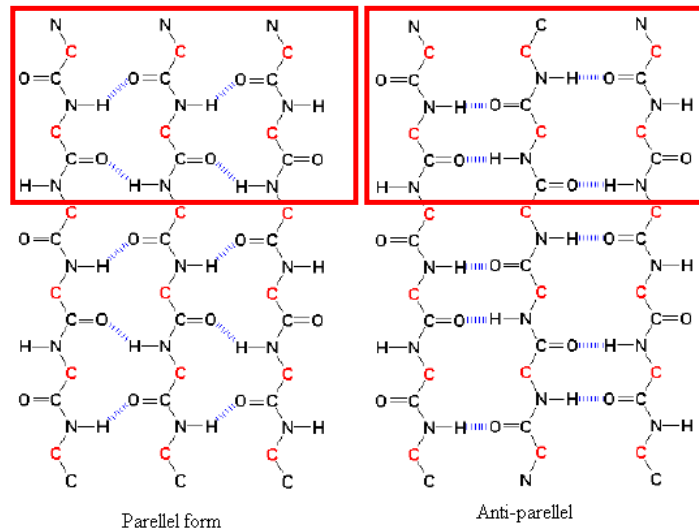
11

# What are the tertiary & quaternary structures of protein?

• Tertiary: It is 3D structures shape of protein. It has a single polypeptide chain "backbone" with one or more protein secondary structures that form the protein domain.

12

What is the tertiary structure of protein?

- Quaternary:

- It is an arrangement of multiple folded protein subunits in a multi-subunit complex.

- It involves at least 2 polypeptides (domains).

- It can be a dimer, tetramer, homo or hetero protein.
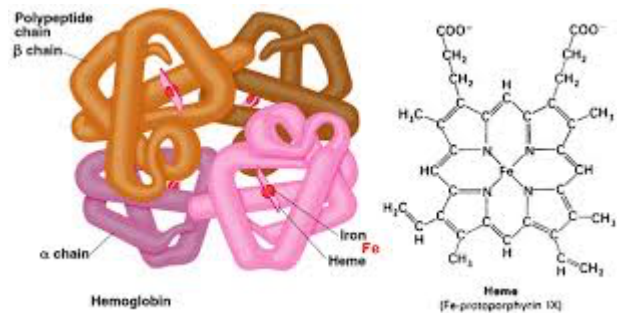
13

what are the "Additional" Structural Levels of proteins?

1. **Coils:** type of 2nd structure that are not helices, sheets, or recognizable turns.

2. **Loops**: type of 2nd structure.

3. **Motifs**: combinations of 2nd structural elements

4. **Domains**: combinations of motifs.

- e.g: Globular proteins are built from recurring structural patterns

- Please look for their structure?

14

What are the types of protein structures based on their folding?

- Folding protein is a process by which a polypeptide chain folds to become a biologically active protein in its native 3D structure.

- Types?

1.   Folded proteins: they are typically stable.

2.   Unfolded proteins: very little.

3.   Partially folded.

4.   *Intrinsically disordered proteins: unstable protein because they are* dynamic.

- Predicting protein structure and function can be very hard *& fun!*

15

Amino acid classification ? Do you remember?

- The amino acid are classified depends on the sidechains which give the protein a specific structure and function.

1. Hydrophobic amino acids.

2. Non-hydrophobic amino acids.

3. Polar amino acids.

4. For example:

- Glycine increases backbone flexibility because they have no R group

- Proline reduces flexibility of polypeptide chain
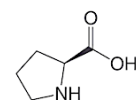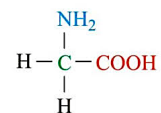
16

Why Globular proteins have a compact hydrophobic core?

• Packing of hydrophobic side chains into interior is main driving force for folding.

• *Problem?* Polypeptide backbone is highly polar (hydrophilic) due to polar -NH and C=O in each peptide unit; these polar groups must be neutralized

• *Solution?* Form regular secondary structures,

• e.g., α-helix, β-sheet, stabilized by H-bonds

17

Which Certain amino acids are "preferred" & others are rare in α−helices? Why?

• Ala, Glu, Leu, Met = good helix formers

• Pro, Gly Tyr, Ser = very poor

• Amino acid composition & distribution varies, depending on on location of helix

in 3-D structure.

Table 2.1 Amino acid sequences of three α helices

1. - Leu - Ser - Phe - Ala - Ala - Ala - Met - Asn - Gly - Leu - Ala -
2. - Ile - Asn - Glu - Gly - Phe - Asp - Leu - Leu - Arg - Ser - Gly -
3. - Lys - Glu - Asp - Ala - Lys - Gly - Lys - Ser - Glu - Glu - Glu -

The first sequence is from the enzyme citrate synthase, residues 260–270, which form a buried helix; the second sequence is from the enzyme alcohol dehydrogenase, residues 355–365, which form a partially exposed helix; and the third sequence is from troponin-C, residues 87–97 which form a completely exposed helix. Charged residues are coloured red, polar residues are blue, and hydrophobic residues are green.

18

What are Loops?

1. **Type of 2nd structure.**

2. **Connect helices and sheets**

3. **Are located on surface of structure**

4. **Are more flexible and can adopt multiple conformations**

5. **Tend to have charged and polar amino acids**

6. **Usually involve active sites such as Calcium binding sites**

7. **Some fall into distinct structural families** (e.g., hairpin loops, reverse turns)
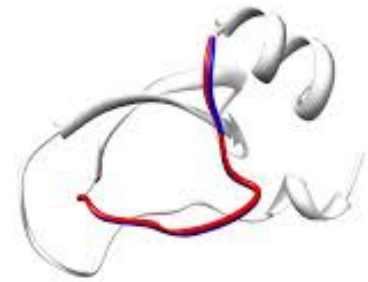
19

What are the most common structural motifs?

1. Helix-turn-helix   e.g., DNA binding **H-T-H**

2. Helix-loop-helix **H-L-H**   e.g., Calcium binding sites in troponin.

3. Coiled-coil: two coils twist over each other.

H-T-H   **H-L-H**
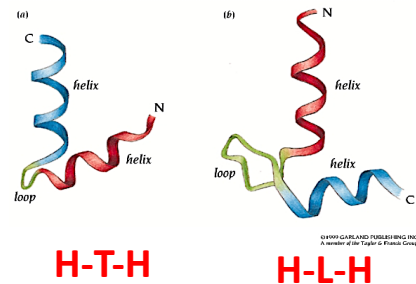
20

Protein structure databases, structural classification & visualization

Please search for types of proteins in the following webs.

1. PDB = Protein Data Bank http://www.rcsb.org/pdb/

2. (RISC) - several different structure viewers
   http://www.pdg.cnb.uam.es/cursos/Barcelona2002/pages/Farmac/CATH/index.html

3. MMDB = Molecular Modeling Database

4. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Structure

   (NCBI Entrez) - Cn3D viewer

5. SCOP = Structural Classification of Proteins http://scop.mrc-lmb.cam.ac.uk

   Levels reflect both evolutionary and structural relationships

6. CATH = Classification by Class, Architecture, Topology and Homology http://www.cathdb.info

The next steps is to predict the protein structure, please follow.....

21