Introduction to Bioinformatics I

Dr Manaf A Guma
University Of Anbar- college of Applied sciences-Hit
Department of applied chemistry

1

What does bioinformatic study?

Bioinformatics and **computational biology** are addressing biological problems with computational methods

It uses the tools and terminology of biology to describe the properties of living organisms and their genes

It answers some questions in genetics fields:

How are we different from others?

2

Bioinformatics:

Genetics:

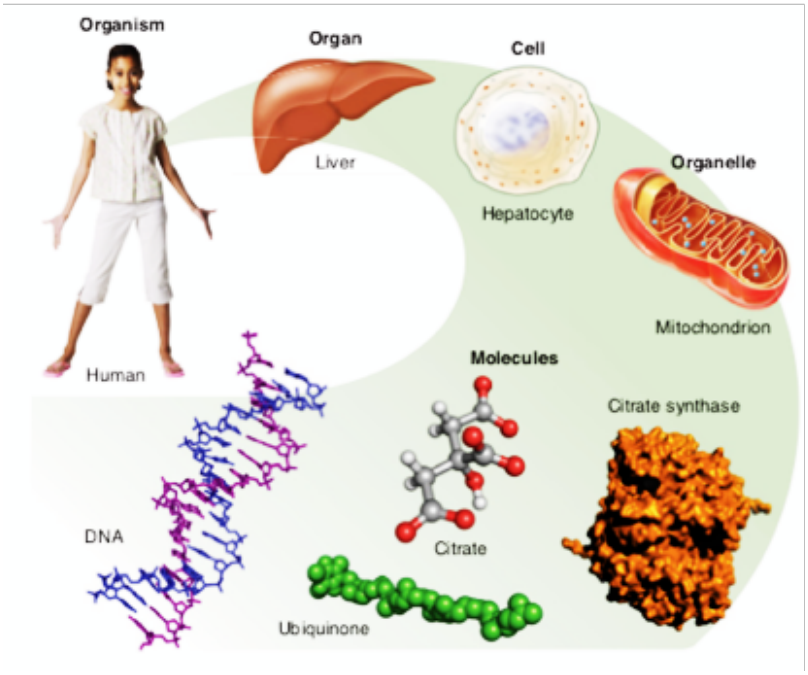
- DNA and RNA analysis and Structure.
- Includes techniques.
- PCR, Cloning and mutagenesis.
- Statistics of different variants of species

Proteins:

- Expression and purification of protein.
- Structural and functional of proteins.
- Chemistry biology of the protein.
- Biophysics.
- NMR, X-ray Crystallography, Electron microscopy and enzymes and others' kinetics
- Interactions and affinities.
- Drug design and discovery

3

What we do consist of ? What are the small molecules that code us?



4

The Storage of Genetic Information: What is the meaning of the terms?

- Bioinformatics are driven by by simple information that are built up from the sequences of DNA, RNA and proteins (which are the **Storage of Genetic Information)**
- DNA consists of four nucleotides that store genetic information.
- The base sequence encodes the necessary information to generate proteins.
- The entirety of genomic DNA in any organism is known as a genome.
- The total pool of mRNA in any organism is referred to as a transcriptome.
- The entire pool of proteins in any organism is referred to as the proteome.
- A genome comprises genes that contain the information to build proteins.
- Genome is the entire DNA sequence of an organism.

5

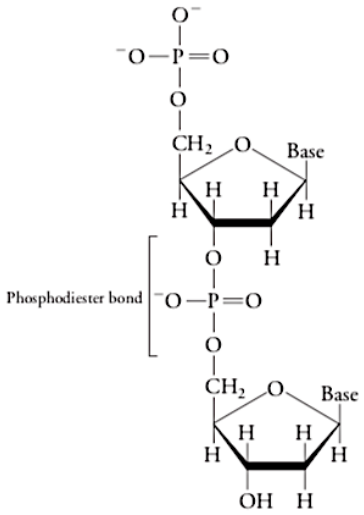
What are Macro (proteins) and micro (DNA)biomolecules ??

- DNA and RNA ?? Which are polymers of nucleotides, each of which consists of a purine or pyrimidine base, deoxyribose or ribose, and phosphate.
- Amino acids ??? Are coded by nucleotides>>>
- Proteins ?? Poly amino acids
- **Note: Micro**- is a prefix which may be applied to word when describing something that is small scale. **Macro**- is a prefix which **means** large.

6

Nucleic Acids:

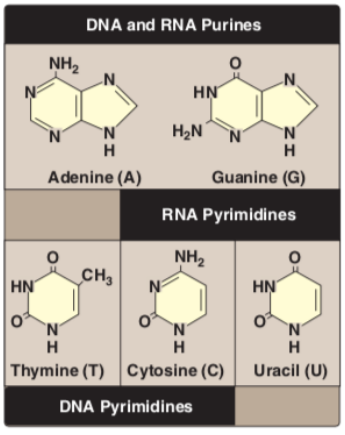
- Polymers of nucleotides are termed **polynucleotides** or **nucleic acids,** better known as DNA and RNA.
- RNA contain the bases adenine, cytosine, guanine, and uracil, whereas the residues in DNA contain adenine, cytosine, guanine, and thymine.
- Polymerization involves the phosphate and sugar groups of the nucleotides, which become linked by **phosphodiester bonds.**



7

Purine or pyrimidine base

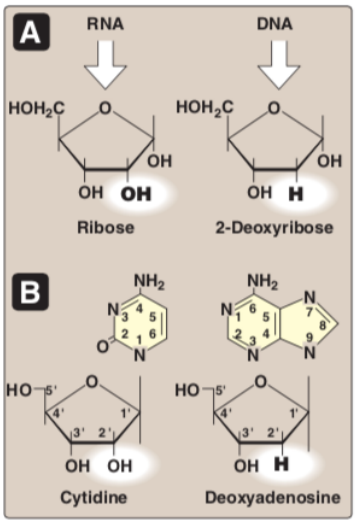
- Nucleotides are composed of a nitrogenous base, a pentose mono-saccharide, and one, two, or three phosphate groups.
- The nitrogen- containing bases belong to two families of compounds: the purines and the pyrimidines.
- Both DNA and RNA contain the same purine bases: adenine (A) and guanine (G). Both DNA and RNA contain the pyrimidine cytosine (C), but they differ in their second pyrimidine base: DNA contains thymine (T), whereas RNA contains uracil (U). T and U differ in that only T has a methyl group



Purines and pyrimidines commonly found in DNA and RNA.

8

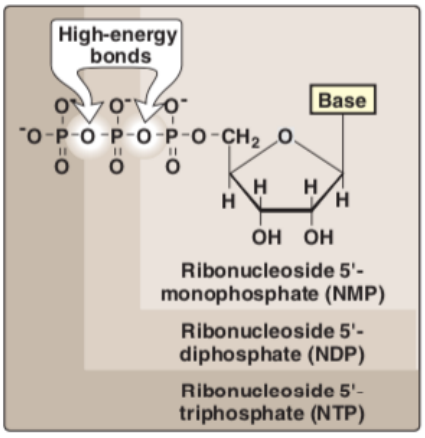
- **Nucleosides**
- The addition of a pentose sugar to a base produces a nucleoside.
- If the sugar is ribose, a ribonucleotide is produced; if the sugar is 2-deoxyribose, a deoxyribonucleoside is produced
- **Nucleotides**
- The addition of one or more phosphate groups to a nucleoside produces a nucleotide.



A. Pentoses found in nucleic acids. B. Examples of the numbering systems for purine- and pyrimidine-containing nucleosides.

9

- The first phosphate group is attached by an ester linkage to the 5'-OH of the pentose. Such a compound is called a nucleoside 5'-phosphate or a 5'-nucleotide.
- The second and third phosphates are each connected to the nucleotide by a "high-energy" bond.

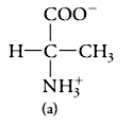


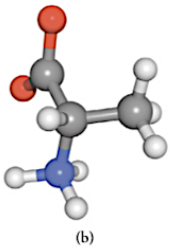
Ribonucleoside monophosphate, diphosphate, and triphosphate.

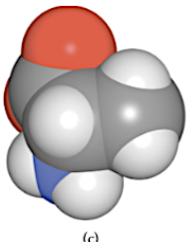
10

Amino Acids:

- They contain an amino group (-NH2) and a carboxylic acid group (-COOH) .
- For example, by a) structural formula, b) ball-and-stick model, or c) space- filling model.



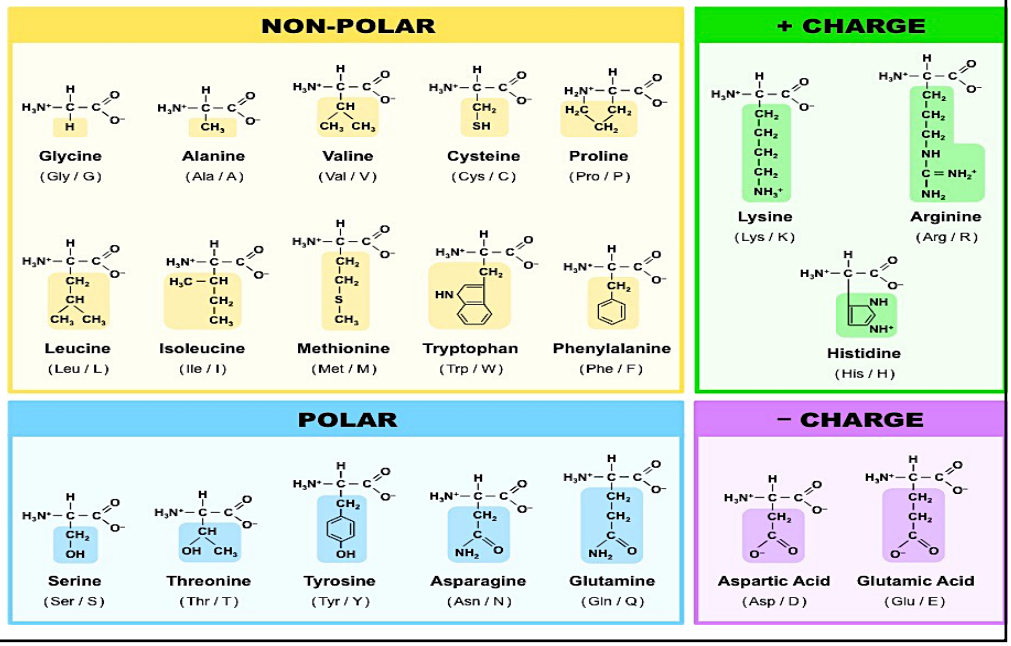




11

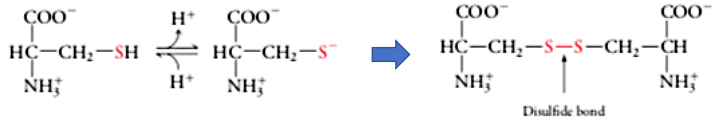
The 20 amino acids have different chemical properties

Note:-
Name:Glycine.
Symbol: G.
3-letters:Gly



12

- **Amino Acids with Hydrophobic Side Chains:** *Sidechains interact very weakly or not with water. e.g.* alanine (Ala).
- **Amino Acids with Polar Side Chains:** *The sides interact with water because they contain hydrogen-bonding groups. E.g.* Serine (Ser). *Cysteine (Cys) has a thiol group: can form a disulphide bond:*

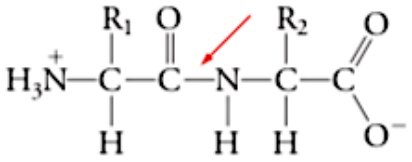


- **Amino Acids with Charged Side Chains** (*4 amino acids only):Side chains are always charged under physiological conditions.*
- (Asp) and (Glu), are acidic (COOH)
- (Lys) and (Arg) are positively charged (NH_3)

13

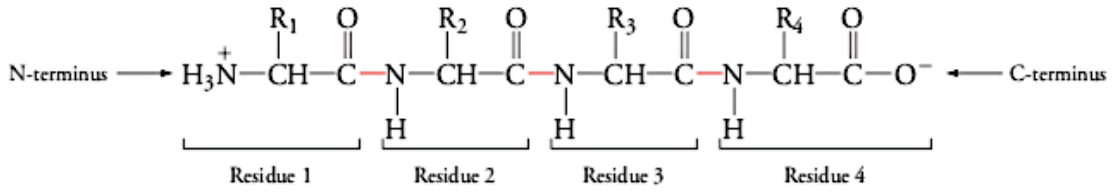
Proteins:

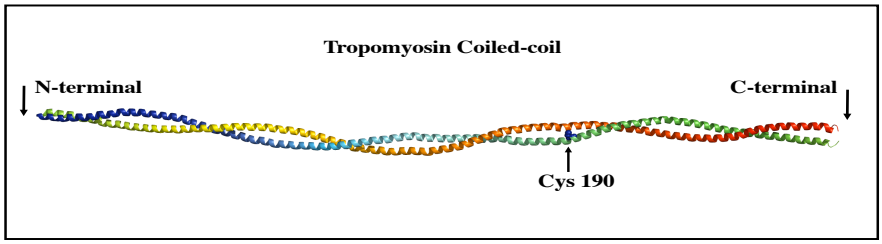
- Polymers of amino acids are called **polypeptides** or **proteins.**
- There are 20 different amino acids make building blocks for proteins.
- proteins may contain many hundreds of amino acid residues.
- The amino acid residues are linked to each other by amide bonds called peptide bo



14

The peptide bond in the ‘’backbone” starts with N-terminus and ends with C-terminus.

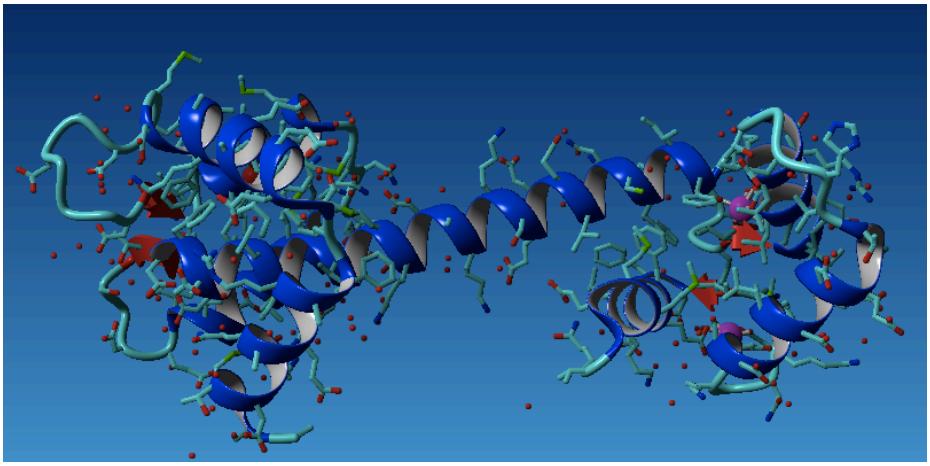




For example: Tropomyosin structure is a protein (with N-& C-terminus.

15

How many atoms are there? How many amino acids ? How many polypeptides?

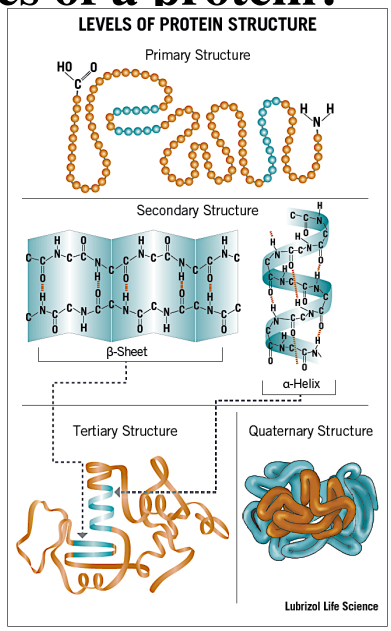


16

What are the different structures of a 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.



17

The Origin and Evolution of Life

- Modern prokaryotic and eukaryotic cells apparently evolved from simpler non-living systems.
- What are the three domains of life? They are bacteria, archaea, and eukaryote.
- What is the ability of the cells? Cells have ability to **replicate** (make a replica or copy of itself) is one of the universal characteristics of living organisms.
- The are two main types of cells:

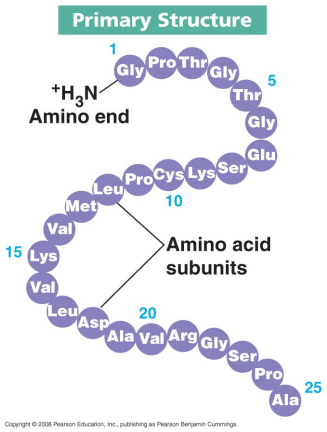
1- Prokaryotes.

2- Eukaryotic.

18

What is the Primary structure of a protein?

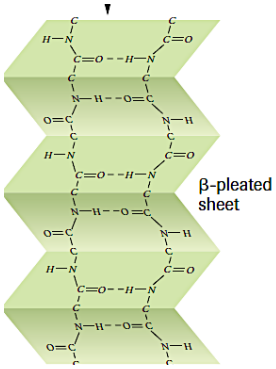
- The Primary structures of protein are:
- Linear sequence of amino acids
- This linear sequence is referred to as a polypeptide chain. The amino acids in the **primary structure** are held together by covalent bonds.

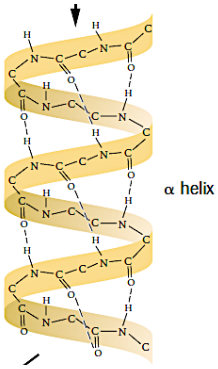


19

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

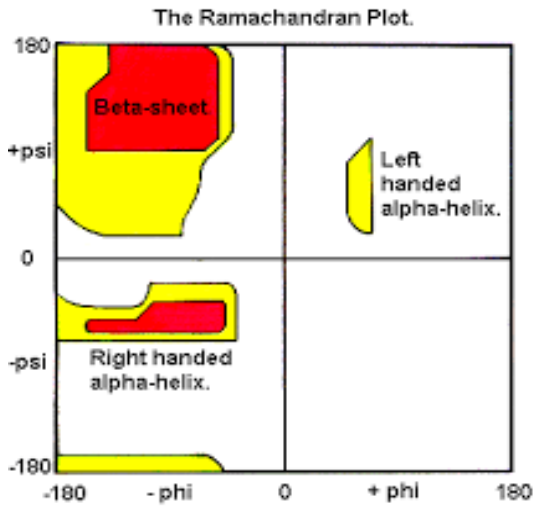




20

The chain conformation of a polypeptide can be determined by?

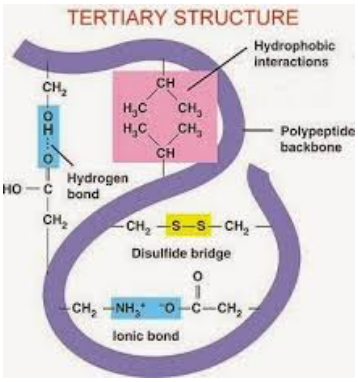
- It can be determined by the torsion angles around the Cα–N binding (ϕ) and the Cα–C binding (ψ) of the constituent amino acid residues.
- A Ramachandran plot is a conformation chart of those values that are sterically possible for ϕ and ψ.
- It determines the alpha helices and the beta sheets contents of a protein.



21

What are the tertiary & quaternary structures of protein?

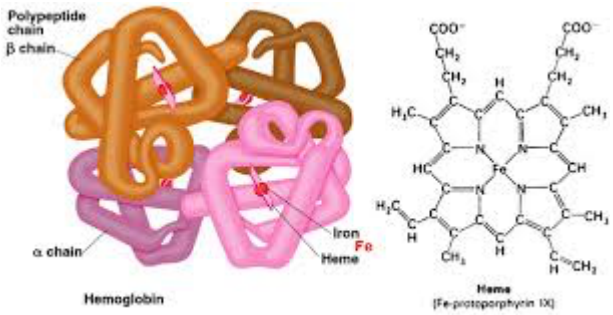
- Tertiary: It is 3D structures shape of protein.
- It has a single polypeptide chain "backbone" with one or more protein secondary structures that form the protein domain.
- Bonds?



22

What is the Quaternary structure of proteins?

- 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.



23

Next lecture, ….

- We will continue explaining biomolecules, the bases bioinformatics study….
- We will talk about DNA and RNA molecules…

24