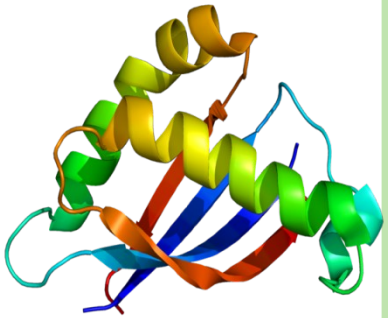
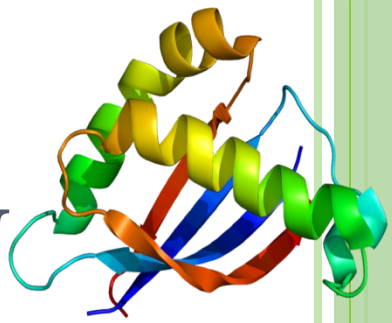


INTRODUCTION TO PROTEIN-PROTEIN INTERACTION

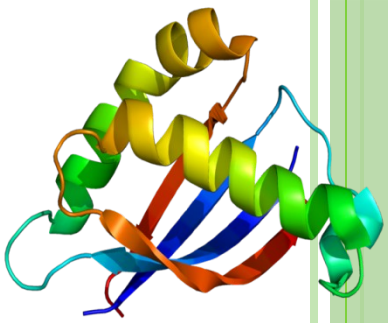
PRATIK DUTTA,
1621CS05,
AI-NLP-ML Group,
CSE Dept, IIT PATNA
Guidance By: Dr. Sriparna Saha

OUTLINE OF THE PRESENTATION

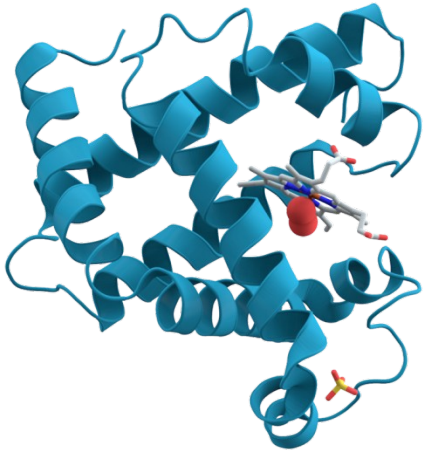


- What is Protein??
- Protein Structure
- RNA/DNA/ Gene/ Protein
- Central Dogma of Biology
- Microarray and Gene Expression Data
- Protein-protein Interaction
- Protein Interaction Network
- Conclusion

WHAT IS PROTEIN

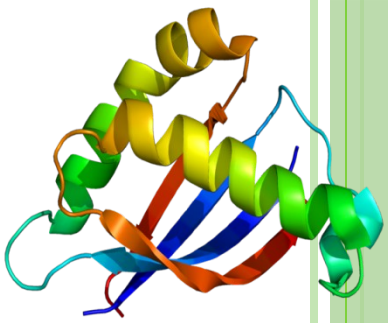


- Protein are large biomolecules that consisting of one or more long chains of amino acids
- A protein contains at least one long polypeptide(chain of amino acids)
- Primarily the sequence of amino acid differs the protein from each other

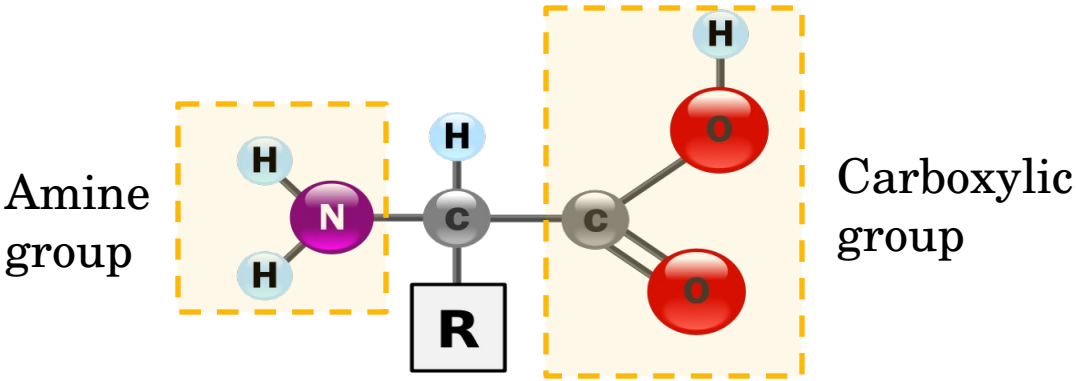


3D structure of protein myoglobin

AMINO-ACIDS AND POLYPEPTIDE



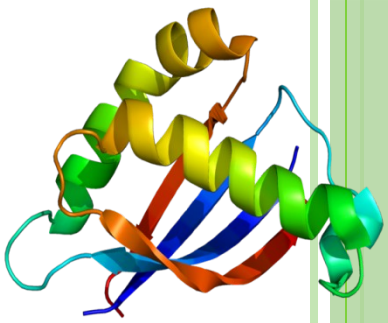
- Amino acid are biologically important organic compound that consists of
  - Amine group($-NH_2$)
  - Carboxylic Acid group (-COOH)
  - and a side chain(- R)



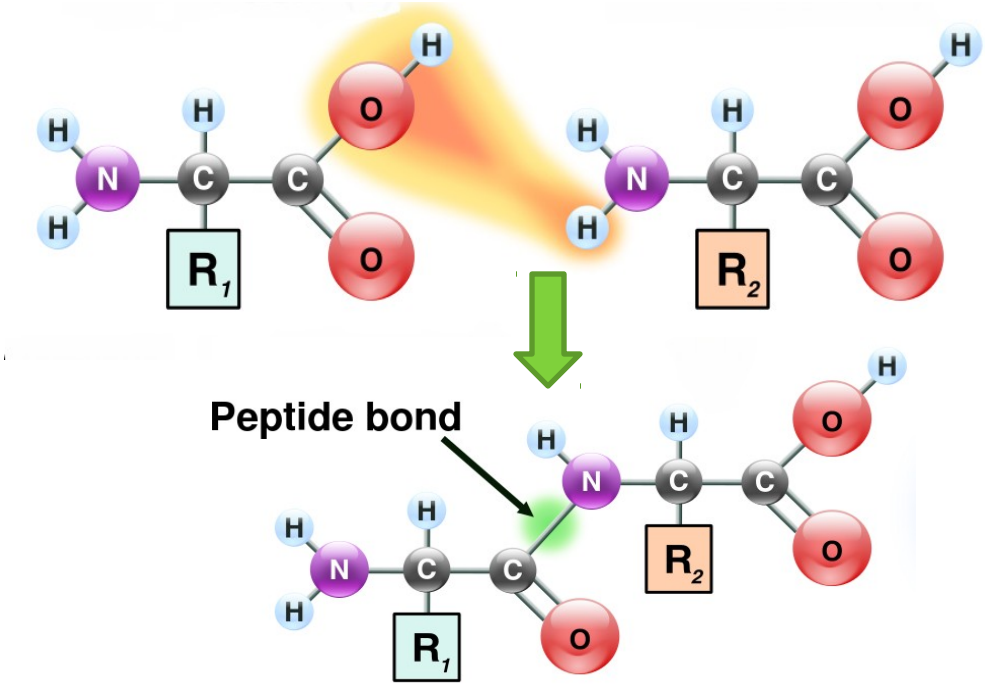
Structure of an α-amino acid

- Therefore the key elements of amino acid are **carbon (C) , hydrogen(H) , oxygen(O)** and **nitrogen(N)**

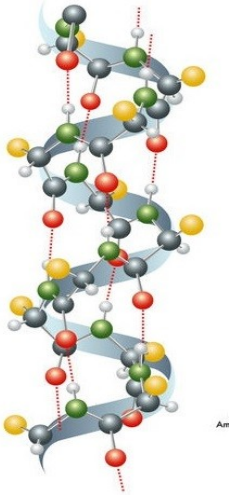
AMINO-ACIDS AND POLYPEPTIDE



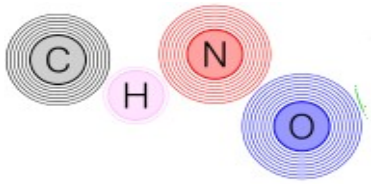
- Though there are about 500 amino acids, only 20 amino acid appear in genetic code
- When two or more amino acid is connected through ***peptide bond*** it is called polypeptide(amino acid chain)



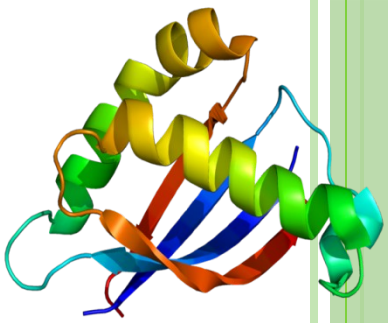
STRUCTURAL ELEMENTS OF PROTEIN

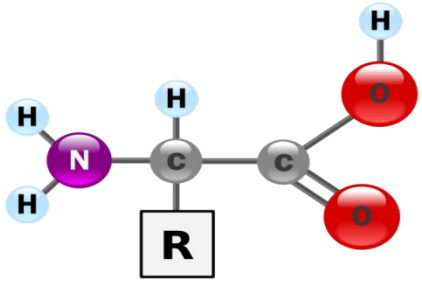


3d structure of protein

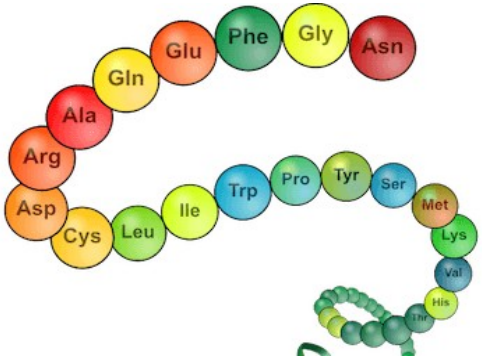


Key elements of protein

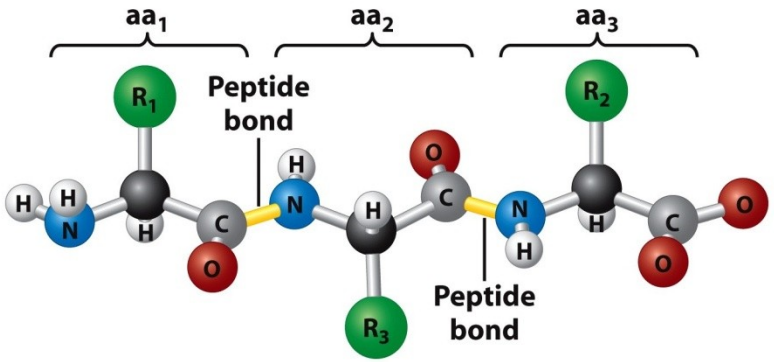




Molecular structure of amino acid

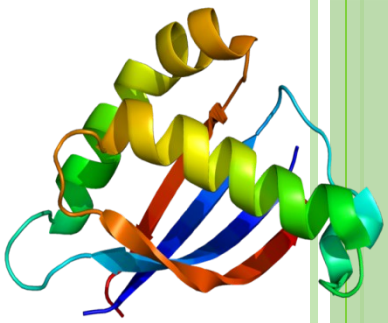


chain of Amino acid

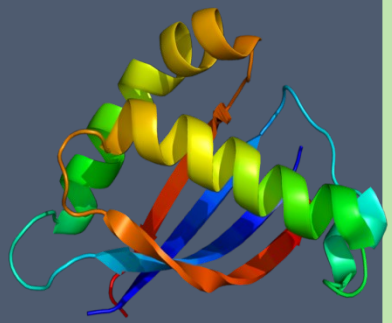


Molecular structure of peptide

PROTEIN STRUCTURE



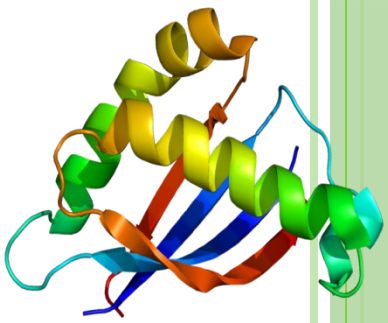
- Protein structure is the three-dimensional arrangement of atoms in a protein molecule
- To perform several biological functions, protein folds into one or more specific spatial conformations
- Several non-covalent interactions are responsible for this confirmations
- Levels of protein structure
  - ***Primary structure:*** *linear sequence* of amino acids in the polypeptide chain
  - ***Secondary structure :*** *helical structure* due to hydrogen bond between the main-chain peptide groups. ***E.g*** α-helix and β-sheets
  - ***Tertiary structure:*** folded into a compact *globular structure* due to hydrophobic interactions



THE CENTRAL DOGMA

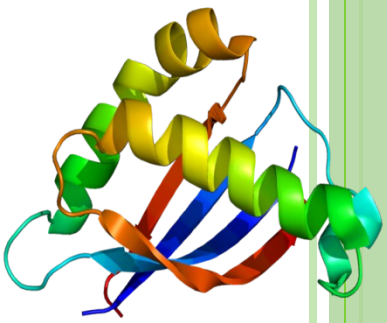
Relation between DNA, RNA and proteins

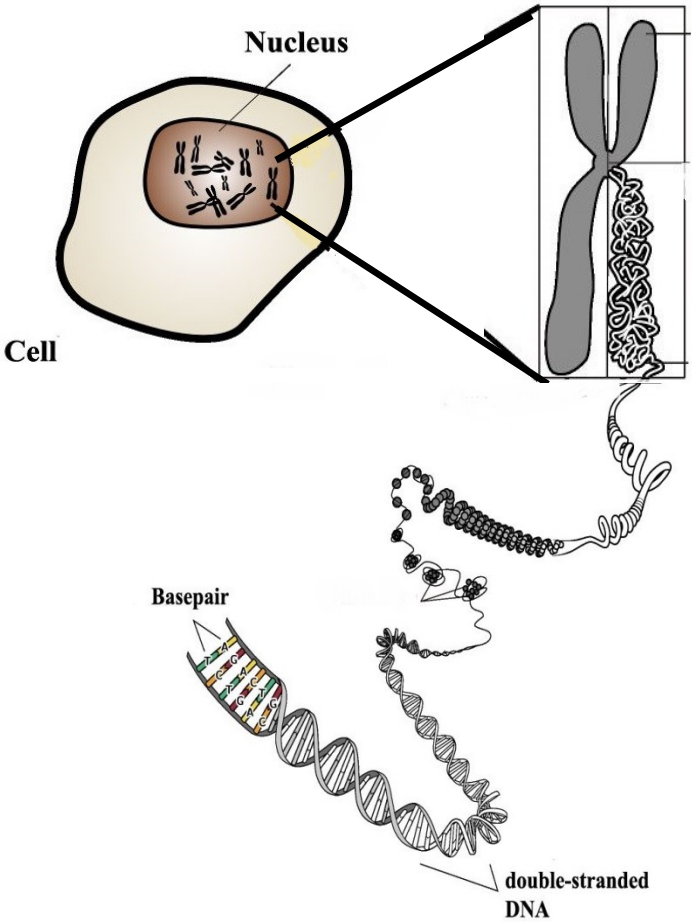
CENTRAL DOGMA

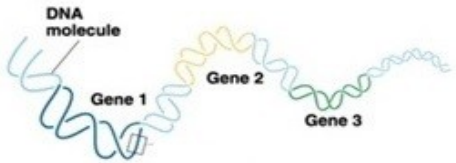


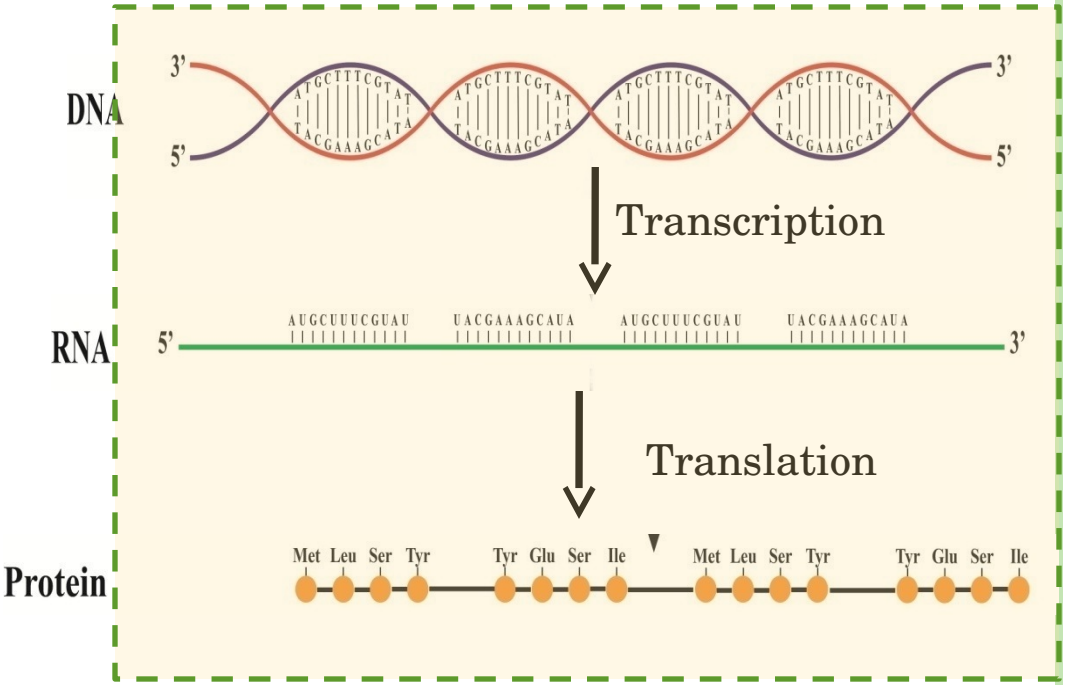
- Explains *the flow of genetic information from **DNA** to **RNA**, to make a functional product **Protein***
- ***Stages of Central Dogma***
  - ***Replication***: Fundamental step of central dogma, make a new DNA from existing DNA by DNA polymerase
  - ***Transcription:*** Make new mRNA from DNA by RNA polymerase
  - ***Translation:*** Make new protein from mRNA by ribosome

CENTRAL DOGMA

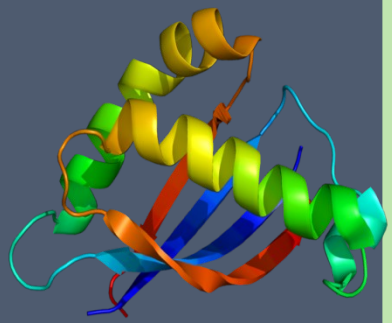






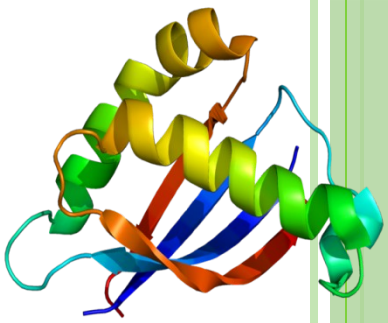


Central Dogma of Molecular Biology



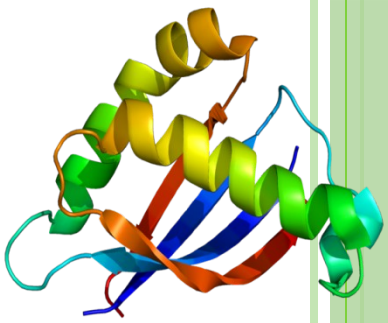
GENE EXPRESSION VALUES

GENE EXPRESSION VALUE

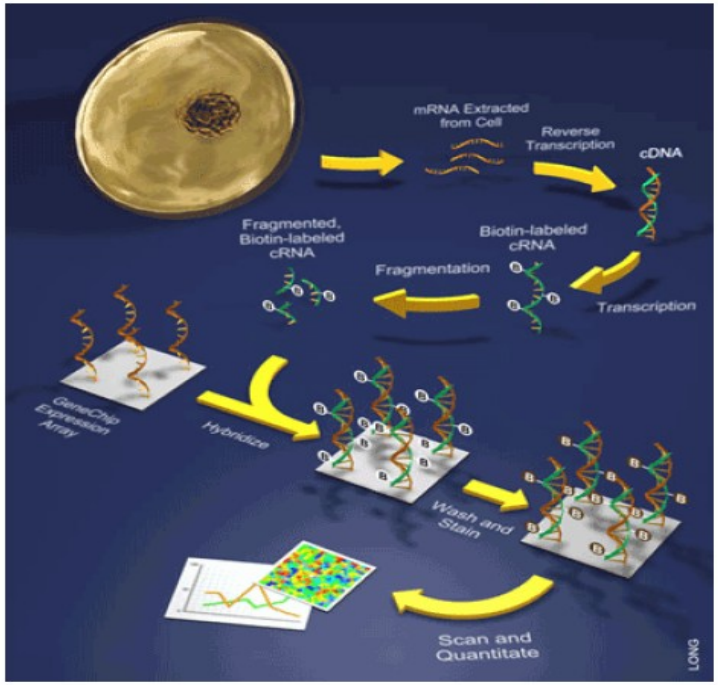


- In the field of molecular biology, ***gene expression profiling*** is the measurement of the activity of thousands of genes at once
- Gene expression values of large genes are measured by microarray
- Gene expression can be quantified ***by measuring either mRNA or protein***
- Gene Expression Measurement Method
  - Single Channel Arrays:
    - Give estimations of the absolute levels of gene expression
    - compare two set of condition for each gene
    - requires two separate single-dye hybridizations
  - Two Channel Arrays:
    - Two hybridized cDNA samples (e.g. diseased tissue versus healthy tissue) and labelled with two different ***fluorescentdyes(Cy5 and Cy3)***

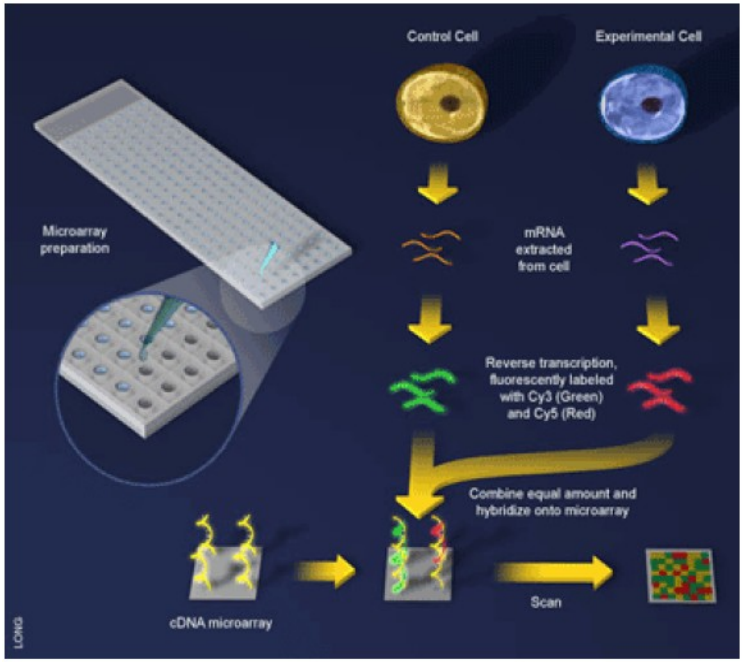
GENE EXPRESSION VALUE



Single channel arrays

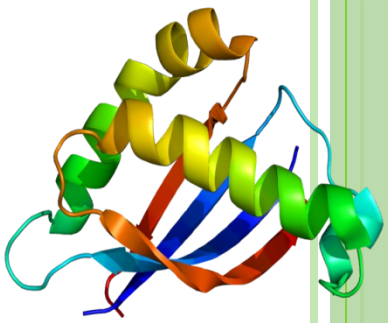


Dual channel arrays



GENE EXPRESSION DATASET

(http://www.ncbi.nlm.nih.gov/sites/GDSbrowser?acc=GDS1388)

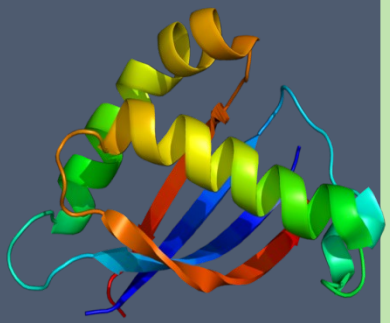


```
!dataset_table_begin
ID_REF  IDENTIFIER  GSM45021    GSM45022    GSM45023    GSM45024    GSM45025    GSM45066    GSM45067    GSM45068    GSM45069    GSM45
GSM45074    GSM45075    GSM45076    GSM45077    GSM45078    GSM45079    GSM45080    GSM45081
1000_at MAPK3   357.7   371.2   404.6   330.7   333.2   241.7   284.6   292.4   268.3   324.2   181.2   407.6   393 280.8   367.8   4
1001_at TIE1    20.9    18.5    10.8    13.5    20.5    10.7    3.7 12.8    25.4    23.1    11.5    34.2    28  17.7    26.3    17.5
1002_f_at   CYP2C19 4.7 18.9    17.6    2.6 5.5 3.3 3.2 1.8 9.8 22.1    26.5    6.4 8.2 6.3 2.5 3.3 10.8    5.5 7.2 3.6 6.1
1003_s_at   CXCR5   174.1   231.7   214.1   201.6   314.5   8.4 7.2 204.6   22.3    109.3   100.4   414 75.7    211.8   232.5   189 2
1004_at CXCR5   257 194.3   221.5   185.4   376.4   355.9   176.8   244.2   186.2   743 268.7   423.2   132.6   157 155.1   216 244.6
1005_at DUSP1   732.9   499.8   1092.3  923.6   287.5   137.9   1033.4  384.2   183.1   184 959.2   5835.6  755.1   150.1   604.2   5
1006_at MMP10   1.1 2.2 11.7    2.6 15.1    1.1 1.1 2.8 3.3 8.6 2   1.8 1.2 1.9 3.7 1.4 2.5 1.1 4.7 5.9 6.9
1007_s_at   DDR1    109.7   174.6   185.9   109.1   215.8   161.9   196.5   162 156.7   140.2   169.1   271.4   177.4   199.3   236.6
1008_f_at   EIF2AK2 887.2   551.8   2038.1  653.5   2216.7  1581.3  677.5   850.3   3090.5  1054.8  2198.8  4194.9  1969.6  1006.7  7
807.6
1009_at HINT1   1145.5  1014.9  1103.1  850.4   886.6   859.3   1228.6  1231.9  1118.3  762.1   1438.9  763.7   1382.4  1008.2  1025.
100_g_at    RABGGTA 169 200.6   196.3   151.4   167.8   111.3   164.9   240.6   155.9   215.2   121.7   277.2   172.7   163.6   194 1
1010_at MAPK11  8.5 10.3    6.6 3.4 11.6    8   7.5 5.7 13.2    25.9    13.3    20.2    48.6    20.3    4.2 3.5 12.4    12.8    9.2 1
1011_s_at   YWHAE   25.9    31.7    43.3    40.9    40.8    26.7    45.1    62  60.4    63.4    38.6    12.5    53.6    72.8    59.6
```

Gene expression level of gene i in mRNA sample j

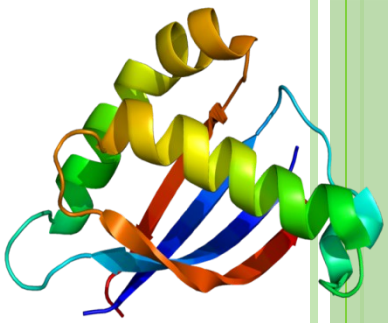
Single Channel Arrays: (normalized) Log2(Intensity)

Double Channel Arrays: $\log_2 \frac{Intensity(Cy5)}{Intensity(Cy3)}$

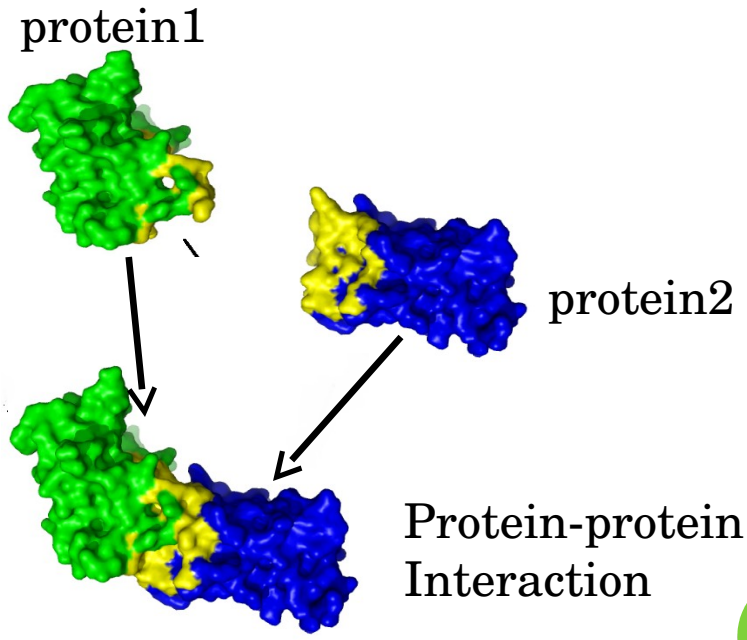


PROTEIN-PROTEIN INTERACTION

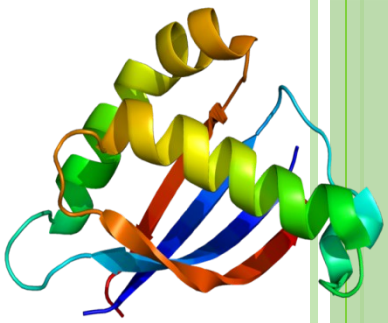
PROTEIN-PROTEIN INTERACTION



- Lasting and specific physical contacts established between two or more proteins for carried out some specific biological activity
- Represents pair wise protein interactions of the organisms
- Example of PPI
  - Muscle Contraction
  - Cellular Transportation

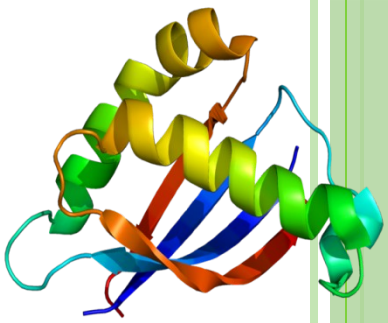


PROTEIN-PROTEIN INTERACTION



- Interaction between two proteins is carried out by several biochemical events
  - ***Electrostatic Forces***:- Force interacting between static electrically charged particles
  - ***Hydrogen bonds:-*** electrostatic attraction between hydrogen(H) and highly electronegative atom(e.g. O, N)
  - ***Van der Waals forces:-*** residual attractive or repulsive forces between molecules or atomic groups
  - ***Hydrophobic interactions:-*** Maximize hydrogen bond

TYPES OF INTERACTION

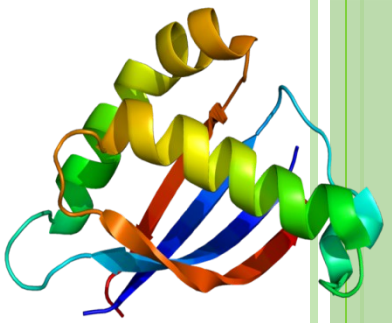


- ***Stability***
  - Stable:- Always stable and active e.g. Hormones, Hemoglobin
  - Transient:- Control the majority of cellular processes, can be strong or weak, fast or slow

- ***Structural***
  - Homo-oligomer :- Same type of subunits e.g. Enzymes
  - Hetero-oligomer:- Different types of subunits e.g. G-proteins

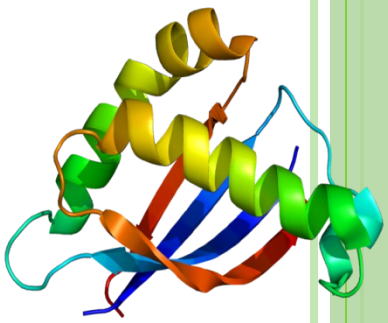
- ***Chemical Bonding***
  - Covalent Bonding: Share electron pairs
  - Non-covalent Bonding: Rather sharing electrons , involves in some electromagnetic forces

IMPORTANCE OF PPI NETWORK

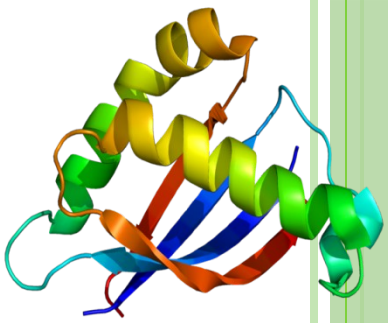


- Useful for isolating groups of interacting proteins that participate in the same biological process
- Helps to understand the mechanism of regulating cell life
- Useful to predict the biological functions of uncharacterized proteins

REFERENCES



- http://www.medicalnewstoday.com/articles/196279.php
- http://www.chemistry.uoguelph.ca/educmat/phy456/456lec02.htm
- http://www.biology.arizona.edu/biochemistry/problem_sets/aa/aliphatic.html
- http://www.particlesciences.com/news/technicalbriefs/2009/protein-structure.html
- http://www.slideshare.net/NaghmehPoorinmohamma/proteinprotein-interaction-ppi
- http://www.clearbiology.com/biochemistry-review-activity/proteins/
- http://worldofbiochemistry.blogspot.in/2013_09_01_archive.html
- http://www.ch.embnet.org/CoursEMBnet/Pages3D07/documents/3Dintro07.pdf
- http://www.ch.embnet.org/CoursEMBnet/PagesTR08/slides/EmbNet_CleanEx_1.pdf
- https://en.wikipedia.org/wiki



THANK YOU