LECTURE-2

Basics of Amino acids and Proteins

HANDOUT

PREAMBLE

Proteins are the most complex and versatile macromolecules comprised of amino acids as the building blocks. There are 20 standard amino acids depending upon the functional group/side chain that is attached to the central asymmetric carbon atom of the amino acid. Each amino acid is encoded from a codon (which is a 3-nucleotide stretch) by the process of translation of the corresponding mRNA transcript, which in turn is generated from the DNA sequence by transcription. Multiple amino acids are joined together in different combinations by means of peptide bond between two amino acids, in order to form a peptide. Multiple peptides are joined together to form a polypeptide, which are eventually joined to form protein. Many proteins are usually composed of different subunit complexes, which may be structurally and functionally different. The hierarchical arrangement can be depicted as:

DNA—mRNA—Codons—Amino acids—Peptides—Polypeptides—Proteins

To understand protein structure and function it is important to understand the basics of amino acids, their structure and function, their properties with respect to the microenvironment that they are usually present. Knowledge of these fundamental aspects can be further extrapolated to gain deeper insights using proteomic approaches.

OUTLINE OF LECTURE

1. Basics of amino acids

A. The structural backbone and classification of amino acids

B. Isomerism in amino acids

C. Significance of amino acid sequence

2. Peptide bond

3. Structural level of proteins

4. Protein folding and misfolding

5. Conclusions

Box: Terminologies

- **Enantiomers**: The dextro and leavo rotatory forms of a given amino acid, which are non-super-imposable mirror images of each other.
- **Amino group:** This consists of an NH_2 group covalently bonded to the central carbon atom. Depending upon the pH of the surrounding medium, it either exists as NH_2 or NH_3^+. Except for proline, which has a secondary amino group, all amino acids have only primary amino groups.
- **Carboxyl group:** A COOH group covalently bound to the central alpha carbon atom, which exists as either COOH or COO^- depending on the pH of the surrounding medium.
- **Peptide bond**: Also known as the amide bond and is responsible for holding two amino acids together. In one bond, the carboxyl group of one amino acid is bound to the amino group of the other amino acid by means of a covalent bond, the formation of which involves loss of a water molecule.
- **Psi) and Phi angles:** Angle of rotation about the bond between the α-carbon atom and carboxyl and amino groups respectively. These angles determine which protein conformations will be favourable.
- **Hydrogen bonds:** Interaction between electropositive and electronegative atoms. They can be formed within a polypeptide chain or between different polypeptide chains.
- **Electrostatic interactions:** Attractive forces existing between oppositely charged groups/atoms, which can stabilize the protein structure.

- **Hydrophobic interactions:** Non-specific interactions between non-polar amino acid side chains, which acts to bury these hydrophobic residues away from a polar environment.
- **Van der Waals forces:** Weak attractive or repulsive forces caused due to changes in polarization.
- **Disulphide bridges:** Specific interaction and oxidation of thiol groups of cysteine residues in different regions of the polypeptide chain(s), which leads to formation of disulphide (S-S) bonds.

1. Basics of Amino acids

A. Structural backbone of amino acids and their classification

An amino acid is composed of an asymmetric or chiral central carbon atom, which is bound to four different groups, namely; a carboxyl group, an amino group, a hydrogen atom and a side chain. The central carbon atom is known as the alpha carbon atom. The amino acids with a side-chain bonded to this carbon are simply referred to as alpha amino acids.

Illustration: Amino acid structures & properties

Amino acids are building blocks or monomers that make up proteins. They consist of a central alpha carbon atom bonded covalently to an amino group, a carboxyl group, a hydrogen atom and a variable side chain, also called the R group. These are basic monomeric constituents of proteins found in varying amounts depending upon the type of protein. They are classified based on the properties of their side chains or R groups that varies in size, structure and charge. The polarity of the side chains is one of the main basis for this classification.

Amino acids having non-polar, aliphatic side chains include glycine, alanine, proline, valine, leucine, isoleucine and methionine. Essential amino acids are those that cannot be synthesized de novo in the organism and therefore must be included in the diet. Non-essential amino acids on the other hand, can be synthesized from various precursors. Serine, threonine, aspargine, glutamine & cysteine consist of polar but uncharged side chains. Lysine, Arginine and Histidine- amino acids have positively charged side chains. Aspartic acid and glutamic acid are polar and negatively charged amino acids. Tryptophan, tyrosine and phenylalanine are all essential amino acids having an aromatic side chain.

B. Isomerism in Amino acids

Amino acids can exist as L-isomers or D-isomers, however, with respect to the conformational stability, all amino acids are usually found in the L-isomeric form.

Illustration: Isomerism in amino acids

All amino acids except glycine contain an asymmetric centre that makes them chiral in nature due to which they can rotate the plane of polarized light. The two enantiomers are designated as D and L, which rotate the plane of polarization in opposite directions. The two enantiomers of amino acids are non-superimposable mirror image due to the spatial arrangement of four different groups about the chiral carbon atom. Rotation of either isomer about its central axis will not give rise to the other isomeric structure.

Illustration: Acid-base properties of amino acids

Amino acids in acidic medium exist in the completely protonated form carrying a net positive charge, which can be confirmed by means of simple paper electrophoresis. The sample solution is applied at the centre of the strip and current is passed through it. The colorless amino acid solution can be detected by spraying the strip with ninhydrin, which gives it a purple color. Migration of the spot towards the negatively charged cathode confirms the net positive charge of the amino acid.

All amino acids exhibit a characteristic titration curve with distinct pK values. 0.1N NaOH is added to the acidic amino acid solution. The cationic form of the amino acid is gradually converted into its neutral or zwitterionic form by loss of a proton from its COOH group. This can again be confirmed by electrophoresis where there is no migration of the sample spot. Number of equivalents of alkali being consumed is plotted against the pH of the amino acid solution to obtain the titration curve. pK_1 of glycine is found to be 2.34 i.e. it starts to lose its carboxyl group proton at this pH. Removal of the proton from the amino group constitutes the second stage of the titration curve. Continued addition of alkali to the amino acid solution gradually converts the zwitterionic form into the anionic form. Migration of the sample spot towards the anode during electrophoresis confirms this. The pK_2 of an amino acid is obtained by continued addition of alkali to the neutral solution of the amino acid. pK_2 of glycine is found to be 9.6. Some amino acids having positively or negatively charged side chains will have pK_1, pK_2 and pK_R, which corresponds to ionization of the side chain. These amino acids have good buffering capacity around 1 pH unit on either side of their pK values.

C. Significance of Amino Acid sequence

The significance of amino acid sequence was elucidated by Anfinsen's experiment, which is described as follows:

ANFINSEN'S EXPERIMENT (Reported in 1973)

- OBJECTIVE: To establish that amino acid sequence dictates protein folding.

- EXPERIMENTAL INSIGHTS:

The roles of compounds like Urea, Guanidine HCl, β-mercaptoethanol, etc. on the refolding of an enzyme Riobnuclease A was elucidated through the experimental study. 8M Urea and Guanidine HCl together act as denaturants and disrupt the non-covalent bonds present in the protein. β-mercaptoethanol on the other hand acts as a reducing agent and disrupts the disulfide bonds, and when applied in large excess converts the disulfides to sulfhydryls. This denatured and reduced form of Ribonuclease A lacks enzymatic activity.

When Urea and β-mercaptoethanol are removed by the dialysis, the denatured and reduced Ribonuclease A spontaneously regains its native conformation and enzymatic activity. The removal of β-mercaptoethanol alone, however did not render the Ribonuclease A active and still lacks its enzymatic activity. It was also observed that trace amounts of β-mercaptoethanol facilitated the correct disulfide bonding and thereby enhanced the formation of native Ribonuclease A.

- CONCLUSIONS: The phenomenon of protein folding is governed by the amino acid sequence, which is the actual repository of all the genetic information.

Illustration: Anfinsen's experiment

Ribonuclease A in its native state has four disulphide bonds between its cysteine residues. When treated with β-mercaptoethanol and 6M Urea, the protein undergoes denaturation and the disulphide linkages are broken. Enzyme activity is lost in the denatured state. It was observed by Anfinsen that removal of urea and β-mercaptoethanol led to the refolding of the enzyme to assume its native state with more than 90% enzyme activity being intact. However, if only β-mercaptoethanol was removed in presence of urea, the formation of disulphide bonds was random, leading to enzyme with only around 1% activity.

2. Peptide Bond

The covalent bond that holds two adjacent amino acid residues together is known as a peptide bond, which is formed between the carboxyl group of one amino acid and the amino group of other amino acid and is accompanied by release of a water molecule. The peptide bond is stabilized by resonance structure. The amide bond exhibits partial double bond character and is planar. In other words, it can exist in “cis” and “trans” form. In the unfolded form of a given protein, the peptide bonds have the liberty to take up either of the two forms; however the folded conformation has the peptide bond in a single form alone. The “trans” form is usually preferred as it’s conformation is stable as compared to the “cis” form (Exception: Proline, which can exist in “cis” as well as “trans” form). The psi and phi are the angles of rotation about the bond between the α-carbon atom and carboxyl and amino groups, respectively. These angles determine which protein conformations will be favourable during protein folding.

Illustration: Peptide Bond formation

Amino acids are the building blocks or monomers that make up proteins. Amino acids are oriented in a head-to-tail fashion and linked together such that the carboxyl group of one amino acid combines with the amino group of another. Two amino acids are joined together by means of such a condensation reaction with the loss of a water molecule forms a dipeptide. Many such amino acids linked together form a polypeptide. The peptide bond is rigid due to its partial double bond character arising from resonance structures. However, the bonds between the α-carbon and amino and carboxyl groups are pure single bonds that are free to rotate.

3. Structural Level of Proteins

A peculiar arrangement of amino acid residues of a protein is essential to render it functionally active and structurally stable. A given protein structure is the organization of different L-alpha-amino acids in a particular way, which is aided by numerous interactions like Covalent interactions, Hydrogen bonding, Vander Waals forces, Ionic interactions and other Hydrophobic interactions. The protein molecule can also undergo reversible changes in its structure, in order to be able to perform its function effectively, and these structural variations are termed as “conformational changes”.

3.1. Primary Structure

The linear sequence of amino acids constitutes the primary structure.

Illustration: Primary structure

Amino acids are joined together in a head-to-tail arrangement by means of peptide bonds with the release of water molecules. This linear sequence of amino acids constitutes the primary structure.

3.2. Secondary Structure

The secondary structure of proteins is mainly a function of the pattern of hydrogen bond in the main chain-peptide groups and their characteristic geometry can be deciphered from the phi and psi angles. Alpha helices and Beta sheets are two most common forms in which proteins can exist in their secondary structure. The alpha helices can be further classified as left-handed or right-handed depending on their orientation. However, most alpha helices are right handed since this conformation is energetically more favorable. Beta sheets on the other hand can be parallel or anti-parallel. Amino acids in parallel β-sheets, which run in the same direction, interact with two different amino acids on the

adjacent strand through hydrogen bonds. Amino acids in anti-parallel strands on the other hand interact with only one amino acid on an adjacent strand.

A compact, globular structure is required for all proteins to attain structural stability. This can be made possible only if there are turns or loops between the different secondary structures. Beta-turns, which are the most commonly observed turn structures, consist of rigid, well-defined structures that usually lie on the surface of the protein molecule and interact with other molecules. A combination of secondary structures such as the helix-turn-helix, which consists of two alpha helices separated by a turn, is also observed and these are known as super-secondary structures or motifs.

Ramachandran plot can be used to understand the relative distribution of alpha helices and beta sheets within a given protein depending on the phi and psi angles of the component amino acid residues. To distinguish different secondary structures, a mapping of the range of phi and psi values to different regions of the plot can be established. The allowable regions are the ones, which represent the conformations wherein atoms of the polypeptide would not have any steric clashes. Disallowed regions generally involve steric hindrance between the side chain and main chain atoms. Glycine stands out as an exception and has no side chain and therefore can adopt phi and psi angles in all four quadrants of the Ramachandran plot.

Illustration: Secondary structure

The folding of the primary structure into the secondary is governed by the permissible rotations about the phi and psi angles. Not all values of these angles lead to sterically favorable conformations. The

Ramachandran plot defines these regions of favorability. Amino acids along the polypeptide backbone interact through hydrogen bonds leading to secondary structures.

The α-helix has intra-chain hydrogen bonds between the 'H' of NH and 'O' of CO in every 4th residue. Most alpha helices are right handed since this conformation is energetically more favorable. The amino acid proline that possesses a cyclic side chain does not fit into the regular alpha helix structure and thereby limits flexibility of the backbone. It is commonly referred to as the "helix breaker" Alpha helices can also wind around each other to form stable structures such that their hydrophobic residues are buried inside, while their polar side chains are exposed to the aqueous environment. α-keratin, the major protein component of hair consists of two such coiled coils forming a left-handed super-helix. Collagen, which is fibrous component of skin, muscle etc. consists of three such coiled alpha helices. It has a characteristic recurring amino acid sequence of glycine-proline-hydroxyproline with glycine appearing at every 3rd residue.

The β-pleated sheet discovered by Pauling & Corey, is another common secondary structure with periodic repeating units. It is composed of two or more polypeptide chains with their side chains oriented above & below the plane. It is an extended structure with hydrogen bonds between the chains stabilizing it. Amino acids in parallel β-sheets, which run in the same direction, interact with two different amino acids on the adjacent strand through hydrogen bonds. Amino acids in anti-parallel strands on the other hand interact with only one amino acid on an adjacent strand.

Protein secondary structures – Turns & Loops: Almost all proteins exhibit a compact, globular structure which is possible only if there are turns or loops between the various regions. β-turns, which are the most commonly observed turn structures, consist of rigid, well-defined structures that usually lie on the surface of the protein molecule and interact with other molecules. A combination of secondary structures such as the helix-turn-helix, which consists of two alpha helices separated by a turn, is also observed and these are known as super-secondary structures or motifs.

Amino acids located far apart on the polypeptide chain interact with each other by means of hydrogen bonds, electrostatic interactions, disulphide bridges etc., allowing the protein to fold three dimensionally in space, giving rise to the tertiary structure. Folding takes place such that the hydrophobic residues are buried inside the structure while the polar residues remain in contact with the surroundings.

3.3. Tertiary structure

The tertiary structure of protein is defined by interactions between amino acid side chains that are located far apart in the polypeptide sequence and result into formation of a three dimensional arrangement of amino acids.

Illustration: Tertiary structure

The tertiary structure of myoglobin, determined by John Kendrew, clearly revealed that the nature of amino acid side chains dictate their location in the tertiary structure. Hydrophobic residues are found buried inside the structure while the polar amino acids are found on the surface. 70% of the main chain of myoglobin is folded into α-helices with the rest being present in the form of turns & loops, which are essential to give it a compact structure.

3.4. Quaternary structure

Many proteins have more than one polypeptide chain, also called as “subunit”, that are assembled together by various interactions like electrostatic, van der Waals, disulphide bonds etc. giving rise to the quaternary structure.

Illustration: Quaternary structure

Different subunits or polypeptide chains interact with one another and are held together by means of ionic, electrostatic, van der Waals etc interactions. Such multi-subunit proteins are said to have a quaternary structure, the final level of protein structure.

Illustration: Structure of hemoglobin

Hemoglobin is a hetero-tetramer composed of two alpha and two beta chains. The alpha globin gene locus resides on chromosome 16 with each gene contributing to the synthesis of the alpha globin chain. The beta globin gene locus resides on chromosome 11 and consists of all genes that are expressed from the time of embryonic development of Hb to that of the adult beta globin gene. The globin chains are synthesized by ribosomes in the cytosol. Every subunit of hemoglobin is bound to a prosthetic group known as heme. This consists of a central iron atom in its ferrous state surrounded by a complex organic ring structure known as protoporphyrin. The heme group is essential for the oxygen binding property of hemoglobin. The iron atom forms six coordination bonds, four of which are to the nitrogen atom of porphyrin rings, one to a His side chain in the globin subunit and the other being the binding site for oxygen. Iron in its Fe^{3+} state does not bind to oxygen.

Table 1: An Overview of Different Structural Level of Proteins

Primary Structure	Secondary Structure	Tertiary Structure	Quaternary Structure
A stretch of amino acids joined together by amide or peptide bonds to form a linear polymer is known as primary structure of the protein. The amino acid sequence can be determined by techniques like Edman Degradation or Mass Spectrometry. The first primary structure that was deduced by Fredrick Sanger was that of insulin.	The secondary structure of proteins is defined by folding of a polypeptide backbone with the help of hydrogen bonds between proximal amino acids and thereby giving rise to a regular arrangement. *Alpha-helices* and *beta-sheets* are the most commonly observed secondary structures of proteins. The amino acid proline is known as the “helix breaker” as it tends to disrupt the helix and is often found at a bend in the structure known as reverse turns or bends.	The tertiary structure of protein is defined by interactions between amino acid side chains that are located far apart in the polypeptide sequence and result into formation of a three dimensional arrangement of the amino acids. The hydrophobic residues compose the core of the protein, whereas, the hydrophilic ones interact with the polar surroundings. The interactions that hold this 3D structure include hydrophobic bonds, electrostatic interactions, hydrogen bonds, etc.	Many proteins have more than one polypeptide chain, also called as subunit, that are assembled together by various interactions like electrostatic, van der Waals, disulphide bonds etc. giving rise to the quaternary structure. The individual subunits of a quaternary structure when present in a peculiar conformation render the protein functional.

4. PROTEIN FOLDING & MISFOLDING

Anfinsen's experiment in 1970 elucidated that amino acid sequence dictates 3D structure. A protein tertiary structure is attained by protein folding which ensures that hydrophobic residues are buried inside the core and the hydrophilic residues are exposed out to interact with the aqueous polar environment. A protein when folded completely and has achieved its native conformation allows it to carry out its intricate biological functions. Improper protein aggregation and misfolding is corrected by the molecular chaperones (heat shock proteins in eukaryotes), which facilitates the process of folding in case of large proteins. While the protein changes from unfolded (high-energy state) to native and folded (low-energy state) form, it passes through a higher free energy transition state. Protein folding process is accompanied by a concomitant increase in entropy of the components and also relies on reducing the number of hydrophobic side chains that are exposed to the outer aqueous environment; which actually acts as a driving force for the process. There is a distinct conformation of the protein that exists in the native state, whereas in the unfolded state, the protein can take up one of the many conformations. The folded proteins are further stabilized by intramolecular hydrogen bonding, the strength of which is highly influenced by the external environment. Protein folding is considered to be significant not only with respect to generating functionally active proteins, but also with respect to other functions like cellular trafficking and control and regulation of cell cycle and cell growth.

The dire consequences of protein misfolding can turn out to be harmful for the cells and the organism in question. Protein misfolding mainly occurs because of mutations or

inappropriate covalent modifications, which in turn lead to the formation of a 3D structure; which is different than the actual native conformation of the protein. The misfolded protein; however, may serve to provide modified functions or in majority of the cases turn out to have toxic after effects. The accumulation of these misfolded proteins or peptide fragments leads to degenerative diseases and is characterized by formation of insoluble protein plaques in liver or brain.

Illustration: Thermodynamics of protein folding

An unfolded polypeptide chain has very high free energy and entropy. Protein folding acts to decrease the free energy of the system by forming favorable interactions and assuming a more stable state. The entropy of the polypeptide chain decreases during this process. As the protein continues to fold in order to assume its stable, low energy native state conformation, the entropy also decreases. While this would seem unfavorable for the system, it must be recalled that the entropy of the surrounding water molecules increases during the process, thereby increasing the overall entropy and making it favorable and spontaneous.

Illustration: Amino acid structure determines 3-D folding

The process of protein folding is governed by the distribution of polar and non-polar amino acid residues in the protein. Hydrophobic amino acids are driven to interact with one another, a process termed as "hydrophobic collapse". They come together and during the process, eliminate water molecules around them. The polar residues remain on the surface and form hydrogen bonds with water molecules while the hydrophobic residues get buried within the core of the protein.

Proteins typically adopt only one characteristic functional native state conformation, which has lowest free energy and is most stable. Folding is limited to one conformation due to properties of the amino acid side chains such as hydrophobicity, size, shape etc. Folding is a highly cooperative process wherein there is progressive stabilization of the intermediates. Although it is theoretically possible to predict protein structure from the amino acid sequence, several long-range interactions often limit these predictions.

Illustration: Molecular chaperones for protein folding

The unfolded protein is bound by DnaJ and then by DnaK, which is an ATP bound protein. The hydrolysis of ATP into ADP and Pi by DnaK is stimulated by DnaJ. The resulting DnaK-ADP remains tightly bound to the unfolded protein. The nucleotide exchange factor GrpE present in bacteria facilitates release of ADP along with DnaJ. This leaves the DnaK bound to the partially folded protein that continues to undergo folding to a more favorable low energy conformation. Once the protein gets completely folded, it gets detached from DnaK which then binds to ATP again, thereby completing the cycle and preparing it for the next round of protein folding. Any protein that may not have been folded completely is then taken over by the GroEL chaperonin system to complete the folding.

Table 2: Diseases caused due to misfolding of proteins

NAME OF DISEASE	PROTEIN AFFECTED
Alzheimer's disease	Misfolded beta amyloid proteins
Cystic Fibrosis	Misfolded CFTR protein
Gaucher's disease	Misfolded beta glucocerebrosidase
Fabry disease	Misfolded alpha galactosidase

Illustration: Protein misfolding diseases

1. Alzheimer's Disease:

- *Structure of certain normal soluble cellular proteins normally rich in alpha helical regions converted into beta strand conformations which further link with each other to form beta sheet aggregates known as "amyloids".*
- *Insoluble amyloid plaques are essentially made up of a single polypeptide chain or fibrils known as "amyloid-b-protein (Ab)".*
- *Observed in the brain of patients with Alzheimer's where dead or dying neurons surround plaques.*
- *Neurotoxicity believed to be caused by the Ab fibrils before they get deposited as amyloid plaques.*
- *The disease presents various symptoms such as memory loss, decreased neuromuscular coordination, confusion and dementia.*

2. Huntington's disease:

- *Neurodegenerative disorder of genetic origin affecting muscular coordination.*
- *Caused by increased number of trinucleotide repeats, CAG, in Huntingtin gene leading to increased number of glutamine residues incorporated in corresponding protein.*
- *This alters the folding of the Huntington protein, which has highest concentration in brain and testes.*
- *Exact function of the protein is unclear but is known to interact with several other proteins.*
- *Mutated protein has also been found to have effects on chaperone proteins, which in turn help in folding several other proteins.*
- *Prominently affects basal ganglia, which plays a key role in movement and behavioural control.*

3. Creutzfeldt–Jakob disease:

- *Initially believed to be caused by viruses or bacteria.*
- *Later discovered to be transmitted by small proteins known as "prions".*
- *Prion proteins composed of beta sheet structures that have been modified from previously existing alpha helices.*

- *Protein aggregates of one abnormal protein sufficient to function as nuclei for other normal proteins to attach themselves to.*
- *Characterized by muscular spasms, loss of muscle control and memory loss.*

4. Cystic fibrosis:

- *Autosomal recessive disorder caused by a mutation in gene for the protein cystic fibrosis transmembrane conductance regulator (CFTR).*
- *CFTR regulates components of sweat, digestive juices and mucus.*
- *Caused by a deletion of three nucleotides leading to the elimination of a phenylalanine residue from the protein and therefore abnormal folding.*
- *Dysfunctional protein gets degraded by the cell.*
- *Disorder can affect several body parts such as the lungs, GI tract and reproductive organs.*

5. Pulmonary emphysema:

- *Progressive disease of the lung causing shortness of breath.*
- *Can be caused by deficiency of the protein "alpha-1-antitrypsin (A1AT)".*
- *A1AT is responsible for protecting the lung tissues from damage by enzyme neutrophil elastase.*
- *Abnormally secreted A1AT gets accumulated in the liver thereby allowing lung tissue damage.*
- *Causes wheezing, shortness of breath, asthma-like symptoms and also liver cirrhosis.*

6. Lathyrism:

- *Regular ingestion of seeds from sweet pea (Lathyrus odoratus) causes disruption of cross-linking in the muscle protein, collagen.*
- *Collagen is an important structural protein having a triple helical structure.*
- *Cross-links formed are due to the oxidation of lysine residues by the enzyme lysyl oxidase to form allysine.*
- *These are essential for proper folding of collagen, giving it the required strength.*
- *β-aminopropionitrile, present in abundance in sweet pea, deactivates this enzyme by binding to its active site*
- *This prevents cross-linking and proper folding of the protein.*
- *Causes muscle fragility and weakness.*

V. CONCLUSION

Amino acids play significant role as building blocks of polypeptides in organisms. The amino acid composition of a given protein is influential in determining the characteristic features of the corresponding protein in question. The amino acid sequence or the primary structure of proteins in other words is responsible for determining the 3D structure of proteins. Proteins do not attain a functionally active state until they have undergone folding to reach a peculiar 3D structure. The different levels of structural organization of proteins when understood vividly can be helpful to extrapolate their functional significance to the cell or organism. Protein folding is a thermodynamically favorable process and occurs until a native, stable conformation is reached. Protein misfolding can be a consequence of mutations or wrong covalent interactions, thereby leading to serious downstream effects on the cell. Therefore, the role of individual amino acids, their special features, their peculiar arrangement to form a polypeptide; the peptide bond geometry; hierarchy of protein structural arrangement and protein folding and misfolding; are all pivotal facets in proteomic studies.

5. REFERENCES

1. A surprising simplicity to protein folding. David Baker. Nature, 2007.

2. From folding theories to folding Proteins: a review and assessment of Simulation studies of protein folding and unfolding. Joan-Emma Shea, Charles L Brooks III. Annu. Rev. Phys. Chem. 2001.

3. Protein folding and misfolding: a paradigm of self-assembly and regulation in complex biological systems .Michele Vendruscolo et al. The Royal Society. 2003.