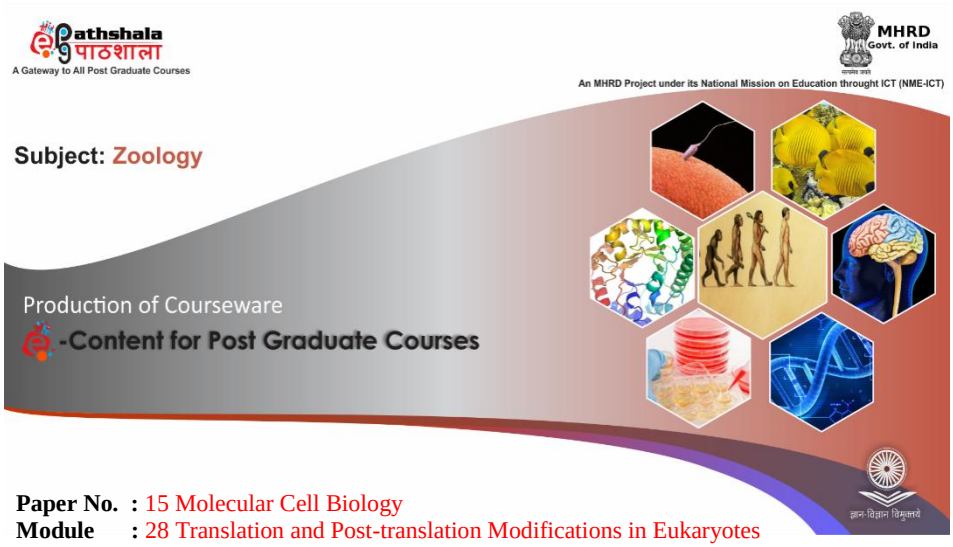


Paper No. : 15 Molecular Cell Biology
Module : 28 Translation and Post-translation Modifications in Eukaryotes

Development Team

Principal Investigator	:	Prof. Neeta Sehgal Department of Zoology, University of Delhi
Co-Principal Investigator	:	Prof. D.K. Singh Department of Zoology, University of Delhi
Paper Coordinator	:	Prof. Kuldeep K. Sharma Department of Zoology, University of Jammu
Content Writer	:	Dr. Renu Solanki, Deen Dayal Upadhyaya College Dr. Sudhida Gautam, Hansraj College, University of Delhi Mr. Kiran K. Salam, Hindu College, University of Delhi
Content Reviewer	:	Prof. Rup Lal Department of Zoology, University of Delhi

ZOOLOGY

Molecular Genetics

Translation and Post-translation Modifications in Eukaryotes

Description of Module	
Subject Name	ZOOLOGY
Paper Name	Molecular Cell Biology; Zool 015
Module Name/Title	Cell regulatory mechanisms
Module Id	M28: Translation and Post-translation Modifications in Eukaryotes
Keywords	Genome, Proteome diversity, post-translational modifications, glycosylation, phosphorylation, methylation

Contents

1. Learning Objectives

- Eukaryotic post-translational events
- Purpose of post translational modifications.
- Why the proteome diversity is much higher than that of the genome.
- How the Post-transcriptional modifications help in maintaining the biological activities within a cell.
- Various diseases are studied by identifying the PTM's responsible to the malfunctioning of the cell.

2. Introduction

Genome has a large number of genes and their expressions are highly regulated with respect to the need of the cell and its surroundings. Transcription and translation are well coordinated processes involved in the transmission of the genetic information from DNA to RNA to proteins and the regulation occurs at each stage. Genetic information in the mRNA is converted to polypeptides through the process of translation by making use of ribosomal machinery. It consists of three phases namely:

1. **Initiation-** Process in which the ribosome binds to the mRNA and initiates polypeptide synthesis is called Initiation. This binding of ribosome and mRNA takes place at the initiator/start codon (AUG) which codes for amino acid methionine.

2. **Elongation -** When the polypeptide is lengthened by adding one amino acid at a time using mRNA as a template and tRNA brings the next nucleotide as coded by the template, it is termed as Elongation Process.

3. **Termination** Synthesis of the polypeptide is curbed on encountering the stop codon (UAG, UAA, UGA) and the entire machinery dissociates (Fig. 1). This is called as Termination. However, the polypeptides produced from translation are not ready to function as discrete protein molecules. This nascent polypeptide needs to be further processed

(folded/modeled) to achieve the desired structure for proper functioning of the protein. These modifications of the polypeptides could take place while they are being still synthesized (co-translational modifications) or once their synthesis is complete (post-translational modifications (PTM)).

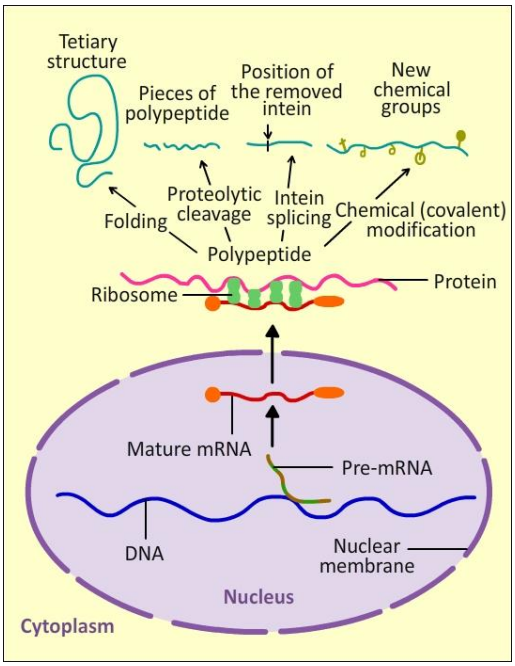


Fig. 1 : Eukaryotic transcription and post-translational modifications (**Source:** Author and Departmental artist)

The newly synthesized proteins are non-functional and require activation under proper physiological conditions by appropriate proteases. These inactivated precursor proteins which can be activated either by peptide cleavage (eg., signal peptide) or addition of certain groups (e.g. Phosphate, methyl etc.) are known as pre-proproteins (Fig. 2).

Certain polypeptides have an inbuilt ability to fold and achieve the required complete, mature conformation without the help of any molecule. However, most polypeptides require chaperones to help them fold properly. Sometimes the modification occurs after the tertiary structure is complete to activate or inactivate the catalytic property influenced by the biological activity of the cell. Misfolded proteins are often tagged so they can be degraded. Also, more than one modification is observed in some proteins by adding various functional groups in a sequential manner to accomplish protein maturation/activation.

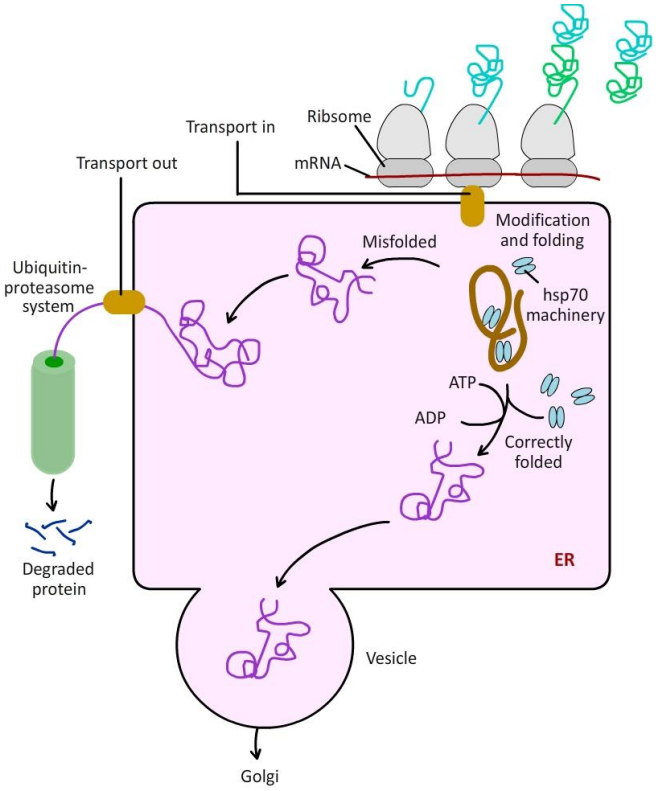


Fig. 2 : Post-translational modifications can give correctly folded or incorrectly folded protein
Source: Author and Departmental artist

3. Purpose of post translational modifications

Post-translational modifications (PTMs) are the various covalent and non-covalent interactions of the polypeptide. These modifications help in converting the polypeptide into a protein molecule having all the functional properties and the required macromolecular 3 dimensional structure. Proteomic diversity is accomplished because of PTMs. PTMs help in protein stabilization, biochemical activity regulation, protein targeting (localization through signal peptide) or protein signaling through protein-protein interaction or cascade amplification.

Human genome is expected to have around 20-25 thousand genes but the proteomic diversity is estimated to be around 1 million proteins. The abundance in the transcriptome relative to the size of the genome is a result of the changes which occur at the transcriptional and mRNA levels. PTM results in an exponential increase in the proteome complexity relatively much higher than the transcriptome and genome (Fig. 3).

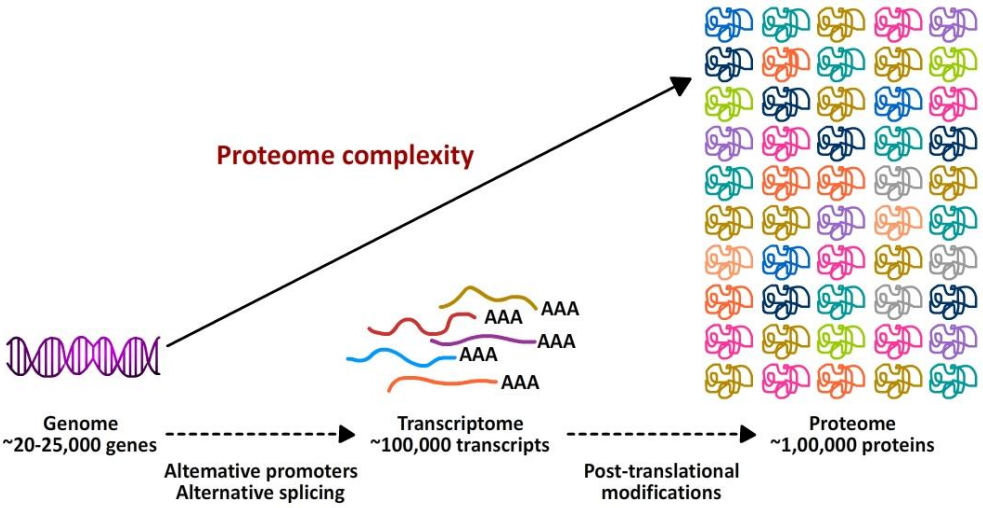


Fig. 3 : Eukaryotic transcription and post-translational modifications **(Source:** Author and Departmental artist)

PTMs are a result of the various cellular activities within a cell in response to an environmental stimuli and leads to the dynamic human proteome. PTM's are very specific and occur at a distinct amino acid side chains, peptide linkages often mediated by kinases, phosphatases, transferases and ligases; which themselves are classified on the basis of their ability to add, remove or transfer a protein, lipid, sugar, phosphate group etc. by the cleavage of specific regulatory units. Even autocatalytic domains have been identified in certain proteins which can modify themselves and are known as autokinase autoprotolytic domains. Reversible PTM have been observed in kinases (enzymes which phosphorylate proteins at specific amino acid side chain for catalytic activation or inactivation) and phosphatase (enzymes which remove a phosphate group from a protein and reverse its biological activity). PTM's are helpful to analysis a specific protein and its role in various diseases like cancer, diabetes, neurodegenerative disorders and heart disease.

4. Post translational modifications

An integral part of the post transcriptional mechanism within a cell is directing a newly synthesized protein molecule to their proper destination by tagging the proteins and the tag is known as the "Signal Peptide". Discovery of signal peptides is attributed to G. Blobel for which he was even awarded the 1999 Nobel Prize in Physiology or Medicine (Fig. 4).

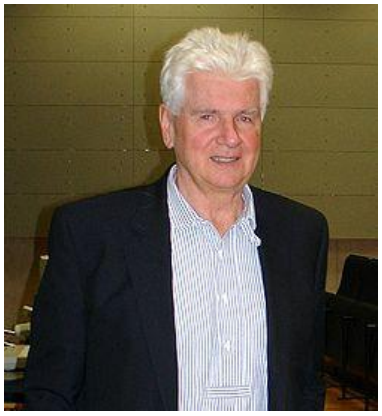


Fig. 4 : G. Blobel: Recipient of 1999 Nobel Prize in Physiology or Medicine
Source: https://en.wikipedia.org/wiki/G%C3%BCnter_Blobel)

Membrane bound proteins, excretory proteins and glycoproteins are synthesized by the ribosomes which are associated with the rough endoplasmic reticulum (ER). A signal peptide/sequence directs a nascent polypeptide to the rough endoplasmic reticulum. Only the proteins that have an N-terminal signal sequence (signal sequence/peptide) can enter the ER. A signal recognition protein (SRP) binds to the signal sequence as it first exits the ribosome and stops the translation process. The SRP-bound ribosome (entire translation complex) attaches to the SRP receptor in the ER (Fig. 5). The SRP receptor is a heterodimeric unit having an alpha (α) and beta (β) subunit. SRP has a translocation channel through which the emerging polypeptide is extruded into the ER lumen. This translocation channel is referred to as the translocon. The signal peptide is cleaved by serine proteases family known as signal peptidases and released in the lumen of ER. Pre-proteins often have a signal peptide with them which is lacking in the proproteins. However, some proteins that are destined for secretion are also further proteolyzed following secretion and are termed Pre-proproteins.

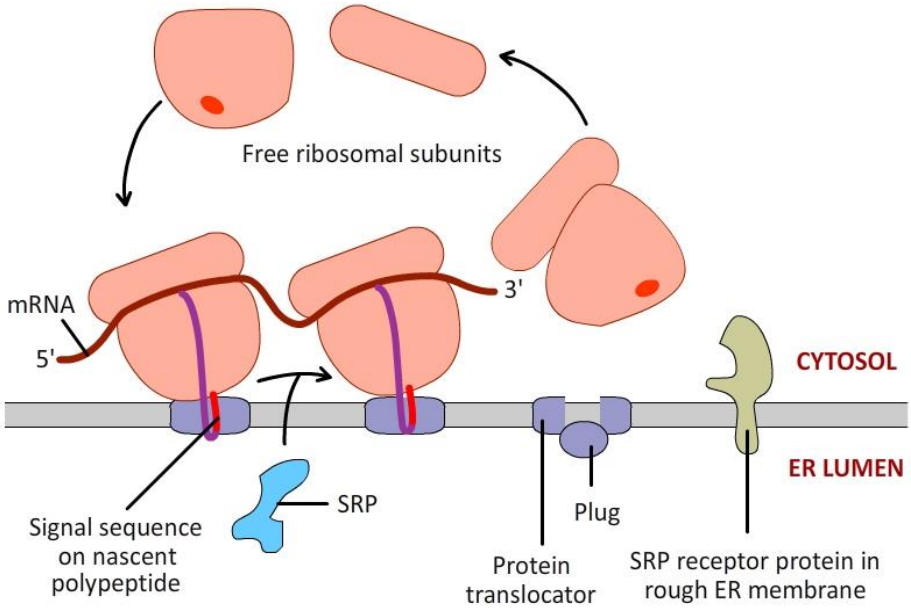


Fig. 5 : Mechanism **of synthesis of membrane bound, secreted, and glycoproteins.** Ribosomes bind with the ER membrane through interaction of the signal recognition particle, SRP in the ribosome with the SRP receptor in the ER membrane. As the protein is synthesized the signal sequence is passed through the ER membrane into the lumen of the ER. After sufficient synthesis the signal peptide is removed by the action of signal peptidase. Synthesis will continue and if the protein is secreted it will end up completely in the lumen of the ER. If the protein is membrane associated a stop transfer motif in the protein will stop the transfer of the protein through the ER membrane. This will become the membrane spanning domain of the protein.
Source: Author and Departmental Artist

PTM's are of various types depending upon the groups which are being added or substituted within the polypeptide chain for the protein activation. Table 1 tells us about the most common amino acid site where the modification occurs, effects of the post-translational modification and whether the modification is of reversible or irreversible type.

Type of protein Modification	Most common amino acid site	Affects	Reversible/ Irreversible
Phosphorylation	Tyrosine, serine, Threonine	Cell cycle, growth, apoptosis, differentiation	Reversible
Acetylation	Methionine	Protects the N-terminus (protein stability) and regulates protein-DNA interaction	Reversible
Amidation	C-terminus of amino acids	Bioactivates the polypeptides	Reversible
Sumolytion	Addition of ubiquitin	Regulation of transcription and programmed cell death	Reversible
Glycosylation	Asparagine-linked (N-linked) or serine/	Intracellular protein movements and helps the	Reversible

	threonine-linked (O-linked) oligosaccharides are major structural components of many cell surface and secreted proteins.	protein molecules to reach the exact cellular destination.	
Methylation	the ε-amine of the R-group of lysine residues and the guanidino moiety of the R-group of arginine, imidazole ring of histidine and the R-group amides of glutamate and aspartate	Chromatin remodelling	Reversible (O-methylation), Irreversible (N-methylation)
Ubiquitination	Proteins are tagged with ubiquitin	Misfolded protein degradation	Reversible
S-nitrosylation	thiol group of cysteine residues	Protein stability and gene regulation.	Reversible
Lipidation	cysteine residues	Protein localization, signal targeting, membrane tethering and acts as a mediator of protein-protein interaction	Reversible

Table 1.: Post-transcriptional modifications: Most common amino acid site, its effect and property of reversibility.

4.1. Phosphorylation, the addition of a phosphate group

Phosphorylation helps in regulating the biological activity of enzymes and proteins. It is one of the most common reversible PTM in animal cells. Specific kinases are responsible for adding or removing the phosphate group. Physiological relevance can be cited in glycogen metabolism in the hepatocytes. Phosphorylation inhibits the activity of glycogen synthase (enzyme involved in the glycogen synthesis) and increases the activity of glycogen phosphorylase (enzymes responsible for the breakdown of glycogen in the tissues) in response to the hormone glucagon secreted from the pancreatic cells. This is helpful in regulating the blood-glucose level of the organism. Kinases add the phosphate group whereas phosphatases remove the phosphate group. The reaction catalysed by protein kinases is given below:

ATP + protein → phosphoprotein + ADP

Common amino acids subject to phosphorylation are serine, threonine and tyrosine in the animal cells. For eg., tyrosine phosphorylation affects the activity of numerous growth factors.

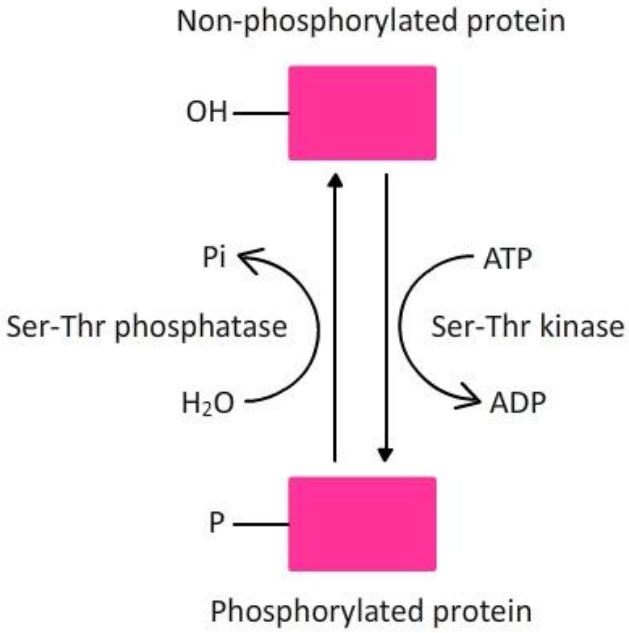


Fig. 6 : Regulation of enzyme activity by addition and of removal of the phosphate group
(**Source:** Departmental Artist and Author)

4.2. Methylation, the addition of a methyl group

Post-translational methylation occurs in the nitrogen or oxygen molecule of the proteins and namely N-methylation or O-methylation, respectively. Activated S-adenosylmethionine (SAM) is the methyl donor. Commonly methylations are observed on the ε-amine of the R-group of lysine residues and the guanidino moiety of the R-group of arginine. Chromatin remodelling is necessary for the transcriptional activity and is regulated by the methylation of lysine residues in histone proteins present in the nucleosome. Activity of numerous transcription factors are modified due to methylation. Humans express 27 lysine (K) methyltransferases (identified as KMT family enzymes) and nine arginine methyltransferases. Methylation of the oxygen of the R-group carboxylates of glutamate and aspartate also takes

place and forms methyl esters. Proteins can also be methylated on the thiol R-group of cysteine.

4.3. Glycosylation, the addition of sugar groups

Glycosylation involves a covalent bonding between the secretory or membrane bound protein and carbohydrate moieties (These moieties are mostly oligosaccharides known as glycans). The sugar group may ultimately be important to the protein function or it may simply act as an address label required to get the protein to its next cellular destination. Importance of this PTM can be emphasized as more than 50% of the proteins are glycosylated. It occurs either in the ER or Golgi body. Asparagine-linked (N-linked) or serine/threonine-linked (O-linked) oligosaccharides are major structural components of many cell surface and secreted proteins. The predominant sugars found in glycoproteins are glucose (Glc), galactose (Gal), mannose (Man), fucose (Fuc), *N*-acetylgalactosamine (GalNAc), *N*-acetylglucosamine (GlcNAc) and *N*-acetylneuraminic acid (NANA or Sialic acid (Sia)).

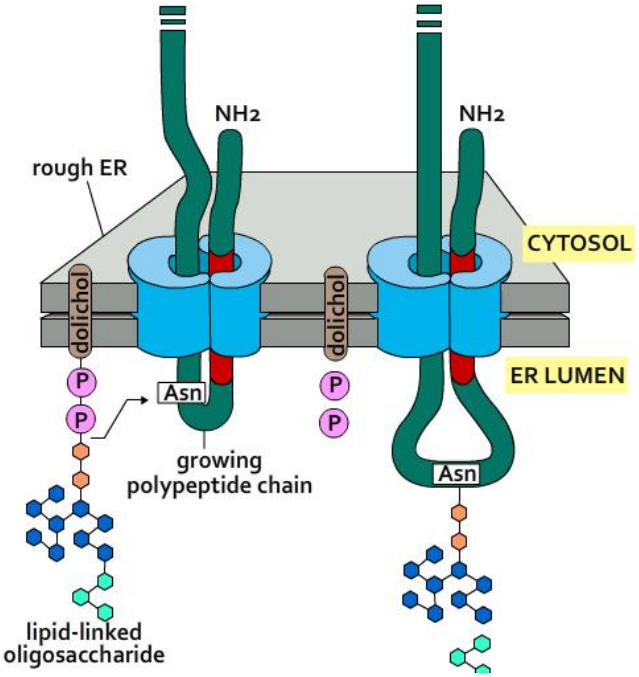
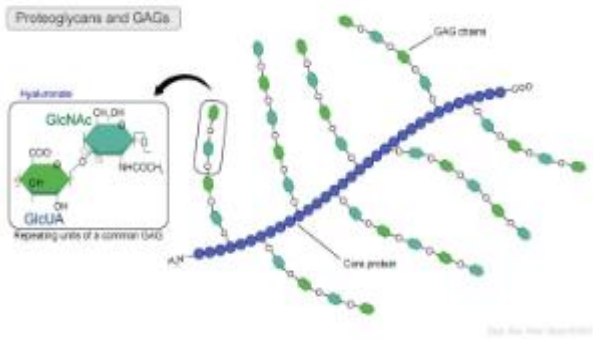
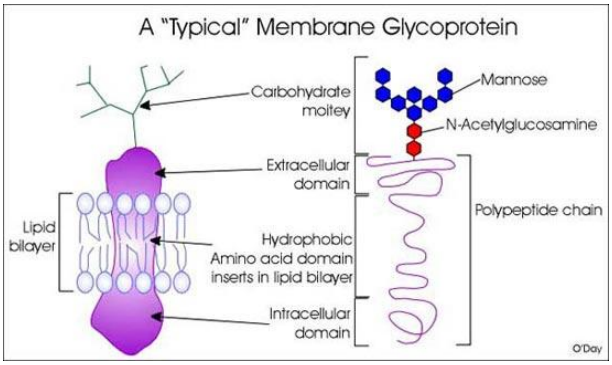


Fig.7: Glycosylation: Attachment of the carbohydrate moiety to the growing polypeptide chain
Source: Author and Departmental Artist

The distinction between proteoglycans and glycoproteins resides in the level and types of carbohydrate modification which are given below in the table-2;

Proteoglycans	Glycoproteins
Proteoglycans also contain the sugar glucuronic acid (GlcA). 	Glycoprotein lacks the sugar glucuronic acid (GlcA). 
The carbohydrate modifications found in glycoproteins are rarely as complex as that of proteoglycans.	The carbohydrate modifications are less complex than that of proteoglycans. The carbohydrates of glycoproteins are linked to the protein component through either *O*-glycosidic or *N*-glycosidic bonds. The *N*-glycosidic linkage is through the amide group of asparagine (Asn, N). The *O*-glycosidic linkage is to the hydroxyl of serine (Ser, S), threonine (Thr, T) or hydroxylysine (hLys). The linkage of carbohydrate to hLys is generally found only in the collagens. The linkage of carbohydrate to hLys is either the single sugar galactose or the disaccharide glucosyl galactose.

Protein sorting, immune recognition, receptor binding, inflammation, and pathogenicity are certain crucial roles in cellular processes which involve glycoproteins. Heterogeneity in the glycoprotein (mass and charge) results due to the diversity of the attached oligosaccharides. Glycosylation helps to maintain the tertiary and quaternary structure of the proteins. The unique structure of the Fab molecule in Immunoglobulin G utilizes as many as 30 glycoprotein interactions. N- glycosylation is the attachment of the glycan towards the nitrogen atom of the protein and in O-glycosylation the glycan attaches to the oxygen atom. The ABO blood groups variability is due to different terminal glycoproteins (Fig. 8). Blood

group O lacks the glycan whereas blood group A and B have N-acetylglucosamine and galactose, respectively.

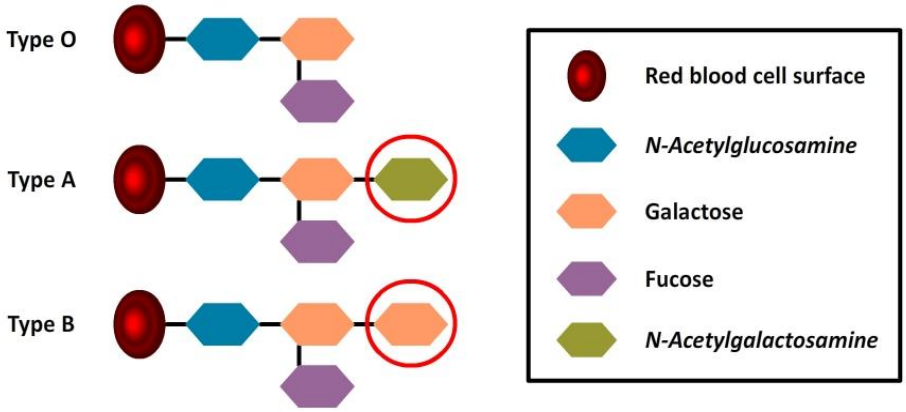


Fig. 8 : Different types of blood groups with its own unique glycoprotein
(Source: Author and Departmental Artist)

4.4. Disulfide bonds, the formation of covalent bonds between 2 cysteine amino acids.

Disulphide bonds (–S–S–) provide stability and hence the functionality to the tertiary / quantenary structure of the protein and involves appropriate folding of the polypeptide chains. These folding are of three types:

a) Covalent modifications of the side chains in the polypeptides
b) Hydrolytic cleavage or isomerization of the peptide bonds
c) Reductive cleavage of the disulphide bond

These bonding help in the stabilization of mature proteins and remain unchanged throughout the proteins life span. These covalent bonds are formed by the oxidative linkage of sulfhydryl groups (–SH), also known as *thiol* groups, on two cysteine residues in the same or different polypeptide chains. However, a functional disulphide bond is of two types: catalytic and allosteric bond. The catalytic bonds help in the enzymatic reactions. The allosteric disulphide bonds may be due to a ligand binding, substrate hydrolysis, proteolysis or oligomeric formation by certain oxidoreductases or by thiol/disulfide exchange. Hence, they control the function of mature protein in which they are present. Protein flexibility is dependent on the

allosteric control, in which a change at the allosteric site influences another site responsible for the enzymatic activity.

Disulphide bonds are present on the secretory protein and the exoplasmic domain of the membrane protein which are synthesized on the ER. The soluble cytosolic proteins which are synthesized on free ribosomes, lack disulphide bonds and depend on the other interactions to have a stable structure. Thus, in eukaryotes disulphide bonds are present in the ER but absent from cytosol. Localization of these bonds within the ER lumen indicates that ER has a favorable redox environment for oxidation of –SH group, which is absent in the cytosol. It has been observed in the yeast cells that a mutation in the ER membrane protein leads to an inability, in the production of disulphide bonds, suggesting the participation of that particular protein in the oxidation of –SH groups in the ER lumen. Glutathione, a major thiol-containing molecule in eukaryotic cells serves to prevent the formation of disulphide bonds in the cytosol (Fig. 9). Glutathione shuttles between the reduced form (GSH) and the oxidized form, a disulfide-linked dimer (GSSG). The GSH: GSSG ratio is over 50:1 in the cytosol; oxidized GSSG in the cytosol is reduced by the enzyme glutathione reductase, using electrons from the potent reducing agent NADPH:

$$NADPH + H^+ + GSSG \longrightarrow NADP^+ + 2\ GSH$$

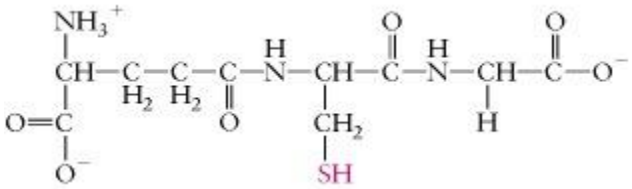


Fig. 9: Tripeptide Glutathione (**Source:** Author and Departmental Artist)

Thus cytosolic proteins in bacterial and eukaryotic cells do not utilize the disulfide bond as a stabilizing force because of the high GSH: GSSG ratio, which would drive the system in the direction of Cys–SH and away from Cys–S–S–Cys.

Proper pairing of the disulphide bond is essential for normal structure and function of the protein. During the synthesis of the Immunoglobulin (Ig) disulphide bonds are formed while the polypeptide is still growing on the ribosome. As a result the first and second cysteine

closest to the N-terminus forms a bond even before the third cysteine is added to the growing polypeptide, automatically ensuring correct pairing of the cysteine molecules. Accordingly third cysteine pairs with the fourth to create the second disulphide bond.

Proinsulin has three disulphide bonds and the first bond formed due to spontaneous oxidation of –SH groups undergo rearrangement to get proper conformation of the protein. Protein Disulphide Isomerase (PDI) is found abundantly in the ER of secretory tissues like liver and pancreas and help in the rearrangement of the disulphide bonds (Fig. 10). The sole purpose of this enzyme is to help the proteins to reach their thermodynamically most stable conformation.

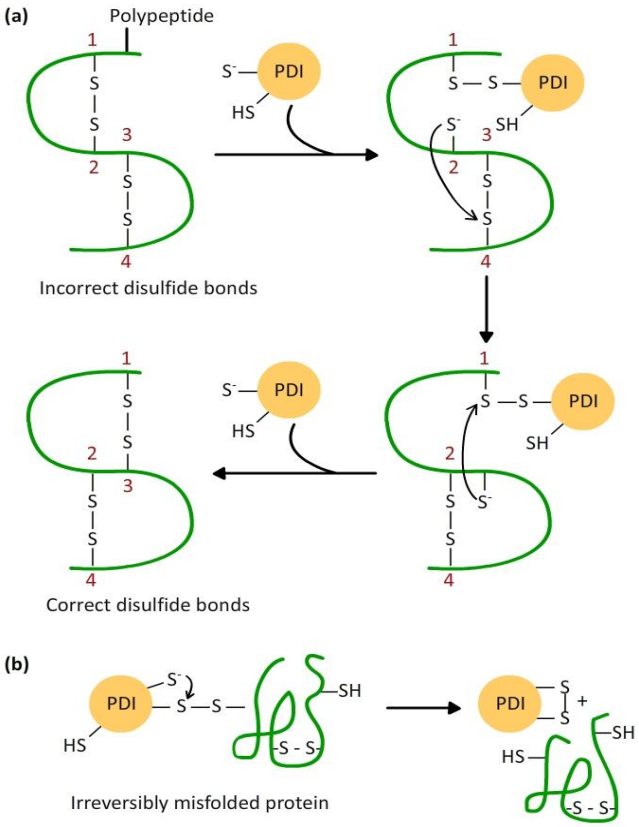


Fig. 10: Rearrangement of the protein disulphide by Protein Disulphide Isomerase
PDI helps in correcting the disulphide bonds; (b) PDI assists in the proper folding of the misfolded protein
Source: Author and Departmental Artist

4.5. Proteolysis/ Proteolytic Cleavage

Cleavage of the nascent polypeptide chain to remove a particular amino acid sequence or a segment of the nascent polypeptide to achieve activation is known as Proteolytic cleavage or Proteolysis. Insulin is a brilliant example which is secreted from the pancreas as a pre-peptide. It undergoes a cleavage at the 24 amino acid signal peptide to yield a protein which folds resulting in the formation of proinsulin. This proinsulin is further cleaved to give active insulin having two peptide chains joined together by disulfide bonds (Fig. 11). Proteins often undergo a cleavage after translation so as to remove the initiator methionine. Some proteins are synthesized as inactive precursors known as Zymogens, which are activated by proteolytic cleavage as observed in the proteins of blood clotting cascade.

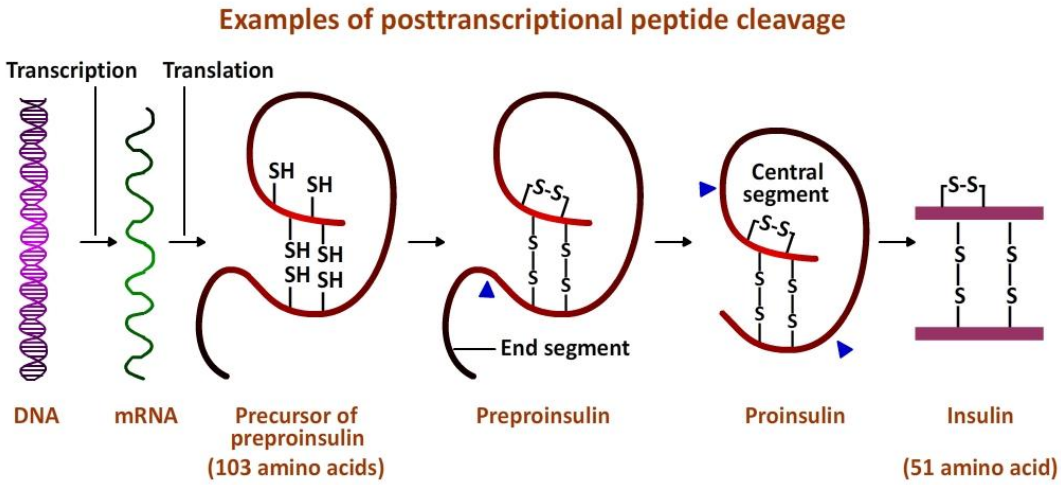


Fig. 11 : Conversion of preproinsulin to insulin (**Source:** Author and Departmental Artist)

4.6. Subunit binding to form a multisubunit protein

Proteins made up of more than one polypeptide are known as the multimeric protein, which are usually assembled in the ER to achieve the functional quaternary structure. Immunoglobulins are made up of two subunits, i.e., a light chain (L) and a heavy chain (H), all linked by S-S bonds (Fig. 12).

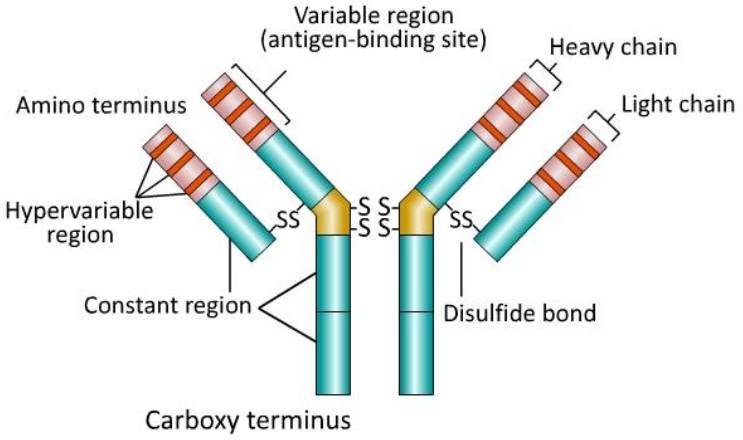


Fig. 12: Structure of Immunoglobulin showing the heavy chain (H) and light chain (L) linked by disulphide bonds (**Source:** Author and Departmental Artist)

Protein folding and subunit assembly is also exhibited by Hemagglutinin (HA) in the ER, a trimeric protein that forms the spike like projections on the surface of the Influenza virus particle. Within the ER of an infected host cell, each spike is formed from three copies of a precursor protein known as HA_0 which has a single membrane-spanning alpha helix. In the golgi complex each copy of HA_0 protein is cleaved to give two polypeptides, HA_1 and HA_2. Thus, each viral spike particle contains three copies of HA_1 and HA_2. The trimer is stabilized by interactions between the exoplasmic domains of the constituent polypeptides as well as by interactions between the three cytosolic and membrane-spanning domains.

4.7. S-nitrosylation

Nitric oxide groups are added to the thiol group of cysteine residues and result in the formation of S-nitrosoprotein (SNO). This is known as the S-nitrosylation. In addition to the protein stability it also regulates gene expression, provides the nitric acid donors and helps in generation, localization, activation, catabolism of S-nitrothiols (SNO's) (formed from nitric acid and free cysteine residues). These SNO's are under high regulation by caspases (enzymes that mediate apoptosis) which can also dinitrosylates it in response to the extra and intracellular cues.

4.8. Lipidation

Covalent attachment of a lipid group to the protein helps in localization within the cell, signal targeting, membrane tethering and acts as a mediator of protein-protein interaction. Eg., Palmitoylation creates a thioester link between long-chain fatty acids and cysteine residues. N-myristorlation of glycine residues which is helpful in membrane targeting and GPI-anchor responsible for linking a glycosyl- phosphatidylinositol (GPI) to an extracellular protein and mediates its attachment to the plasma membrane.

4.9. Acetylation

Acetylation takes place in almost all the eukaryotic cells and involves the transfer of an acetyl group to nitrogen atom of proteins. It helps in protein stability by protecting the N-terminus and regulating protein-DNA interactions in the case of histones and has both reversible and irreversible mechanisms. N-terminal acetylation is done by Methionine aminopeptidase (MAP) which results in the cleavage of N-terminal methionine before replacing the amino acid with an acetyl group from acetyl-coA by the enzyme N-acetyltransferase.

4.10. Ubiquitylation

Aaron Ciechanover, Avram Hershko and Irwin Rose received a Nobel Prize in 2004, for revealing the protein degradation and associated biological processes responsible for controlling the cell cycle, gene transcription and immunity in eukaryotes. Damaged proteins are required to be removed from the cell to prevent aberrant activities of the defective protein**.** Cellular proteins are destroyed within the proteasome found in the cytosol and the nucleus. Proteasome has a sedimentation coefficient of 26 S comprising a 20*S* barrel-shaped catalytic core which has 19*S* regulatory complexes at both ends. Degradation of proteins within the proteasome is an ATP dependent mechanism. Autophagy is another major protein (and other cellular constituents) degradation pathway. Autophagy involves the sequestration of targeted cytoplasmic constituents into a double membrane vesicle termed the autophagosome. Following its formation the autophagosome fuses with the lysosomal machinery of the cell and the contents are degraded.

Misfolded proteins are tagged with Ubiquitin an 8 kDa polypeptide consisting of 76 amino acid. Deubiquitylating proteins are the proteins which are not destined to be degraded but are tagged for some specific purpose of regulation and not degradation. They are not destined for the proteosome degradation. Deubiquitylating enzymes (DUBs); a family of isopeptidases performs the vital function of removing ubiquitin from proteins which are modified for some specific regulatory purposes. Recycle of ubiquitin is also carried by DUB's by cleaving the polyubiquitin chain to yield monoubiquitin. The DUB enzymes are all members either the large family of cysteine proteases or the large family of metalloproteases.

For entering the proteasome the proteins are first tagged by attachment of multimers of the 76 amino acid protein ubiquitin, a process termed ubiquitylation or ubiquitination. Degradation via 26S proteasome is helpful in cell cycle regulation, cell proliferation and differentiation, programmed cell death (apoptosis), DNA repair, immune and inflammatory processes, and organelle biogenesis. This is a two-step degradation process:

1. Ubiquitylation of the target protein involves addition of multiple units of ubiquitin in a sequential manner so that the target protein is polyubiquitinated before entering the proteasome (Fig. 13). The process is explained below in detail:

    a) First the Ubiquitin activating enzyme (E1) activates ubiquitin by utilizing an ATP so that the ubiquitin binds to E1 via a high energy thiol ester.

    b) Ubiquitin conjugating enzymes UBE or Ubiquitin carrier proteins (E2), transfers the ubiquitin via an E2 thiol ester intermediate to the substrate protein. The substrate proteins are recognizable by E2 as they are bound to Ubiquity protein ligases (E3).

    c) Ubiquitin protein ligases (E3), helps in attaching the ubiquitin to the substrate protein. Usually the ubiquitin is transfer takes place at the ε-amino group of an internal lysine residue within the substrate protein.

2. Degradation by the proteasome complex involves the release of ubiquitin monomers that can be re-used to tag additional proteins. If the proteasomes are non-functional or

deregulated it leads to many human diseases such as cancer, myeloproliferative diseases, and neurodegenerative diseases.

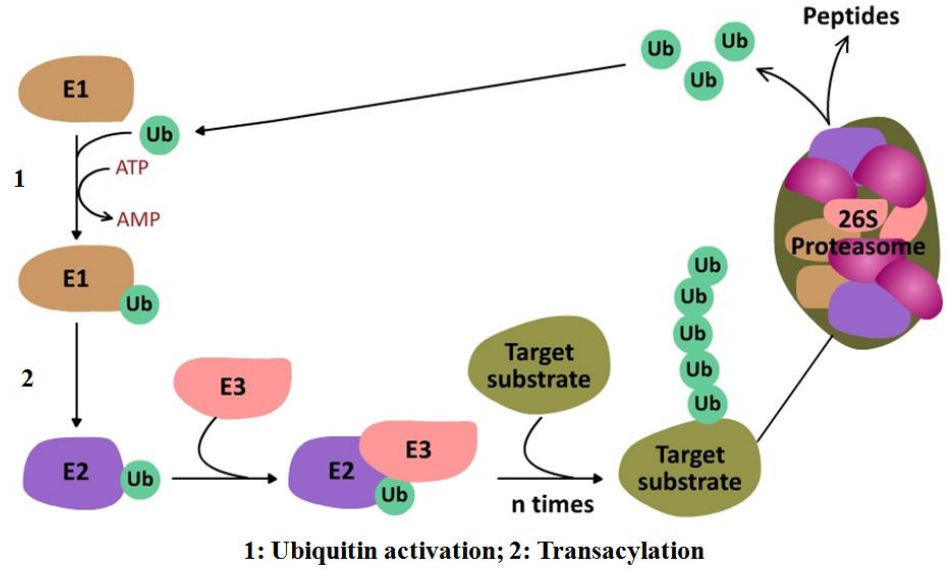


Fig.13 : Process of ubiquitylation and proteasome-mediated protein degradation.

a. Begins with the attachment of an ubiquitin monomer to a member of the ubiquitin activating enzyme (E1) family.

b. This "activated" ubiquitin is then transferred to a member of the ubiquitin-conjugating enzyme (E2) family and this process is known as transacylation.

c. This E2-ubiquitin complex is further targeted by a member of the ubiquitin ligase (E3) family of enzymes which transfers the ubiquitin to the target substrate protein.

d. Once ubiquitylation is completed the tagged protein can be degraded in the proteasome.

4.11. SUMOlytion

Small ubiquitin related modifier (SUMO) is ubiquitin like proteins but unlike ubiquitin it does not participate in the protein degradation. SUMO proteins have small ubiquitin modifier proteins attached to them and help in stabilization of proteins, cell cycle progression, and regulation of transcription and programmed cell death (apoptosis).

4.12. Vitamin C-Dependent Modifications

Vitamin C acts as a cofactor for modification of certain protein having proline and lysine undergo hydroxylation and carboxy terminal amidation. The hydroxylating enzymes are identified as prolyl hydroxylase and lysyl hydroxylase. Glycine acts as the donor of the amide for C-terminal amidation. Collagens are the most important hydroxylated proteins. C-terminal amidation is observed in several peptide hormones such as oxytocin and vasopressin.

4.13. Vitamin K-Dependent Modifications

Vitamin K is a cofactor in the carboxylation of glutamic acid residues catalyzed by the enzyme gamma-glutamyl carboxylase (γ-glutamyl carboxylase). The result of this type of reaction is the formation of a γ-carboxyglutamate (gamma-carboxyglutamate), referred to as a *gla* residue.

4.14. Selenoproteins

Proteins containing the selenium as an additional moiety (in the 21st amino acid) are known as selenoproteins. The cysteine associated with the selenium is known as the selenocysteine. Their synthesis is dependent on the dietary intake of the selenium,. Several muscular and cardiovascular, neurological disorders, immune dysfunction, cancer and endocrine functions are impaired due to the selenium deficiency, mutations or polymorphism in the selenoprotein genes. Selenoproteins also have an essential role in a variety of cell processes and diseases.

4.15. Myristoylation

Myristoylation is a PTM which occurs co-translationally and post translationally and helps to achieve higher level of proteome diversity by developing complexcity of the cellular function (Fig. 14). It involves an attachment of a 14-carbon saturated fatty acid (myristic acid), to the alpha-amino group of an N-terminal glycine residue of protein chain. It's a kind of lipid modification involving mystic acid a saturated fatty acid (14:0) 14 carbon with the systematic name of n-Tetradecanoic. The reaction (i.e. addition of the mystic acid) is catalysed by N-myristoyltransferase (NMT) which occurs in the cytoplasm of the cell. When less than 100 residues have been polymerized to the growing chain while still attached to the ribosomes it occurs co-translationally, by removing the leader methionine residue by a methionine

aminopeptidase to expose the N-terminal of a glycine. Post-translational myristoylation involves the exposure of an internal glycine by caspases in the apoptotic cascade by cleaving the pro-apoptotic proteins and occurs in approximately 80% of myristoylated proteins. Myristoylation helps in stabilization of the protein structure, protein-protein interaction, protein localization and enhancing the protein interaction with the organelles or the plasma membrane. Myristoylation is also referred to as a "molecular switch" as it not only diversifies the function of a protein molecule but also adds layers of regulation.

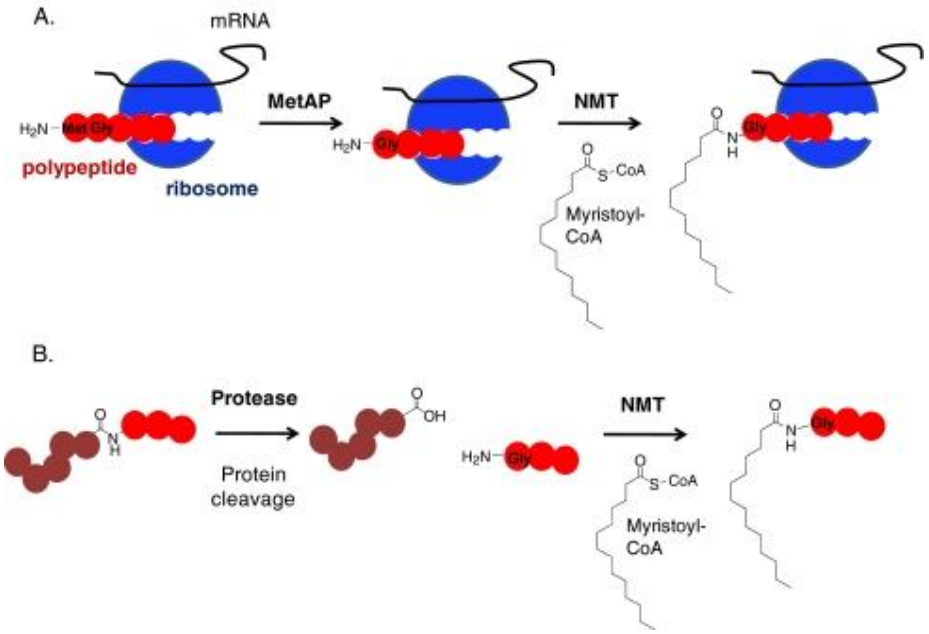


Fig. 14: Myristoylation:A. Co-translational and B. post-translational myristoylation. *MetAP* methionine aminopeptidase
Source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2816741/bin/12154_2009_32_Fig1_HTML.jpg

5. Chaperones: Role in PTM and mechanism

Chaperones are the proteins which help in the formation of covalent bonds leading to protein folding. Chaperones bind to hydrophobic regions of the polypeptide and shield them from the aqueous environment until the entire polypeptide is translated. Then the chaperones help the protein to fold into its proper shape, with the hydrophobic R groups in the interior of the protein. Heat Shock Proteins (HSP) is produced under stressful condition or when the temperature of the cell is high and they acts as Chaperones that aid in protein folding. HSP

are named according to their molecular weight eg., hsp70, hsp 90 etc. Within the ER lumen a large number of HSP's help in the folding of newly synthesized proteins. PDI is a catalyst which helps the chaperone in folding the polypeptide chain. Hsp70 is a cytosolic protein and this ER chaperone transiently binds and folds the protein to achieve a functional conformation. ER has a large number of chaperons, two homologous lectins calnexin and calreticulin, bind to specific carbohydrates which are attached to the newly synthesized proteins and aid in protein folding (Fig.15).

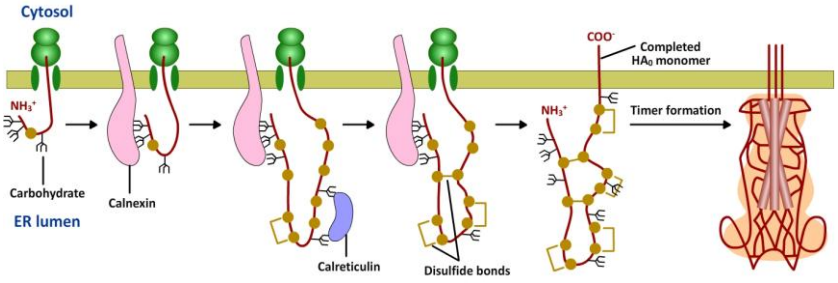


Fig.15 : Role of chaperones in the folding of polypeptides while the nascent chain is still growing, calnexin and calreticulin the protein-folding catalysts, associate with the growing chain and three disulphide bonds are formed in the globular head domain. On completion of translation, three additional disulfide bonds are formed which help in rearranging the monomer. Three HA0 chains then interact with each other, initially via their transmembrane α helices.This association apparently triggers the formation of a long stem containing one α helix (dark rod) from the luminal part of each HA0 polypeptide. Finally, interactions between the three globular heads occur, generating the mature trimeric spike. Calnexin and calreticulin bind to N-linked oligosaccharides with a single glucose residue on unfolded protein segments, thereby promoting the proper folding and assembly of newly synthesized glycoproteins such as HA.
Source: Author and Departmental Artist

IRE1 is a transmembrane multifunctional protein (green) in the inner nuclear membrane which is in continuation with the ER membrane. IPR1 has a binding site for unfolded proteins (blue) on its luminal surface; its nuclear-facing domain contains a protein kinase of unknown function and a specific RNA endonuclease (green). Binding of unfolded proteins in the ER lumen dimerizes the receptor and somehow activates the endonuclease, which cleaves the unspliced mRNA precursor encoding the transcription factor HAC1 (Fig:16). The two exons of HAC1 mRNA then are linked together by tRNA ligase, which usually splices tRNA precursors, forming a functional HAC1 mRNA. Following its synthesis in the cytosol, HAC1 protein moves back into the nucleus and activates the transcription of genes, encoding several chaperones and other proteins that assist in folding unfolded proteins in the ER lumen.

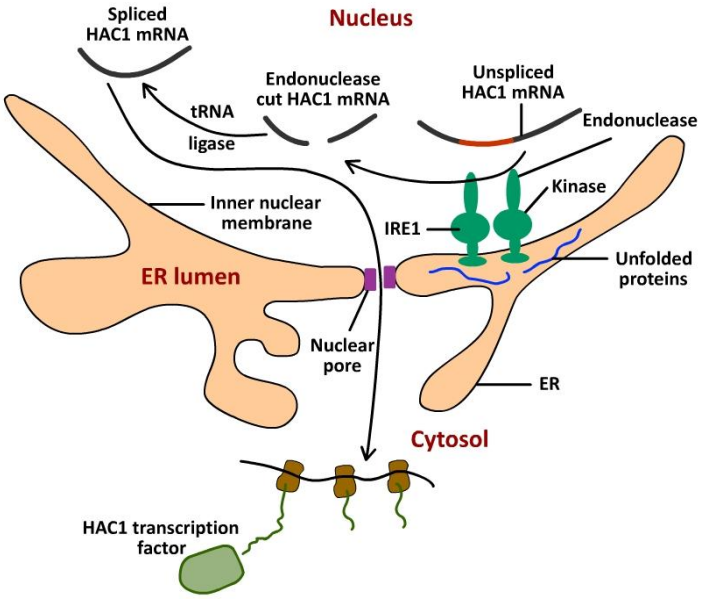


Fig. 16: The unfolded-protein response to IPR1 (**Source:** Author and Departmental Artist)

6. Role of PTMs in diseases:

PTMs play an essential role in disease pathology, cellular homeostasis and are the most actively sought after molecular targets for developing drugs and chronic disease therapy. Proteins play a major role in gene transcription, translation, cell signaling cascades, enzyme activities etc. abnormal PTMs lead to the production of disorderly proteins which are non-functional or dysfunctional in their activities. PTMs like phosphorylation, acetylation of proteins and carbohydrates play a key role in expanding the avenue of translational medicine for heterogenous disease like cancer. Lysine acetylation of tau proteins results in "tau tangles" in the case of dementia whereas lysine hyperacetylation of β-amyloid peptide results in impaired cognition in Alzheimer's disease. Age dependent memory impairment has also been observed due to alterations in histone acetylation in mouse models. In many mitochondrial, neurological, and cardiovascular diseases the carbonylation pattern has been an informative tool for the identification of stress levels.

Improper or incomplete glycosylation in the Fc receptor for immunoglobulin A has been shown to impact the IgA-mediated immune response which in turn affects many diseases including HIV, alcoholic liver cirrhosis, and other neuropathies. A many cardiovascular

diseases as well as neurological disorders like Parkinson's disease are due to abnormal levels of homocysteine (a methylation product of methionine). High levels of nitric oxide within the cell lead to inflammation, organelle damage laeding to cell death mainly by apoptosis. In many diseases including autoimmune response, HIV infection, lung diseases, multiple sclerosis, and cellular enzyme regulation tyrosine sulfation plays an important role.

7. Detecting and Quantifying Post-Translational Modifications

Post-transcriptional modification studies are facing a lot of challenges with the development of specific detection, purification and quantification methods. The advents of various refined methods of proteomics are helpful in overcoming the various challenges. These methods include the use of MALDI/MS (Matrix-assisted Laser Desorption/Ionization Mass Spectrometry, the combination of ECD (Electron Capture Dissociation) and peptide fragmentation with new generations of high-sensitivity FTMS (Fourier-Transform Mass Spectrometry).

Also, the field of bioinformatics has become a promising area of development and helps to predict modification sites in-silico. The development of databases of protein modifications with reference to the already existing protein and genomic database is thought to be a significant way in which the detection and quantification of PTMs could also be improved.

8. Significance

Field of proteomics, gene regulation and protein function helped in understanding the concept of post translational modifications. These PTMs studies are been used to develop and determine advance methods for some diseases such as heart diseases, cancer, neurodegenerative diseases and diabetes.

9. Summary

- Translation in eukaryotes takes place in 3 steps viz, initiation, elongation and termination. Expression of genes is a highly regulated process and depends upon the cellular environment.

- Polypeptides produced from translation needs further processing/modifications for their activation to perform various cellular/enzymatic functions. These modifications of the polypeptides could take place while they are being still synthesized (co-translational modifications) or once their synthesis is complete (post-translational modifications (PTM)).

- Post-translational modifications (PTMs) are the various covalent and non-covalent interactions of the polypeptide which help in converting the polypeptide into a protein molecule having all the functional properties and the required macromolecular 3 dimensional structure.

- PTM's are responsible for the proteomic diversity and helps in protein stabilization, biochemical activity regulation, protein targeting (localization through signal peptide) or protein signaling through protein-protein interaction or cascade amplification.

- PTM results in an exponential increase in the proteome complexity relatively much higher than the transcriptome and genome. Human genome is expected to have around 20-25 thousand genes but the proteomic diversity is estimated to be around 1 million proteins.

- PTM's are very specific and occur at a distinct amino acid side chains, peptide linkages often mediated by kinases, phosphatases, transferases and ligases; which themselves are classified on the basis of their ability to add, remove or transfer a protein, lipid, sugar, phosphate group etc by the cleavage of specific regulatory units.

- A signal peptide helps in the binding of ribosomes and ER.

- Phosphorylation helps in regulating the biological activity of enzymes and proteins by adding or removing the phosphate group. Common amino acids subject to phosphorylation are serine, threonine and tyrosine in the animal cells eg., tyrosine phosphorylation affects the activity of numerous growth factors.

- Chromatin remodelling is necessary for the transcriptional activity and is regulated by the methylation of lysine residues in histone proteins present in the nucleosome. Activity of numerous transcription factors are modified due to methylation.
- Glycosylation is the most common type of PTM. It involves a covalent bonding between the secretory or membrane bound protein and carbohydrate moieties
- Disulphide bonds (–S–S–) provide stability and hence the functionality to the tertiary/quantenary structure of the protein and involves appropriate folding of the polypeptide chains which remains unchanged throughout the lifespan of protein.
- Insulin is produced as a preproprotein and later on cleaved at specific sites to become active.
- S-nitrosylation is the addition of Nitric oxide groups to the thiol group of cysteine residues to attain protein stability. Acetylation and lipidation involves the attachment of specific acetyl and lipid group to the polypeptide chain.
- Misfolded proteins are degraded by ubiquitination; it involves tagging the protein with ubiquitin and then cleaving it the proteasome. Deubiquitinating enzymes help in removing the ubiquitin from the proteins for specific regulation purposes. SUMOlytion also resembles ubiquitination and helps in Apoptosis.
- Chaperones bind to hydrophobic regions of the polypeptide and shield them from the aqueous environment until the entire polypeptide is translated and then help in protein folding. Heat Shock Proteins (HSP) is produced under stressful condition or when the temperature of the cell is high and they acts as Chaperones that aid in protein folding.
- IRE1 is a transmembrane multifunctional protein and helps in the binding and folding of unfolded proteins in the ER lumen.
- Various techniques being used to study PTMs are MALDI/MS (Matrix-assisted Laser Desorption/Ionization Mass Spectrometry, the combination of ECD (Electron Capture

Dissociation) and peptide fragmentation with new generations of high-sensitivity FTMS (Fourier-Transform Mass Spectrometry).

- The study of post-translational modifications is also being benefitted by the growing bioinformatics field especially in the in-depth study of some diseases such as heart diseases, cancer, neurodegenerative diseases and diabetes.