BT2062 - Analytical Techniques in Biotechnology

Assignment 2 - Posttranslational Modifications of Proteins

Submitted by Sahana Gangadharan (BE17B038)

1 Introduction

Regulation of genes in different cells happens by various strategies which could control the gene expression, thereby affecting the production (increased/decreased quantity) of the corresponding mRNAs & proteins. The several key stages in the regulation of gene expression are as follows -

1. Transcription regulation.
2. mRNA processing - including the multiple post-transcriptional modifications like capping and tailing of mRNA as well as RNA splicing.
3. Translational regulation.
4. Post-translational modifications.

Since gene regulation can happen at every stage of this process, this allows for a vast spread of different proteins with varied functionalities, from a very smaller set of genes or the genome itself.

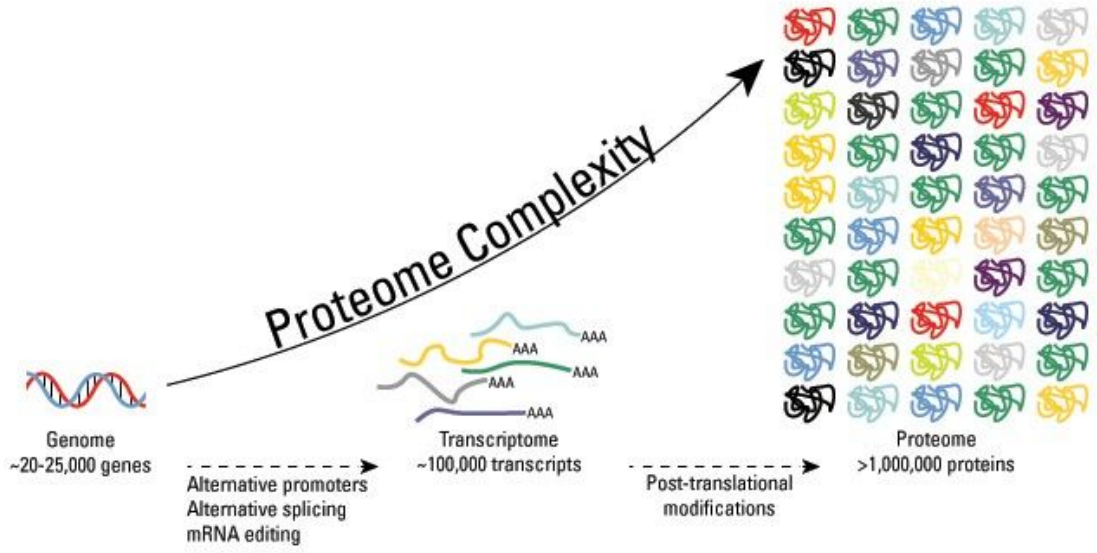

Fig 1: Proteomo complexity. **Source**: sandwalk.blogspot.com/the-protein-compexity-myth.html

2 Types of modifications

As it is evident from figure 1, post-translational modifications (PTMs) have a much higher impact in creating a diverse set of proteins than the other regulating methods. In general, PTMs involve modification to amino acid residues and addition of new functional groups to the proteins either through enzymatic or non-enzymatic processes. [1]. Some of the well-characterized PTMs are methylation, acetylation, phosphorylation, lipidation, glycosylation, and ubiquitination. PTMs can also involve the modification of residues, such as citrullination of arginine, deamidation of glutamine and asparagine, and elimination of serine and threonine. Non-enzymatic addition of functional groups occurs in glycation and carbamylation.

Apart from the above-said methods, some proteins get modified immediately after translation to better aid in protein folding or to direct the freshly-made nascent polypeptide to specific cellular compartments. The other methods happen after the protein has been localized in its specific region. These modifications decide the catalytic activity or other functionalities of the protein. Some PTMs also involve adding a tag to the protein, as an act of marking it for degradation.

Protein PTMs can also be reversible depending on the kind of modification that has occurred. For example, a common method of catalytic (in)activation is through phosphorylation by kinases at specific amino acid side chains. On the other hand, phosphatases hydrolyze the phosphate group to remove it from the amino acid residue and reverse the biological activity. Disulfide bond formation is also sometimes considered to be a type of post-translational modification [2]. These modifications affect the functioning of proteins and are also often termed gene regulation, although they do not have any direct effects on the gene.

2.1 Phosphorylation

Phosphorylation usually occurs on the following residues - serine, threonine or tyrosine residues. Phosphorylation is one of the well-studied PTMs owing to its prevalence in eukaryotic signalling, the availability of reagents for detection and enrichment, and by historical virtue of being one of the first identified PTMs. In eukaryotic organisms, enzyme - kinases are responsible for this PTM. They add a phosphoryl group with a -2 charge, in the hydroxyl groups of their side-chains, at the above residues in physiological pH. The presence of this group can have a higher impact on the protein folding, function and stability.

Dysregulation of phosphorylation by mutations in the Ser/Thr/Tyr residues that do not allow the kinase to add a phosphoryl group or dysregulation in the kinase activity can contribute to the progression of cancer or other diseases. On the other hand, the hyperphosphorylation of certain proteins also leads to diseases. For example, the hyperphosphorylation of tau protein is associated with Alzheimer's disease [3].

Phosphorylation has wide-ranging effects in the activation, deactivation, and modulation of protein functioning and plays an integral role in most signal transduction pathways [4]. As mentioned above, this PTM is reversible, i.e., the phosphoryl moiety can be removed by the action of the phosphatase enzyme. Apart from its vital role in cell signalling and its role in modulating chromatin structure, phosphorylation has also been employed in regulating membrane transport, protein-protein interactions and the activation and inhibition of enzymes.

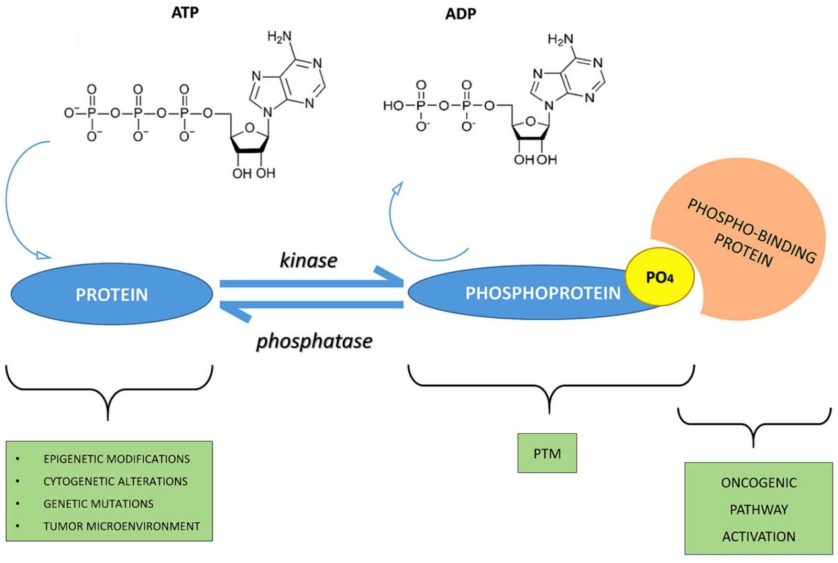

Figure 2: Phosphorylation of proteins as a PTM. **Source**: (Ardito et al., 2017) [5]

2.2 Glycosylation

Glycosylation involves the attachment of a carbohydrate group to an amino acid residue in the protein. The carbohydrate moiety that is added to the protein can range from being a simple monosaccharide (added in modifications of nuclear transcription factors) to highly complex branched polysaccharides (added in modifications of cell surface receptors).

Types of Glycosylation:

This PTM usually occurs in the nitrogen of arginine and asparagine residues, where they are N-linked and at the hydroxyl oxygen of serine, threonine, tyrosine, hydroxylysine or hydroxyproline residues, where they are O-linked. C-linked glycans are added at the carbon of tryptophan residues. The other forms of Glycosylation are Glypiation and Phosphoglycosylation. Glypiation is done by covalently attaching glycosylphosphatidylinositol anchors to the protein, and is very helpful in localizing proteins to the cell membranes. Phosphoglycosylation is done by attaching oligosaccharides serine or threonine in a peptide backbone via phosphodiesters [6]. Endopeptidase *Proteinase I*, isolated from *Dictyostelium* was the first protein that was shown to have GlcNAc-1-PO4 linked to a serine residue.

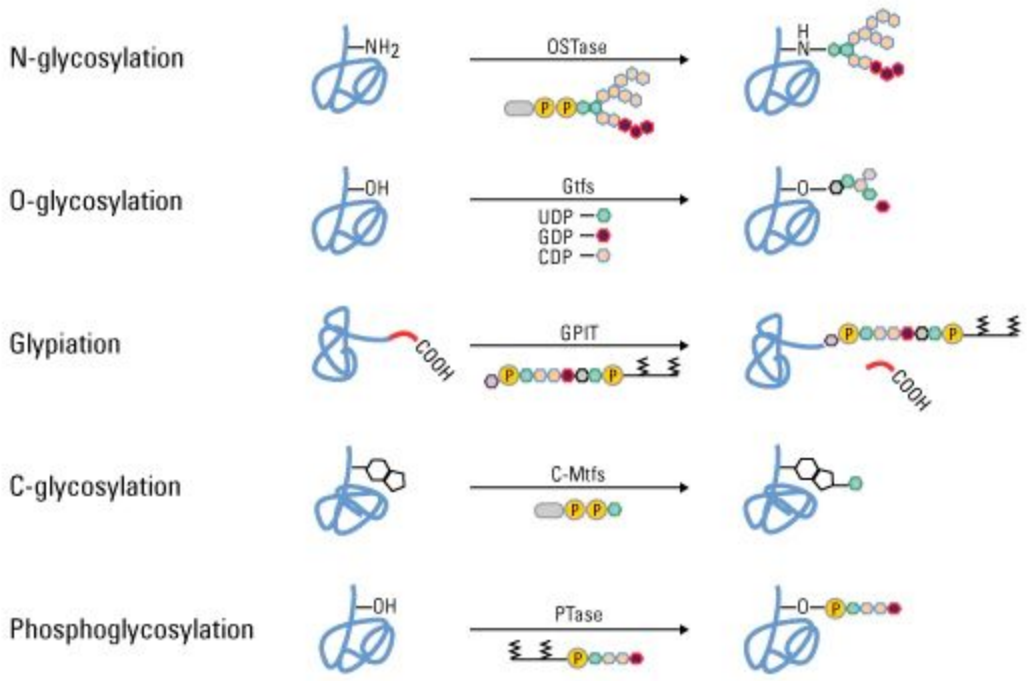

Figure 3: Types of Glycosylation PTM. **Source**: Thermofisher/Overview of PTMs

2.3 Ubiquitination

Ubiquitin a small polypeptide chain (size 8-kDa). It consists of 76 amino acids. Ubiquitination is the process of appending this small protein to a large protein at NH2 of Lysine residues. The C-terminal end of ubiquitin, i.e., C-term Carbon in the Glycine residue forms a bond with the Lysine residue of the protein that is being modified. Initial addition of a single ubiquitin, termed as monoubiquitination, can be followed by the formation of a ubiquitin polymer. The degradation of ubiquitinated proteins and the recycling of ubiquitins are taken care of by 26S proteasomes that specifically recognize polyubiquitinated proteins. Ubiquitination is the best method for marking proteins for degradation [7]. However, ubiquitination and proteasomal degradation are not entirely congruent, since ubiquitination functions differently from targeting proteins for degradation [8]. Other commonly observed functions of post-translational modification using ubiquitins are DNA repair, subcellular localization and modulating protein-protein interactions [9]. Ubiquitin is expressed as a fusion protein and is recycled by ubiquitin-specific hydrolases. Hence this is also a reversible PTM.

Mechanism:

Ubiquitin is activated and transferred by a cascade of enzymes. The first enzyme in this cascade, E1 activates ubiquitin by consuming ATP and by forming a thiol ester with the C-terminal glycine of ubiquitin. The ubiquitin moiety is then transferred to an E2 enzyme by transesterification. The E2 enzyme can now act in two ways. It can transfer the ubiquitin to an E3 enzyme, which will later transfer it to the final target. Or, it ubiquitinates the target by direct interaction with an E3 enzyme. In some cases, a novel class of E4 enzymes exists that is responsible for polyubiquitination.

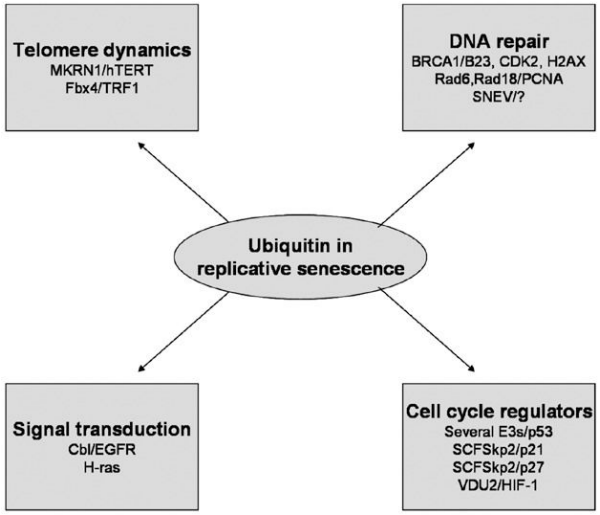

Figure 4: Implications of Ubiquitin system on different pathways **Source**: Grillari J et al., 2006

Small Ubiquitin-like Modifiers (SUMOs) are part of the ubiquitin-like protein family and are also attached to lysine residues of the substrate protein. Sumoylation, however, unlike ubiquitins, does not mark proteins for degradation. Instead, it has been shown to play important roles in protein localization, transcription inhibition, mitosis, and multiple DNA repair pathways [10]. Telomerase activity is dependent on Ubiquitination. Furthermore, the telomeric repeat binding factor 1 (TRF1), a negative regulator of telomere length, is targeted to ubiquitin-dependent degradation.

2.4 S-nitrosylation

S-nitrosylation is the process by which an NO group is added to a cysteine thiol residue to form an *S*-nitrosoprotein. (SNO-Protein). Regulation by S-nitrosylation impacts an extensive range of signal transduction pathways mediated by phosphorylation/dephosphorylation. Nitric oxide plays a major role in cellular signal transduction, some of this through S-nitrosylation. It plays a vital role in the autoregulation of blood flow [11]. It has been estimated that nearly 70% of the human proteome undergoes nitrosylation at conserved residues [12].

It is important to note that S-nitrosylation is not a random event, and only specific cysteine residues are S-nitrosylated. Also, due to the presence of other cysteine residues and due to the labile nature of SNOs, it is pretty difficult to detect S-nitrosylated cysteines and distinguish them from non-S-nitrosylated amino acids. The biotin switch assay, developed by Jaffrey et al., [13] is a common method of detecting SNOs. All the free cysteines are blocked. The remaining Cys residues are expected to be de-nitrosylated. The free thiol groups are then biotinylated. The biotinylated proteins are detected by SDS-PAGE followed by Western blot or mass spectroscopy analysis.

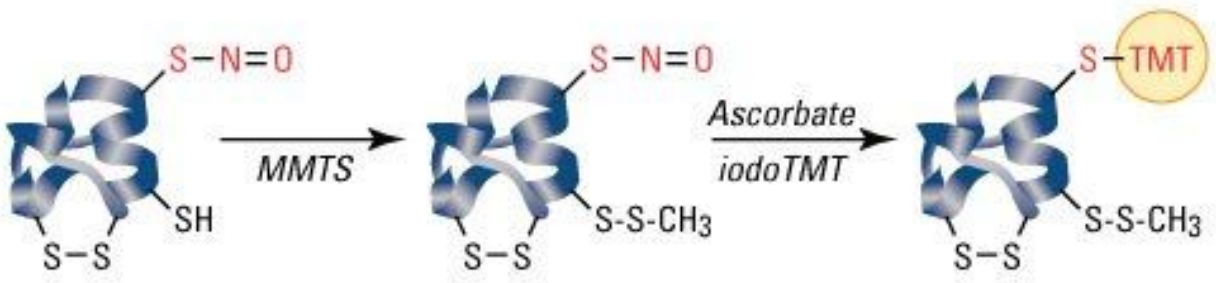

Figure 5: Detection of SNO proteins using the biotin switch assay. **Source**: Thermofisher/Overview of PTMs

2.5 Methylation

The addition of a methyl group to either a Nitrogen or Oxygen atom (N-methylation and O-methylation respectively) on an amino acid residue also regulates gene expression [14]. The addition of the CH3 moiety increases the hydrophobicity of the protein and can neutralize a negative amino acid charge when bound to carboxylic acids. This process is mediated by methyltransferases, with S-adenosyl methionine (SAM) being the primary methyl donor. The methyl group, donated by SAM, is transferred by esterification to the γ-carboxyl group of one or more glutamic acid residues.

Methylation is a very frequently adopted PTM, that SAM is suggested to be the most-used molecule in enzymatic reactions, after ATP. It is important to note that N-methylation is irreversible, while O-methylation can be reversed. Protein methylation, mediated by signal transduction, is involved in cellular stress responses and also the aging and repair of proteins. In general, methylation occurs at lysine and arginine residues. A commonplace where methylation is frequently observed is at the nucleosomes, in specific, the histone proteins. Methylation is a well-known type of epigenetic regulation, as histone methylation/ demethylation is responsible for the availability of DNA for transcription and translation.

2.6 N-acetylation

N-acetylation, i.e., the transfer of acetyl groups to nitrogen in amino acids, occurs via both reversible and irreversible mechanisms. N-terminal acetylation, for instance, is common and is estimated to occur in around 85% of all human proteins. N-terminal acetylation is known to play a part in protein stability, localization to the Golgi apparatus, p53-dependent apoptosis and global protein synthesis. Lysine acetylation plays an important regulatory role in the functioning of p53, tubulin and STAT3. N-terminal acetylation happens by the cleavage of the N-terminal methionine residue. This is in turn, mediated by methionine aminopeptidase before replacing the amino acid with an acetyl group from acetyl-CoA by N-acetyltransferase enzymes. This type of acetylation happens alongside translation,i.e., N-terminus is acetylated on growing polypeptide chains that are still attached to the ribosome.

Acetylation at the ε-NH2 of lysine on histone's N-termini is a common method of regulating gene transcription [1]. Histone acetylation, as mentioned above, is a reversible event that reduces chromosomal condensation to promote transcription. The acetylation of these lysine residues in histones is regulated by transcription factors that contain histone acetyltransferase (HAT) activity. A group of NAD-dependent deacetylases that target histones are referred to as Sirtuins (silent information regulator). They maintain gene silencing by hypoacetylated histone and also aid in the maintenance of genomic stability.

2.7 Lipidation

Lipidation involves the covalent addition of lipid moieties to protein residues. This process is primarily observed in the localization of proteins to the cell membrane and protein trafficking. The commonly seen types of PTMs through lipidation are as follows [1]:

1. **N-myristoylation** is the adding of myristate to an N-terminal glycine residue. This gives a hydrophobic handle to the membrane proteins that are being modified. N-myristoylation can act as a conformational localization switch, where conformational changes in proteins influence membrane attachment. Because of this conditional localization, signal proteins that selectively localize to the membrane, such as Src-family kinases, are N-myristoylated.
2. **S-palmitoylation** is the attachment of thioester linkage of long-chain fatty acids to cytoplasmic cysteine residues. C16 palmitoyl group from palmitoyl-CoA is added to the cysteine residues via palmitoyl acyltransferases. S-palmitoylation also works around membrane localization, but this process acts like an on/off switch (reversible in nature). S-palmitoylation can also act as a strengthening agent for other lipidation modifications.
3. **Prenylation** is the thioether linkage of a protein and farnesyl (C-15) moiety or geranylgeranyl (C-20) group. These functional groups are usually attached at specific cysteine residues that are about 5 residues away from the C-terminus. This addition is mediated by farnesyl transferase (FT) or geranylgeranyl transferases (GGT I and II). Prenylation occurs in the endoplasmic reticulum. It is a part of a larger PTM process that is followed by proteolytic cleavage by Rce1, and methylation via isoprenyl-cysteine methyltransferase (ICMT).

2.8 Proteolysis

Sometimes, when certain genes are transcribed together, the transcripts are, in turn, translated together. In the case where the proteins have roles in different parts of the cell, this would mean that in order for the proteins to function properly, they need to get cleaved. Such modifications after translation are called ‘proteolysis’. Peptide bonds are very stable in physiological conditions. Therefore, this will require the involvement of certain proteins/enzymes known as ‘proteases’ to carry out this function. Proteolysis is a thermodynamically favourable and irreversible reaction.

There are different families of proteases that specifically cleave certain peptide bonds that are flanked by certain active groups. The commonly known protease families are - serine proteases, cysteine proteases, aspartic proteases, threonine proteases and metalloproteases. Another way of classifying proteases is based on the site of action, that is aminopeptidases and carboxypeptidase, which cleave at the amino or carboxy terminus of a protein, respectively.

3 Histone Modification - A case study

Gene regulation is also known to exist at a Transcriptional level via Post-translational modifications. Histone modification is a covalent PTM that happens via other established PTMs such as methylation, phosphorylation and acetylation [15]. The histone modification regulates gene expression by modifying chromatin structure (depends on the process of PTM involved), i.e., by exposing or not exposing certain genes to other biomolecules.

The double-stranded DNA is wound around the eight histone proteins which package itself throughout the chromosome. Certain histone modifications are passed on through cell division and cell cycle. Therefore, PTM via histone modifications acts in epigenetic regulation of various cellular processes and the overall development of the organism. Diverse biological processes such as chromosome packaging, transcriptional activation/ inactivation and DNA repair/damage are mediated by this PTM. In most species, histone H3 is primarily acetylated at lysines found in position 9, 14, 18, 23, and 56. They are methylated at arginine position: 2, and lysines at position 4, 9, 27, 36, and 79, and phosphorylated at Serine positions 10, 28 and Threonine positions 3, 11. Histone H4 is primarily acetylated at lysines positions 5, 8, 12 and 16, methylated at arginine 3, and lysine 20, and phosphorylated at serine 1.

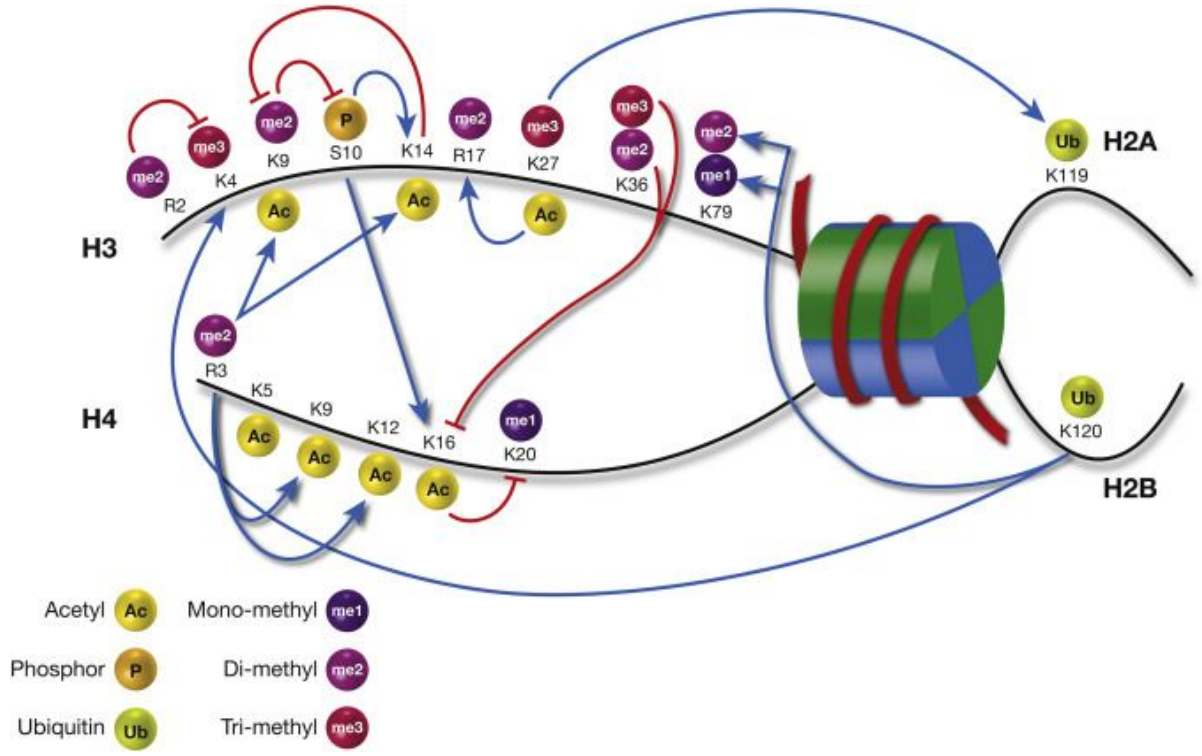

Figure 6: Various modifications in histones **Source**: Simonet et al., 2016

Types of Histone modifications

Acetylation: Reported first in 1964 by Allfrey et al., [16] acetylation of lysines have proven to be highly dynamic and have serious impacts. This process is reversible and is mediated by histone acetyltransferases (HATs) and histone deacetylases (HDACs). HATs function dependently on acetyl CoA and transfer acetyl groups to ε-amino groups of lysine side chains. This transfer neutralizes the charge on lysine residue. The introduction of acetyl group reduces the attraction between histone proteins and the DNA, thereby reducing DNA condensation. This, in turn, opens up the DNA (genes, basically) to transcription factors and other proteins for its expression. Thus, acetylation aids in gene expression.

Methylation: Lysine methyltransferases (HKMT) are the enzymes that mediate methylation in histone proteins. Lysines and arginines that are close to N-terminus (called the SET domain) are prone to methylation. A single (monomethylation) methyl group, double (dimethylation) methyl groups and triple (trimethylation) methyl groups can be added. A common observation is that monomethylation enhances gene expression, while dimethylation and tri-methylation down-regulate and inhibits gene expression. This is also a reversible modification where enzymes known as histone demethylase remove methyl groups from the lysine/arginine residues.

Phosphorylation: This mechanism in histones is highly dynamic. Residues such as serines, threonines and tyrosines get phosphorylated by the histone kinase enzyme. Histone kinases transfer a phosphate group to the hydroxyl group of the amino acid residue from ATP. This modification instils a double negative charge on the protein, which in turn affects the structure of the chromatin. Not many studies have explored histone phosphatases (the enzymes that hydrolyse the phosphate groups in the protein). However, phosphorylation of core histones at the DNA entry-exit regions to nucleosomes has been shown to increase chromatin accessibility, especially in conjunction with acetylation [17].

4 Conclusion

In general, post-translational modifications affect the structure, stability, properties and functions of the nascent polypeptide chains. These modifications equip the proteins for their functional roles inside the highly complex working machinery, the cell. Also, PTMs can affect gene regulation at various stages. These multiple levels of modifications are fine-tuned and coordinated in a healthy cell, and their dysregulation has been linked to several cancers [18].

5 References

1. Overview of Post-Translational Modification - IN.
2. . 2020. Post-translational modification. *Wikipedia*
3. Phosphorylation Bioinformatics. *Nov. Biol.*
4. **Olsen JV, Blagoev B, Gnad F, Macek B,** et al. 2006. Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell* **127**: 635–48.
5. **Ardito F, Giuliani M, Perrone D, Troiano G,** et al. 2017. The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review). *Int. J. Mol. Med.* **40**: 271–80.
6. **Haynes PA**. 1998. Phosphoglycosylation: A new structural class of glycosylation? *Glycobiology* **8**: 1–5.
7. **Glickman MH**, **Ciechanover A**. 2002. The Ubiquitin-Proteasome Proteolytic Pathway: Destruction for the Sake of Construction. *Physiol. Rev.* **82**: 373–428.
8. **Grillari J**, **Grillari-Voglauer R**, **Jansen-Dürr P**. 2010. Post-Translational Modification of Cellular Proteins by Ubiquitin and Ubiquitin-Like Molecules: Role in Cellular Senescence and Aging. In Tavernarakis N. ed; *Protein Metabolism and Homeostasis in Aging*. Springer US. p 172–96.
9. **Mukhopadhyay D**, **Riezman H**. 2007. Proteasome-Independent Functions of Ubiquitin in Endocytosis and Signaling. *Science* **315**: 201–5.
10. **Jalal D**, **Chalissery J**, **Hassan AH**. 2017. Genome maintenance in Saccharomyces cerevisiae: the role of SUMO and SUMO-targeted ubiquitin ligases. *Nucleic Acids Res.* **45**: 2242–61.
11. **Hess DT, Matsumoto A, Kim S-O, Marshall HE,** et al. 2005. Protein S -nitrosylation: purview and parameters. *Nat. Rev. Mol. Cell Biol.* **6**: 150–66.
12. **Stomberski CT**, **Hess DT**, **Stamler JS**. 2019. Protein S-Nitrosylation: Determinants of Specificity and Enzymatic Regulation of S-Nitrosothiol-Based Signaling. *Antioxid. Redox Signal.* **30**: 1331–51.
13. **Jaffrey SR**, **Snyder SH**. 2001. The biotin switch method for the detection of S-nitrosylated

proteins. *Sci. STKE Signal Transduct. Knowl. Environ.* **2001**: pl1.

14. **Aletta JM**, **Cimato TR**, **Ettinger MJ**. 1998. Protein methylation: a signal event in post-translational modification. *Trends Biochem. Sci.* **23**: 89–91.
15. **Bannister AJ**, **Kouzarides T**. 2011. Regulation of chromatin by histone modifications. *Cell Res.* **21**: 381–95.
16. **Allfrey VG**, **Faulkner R**, **Mirsky AE**. 1964. Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. *Proc. Natl. Acad. Sci.* **51**: 786–94.
17. **Brehove M, Wang T, North J, Luo Y,** et al. 2015. Histone Core Phosphorylation Regulates DNA Accessibility. *J. Biol. Chem.* : jbc.M115.661363.
18. **Noberini R, Osti D, Miccolo C, Richichi C,** et al. 2018. Extensive and systematic rewiring of histone post-translational modifications in cancer model systems. *Nucleic Acids Res.* **46**: 3817–32.