Eukaryotic Gene Expression

Prof.P.N.Rangarajan

Department of Biochemistry

Indian Institute of Science, Bangalore

Module No.# 03

Lecture No.# 08

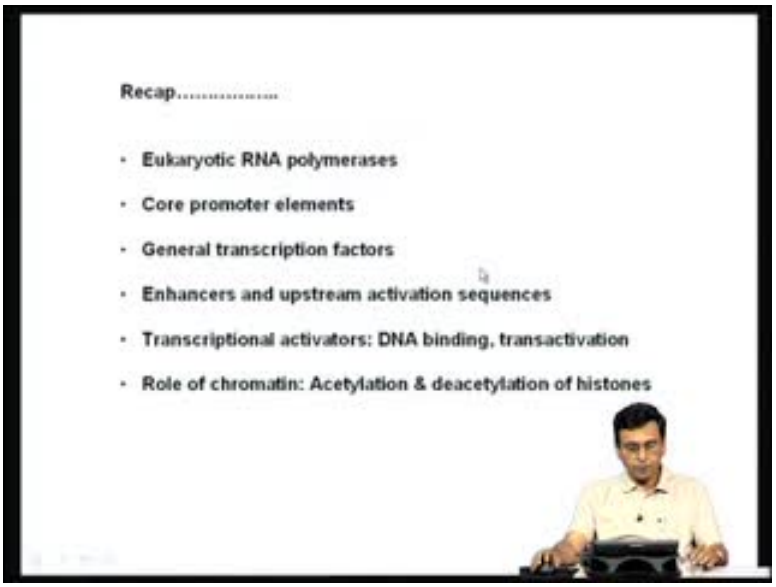
Eukaryotic Gene Regulation: Post Translational Modifications of Histones

Welcome to this eighth lecture in this course on eukaryotic gene expression. In the last class, we talked about importance of chromatin in the regulation of gene expression, and we discussed about how covalent modification of histones can affect gene regulation. And we specifically discussed one particular post translational modification of histones, that is, acetylation and deacetylation of histones, how they influence gene regulation.

And I told you, very clearly, that there are enzymes called as histone acetyltransferases, which add a acetyl group on histones, and this acetylation of histones results in a positive regulation of gene expression or activation of gene expression. And there are also enzymes called histone deacetylases or HDACs, which remove the acetyl group from the histones, and this results in the repression of gene expression.

So, today, what we will do is, we will go and then discuss one more post translational modification of histones, namely, methylation, and then see, how methylation of histones affects gene expression, and we will also talk about other histone post translational modifications, especially phosphorylation, and so on and so forth.

(Refer Slide Time: 01:26)



So, just to recapitulate what we have discussed, so far, in this course, we began with discussing about eukaryotic RNA polymerases. Then, we discussed about some of the core promoter elements like the TATA box, the initiator, and so on and so forth.

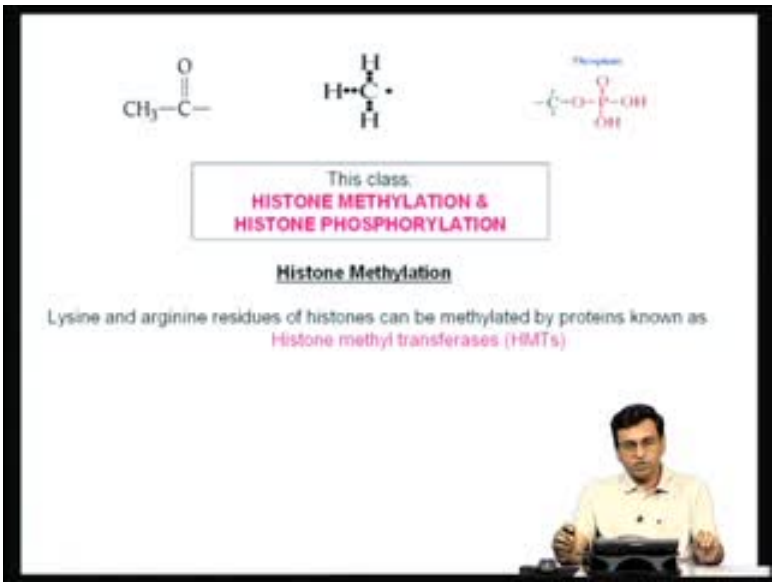
Then, we talked about general transcription factors, which come and assemble in the core promoter elements, and we, in fact, discussed how variations in general transcription factors itself can bring about differential gene regulation, and variations in the co-promoter elements can also bring differential gene regulation.

Then, we moved little bit further upstream, and then discussed about enhancer sequences and upstream activation sequences. Then, we discussed, how transcriptional activators, which are sequence-specific DNA binding proteins, go and bind through their DNA binding domains, specifically to specific sequence in the upstream regions, and then, through their trans activation domains, can influence the rate of transcription initiation.

And just recently, we discussed about DNA is not naked inside the cells. DNA is actually organized the form of chromatin. You have nucleosomes, and therefore, the first thing that has to happen, if transcription initiation has to take place, the DNA has to be unwound, and the histones have to be displaced, and this is where covalent modifications of histones have become a very important player in regulation of gene expression. And in

the last class, we discussed about how acetylation and deacetylation of histones regulates gene regulation.

(Refer Slide Time: 02:45)

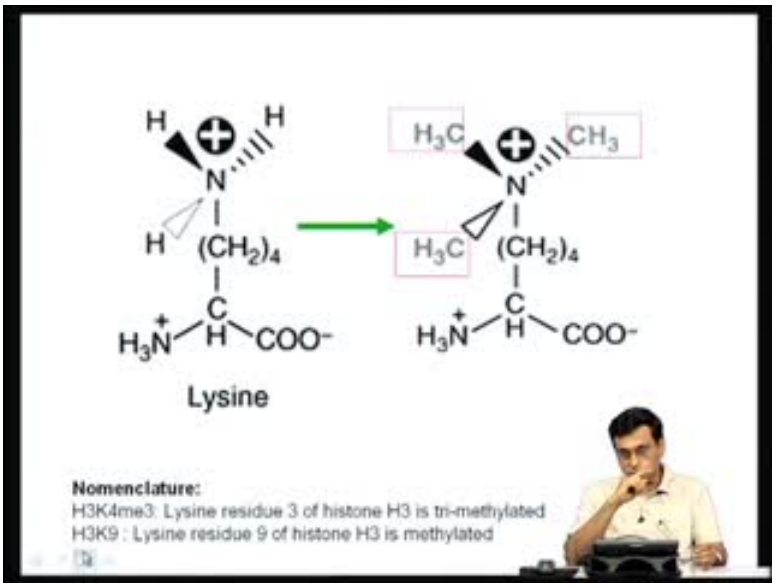


So, acetylation is one major mechanism by which gene expression, gene transcription can be activated in eukaryotic cells, and today, how we will discuss how addition of a methyl group to histones bring about changes in the gene regulation, and we will also talk little bit about phosphorylation of histones.

So, the enzymes which will add acetyl group are called histone acetyltransferases; the enzymes which will add methyl group, which we are going to discuss today, they are called as histone methyltransferases, and enzymes which phosphorylate histones, they are called as histone kinases.

So, as I now just mentioned, we will, let us discuss about methylation of histones and their role in gene regulation, and the lysine and arginine residues of the histones can be methylated by proteins known as histone methyltransferases. So, the 2 amino acids lysines, which are both are basic amino acids, these are the targets for methylation, and the enzymes which will add methyl group to these histones, they are known as histone methyltransferases.

(Refer Slide Time: 03:38)

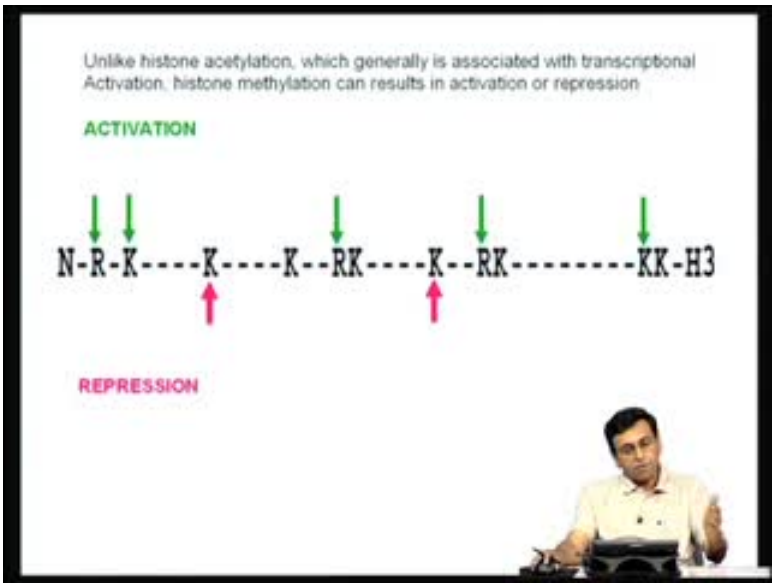


So, this is just a cartoon, just tells you how does the lysine residue of a histone look like, and you can see, you have a positively charged, positive charge here. That is why, they are called basic amino acids, and if you add one methyl group to this lysine, then becomes a mono methyllysine. So, the hydrogen here is replaced by methyl group. Now, if you have two methyl groups, then becomes dimethyllysine, and if you have three methyl groups, then becomes trimethyllysine.

So, you have histones. If you have only methyl at 1 methyl group, it will becomes a mono methyl histone, di-methyl histone, and tri-methyl histones. Accordingly, depending upon the number of methyl groups you have been added to the histones. There is a nomenclature; I just got an example here. For example, if you say H3K4me3, it actually means, in the histone H3, the lysine 4 residue is tri-methylated.

So, lysine residue 3 of the histone H3 is tri-methylated, and you denote this as H3K4me3. Similarly, if you say H3K9, it actually means the lysine 9 residue of histone H3 has been methylated. So, there are enzymes which will add only one methyl group, two methyl groups, or three methyl groups, and accordingly, you have mono methyl, di-methyl, or tri-methyllysines in the histones.

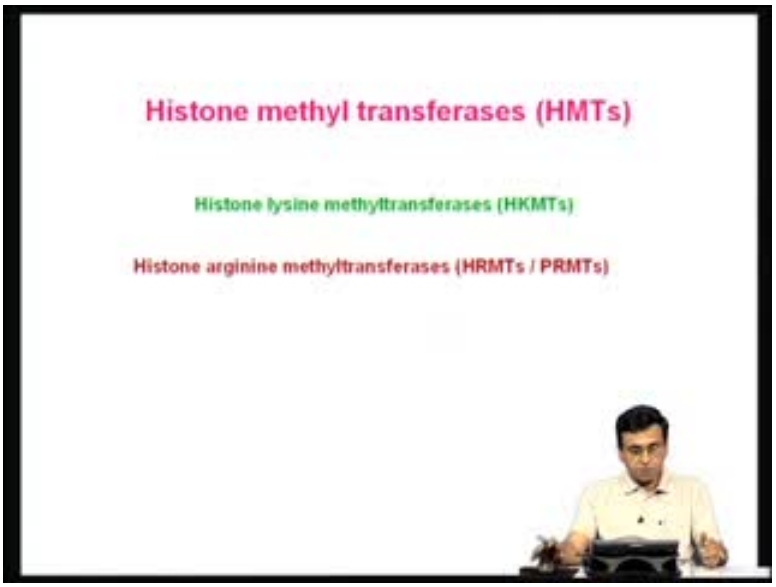
(Refer Slide Time: 04:52)



Now, the important distinction between the histone acetylation and histone methylation is that histone acetylation, most often, results in transcriptional activation. They, of course, in biology, there are always exceptions. When I say the histone acetylation results in transcription activation, there may be one of the instances where histone acetylation may also lead to repression of transcription.

But, in general, acetylation of histones results in activation of transcription. That is what happens in common, but unlike histone acetylation, the histone methylation can either result in activation, or can result in repression of transcription. And here, for example, I give an example where methylation of this particular arginine or lysine in the histone H3, can actually result in activation of transcription. On the other hand, if you methylate these lysines, it can actually result in repression of transcription.

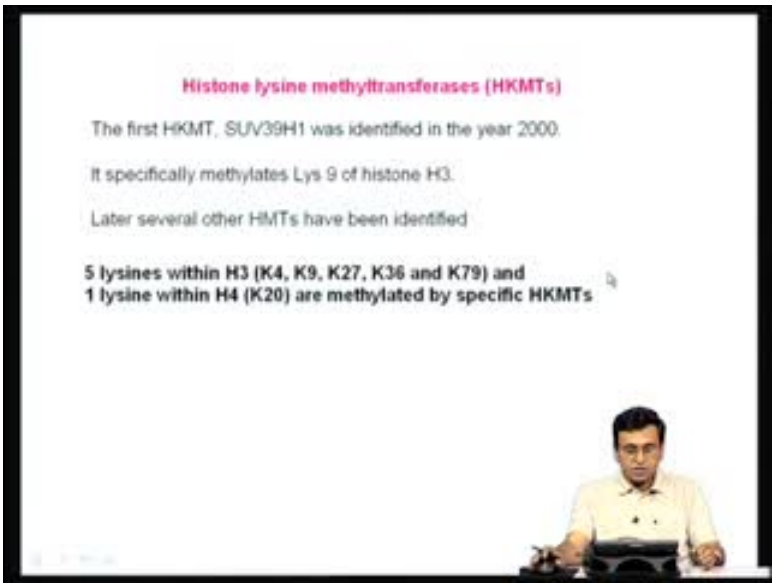
(Refer Slide Time: 05:52)



So, depending upon what kind of residues, which residue is getting phosphorylated, methylated, the… it can either result in activation of transcription, or it can result in repression of transcription. Now, let us study the enzymes, which actually add this methyl group to histones. So, these enzymes, which add a methyl group to histones, they are known as histone methyltransferases.

Now, just as I told you that there are either a lysine can be methylated or an arginine can be methylated, and accordingly, you have two classes of histone methyltransferases. They are called histone lysine methyltransferases, HKMTs, or histone arginine methyltransferases, or HRMTs, or in general, protein arginine methyltransferases.

(Refer Slide Time: 06:23)

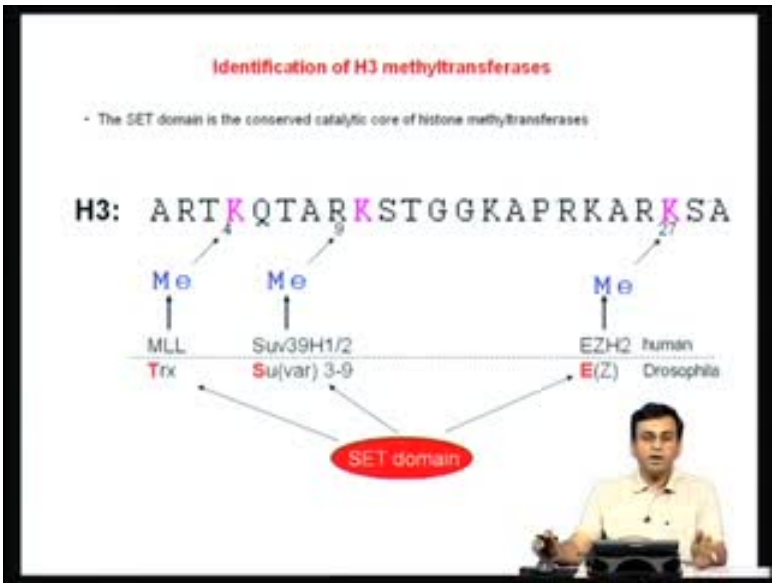


Let us now look at first histone methyl lysine methyltransferases and see, what kind of methyltransferases are there, which methylate the lysine residue of histones. Now, the first histone lysine methyltransferase, which was discovered in Drosophila, is called SUV39H1, somewhere in the year 2000. So, this was the first histone methyltransferase to have been discovered, and this methyltransferase was soon found to methylate, specifically, the lysine 9 of histone H3.

So, this particular methyltransferase– the SUV39H1, specifically methylates the lysine 9 residue of histone H3, and once people realized that the methylation of histones can actually take place, and there are actually enzymes which can methylate histones, people started looking for such methyltransferase activities, and several other methyltransferases were later discovered, and we will discuss some of them in the next few minutes.

Now, so all these studies I have actually shown, there are at least five lysine residues in the histone H3, which can be methylated by different enzymes, and these include lysine 4, lysine 9, lysine 27, lysine 36, and lysine 79. So, at least five lysine residues in the H3 are targets for eight different methyltransferases in the eukaryotic cells, and in the case of H4, at least one lysine residue, that means, the lysine 20, can be methylated by a specific methyltransferase.

(Refer Slide Time: 07:41)

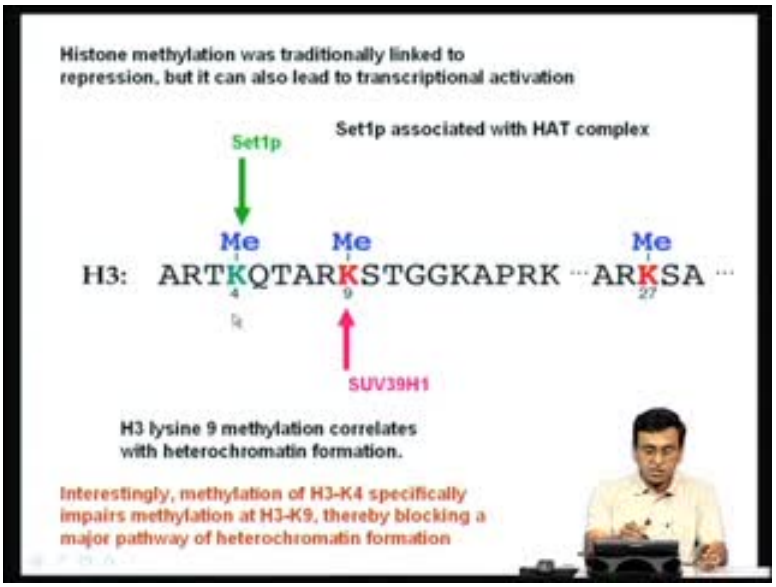


So, you can see here, I actually designated the three methyltransferases, which can methylate the specific histone residues in histone H3. The lysine residue, showed in pink colour here, and you can see a methyltransferase called Trx in the case of Drosophila, and its homolog, called MLL, methylates the lysine 4 residue of histone H3. Similarly, a the methyltransferase called SUV39H1 in the in case of human, or SU329 in the case of Drosophila, they can methylate the lysine residue in the… of H3, and the methyllysine 27 of histone H3 can be methylated by EZH2 in the case of humans, or its Drosophila homolog called EZ.

So, there are specific histone methyltransferases which can specifically methylate specific residues in the amino terminal regions of the histone H3. Now, just as I told you in the last class that the acetylated histones are recognized by proteins through what are called as bromodomains, right? The same way, the catalytic domains involved in methylation of histones in the case of histone methyltransferases are called, is known as a SET domain, and a SET actually stands for the catalytic domain present in these 3 proteins– S, E, T.

So, when you say SET domain, SET domain is nothing but a catalytic domain in…, which is present in the histone methyltransferases that is involved in methylation of lysine residues of histones. This is called as a SET domain.

(Refer Slide Time: 09:11)



Now, histone methylation is traditionally linked to repression, but it can also lead to transcription activation. That is what I told you in the previous slide, just as histone acetylation is most of most commonly results in transcriptional activation, in the case of histone methylation, although traditionally, methylation involves, results in transcription repression, there are instances where histone methylation, lysine methylation, can actually to transcriptional activation. For example, I have given one example here. Lysine 4 methylation by a histone methyltransferase is called Set1p can actually result in activation of transcription, whereas lysine 9 methylation by SUV39H1 can result in repression of transcription.

So, histone methylation can either will lead to activation or repression of transcription. In fact, when they looked at this Set1p, which is a lysine methyltransferase, very often, the Set1p was found to be associated with a HAT complex, that is, the histone acetyltransferase complex, and now you can see, how acetylation by Set1p can results in transcriptional activation.

So, what could be happening is, when this particular methylation takes place, this actually becomes signal for histone acetylation, because Set1p is part of histone acetylation. So, Set1p, probably, will be bringing the HAT complex to the vicinity of the methylated residue, and then the HAT will now methylate other lysine residue, other residues in the histone H3 that can result in the activation of transcription.

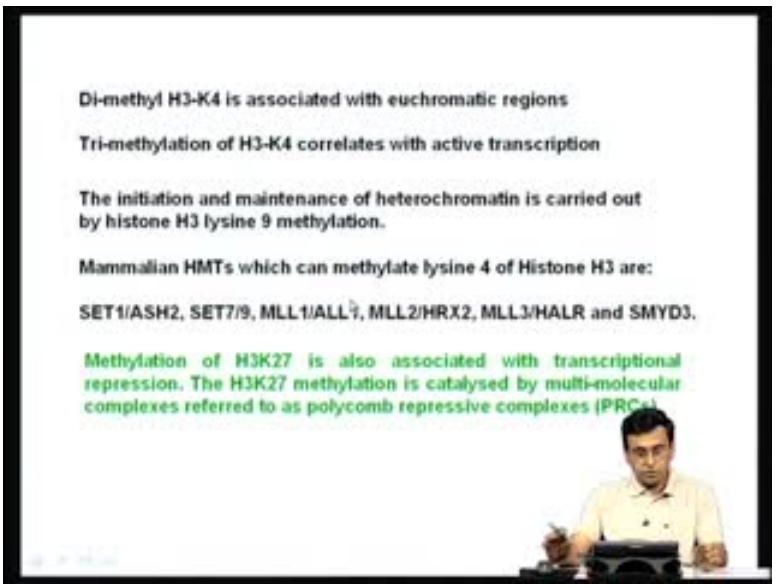
Very interestingly, the lysine 9 methylation, in the case of histone H3, often results in transcription repression, and, in fact, it is found to be one of the major mechanisms by which euchromatin is converted to heterochromatin. So, in most of the cases, in the heterochromatin, the lysine 9 of histone H3 was found to be methylated.

And, very interestingly, methylation of H3 and lysine 4 residue of histone H3 impairs methylation at lysine 9. That means, when lysine 4 gets methylated by Set1p, the chances of lysine 9 get methylated by SUV39H1 is reduced. So, methylation of this residue actually affects the methylation at the adjacent residue.

So, therefore, then, most probably, that is why you always find lysine 4 methylation mostly in the euchromatic regions, because in the lysine 4, when it gets methylated, there are very rare chances that the lysine 9 will get methylated, and therefore, very rarely, you find heterochromatin when lysine 4 is methylated, because this actually reduces the efficiency of methylation of lysine 9 residue.

So heterochromatin usually does not contain lysine 4, and therefore, lysine 4 methylation is usually found in euchromatin regions, whereas lysine 9 usually found in the heterochromatin regions, and methylation of lysine 4, actually, negatively regulates methylation of lysine 9 residue.

(Refer Slide Time: 12:00)



Now, people have now looked at these methylated histones in a number of systems, eukaryotic systems, like Drosophila, mouse, and so on and so forth, and they have actually drawn certain conclusions, and some of these conclusions are, di-methyl H3-K4 was shown to be associated with euchromatic regions.

So, when with the presence of two methyl rows in the lysine 4 is often found to be present in the euchromatic regions, and similarly, tri-methylation of the lysine 4 residue of H3, often, is a signal that these genes are going to be actively transcribed.

So, it is not only this specific residue of lysine, but also depending upon the number of methyl group that is added, can be a signature for a specific physiological response. So, di-methyllysine is often is found in euchromatic regions, and the presence of one more lysine, and the tri-methyllysine, usually signifies that these promoters, in which the particular histone has been…, contains a tri-methyllysine at lysine 4, these genes are likely to be actively transcribed, and the initiation and maintenance of heterochromatin is carried out by histone H3 lysine 9 methylation.

So, remember, whenever you talk about histone H3 methylation, there are going to be two residues which are very important– one is lysine 4, another is lysine 9; and lysine 9 methylation, often, results in the activation of transcription; lysine 9 residue results in the repression and heterochromatinization of the chromatin.

So, mammalian histone methyltransferases, which can methylate lysine 4 of histone H3, a number of such enzymes have been discovered and some of them are mentioned here, and the names are right here. So, a number of histone methyltransferases are capable of adding a methyl group in the lysine 4 residue of histone H3.

And, very interestingly, methylation of the lysine 27 has also been shown to be involved in transcription repression. So, in addition to lysine 4 and lysine 9, lysine 27 is also plays a very important role in the regulation of gene expression and chromatin structure, and lysine 27 methylation can also result in transcription repression.

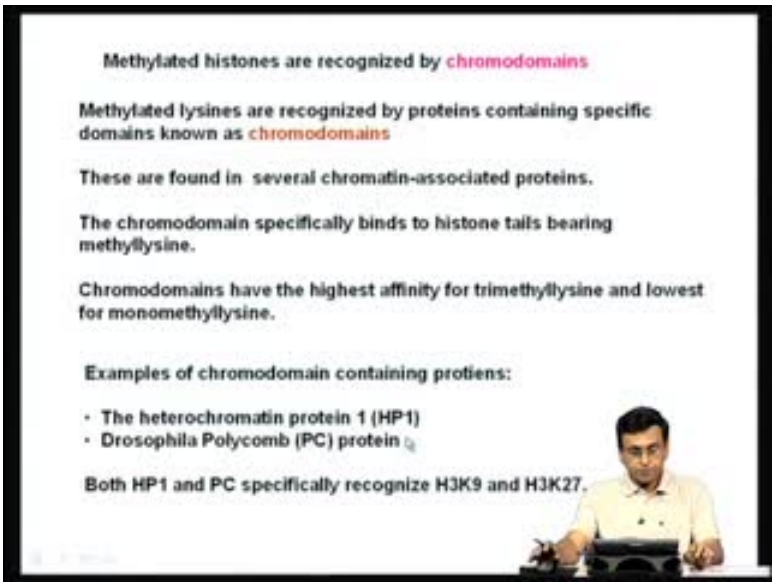
So, remember, when somebody asks you about how heterochromatinization, how histone methylation affects heterochromatin formation, you have to remember that methylation

of lysine 9 or lysine 27 in histone H3 can result in heterochromatin formation and transcription repression.

So, methylation of these two residues can actually go into formation of heterochromatin. The and H3 lysine, the lysine 27 methylation because of H3 histone is actually catalyzed by a multi molecular complex, and is often referred to as the polycomb repressive complexes. Polycomb is name of a protein, and this protein is actually involved in transcription repression and heterochromatin formation.

We will discuss some more details about this in the next few more classes, later, but just remember here, that lysine 27 as well as lysine 9 methylation signals heterochromatinization, and methylation of lysine 27 is actually carried out by a huge multi protein complex, and one of the members of this multi protein complex is a protein called polycomb.

(Refer Slide Time: 15:04)



Now, just as the acetylated histones are recognized by proteins which contain a specific domain called as bromodomains, similarly, methylated histones are recognized by proteins, which contain specific domains called as chromodomains.

So, remember, whenever you talk about histone modifications, chromatin, and gene regulation, acetylated histones are recognized by proteins through their bromodomains, whereas methylated histones are recognize proteins through their chromodomains.

So, methylated lysines are recognized proteins containing specific domains known as chromodomains, and these are found in several chromatin associated proteins. So, you can see here, the way that the whole system works, is the moment an acetyl group is added, certain proteins containing bromodomains come and interact with these acetylated residues, and that is a signal for recruitment of other protein complexes, resulting in transcription activation.

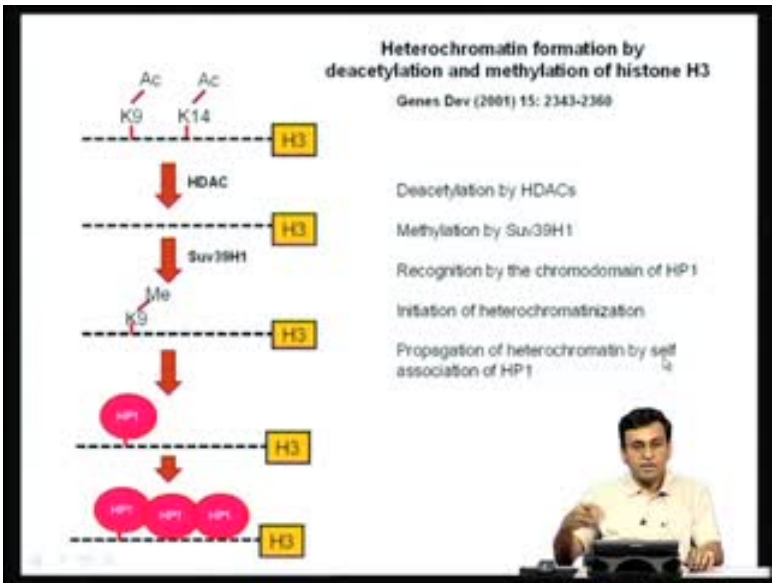
Similarly, methylation of lysine residues, in the case of histone, recruits proteins, which contain what are called chromodomains, and these proteins, then, bring other players into the field, and then depending upon which kind of lysine is methylated, it can either result in transcription activation or transcription repression.

So, the chromodomains specifically binds the histone tails bearing methyllysine. So, it is very specific for methyllysine residues, and the chromodomains have the highest affinity for tri-methyllysine, and the lowest for monomethyllysine.

So, when you have three methyl groups of a lysine, the chances of each attracting a chromodomain-containing protein are much higher than when only one or two lysine residues are methylated. A number of such proteins which contain chromodomains have been identified, and some of the most well studied are: a protein called a heterochromatin protein 1, now abbreviated as HP1, and in the case of Drosophila, a protein called polycomb protein plays a very important role in gene silencing and heterochromatin formation.

In the later stages of this course, when we talk about gene silencing and other heterochromatin, so and so forth, we will discuss the exact mechanism why is these proteins actually bring about gene silencing, heterochromatinization, especially things like H chromosome activation, and so on so forth, but at this time point, just remember, proteins, chromodomain-containing proteins, the two very well studied examples are heterochromatin protein 1 or HP1, and polycomb protein or PC.

(Refer Slide Time: 17:31)



Both HP1 and PC specifically recognize the lysine 9 and lysine 27 of histone H3. So, the recognition of these proteins of these particular methylation groups result in transcriptional repression and heterochromatin formation. So, I just now given a cartoon here, just to give an idea how exactly heterochromatin formation is, actually, initialized, because this results, the usually, heterochromatin involves more than one type of histone post translational modification.

We should, let us say, for example, there is a euchromatin region, and I told you very clearly, in euchromatin region, usually, the lysine residues are in acetylated form. So, let us see, for example, you have a euchromatin in which the histone H3, the lysine 9, and lysine 14, are acetylated by a specific HAT histone acetyltransferase.

Now, if a heterochromatinization has to be initiated in this particular locus, what happens? The first, the acetylated groups have to be removed. Therefore, a histone deacetylase now removes the acetyl groups from this particular histone lysine 9 and lysine 14 residues, and these deacetylated lysine residues now serve as the signal for recognition by a histone methyltransferase called SUV39H1.

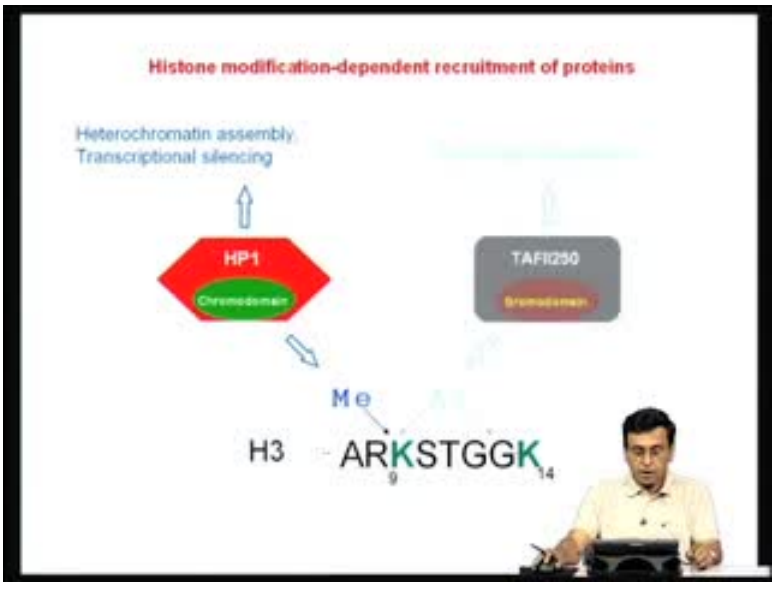
So, the deacetylated lysine 1 and deacetylated lysine 14, now, is recognized by a histone methyltransferase, and this histone methyltransferase is now specifically methylates lysine 9, and methylation of this lysine 9 now attracts its chromodomain-containing

protein called HP1. Now, HP1 comes and binds this lysine 1, and then, through subsequent methylations and multimerization of HP, heterochromatinization now spreads.

So, this is how entire chromosomes can be inactivated. So, you can see, the initiating event in heterochromatinization is deacetylation, followed by methylation, and then recognition of this methylated histones by chromodomain-containing proteins resulting in a chain of events, leading to heterochromatinization of the particular region.

So, you can see, how two specific to post translational modification act in a coordinated manner to bring about a specific physiological response. In this case, deacetylation, followed by methylation of specific lysine residue, can result in heterochromatinization. So, I just listed here the heterochromatin formation, in one particular case, results, first involves the deacetylation by histone deacetylases, then methylation by a histone methyltransferase, then recognition of these methylated lysines by a chromodomain-containing protein like HP1, then initiation of heterochromatinization, and then, further propagation of this heterochromatinization by self-association of HP1.

(Refer Slide Time: 19:52)



So, I just mentioned in another cartoon, the same example, where to actually show how specific histone post translational modifications can either result in transcriptional activation, or transcriptional gene silencing, or heterochromatin formation, depending

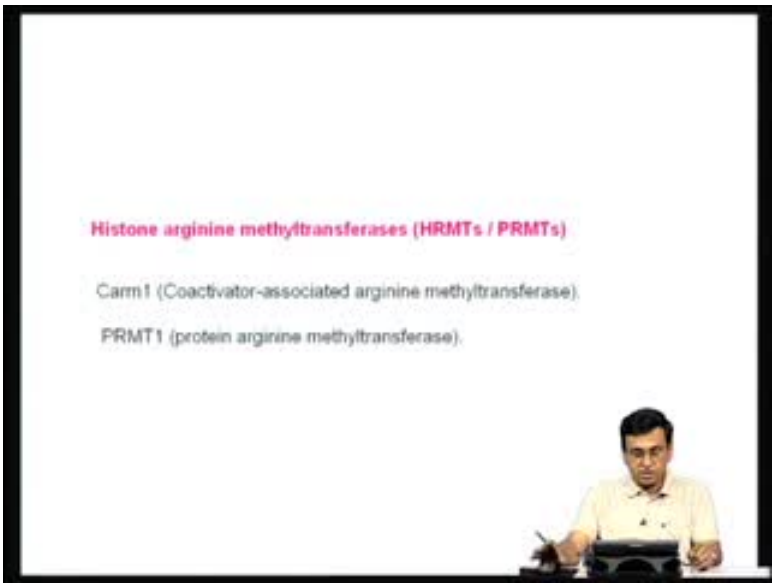
upon what kind of post translational modification is involved. In this particular example, for example, lysine 9, lysine 14 acetylation can actually result in the recruitment of a specific TAF, subunit of TAF.

The TBP-associated protein subunit 250, TAFII250, which, through its bromodomain, recognizes the acetyl groups of these 2 lysines, and this now recruits other components of the general transcription factors and RNA polymerase, and transcription activation takes place.

On the other hand, if this lysine 9 is actually methylated, instead of recruiting a bromodomain-containing protein, now, this now becomes a signal for a chromodomain-containing methyltransferase, and then this chromodomain-containing HP1 protein now recognize this methyl containing lysine, and then it brings about heterochromatin formation and gene silencing. So, in one instance, the same lysine residue, when acetylated, now becomes a target for recognition by a bromodomain-containing transcription activation, and the same lysine residue, when it is methylated, becomes a target for a chromodomain-containing protein, resulting in gene silencing and heterochromatin formation.

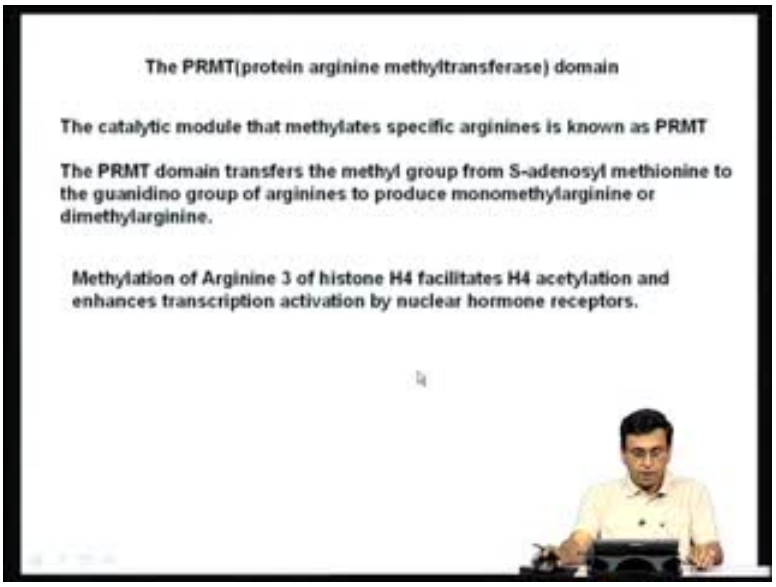
So, you can see how, just modification of histone can, actually, either result in transcription activation or transcription repression, depending upon what kind of proteins are recruited to this particular modified lysine residue of the histones. So, this is where the chromatin structure and histone post translational modification can have a very significant influence on the regulation of gene expression in chromatin templates.

(Refer Slide Time: 21:35)



So far, we discussed about histone lysine methyltransferases, and how they are affect gene regulation. Now, let us talk briefly about histone arginine methyltransferases, which specifically methylate arginine residues of histones. At least two well-known histone arginine methyltransferase have been characterized; one of them is called as the coactivator-associated arginine methyltransferase, another is called as the protein arginine methyltransferase 1 or PRMT1. This have been very well studied.

(Refer Slide Time: 22:03)



Now, all these arginine methyltransferases or protein arginine methyltransferases contain a particular domain called as PRMT domain, which refers to protein arginine methyltransferase domain. The catalytic module that methylates specific arginines is known as PRMT, and this PRMT domain actually transfers the methyl group from s-adenosyl methionine to the guanidine group of arginines, to produce monomethylarginine or di-methylarginine.

Now methylation, for example, that has been shown that methylation of the arginine 3 residue of histone H4 can facilitate H4 acetylation and enhance transcription activation by nuclear hormone receptors. Now, remember, here, in this initial series of lectures, I am actually introducing you to specific concepts. Now, later, in the classes, we will revisit each one of these post translational modifications. We are going to take specific example, for example, when I say nuclear hormone receptors, we are going to spend at least 1 or 2 hours, in the later stages of this course, to actually discuss about how transcription activation is brought about nuclear hormone receptors.

How on a nuclear hormone receptor binds to specific target sequence of DNA? What kind of histone modifications take place? What kind of proteins are recruited? And, what is the exact mechanism by which transcription activation is affected by nuclear hormone receptors, and so on and so forth.

So, we will discuss in more and more detail, the specific mechanisms by which the transcriptional activators bring out histone covalent modifications, leading to either activation or repression of transcription. But, what you are now, right now, studying is actually kind of a foundation for you to understand some more advanced concepts of eukaryotic gene regulation.

(Refer Slide Time: 23:38)

Now, I am giving you one more interesting example, just to give you an idea how important is these histone methyltransferases can be in certain specific physiological responses. Now, let me give one example, where people have actually shown that histone methylation plays a very important role during development, especially early embryonic development.

For example, there is a mammalian histone methyltransferase known as G9a, which specifically methylates the lysine 9 of histone H3, and it has been shown, the G9a actually represses transcription of many genes involved in early embryonic development. So, methylation of lysine 9 of histone H3 by G9a in the promoters of can actually bring about repression of a number of genes, in the case of which are, actually, play very important role in early embryonic development.

So, if you want to now really ask the question how important is this particular methyltransferase, what is the physiological significance of methylation of lysine 9 of all these genes? Because many of these experiments are actually done in vitro, so just because a protein is getting methylated or phosporylated in vitro by a specific enzyme, does not really mean that it has a really physiological response.

So, to really understand whether this methylation by this particular enzyme really has any meaning or any physiological relevance, people actually knocked out this particular

enzyme, this particular histone methyltransferase, and when you knock out this particular methyltransferase, you would have to look in such kind of a mice, so severe growth retardation and early lethality. So, they die at a very early stage of embryonic development. So, they are embryonic lethals.

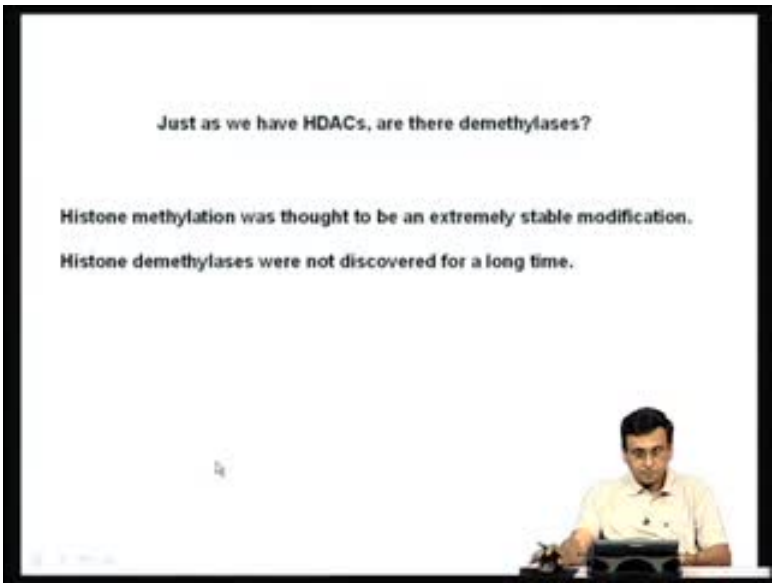
And similarly, if you knock out this G9a in embryonic stem cells, embryonic stems are very important. They are the pluripotent stem cells, which ultimately give rise to all kinds of tissues during the development, and if you knock out this G9a methyltransferase embryonic stem cells, those stem cells should, actually, can demonstrate there is a very reduced lysine 9 methylation of the histone H3, indicating that G9a is the dominant methyltransferase that specifically methylates the lysine 9 of histone H3 in vivo.

So, it is a major enzyme. You have the major histone methyltransferase that specifically methylate this particular lysine residue, and if you knock out this particular enzyme, there is a drastic reduction in the lysine 9 methylation of histone H3, clearly indicating that this enzyme has a very important physiological role, and if you knock out this, the mice have a very severe phenotype.

So, in the absence of G9a, H3-K9 methylation was drastically reduced in the euchromatic regions, indicating that euchromatic histone H3-K9 methylation regulated by G9a is actually essential for early embryogenesis.

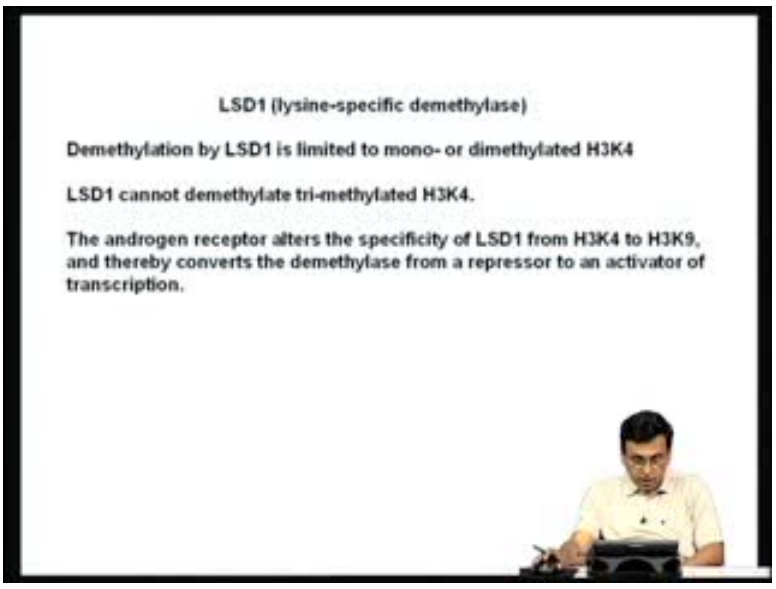
So, you can see, how both are by using a combination of in vitro experiments, where you actually demonstrate the enzyme specifically methylates a particular residue, and by actually doing experiments in whole animal models by knocking out, you can clearly demonstrate that this particular histone modification by a specific enzyme has a very high physiological significance, because you get a specific phenotype, and this entire information is actually taken from this particular journal.

(Refer Slide Time: 26:44)



So, the question, now, is– just like we have histone deacetylases, which remove the acetyl groups added by histone acetyltransferases or HATs, similarly, if you are adding a methyl group histones, there should also be enzymes which remove the methyl group. So, just I have deacetylases, are there demethylases?

(Refer Slide Time: 27:18)



Now, for a long time, people did not identify histone demethylases. In fact, many people thought histone methylation is a very stable post translational modifications, and probably, it is a kind of a permanent phenomena. But soon, histone demethylases were

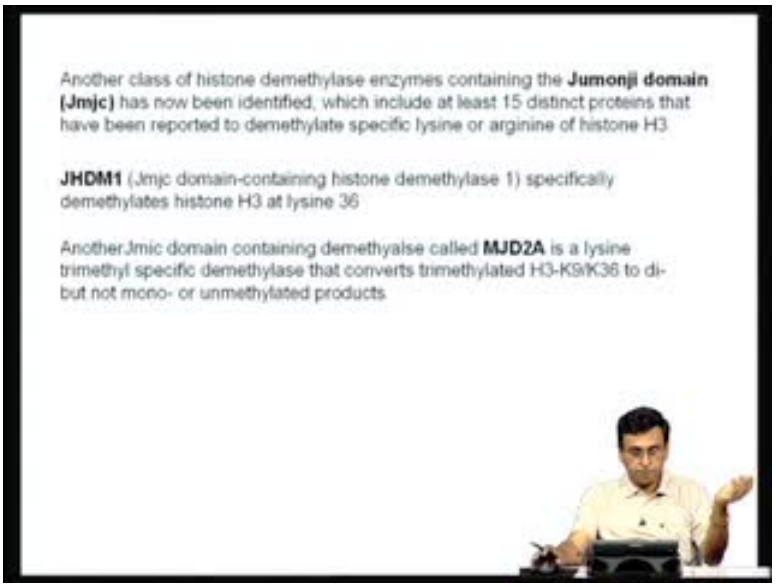
indeed discovered. In fact, the first histone demethylase which was discovered is an enzyme called LSD1, which stands for lysine specific demethylase.

So, demethylation by LSD1 is limited to either mono or di-methylated lysine 4 of histone H3, H3K4, and LSD1 cannot demethylate tri-methylated H3K4. So, we still not know once a tri-methyllysine is formed, how is it demethylated, we do not know still. This particular enzyme can only demethylate mono or di-methylated lysine 4 of histone H3, and there are examples to actually show this dmethylation by this enzyme can actually have a important role in gene regulation.

For example, in the case of androgen receptor, which is a member of the severe hormone receptor family and is a transcriptional activator, androgen receptor goes and binds to specific sequences called androgen receptor response elements, and binding of this androgen with response elements results in activation of specific genes. And androgen receptor was shown to alter the specificity of LSD1 from H3K4 to H3K9, and by this particular mechanism, it converts a demethylase from a repressor to an activator of transcription.

So, in general, demethylation of LSD1, demethylation by the LSD1, often, can actually result in repression, but in the presence of androgen receptor, the LSD1, instead of demethylating H3K4, now demethylates H3K9, and this actually results in activation of transcription, rather than repression of transcription.

(Refer Slide Time: 29:06)



So, you can see, the same enzyme can change in substrate specificity in the presence of another protein. Normally, it only demethylates H3K4, but in present of androgen receptor it can actually demethylate H9, and these two can easily elicit totally different kind of physiological responses. Now, if you start now looking at some of the literature and journals, you will find there are number of such examples where, although we discussed here only about some basic phenomena, there are variations of many of these phenomena. The same enzyme, depending upon the physiological situation, depending upon it is interacting proteins, can change its specificity and can bring about different physiological responses.

So, in addition to LSD1, once LSD1 was identified as a histonedemethylase, many other demethylases were also discovered. For example, a new class of histone demethylase enzymes containing what is known as, say, jumonji domain, has now been identified.

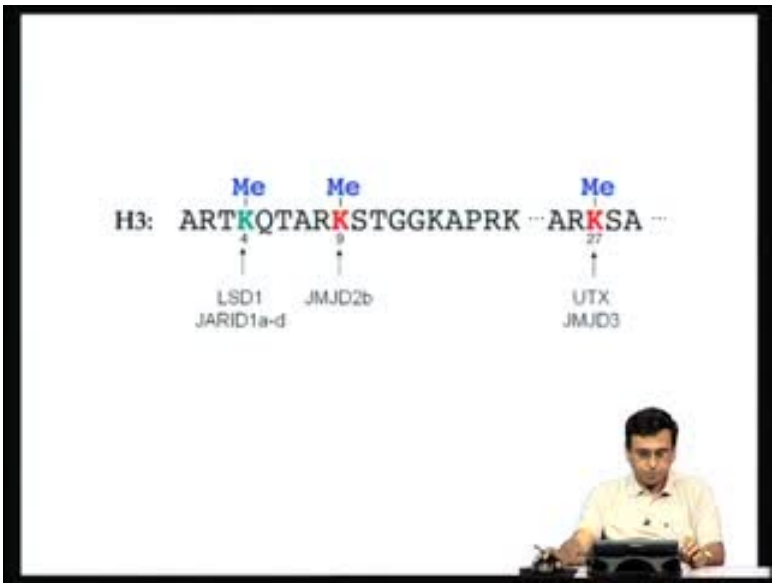
So, you have to now remember, some of these nomenclatures are very important. Bromodomain, chromodomain, SET domain, jumonji domains– all these domains are very important to have very important functions. Some are for recognition specific histone modifications, or some of them are very important catalytic domains of key enzymes involved in histone modifications.

So, a new class of histone demethylase enzymes containing a new catalytic domain called jumonji domain have also been identified. There are at least 15 such proteins which contain this particular catalytic domain, and they specifically demethylate either a lysine or arginine residues of histone H3.

So, one of such enzymes called JHDM, which actually stands for Jmjc containing jumonji domain-containing histone demethylase 1, specifically demethylates H3 at lysine 36, and another such histone demethylase, called MJD2A, is a lysine tri-methyl specific demethylase, that converts tri-methylated HK H3K9 or K H3K36 into di-methyl, but not to mono or unmethylated products.

So, a trimethyllysine is converted to dimethyllysine, but whether this dimethyllysine will further get converted to mono, or methyl, or unmethylated, is not known. So, there are demethylases– just as we have deacetylases which remove the acetyl group hired by histone acetyltransferases, there are demethylases which specifically remove the methylation methyl groups added by histone methyltransferases.

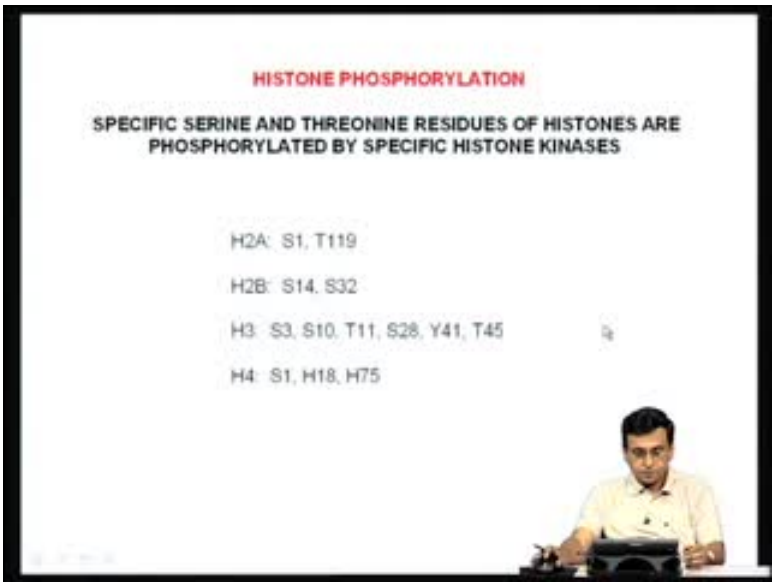
(Refer Slide Time: 31:04)



So, I just given the… whatever I told you, so far, I will show you in form of cartoon. For example, in the case of histone, the residues that are normally methyl are H4, lysine 4, K9, and K27, these are residues which are normally undergoes methylation, and the

enzyme which demethylates lysine 4 is called LSD1; the enzyme which demethylates lysine 9 is JMJD2b; and enzyme which demethylates lysine 27 is JMJD3.

(Refer Slide Time: 31:34)



So, we have, so far, discussed two important post translational modifications that happen for histones. One is the histone acetylation, another is the histone deacetylation, sorry, histone methylation. And we also discussed, there are enzymes which actually remove acetyl groups, and these are called HDACs or histone deacetylases. There are enzymes which remove the methyl groups, and these are as called histone deacetylases.

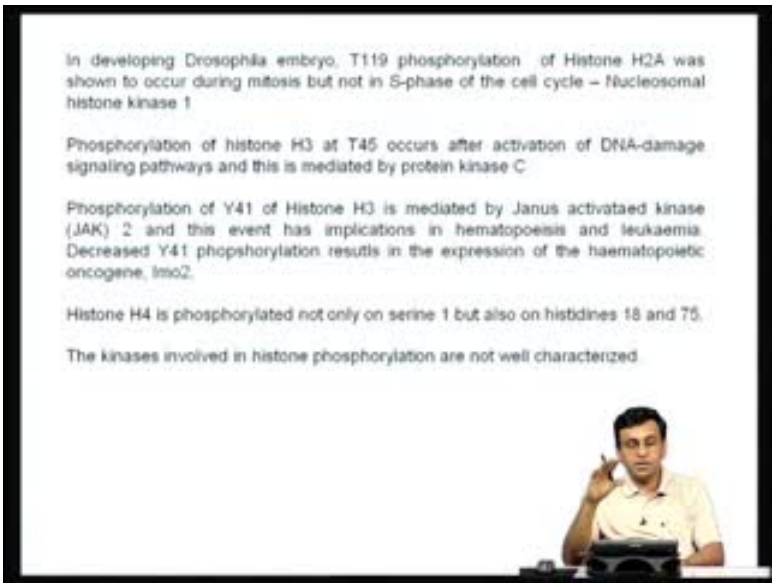
Now, we will now talk about one more histone covalent modification, namely histone phosphorylation. So, specific serine or threonine residues of histones are phosphorylated by specific histone kinases. So, just as people discovered histone acetyltranferases and histone methylases, people have also identified histone kinases, which specifically add a phosphate group to specific serine residues, in various histones. And I just listed here some of the serine and threonine residues, which are found to be phosphorylated in histones.

For example, in the case of histone H2A, the serine 1 and threonine 119 have been shown to be phosphorylated. Similarly, in the case of histone H2B, serine 14, and serine 32 are can be phosphorylated in vivo. Similarly, in the case of H3, serine 3, serine 10,

threonine 11, serine 28, tyrosine 41, threonine 45, remember, there are 3 amino acids in proteins, which can be phosphorylated serine, threonine, and tyrosine.

So, the enzymes which phosphorylates serine is called serine kinases, serine protein kinases, serine protein kinases, and there enzyme which phosphorylate tyrosine residue called tyrosine protein kinases, since, similarly, because of H4 S1, and very rarely, histidine also can be phosphorylated. And in one case, for example, the histidine 18 of H4 was also found to exist as a phosphorylated state, and so is histidine 75.

(Refer Slide Time: 33:26)



So, histidine, serines, and threonines in selective histones have been showed to be phosphorylated, but what are the actual enzymes which phosphorylate this histones have not been identified, or not been characterized very well. So, the histone acetylation and histone methylation, and the enzymes doing these functions have been very well characterized, but the kinases, which are actually phosphorylate histones, have not been very well studied.

Now, just to tell you that histone phosphorylation also has very important physiological significance, I am just giving you couple of examples here. For example, in the case of Drosophila embryo, during development, threonine 119 phosphorylation of histone H2A was shown to occur during mitosis, but not in S phase of the mitotic cycle, and this

threonine 119 phosphorylation was shown to be brought about by an enzyme called nucleosomal histone kinase 1.

So, histone phosphorylation has been documented in specific stages of cell cycle in Drosophila embryonic development. Similarly, in the case of H3 phosphorylation at T45, this was shown to occur after activation of DNA damage in signaling pathways, and this is mediated by protein kinase C. So, histone phosphorylation is not only involved in gene regulation. It is also known, a number of other physiological responses like cell cycle regulation, DNA damage signaling, and so on and so forth.

Similarly, phosphorylation of tyrosine 41 of histone H3, which is actually carried out by an enzyme called janus activated kinase or JAK2, and this has implications in hematopoiesis and leukemia. This is just an example to tell you that phosphorylation of these particular histones can actually result in the modulation of expression of genes involved in hematopoiesis, because if you now tinker with this particular phosphorylation, it can actually result in leukemia.

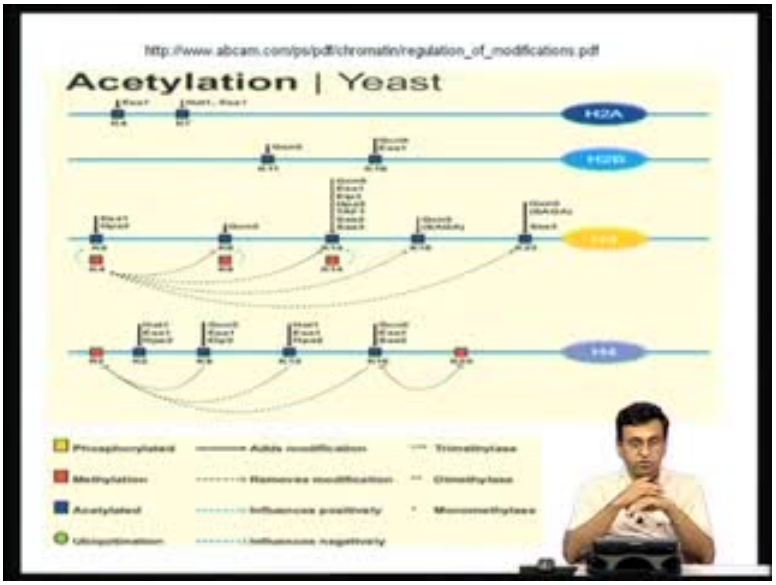
In fact, people have actually shown that if there is a decreased phosphorylation of tyrosine residue of histone H3, this actually results in the expression of an oncogene called lmo2. Now, lmo2 is a very important oncogene, and if you over express lmo2, it results in leukemia.

In fact, in some of these gene therapy trials, when they try to actually cure a disease called severe combined immunodeficiency syndrome for… by what is called as retroviral gene transfer, people have actually found out that this gene actually goes and integrates near this lmo2 gene, and results in activation of this lmo2 gene, and as a result, some of these patients, which actually underwent gene therapy for severe combined immunodeficiency syndrome, they actually develop leukemia.

So, this lmo2 is a very important oncogene, and if you activate this oncogene or if you over express the oncogene, it results in leukemia, and histone phosphorylation has been shown to play a very important role in the modulation of expression of this lmo2, and this lmo2 gene has a very important role in normal haematopoiesis.

Similarly, histone H4 is phosphorylated on serine 1, but also on histidine 18 and 75, but unlike the acetylation, deacetylation, methylation, and demethylation, somehow, the histone phosphorylation has not been very well studied, and the actual kinase, which actually involved in this histone phosphorylation, are not that well characterized compared to the other post translational modifications.

(Refer Slide Time: 36:28)



Now, if you actually go into the literature, you realize that all these post translational modifications, what I told you, whether it is acetylation, deacetylation, phosphorylation, methylation, or deacetylation, they have all been very well studied. In fact, today, the literature is flooded with examples of how kinds of histone modifications are taking place, and what are the important physiological responses. In fact, there are very nice websites; in fact, there are companies which actually sell antibodies, which specifically recognize a specific post translational modification of a specific histone.

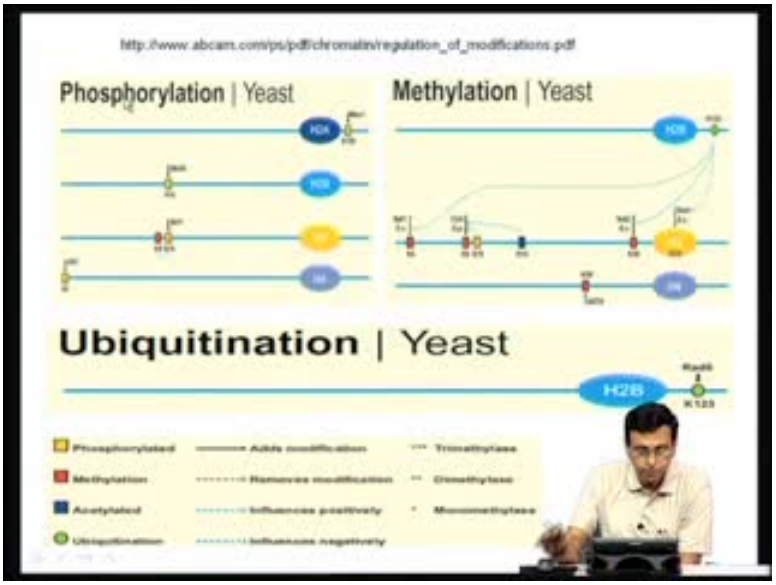
For example, there are antibodies which will specifically recognize proteins, in which the lysine 4 is acetylated, or lysine 27 is acetylated, or a specific lysine residue is methylated, or a specific lysine residue is phosphorylated. So, using these antibodies, people have actually identified what kind of genes are associated with these particular post translational modifications.

For example, there is a company called abcam, which actually sells many of these antibodies. Again, these peptides, which contain specific post translational modification of histones, and they give very nice charts, what is called as a histone modification charts. They actually telling, and these are constantly updated, and this…, these charts actually tell you what is the important post translational modifications in histones have been studied so far, and what are the enzymes which are actually involved in this, and what kind of transcription factors have is to be shown to be associated with this particular histone post translational modification.

For example, in this particular chart, it actually shows you what are all the acetylations of histones which have been reported, so far, in the case of yeast, for example, H2A, H2B, H3, and H4. You can see here, lysine 14 we have, which are all the transcription factors, are co-activators, which are known to be recognized involved in the lysine 14 acetylation.

Similarly, gcm 5, which is being shown to be involved in the recognition of lysine 18, and so on and so forth. So, I am not going to the details, basically, to tell you that there are now information available; literature, where you can actually look up, and then see which are all the residues and various histones, which are acetylated in species like yeast, mice, or humans.

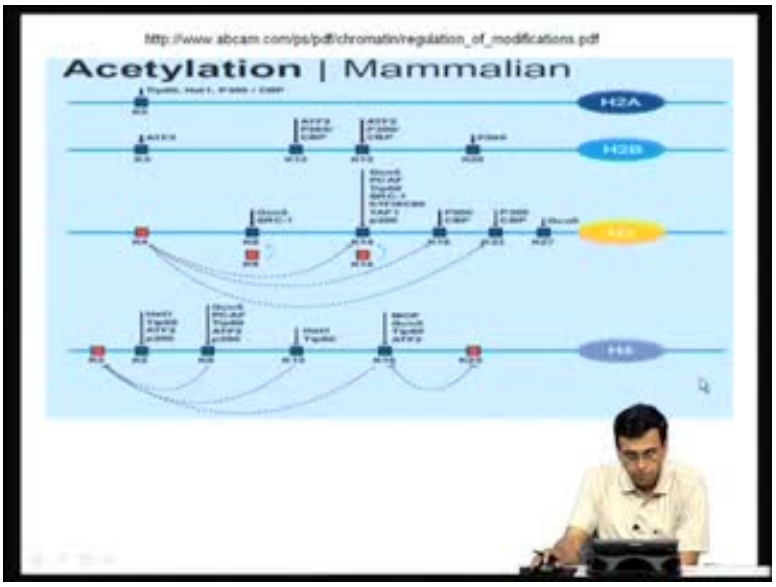
(Refer Slide Time: 38:42)



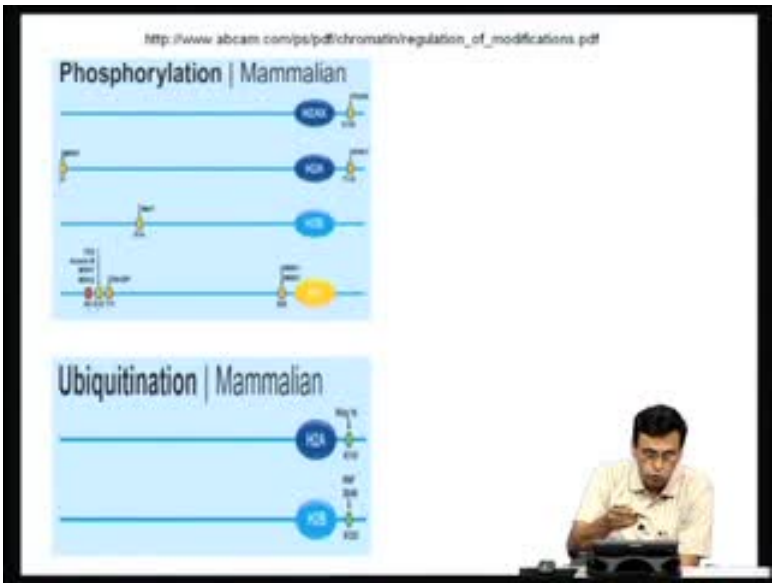
This chart just tells you what kind of residues are acetylated in yeast. Similarly, residues which are phosphorylated in the case of yeast histones are shown in this chart; residues which are methylated of different histones are shown in this chart; and residues which are ubiquitinylated. Ubiquitin, again, ubiquitinization is a signal for degradation of proteins.

And ubiquitination of histones also play a very important role in regulation of gene expression, and residues involved in this, which are specifically ubiquitinated, are also shown in this particular chart.

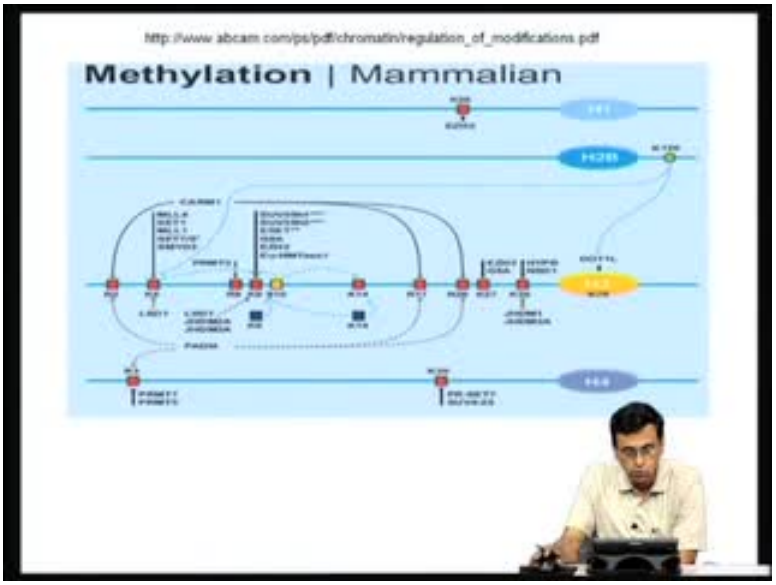
(Refer Slide Time: 39:09)



(Refer Slide Time: 39:18)

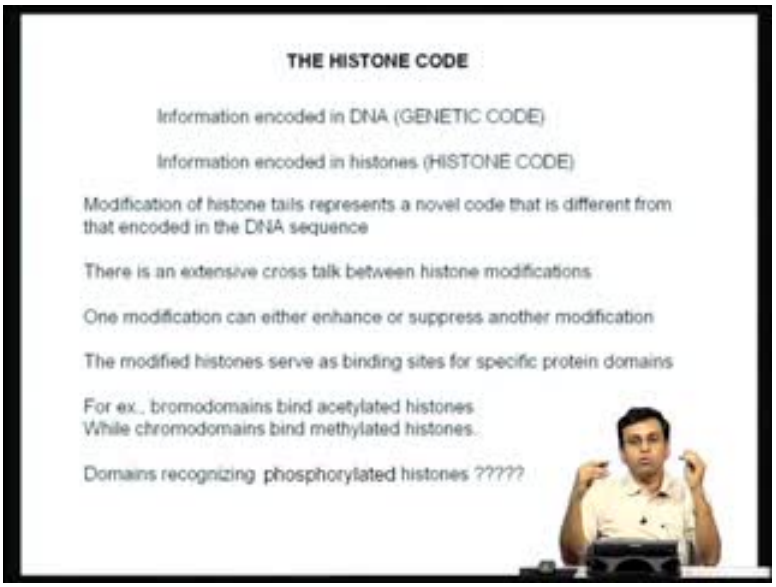


(Refer Slide Time: 39:23)



Similarly, there are charts which clearly tell you which are all the acetylated residues, in the case of H2A, H2B, H3, and H4, in the case of mammalian histones. Similarly, phosphorylated histone residues in mammalian histones, ubiquitinated residues in mammalian histones, and which are the residues are methylated in mammalian histones, and so on and so forth.

(Refer Slide Time: 39:28)



So, what I have been discussing, so far, is, at least, I have discussed about three major mechanisms, by which the regulation of gene expression can take place. We began telling you that chromatin structure can be modified by specific post translational modifications of histones, and post translational modification of these histones can have a very important role in regulation of gene expression.

And I mentioned to you that one of the most important, or one of the first histone post translational modification have been identified, is the histone acetylation. And, in fact, one of the first histone acetyltransferases which have been discovered is a protein called dcm5p, and histone acetylation is brought about by a group of enzymes called histone acetyltransferases, and acetylation of histones can actually result in, generally results in the activation of transcription.

Similarly, one acetylated residues, histones are acetylated, there are a group of enzymes called histone deacetylases, which actually remove this histone acetyl groups on this histone, and this, somehow, has HDACs or histone deacetylases, and these HDACs, actually, play a very important role in deacetylation or repression of transcription, and then, we discussed one more post translational modification of histone, namely, histone methylation. And I told you, there are enzymes called a histone methyltransferases, which specifically recognize lysine residues or arginine residues of histones, and then add 1, 2, or 3 methyl groups.

Accordingly, you have a monomethyllysine, dimethyllysine, or trimethyllysinelysine–containing histones, and methylation, unlike acetylation, can either result in activation of transcription or repression of transcription, and there are also enzymes called demethylases, which remove these methyl groups, and again, depending upon situation, methylation or demethylation can either result in activation or repression of transcription.

And histone methylation plays a very important role in heterochromatin formation or gene silencing, and we also recognized that there are proteins; there are specific sets of proteins which specifically recognize either acetylated histones or methylated histones, or the proteins which recognize acetylated histones have very important domains called as bromodomains, whereas proteins which recognize the methylated lysines, methylated residues of histones, contain what are called as chromodomains.

So, the histone post translational modifications serve as signals for recognition of specific proteins, through… which contains specific domain, and we also discussed specific catalytic domains like SET domain, and so on and so forth, which play a very important role in their enzyme activity.

Then, I discussed about histone phosphorylation, and how histone phosphorylation can also play very important role. Now, what all these things, actually, told you, is that histone covalent modifications has now brought in a new paradigm into the regulation of gene expression, and all of you, so far, must have studied what is known as a genetic code.

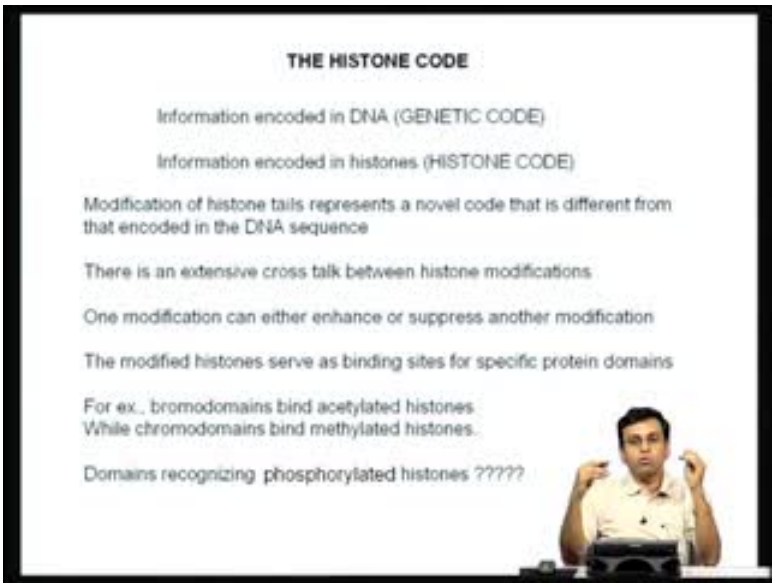
Now, what is genetic code? Genetic code is nothing but information, which is encoded with the DNA, the 4 alphabets A, T, G, C, and the way the codons are organized, the way this A, T, G, C sequence is read in triplets, is what is known as a genetic code, and this genetic code is actually forms the basis for what we are today. That is the major mechanism by which information is transmitted from one generation to another generation.

Now, once people have realized that in addition to the genetic code, in addition to the DNA sequence information, there also seems to be certain signatures, or certain information which is encoded in histones. And I can all that we have discussed in the last two classes is, basically, specific covalent modifications of histones can, actually, alter

gene expression, and therefore, they seems to be some kind of a code which is also embedded in the histones, and this is what is known to be known as histone code.

So, just as we have a genetic code, which is determined by the linear sequence, DNA sequence, now, we have now, people now begin to advocate a new theory, that just as we have a genetic code, there may be probably be a histone code, which actually specifies the specific post translational modifications that can ultimately have specific signature patterns of regulation of gene expression.

(Refer Slide Time: 39:28)



So, just as a genetic code, there may be a histone code. So, what is this histone code? This histone code basically tells you that modification of histone tails represents a novel code that is different from that encoded in the DNA sequence. So, this information is encoded in the DNA sequence is called as a genetic code, but covalent modification of the histones may also be a new code, and depending about what kind of covalent modification takes place, you can actually specify, or say, what kind of genes may be regulated, what kind of gene expression changes may take place, or what kind of physiological responses may be elicited.

So, there is an… and from one of the examples I have given, clearly, it is very clear that there is an extensive talk between covalent modifications. There are examples where I told you, methylation of one particular residue affects the methylation of an adjacent

residue, and there have been examples where, and I also told you, for example, in the case of heterochromatin formation, first, the acetyl group has to be removed, and the deacetylated lysines, now, serve as a signal for the recruitment of a methyltransferase, and then, this methyltransferase now acts a methyl group to this deacetylated lysine, which is then recognized by chromodomain-containing protein, and this way, ultimately, results in the gene silencing and methylation, and subsequent methylation and spreading of heterochromatin.

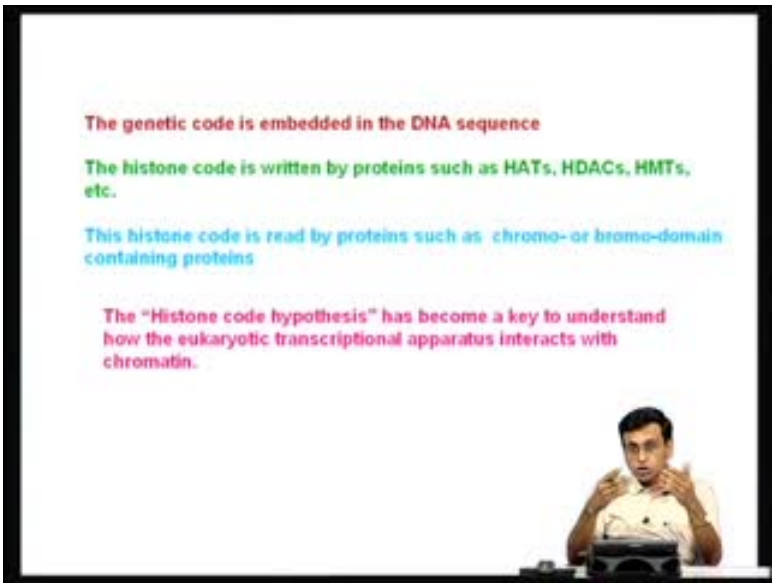
So, many a time, a combination of histones covalent modifications, together, can lead to a specific physiological, say, response, or a specific signal for activation or repression of gene expression. So, one modification of histone can either enhance or suppress another modification. This is what I just discussed.

Now, many times, the deacetylation, acetylation, methylation, demethylation, phosphorylation, they are all linked together. One modification of histone can actually lead to another modification, or one histone modification serves as a signal for another histone modification to take place, and it is the combined effect of all these post translational modifications that may serve as a signal for a specific event. It can, ultimately, it can result recruitment of either activators, or activation of transcription, or it can result in the repression of transcription.

And I also told you, very clearly, that the modified histones has served as binding sites for specific protein domains. So, once histones undergo post translational modifications, these modified histones now serve as signals to attract specific proteins, which interact with the specific modified histones through specific interacting domains, and I gave you two examples, where proteins which contain what are called as bromodomains, they specifically identified acetylated histones, and this usually results in the recruitment of histone acetyltransferases, which results in the activation of transcription.

Similarly, there are proteins which specifically recognize the methylated histones, and these proteins, which recognize methylated histones, contain what are called as the chromodomains. Now, although we discussed these things, and although we discussed, very briefly, about enzymes which phosphorylate histones, the exact mechanism by which these phosphorylated histones are recognized proteins is not very clear step.

(Refer Slide Time: 47:04)



So, what kind of kinases are actually involved in histone phosphorylation, and what kind of proteins have to recognize these phosphorylated histones is not very clear. So, there are three important things that I wanted to remember when we talk about histone covalent modifications: one is, the genetic code is embedded in the DNA sequence of an organism. It is the sequence of four letters A, T, G, C, namely, the DNA sequence, which constitutes the genetic code.

Whereas, in the last two classes, we have discussed what is called as the histone code, and this histone code depends on the post translational modification of histones, and this histone code is written by proteins, which can be histone acetyltransferases, histone deacetylases, histone methyltransferases, histone demethylases, or histone kinases.

So, just as the genetic code is written by the linear sequence of DNA, the histone code is written by enzymes which post translationally modify, which bring about post translational modification of histones and the histone code. Once you have these modified histones, once a histone code is written, this code is now read by specific proteins, and these proteins include the proteins such as those which contain chromodomain proteins, or proteins which contain the bromodomain proteins.

So, this kind of histone code hypothesis, now people, now this is still not very conclusively proven. Genetic code is very conclusively proven, but the histone code is

still at a very early stage, and people think that just as we have a genetic code, and depending upon what kind of a linear sequence, you can actually predict what kind of a protein it codes, what kind of a physiological function it may have, and so on and so forth.

Similarly, people are now advocating there is something called a histone code. So, people who are now saying, depending upon what kind of histone is getting acetylated, or what kind of residue in histones has been phosphorylated or it can be methylated, you can actually, it can have a specific physiological function, and this particular post translational modification of histones can actually serve as a code for a specific physiological response.

So, this histone code hypothesis has now become a key to understand how eukaryotic transcription apparatus interacts with chromatin. So, you can see, now, you have come a full circle, which started looking at gene regulation as if DNA is naked, and we started looking at what is happening at a core promoter sequences, what is happening at the upstream promoter sequences.

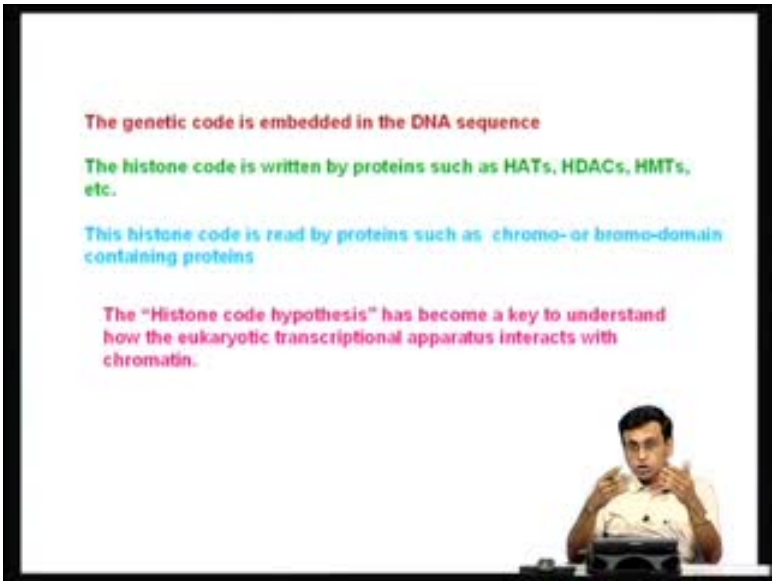
Then, we discussed about what kind of protein factors go and bind in the core promoter elements, what kind of protein factors go and bind to the upstream activation sequences, and then, I very briefly discussed about the link between activators and the general basal transcription machinery, and I told you, there could be proteins called as co-activators, which may act as a bridge between activators and the RNA polymerase machinery.

Then, I deliberately did not discuss this co activator concept in detail, because I thought I should bring in the chromatin concept first, because the major mechanism by which all the gene expression changes are taking place, the chromatin seems to be playing a very important role, and now, in the last two class, I have described how covalent modification of histones have become a very important player in the regulation of gene expression, and we have now reached a stage where people are now proposing that there is actual something called histone code.

So, just as you have a genetic code, probably, there is a histone code, and depending on what kind of post translational modification is taking place of a particular histone, a specific physiological response may be emanating, and I gave you at least half a dozen

examples, where, if you block a specific post translational modification of histone by knocking out a specific histone methyltransferase, or knocking off a specific histone acetyltransferase, you can actually bring about specific changes, and, in fact, people, see, the topic have chosen for this particular course is eukaryotic gene expression basics and benefits.

(Refer Slide Time: 47:04)



Now, this histone modification has now become the basis for the founding of a number of pharmaceutical companies. The number of pharmaceutical companies are now working to see whether you can develop actually inhibitors, which can specifically block the specific post translational modification of histones.

For example, I will come back and revisit some of the topics which I mentioned earlier. There are companies now, for example, which are working to develop and small molecules, which specifically block the activity of histone deacetylases, and, in fact, people have actually shown that many of this histone deacetylases are over expressed in a number of cancers.

And if you block using these molecules, if you block the expression of some of these histone deacetylases, you can actually prevent self-proliferation and smaller some of these inhibitors of HDACs– these are called as HDAC inhibitors– they can actually survive as anti-cancer agents, and, in fact, very surprisingly, some of these molecules

have even been approved now for human use, for certain treatment of certain kind of cancers.

So, this concept that histone, covalent modification of histones can affect gene regulation has now led to the prolification of a number of pharmaceutical companies, which are now trying to develop either agonists or antagonists, which can either affect the histone modification.

People are trying to modify histone acetylases, HATs. People are trying to develop inhibitors of HDACs, and depending upon what kind– whether HATs are active or HDACs are active in a specific kind of cancer, and by using such inhibitors, people can actually find about now thinking, you can actually use this basic research, or the knowledge arising out of the histone covalent modification, for the benefit of mankind.

And you can actually develop molecules which specifically inhibit some of the histone post translational modification, and this entire concept that histone modifications can, actually, which is now, say, which is the histone modifications are at a level which is different from that happening in the DNA sequence. This is different from a genetic code.

And this entire area, in fact, as now led to a new area called epigenetic regulation of gene expression. In fact, in next class, we are going to discuss what exactly we mean by epigenetic regulation of gene expression, and I am going to tell you that epigenetic regulation of gene expression is not really dependent on the linear sequence of DNA. It is actually depend more on the histone code, and there is one more important event that plays a very important role for epigenetic regulation gene expression, that is, DNA methylation

So, what we are going to discuss in the next class is that how histone modification and DNA methylation can, actually, play a very important role in epigenetic regulation of gene expression, and at the end of the next class, we are going to realize that all these things actually act in concerted manner.

We have histone modifications, DNA methylations, all these things in concert can actually determine whether a chromatin should become a euchromatin, or it should

become heterochromatin, or whether genes have to be activated, or genes have to be silenced. I think I will stop here.