Eukaryotic Gene Expression

Prof. P. N. Rangarajan

Department of Biochemistry

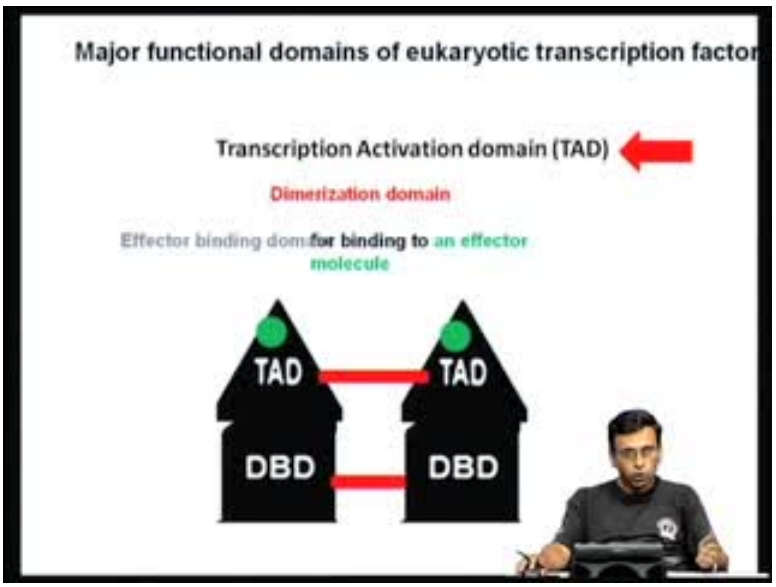
Indian Institute of Science, Bangalore

Lecture No. # 06

Eukaryotic Transcription Factors: Transcription Activation Domains

In today's lecture, we are going to discuss about eukaryotic transcription factors. I think in the previous lecture, I basically told you about the DNA binding domain; I told you there are two major domains of eukaryotic transcription factors, namely the DNA binding domain and transcription activation domain. We discussed very briefly what kinds of DNA binding domains are actually present in various eukaryotic transcription factors. We discussed about helix-turn-helix motif, zinc finger motif, leucine zipper motif and the helix-loop-helix motif, to mention a few. And then, we basically also discussed about the various experiments that people actually used to identify transcription factor binding sites in various promoters; and also, to identify the various functional domains in the transcription factors. So, today, we are going to focus about the other important domain in the transcription factors, namely the transcription activation domain.

(Refer Slide Time: 01:13)



Let us just recapitulate what we studied in the last class. What basically I told you in the last class is that eukaryotic transcription factors basically contain a DNA binding domain. And, through this DNA binding domain, these factors are able to go and bind to specific sequences in the upstream region of various promoters. And, this sequence-specific binding of a transcription factor to the promoter is the first step in the eukaryotic gene expression in the transcription activation process.

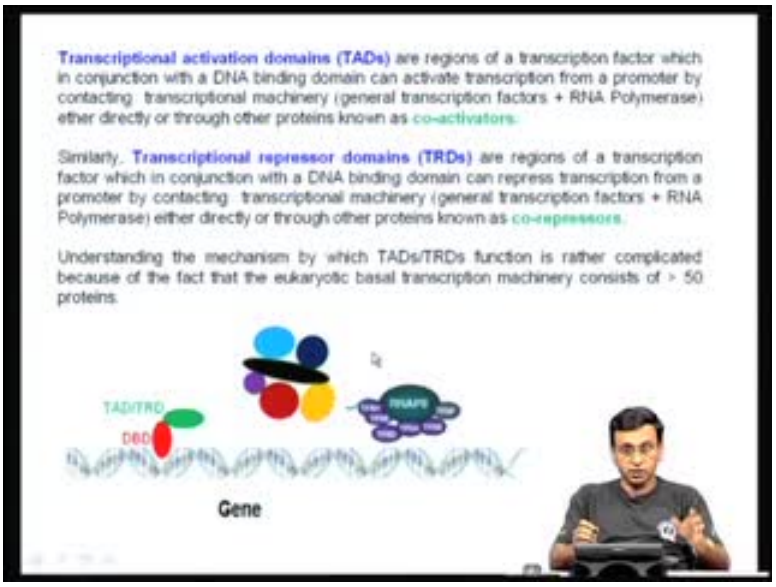
Today, we are going to focus about the transcription activation domain, which is actually another important key domain in the eukaryotic transcription activators. And, we also mentioned in the last class that in addition to these two domains, there are also many other important functional domains in transcription factors, for example, they have what is called as the dimerization domain; in some transcription factors, the dimerization domain may be present in the DNA bind domain; whereas, in some cases, it may be in the transcription activation domain. And, it is these DNA binding domain, which actually brings the two monomers. Many a times, the transcription factors function as dimmers, and through this dimerization domain, the two monomers come together. And, motifs like leucine zipper, helix-loop-helix are actually trying to bring these two monomers together, and then project the DNA bind domain in such a way that it is now able to go and then recognize this specific response element in the DNA.

And, I also told you that certain transcription factors like glucocorticoid receptor and other steroid hormone receptors also have a domain for binding to certain effector molecules like small molecules. For example, in the case of steroid hormone receptors, it is the steroid hormones like glucocorticoid receptor; it is glucocorticoid hormone; estrogen receptors – it is the estrogen hormone; and so on and so forth. So, you need to have some of these transcription factors have a domain for binding to the effector molecules or the small molecules to which small molecules like steroid hormones can actually bind. And, binding all these steroid hormones brings about a conformational change, and therefore, now this receptor or this transform factor can go and then carry out its functions.

Since we discussed in the last class, extensively about the DNA binding domain, we are going to skip that today and focus primarily on the transcription activation domain, and also, little bit about dimerization domain and effector binding domain, and so on and so forth. So, let us now really see what has been done and what kind of experiments were

actually done to understand the functions of various transcription activation domain in eukaryotic transcription factors.

(Refer Slide Time: 03:39)



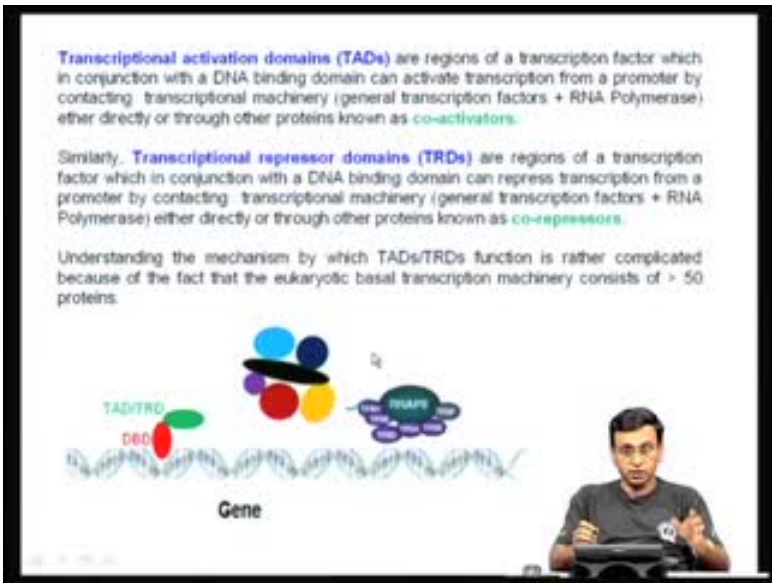
Now, let us first try to see how we can define a transcription activation domain. Now, transcription activation domains have actually abbreviated as TADs. They are actually regions of a transcription factor, which in conjunction with a DNA binding domain can activate transcription from a promoter by contacting transcriptional machinery; when I say transcription machinery, it means the general transcription factors on the RNA polymerase either directly or through other proteins known as co-activators. So, the function of a transcription activation domain is basically to contact the preinitiation complex in the core promoter region. And, this can happen either through direct interactions or through certain other proteins, which are called as co-activators. So, they actually act as mediators between the eukaryotic transcription factors and the basal transcription machinery.

Just as we have transcription activators, there are also transcription factors, which when actually bind to certain specific sequence, the promoter bring about repression of transcription. So, the regulation can be either positive or negative, and such transcription factors which actually result in the decrease in transcription activation is called as transcriptional repressors. And, such transcription repressors instead of a transcription activation domain, they actually possess what is called as a transcription repression

domain. So, transcription repression domains just as the transcription activation domains are regions of a transcription factor, which in conjunction with a DNA binding domain can repress transcription instead of activation from a promoter again by contacting the general transcription machinery or the RNA polymerase. And, in this case, again the interaction can be either direct or it may be mediated by molecules; in this case, they are called as co-repressors. So, the transcription activation domain or the transcription repression domain either can activate or repress transcription either directly by interacting with general transcription machinery or through molecules called co-activators or co-repressors.

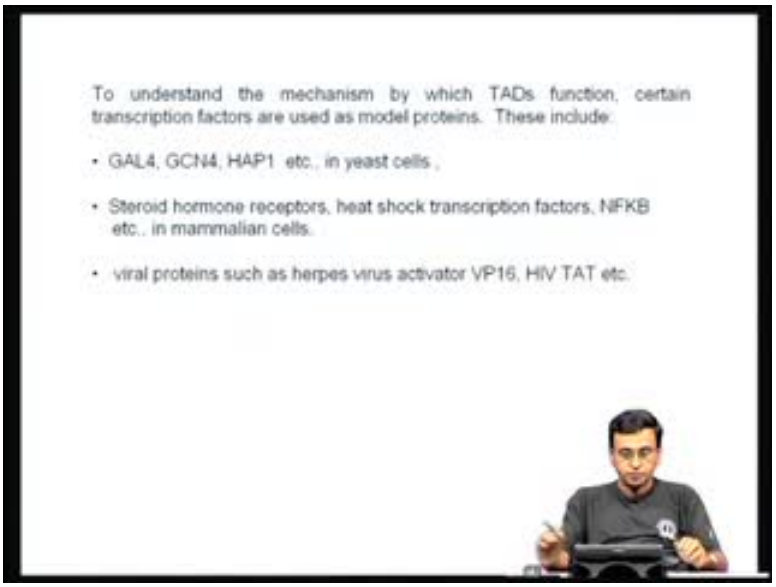
Now, the lot of effort has actually gone on to understand how actually this transcription activation domains or transcription repression domains actually function. No, it is not very easy to understand this, because see we know from the previous lectures, the transcription initiation requires at least 50 odd proteins. We have RNA polymerase, which itself is a multi-subunit protein; we have general transcription factors, which again each one is a multi-subunit protein. So, at least about 50 to 60 proteins are actually involved in transcription initiation in the co-promoter region. So, we are now talking about certain other transcription factors or transcription activators or transcription repressors, which can interact with any one of these 50 odd proteins. So, if you have to understand how these transcription activators are functioning, you need to understand which component of the general transcription machinery they are interacting with. So, the studies of these kinds of mechanistic insights are not very easy. And, lot of effort has actually gone on to understand the mechanism by which this transcription activation domains actually function.

(Refer Slide Time: 03:39)



So, what I (( )) showed in this cartoon is basically to summarize what I have told you so far. You have transcription factors, which contain a DNA binding domain and either a transcription activation domain in the case of transcription activators; or transcription repression domain in the case of transcription repressors. And, when they bind to specific sequence in the promoter region of a DNA, they can either directly contact the general transcription machinery or there could be what are called as co-activators or co-repressors, which again are multi-subunit complexes. And, because of these interactions that these transactivators or transrepressors interact with its mediators. These mediators in turn interact with the general transcription machinery and result in enhancement of transcription of that particular gene.

(Refer Slide Time: 07:17)



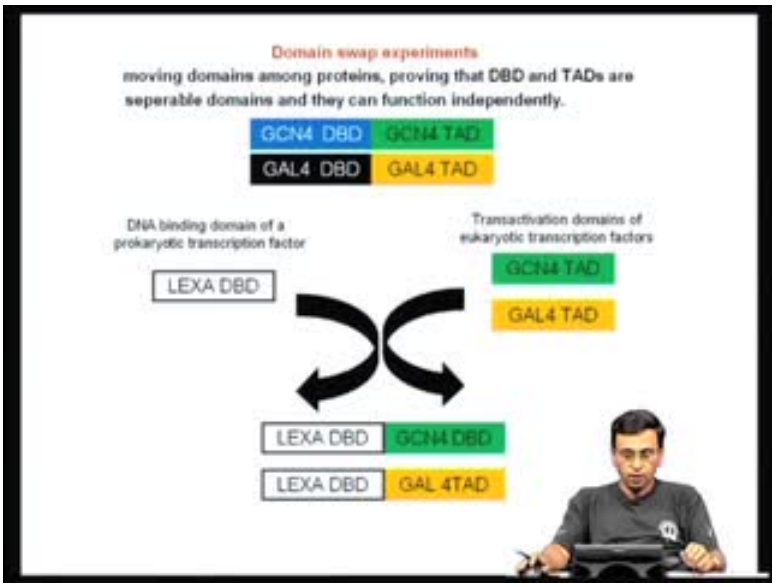
Now, to understand the mechanism by which these transcription activation domains function, we need to do… As I said, there are more than 2000 genes-specific activators that are been discoursed so far. So, what has been done is to take one or two of these transcription activators as examples, and then see how the trans activation domain in these particular trans activators actually function.

During the late 1990s or late 80s and early 90s, certain classical transcriptional activators have been chosen model system to understand the mechanism by which they function as transcriptional activators; to actually study what kind of transcription activation domains these proteins actually contain. So, for example, some of the classical examples, which people have been using for understanding the mechanism of transcription activation domains; or, if the yeast proteins like GAL4, which actually is involved in the upregulation of galactose metabolism genes; GCN4, again a very important transcription factor, which is actually involved in the upregulation of genes involved in general amino acid control in the case of yeast, because general amino acid control means when the yeast cells are grown in a medium rich in amino acids, none of the amino acids biosynthetic genes are turned down, because there is no need, because the amino acids are present in the medium. But, when the amino acids are not there in the culture medium, then all these amino acids biosynthetic genes are turned down. And therefore, the organism starts making their own amino acids and a transcription factor of GCN4 plays a very important role in this general amino acid control. Similarly, there is a

protein called heme activator protein. Here, just like in the case of steroid hormone receptors, the steroid hormone actors in activator; in the case of heme, there is a protein called heme activator protein or HAP1, whose activity is actually modulated by heme. Heme binding to this transcription factor modulates its activity. So, some of these transcription factors were actually used as model systems to understand the mechanism by which their trans activation domains actually function.

When you come to mammalian cells, the classical examples have been steroid hormone receptors, heat shock transcription factors, nuclear factor kappa-B. These are the proteins. Some of these proteins we have already discussed in the previous class very briefly and these again have been taken as examples to understand the mechanism by which some of these transcription factors are actually activating transcription. It is not only the eukaryotic transcription factors, people have also realized, there are many viruses, which actually infect these eukaryotic cells; many animal viruses; for example, you have herpes virus, HIV, Adeno virus, and so and so forth. It turns out, even these viruses actually contain transcriptional activator proteins, or even co-activator proteins. And therefore, even these viral proteins were actually used as model system to understand the mechanism by which these viral proteins are able to activate transcription of the viral genomes. And, some of the very well-studied viral transcription activators, for example, herpes virus activator, VP16, which is a very good model system to study some of the early work on transcription activation (( )) using the herpes virus VP16 protein. And, more recently, the human immunodeficiency virus of the HIV also contains a very pore turn transcription factor called TAT protein. And, this also has been now used as a model system to understand the mechanism by which this TAT protein actually enhances the rate of transcription of the HIV viral genome. So, a number of transcription factors of eukaryotic origin are used as the model system to understand the mechanism by which the transcription activation domains actually function.

(Refer Slide Time: 10:47)



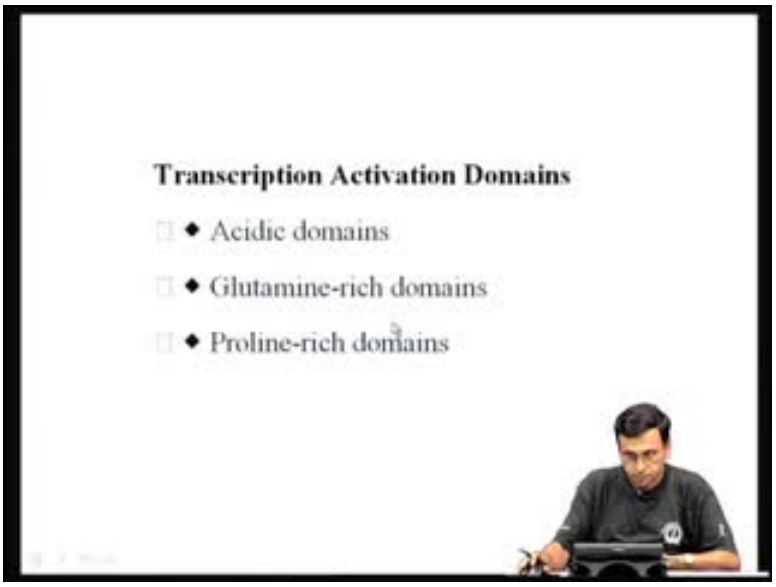
Now, few of the very interesting earlier experiments were done in these studies were actually being pursued, and one of the most interesting experiments that were done are what are called as domain swap experiments. Now, what do we mean by this? Now, I have already mentioned to you that in many of these eukaryotic transcription factors, they contain a DNA binding domain and a transcription activation domain, and these are modular in nature. And, it turns out these domains can actually function independently of each other. So, to actually demonstrate that these two domains, that is, the DNA binding domain and the transcription activation domain can actually function independently, what kind of experiments were actually done.

The most popular experiment is called as domain swap experiment, wherein you take the transcription activation domain of one transcription factor and link it to the DNA binding domain of another transcription factor. That is what I have shown in this particular cartoon. For example, you have the GCN4 and GAL4, which are very portent transcriptional activators in yeast cells and each of them have their own DNA binding domain and transcription activation domain, which I have shown in different colors. Now, what has been very interestingly done is the people have actually taken and the transcription activation domain of either GCN4 or the GAL4, which are of eukaryotic origin are actually fused it to the DNA binding domain of a transcription factor called LEXA, which is of prokaryotic origin. So, you have a DNA binding domain from a prokaryote and transcriptional activation domain from a eukaryote, and when you fuse

these things and then make this kind of a chimeras, where the LEXA DNA binding domain is fused to the GCN4 – sorry, this is not DBD; this should be transcription activation domain. Similarly, LEXA DNA binding domain is fused to GAL4 transcription activation domain.

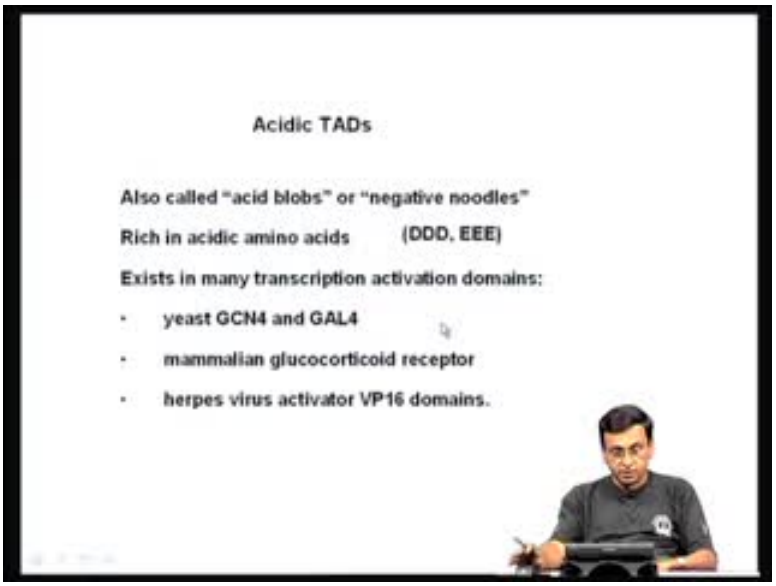
These chimeric proteins instead of activating from a GCN4 DNA recognition element or a GAL4 recognition elements, will now start activating transcription from a LEXA recognition segments; clearly telling that these transcription activation domains can function irrespective of what kind of DNA domain is attached to them. If these transactivation domains are attached to a GAL4 DNA binding domain, they will activate from a GAL4 response element, promoter of containing GAL4 response element. If the transactivation domain is attached to a GCN4 DNA binding domain, they can activate transcription from a GCN 4 recognition element in the promoters; or, if they are attached to even a prokaryotic DNA binding domain, they can actually activate transcription from the sequence recognized by the prokaryotic DNA binding domain. So, using these kinds of studies, people actually demonstrated very clearly that eukaryotic transcription factors contain the DNA binding domain in transcription factors, which are modular in the structure, modular in nature. And, these domains can be swapped between transcription factors, and they can function independently of each other.

(Refer Slide Time: 13:29)



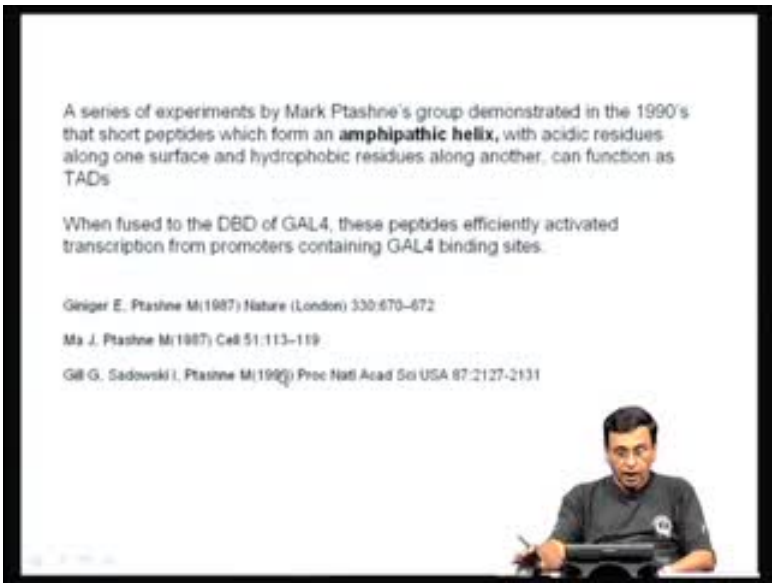
Using some of these domain swap experiments and many other deletion studies (( )) and so on and so forth, people have started understanding how exactly transcription activation takes place; how exactly this TAD actually function. The outcome of all these experiments is that at least three major transcription activation domains were actually identified in many of the eukaryotic transcription factors, and these were classified as acidic domains, glutamine-rich domains and proline-rich domains.

(Refer Slide Time: 14:01)



Now, what do I mean by this? What I mean by this is that the acidic transactivation domains are actually characterized by the presence of what are called as a negative charge. That means they are highly rich in acidic amino acids, such as the glutamic acid and aspartic acid. So, in transcription factors like yeast GCN 4, GAL 4, mammalian glucocorticoid receptor or the herpes virus VP16 transactivation domain, the transcription activation domain is highly rich in acidic amino acids, and therefore, these are called as acidic transcription activation domains.

(Refer Slide Time: 14:34)

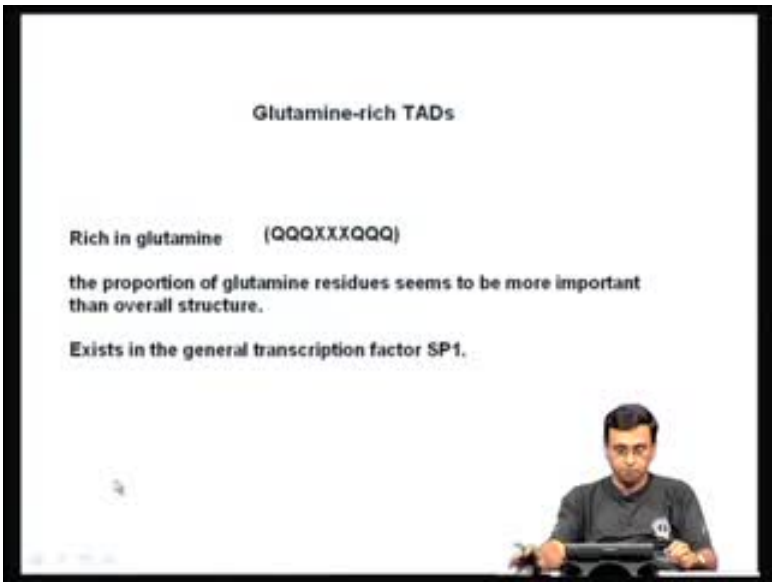


And, very interesting experiments were actually done in the late 80s or early 90s especially from the laboratory of Mark Ptashne, where he actually extensively studied some of these transcription factors, which contained the acidic activation domain. And, he actually created an artificial transcription activation domain, which contains what is called as an amphipathic helix. This is not a natural domain; it is constructed as an artificial transcription activation domain; it is basically a short peptide, in which the acidic residues are present on one surface of the helix and hydrophobic residues are present on the other surface of the helix.

Now, when you make this kind of amphipathic helix with acidic residues on one side and hydrophobic residues on another side, and this kind of amphipathic helix is the gene coding for this peptide, when you link it to heterologous DNA binding domain, it actually functioned as a transcription activation domain. So, this is actually to prove that many of the acidic amino acids present in the natural transcription activation domains, which are showed in the previous slide like the GCN4. GAL4, these acidic activation domains can actually mimicked by a synthetic transcription activation domain, which actually contains acidic amino acids on one side, hydrophobic amino acids on one side, clearly indicating that transcript presence of acidic amino acids is one of the major characteristic features of transactivation domains of eukaryotic transcription factors.

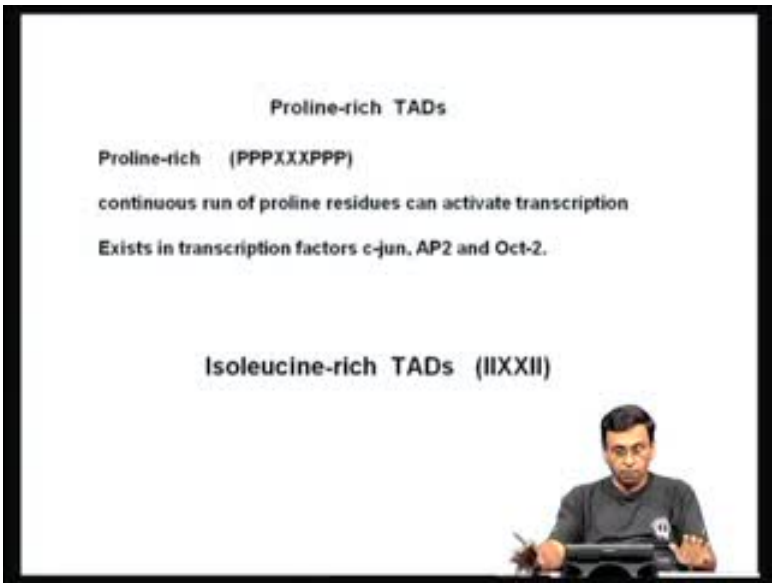
When fused to the DNA binding domain of GAL4, these peptides which contain amphipathic helix with acidic residues on one side and hydrophobic residues on other side, efficiently activated transcription from promoters containing GAL4 binding sites. These are all very interesting sites and those terms are very exciting results, because you can actually make synthetic peptides and demonstrate they can actually function as transcription activators in a cell-free transcription system or in a transcription system. So, those are all very exciting experiments, which were done in those days. If you want to really study more about those, you should have actually studied some of these original papers published from Ptashne's book; very interesting experiments were done.

(Refer Slide Time: 16:31)



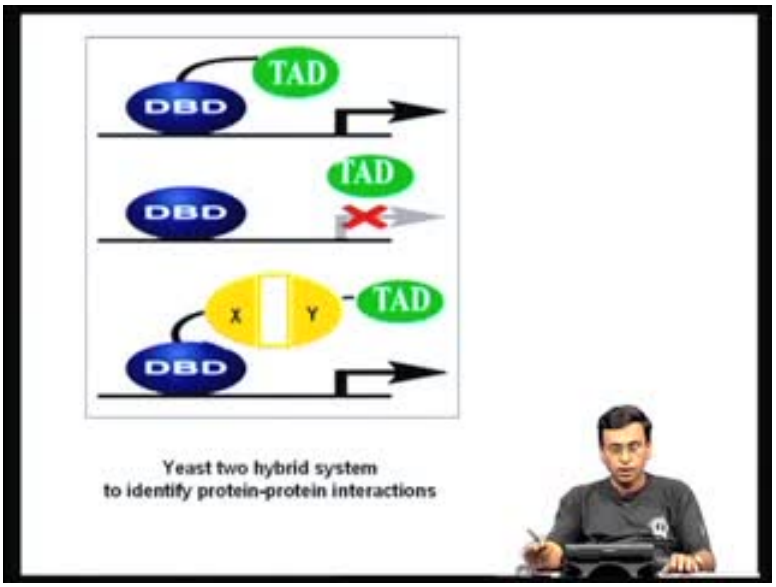
Similarly, certain transcription factors like SP1 actually contain a transcription activation domain, which are rich in glutamine. So, the proportion of glutamine residues seems to be more important than the overall structure. So, just as some factors have glutamic acid or aspartic acid or acidic amino acids, there are certain transactivation domains, which are highly rich in glutamine, amino acid.

(Refer Slide Time: 16:54)



And, there are many other transcription factors like c-jun, AP2 or Oct-2, which actually contain a proline-rich transcription activation domain. These are just examples I gave you. As people started studying more and more transcription factors, more interesting transcription activation domains were discovered. For example, there are some transcription factors, which contain isoleucine-rich region. So, we will not discuss all these transactivation domains, but just remember that very interesting outcomes came from some of these mechanistic insights and transcription activation domain functions, there are domains, which contain acidic residues; there are domains, which contains globulin residues; there are some, which are proline-rich residues. They all function as transcription activation domain when linked to heterologous DNA binding domain. And, you aside the active area of a promoter either in a cell-free system or in a transfection system in vivo.

(Refer Slide Time: 17:45)



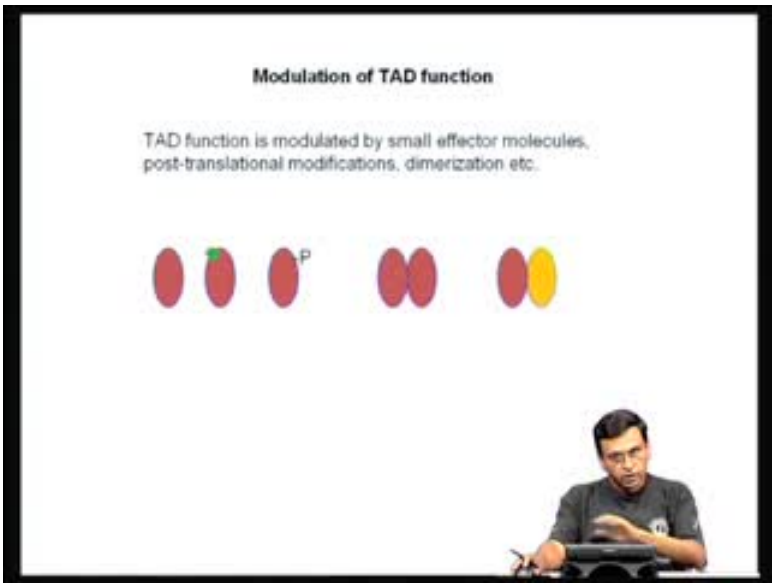
The outcome of some of the important experiments, which are actually done to understand a mechanism of transcription activation, there are also some very interesting benefits that came out of the studies, because wherever possible, I also mentioned some of the benefits that came out of the basic research. For example, one important benefit that came out of understanding that transcription factors actually have a modular structure is, we know, for example, from what I would have told you that when you have a DNA binding domain and when it is linked to a transcription activation domain, it can activate transcription. And, if these transactivation domain and DNA binding domain are not linked, then they cannot activate transcription; they have to be in physical proximity, only then transcription activation can takes place.

Now, using this particular principle, people have developed what is called as a yeast two hybrid assay to actually study protein-protein interactions. Now, what do you mean by that? Suppose for example, I am interested in a protein, let us say X, and I want to know are the proteins, which actually interact with these particular protein X in a particular cell. Now, understanding these protein-protein interactions are very important and as we study more and more about transcription regulations, you realize that the entire regulation of gene expression hinges on two things: one is DNA protein interactions and protein-protein interactions. It is these two interactions, which actually got the entire gamete of eukaryotic gene regulation.

We want to now study the mechanism by which transcription factors actually function, you need to identify what kind of proteins are actually interacting. For example, if I have to now understand what kind of protein the glucocorticoid receptor is interacting, because ultimately it has to do a transcription. I do a simple yeast two hybrid assay, where I take the transcription activation domain of the glucocorticoid receptor and fuse it to the DNA binding domain of a well-characterized transcription factor. Then, I have a yeast cell library, in which each cell expresses different kind of a protein, which is actually linked to the transcription activation domain; these proteins are actually link to the transcription activation domain of a well characterized transcription factor.

Now, as I told you, if the DNA binding domain has to activate transcription, the transcription activation domain has to be brought in close physical proximity of the DNA binding domain. So, when you actually transfect your protein of interest, in this case, for example, a glucocorticoid receptor linked to DNA binding domain or a transactivation domain of glucocorticoid linked to DNA binding domain, and if I introduce them due to yeast cells, and if there is a protein in yeast cell, which actually interacts the glucocorticoid receptor or protein X, then because of this interaction, the transcription activation domain will be brought in close proximity to the DNA binding domain and as a result, transcription activation takes place. And, you can actually use reporter genes like beta galactosidase and you can actually demonstrate blue color, indicating that you have actually identified a particular protein, which is actually interacting with this protein. So, this yeast two hybrid system became a very popular assay to study protein-protein interactions in eukaryotic cells. So, one of the major outcomes of demonstrating that the eukaryotic transcription factor possess a modular structure with a DNA binding domain and transcription activation domain, they can be loosely (( )) to each other and that is sufficient for transcription activation, used to understand and identify protein–protein interactions in eukaryotic cells. So, this is one of the important benefits that actually came out of all these studies.

(Refer Slide Time: 21:12)



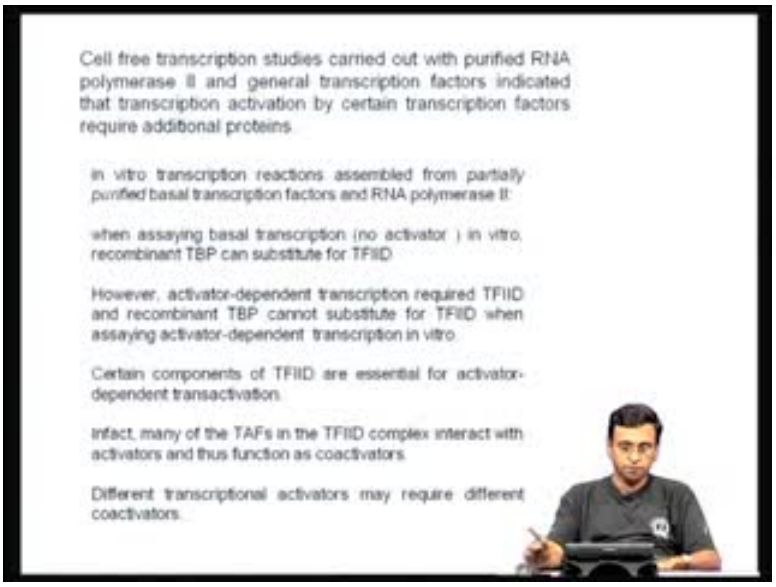
Now, as people started studying more and more about the transcription activation domain function in eukaryotic cells, it became very clear, there is nothing like a universal mechanism. Depending upon what kind of you are studying and what kind of cell type you are studying, the regulation may be different. For example, the function of a transcription activation domain could be modulated by binding of a small effector molecule. As I already told, in the case of proteins like glucocorticoid receptor, the binding of a steroid hormone actually modulates the activity of the glucocorticoid receptor. Now, the steroid molecule brings about a conformational change, dissociate the hsp90 in the (( )), results a nuclear translocation, and activation of transcription.

There are actually instances where the transcription factor or the transcription activation domain function is actually modulated by specific post-translational modifications. It could be phosphorylation, acetylation, sumoylation, ubiquitination, and so on and so forth. There are number of post-translational modifications, and when these post-translational modifications happen, the transactivation domain function either can be activated or inhibited. We will discuss some of the examples later in this course.

There are also very important regulation, where dimerization itself may play a very important role in the activity of the transcription factor. And, this dimerization can happen between two same monomers when you call the result in the formation of homodimers or you can actually form in the form of heterodimers. Example can be the

glucocorticoid receptor, for example, functions as a homodimer. So, two monomers actually come together and through the dimerization interfaces, they interact and the dimer is actually goes and binds to the glucocorticoid response element in activate transcription. But, in case of transcription factors known as c-jun, c-phos, or even last time we told NF kappa-B, which actually contains two different proteins, which we discussed in the last class, actually two monomers are not same, they are different. So, if the dimerization can happen between two monomers of the same type or two different monomers resulting in the formation of either homodimers or heterodimers, and that can also result in the activation of the transcription activation domain. So, there are number of mechanisms by which the function of a transcription activation domain can be modulated.

(Refer Slide Time: 23:24)

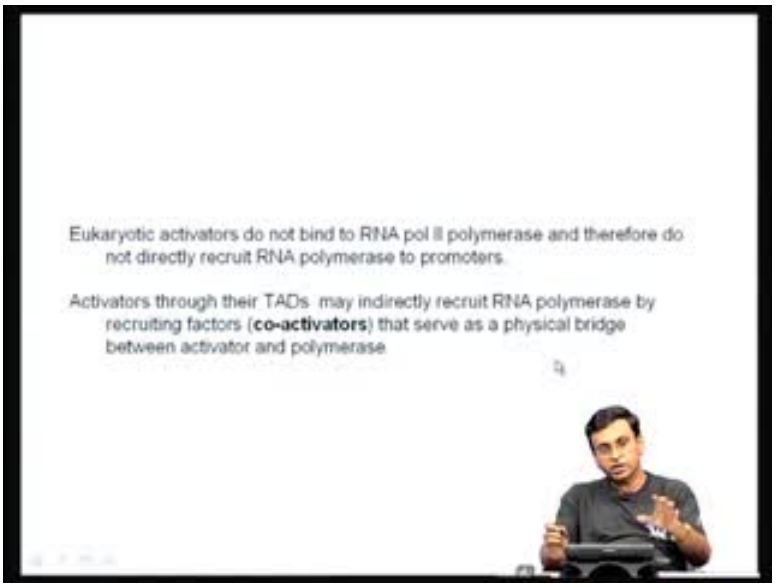


Another interesting aspect that actually emerged when people started making all these chimeric transcription factors and making recombinant transcription activation domains and started looking at the function especially in cell-free transcription systems, a very important observation that came out of these studies is that when cell-free transcription studies were carried out with purified RNA polymerase II and general transcription factors, it became very clear, certain transcription factors or certain transcription activation domains required additional proteins. That is, in addition to the RNA polymerase II and general transcription factor, certain other proteins are actually required for the efficient transcription activation by certain eukaryotic transcription factors. For

example, when people started purifying RNA polymerase II, and people also started purifying the general transcription factors, and people in fact (( )) proteins like TATA binding proteins, TF2D, and so on and so forth, a very interesting experiment was done; people started doing what is called as reconstituted in vitro transcription systems. We studied this the couple of classes ago that you can either you crude nuclear extracts or you can actually use reconstitutive systems, where if the components are either purified to almost 80 to 90 percent purity from various cell extracts; or, you can even use recombinant general transcription factors and then as a promoter activity in vitro. When they did these kinds of studies, it became very clear, when you want to assay basal transcription, you can actually substitute TFIID with TBP. Now, as I told you, TFIID is nothing but TBP and TBP-associated factors or TAFs.

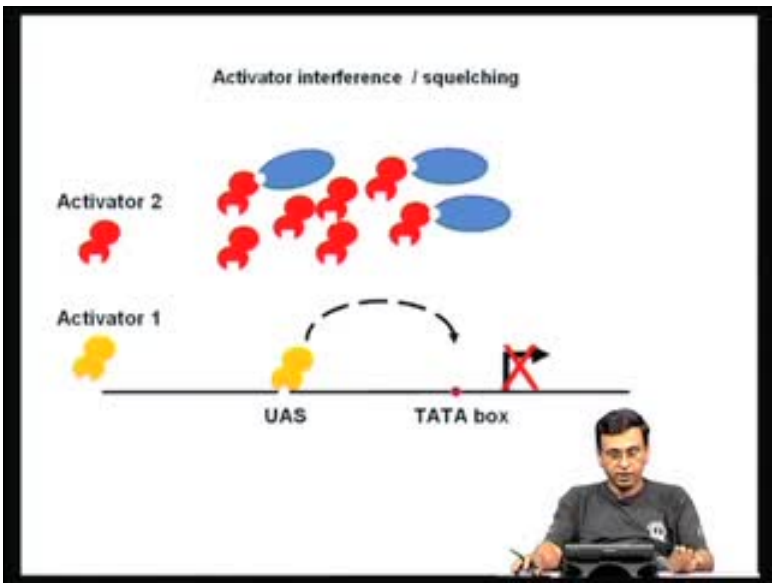
Now, if you want to just assay basal activity of a promoter, if you just have a recombinant TBP instead of TFIID, you can demonstrate basal activation. But, in such an assay, if you want to put a transcriptional activator and then ask whether the transcription activator can active transcription, the TBP could not sustain or could not promote activator-dependent transcription, which means very clearly that the TFIID, which actually contains TBP and TAFs, the TAFs have a very important role in facilitating activator-dependent transcription in cell-free transcription assay systems. So, it became very clear, the TAFs, which are the key components of TFIID in addition to TBP, may actually function as co-activators. So, this is the first thing that actually came that one of the major mechanisms by which this general transcription factor facilitate is that, many components of the general transcription factors actually interact with certain eukaryotic activators, and this interaction actually promotes in rapid recruitment of RNA polymerase to the promoter, resulting in enhancement of transcription. And, it also became very clear that different transcription activators may require different co-activators. For example, the TFIID itself consists of a number of TAFs. And, in some cases, a particular TAF may serve as a co-activator, but in other transcription activators, a different TAF subunit may serve as a transcription activator. So, the outcome of what I told so far is that to demonstrate activator-dependent transcription and cell-free transcription assays, it became essential that TBP is not sufficient, you require TF2D; clearly telling that for certain transcriptional activators, TAFs, which are components of TFIID, actually function as co-activators, and they are actually required to facilitate interaction between the RNA polymerase II and the activator.

(Refer Slide Time: 26:53)



Eukaryotic activators do not bind to RNA polymerase, and therefore, do not directly recruit RNA polymerase to the promoters. Activators through their transcription activation domains may indirectly recruit RNA polymerase to the promoters by recruiting certain factors, which I just now called as co-activators, and therefore, these co-activator serve as physical bridge between activators and RNA polymerase. So, this is one important paradigm that emerged from the mechanistic studies on transcription activation domains of a number of eukaryotic factors.

(Refer Slide Time: 27:24)



The other interesting experiment especially from the group of Mark Ptashne in late 80s and early 90s emerged is that – I told you in the last slide that different eukaryotic activators may actually require different co-activators. But, there are also many examples, where two different activators may require the same co-activator. Now, how do you demonstrate that? Some very interesting experiments were actually done and these are actually called as activator interference or squelching experiments.

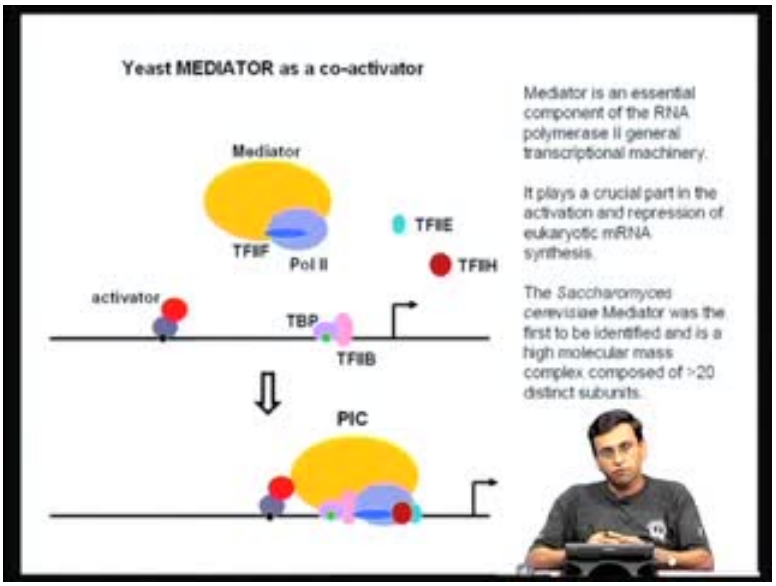
Let me just explain to you what actually you mean (( )) Let us say, for example, there is an activator 1 and activator 2. These two activators bind to 2 different sequences. So, this is a DNA binding domain; transactivation domain for activator 1; DNA binding domain and transcription activation domain of transcription transform factor 2. Let us say for example, this is GAL4 and this is GCN4 (Refer Slide Time: 28:13). They bind two different sequences.

Now, in a cell-free transcription system, if you now put purified GAL4, it will now go and bind to the GAL4 response element and activate transcription. And, let us assume that this transcription activation actually involves a specific co-activator, which I have shown in blue here (Refer Slide Time: 28:31). So, when the activator 1, in this case, for example, GAL4 binds to its sequence through the mediation of a specific co-activator, it now enhances the recruitment RNA polymerase; results in activation of transcription. I remember, activator 2 cannot bind to the response element here. So, the activator 2 cannot bind to the promoter and activate transcription, because it has no binding domain; the DNA binding domain here is different. But, if you now add excess amount of the activator 2, this activator 2 also requires the same co-activator, which is required by activator 1. And therefore, the co-activator being in limiting amounts got completed out by this excess co-activator, and as a result, the co-activator is not available for transcription activation by activator 1. This actually resulted inhibition of transcription. What actually it means is that addition of excess co-activator here actually squelched the transcription activation by another transcriptional activator.

Now, these kinds of experiments actually told you that two different co-activator proteins may actually be comparing for the same co-activator proteins. So, there are instances where each transcription factor may require a different co-activator; there are also instances where both factors may be comparing for the same co-activators. And, these

kinds of squelching experiments gave very new mechanistic insights into the co-activator requirements for very different transcription factors.

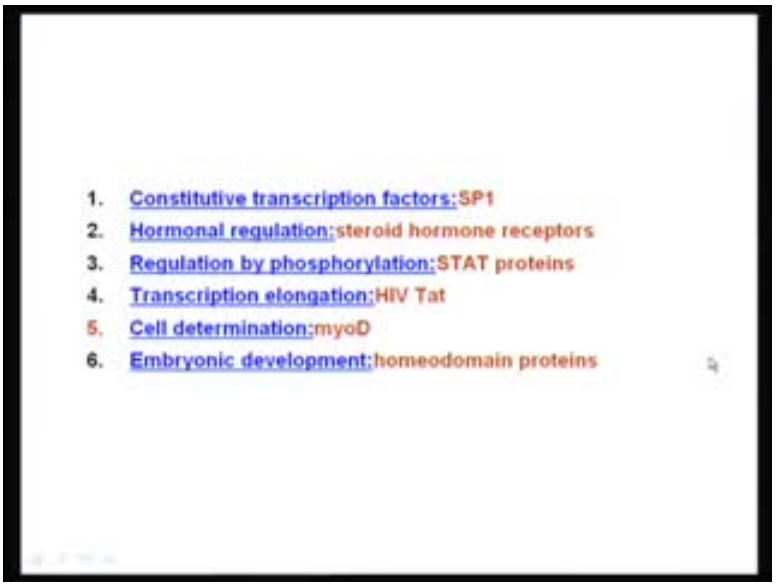
(Refer Slide Time: 29:59)



It became very clear there are a wide variety of co-activators. If we now look at the history of eukaryotic gene regulation, initially, people started looking at general transcription factors, RNA polymerase and then core promoter elements. Then, they realized that this alone is not sufficient, you require further upstream sequences either in the form of enhancers or distal promoter elements or proximal promoter elements. Then, people started looking at protein factors that are actually binding to this. A wide variety of transcription activators started discovered. Then, people looked at the mechanism of transcription activator function. People identified that they have different DNA binding domains, different transcription activation domains. And now, people have realized, in addition to all these things, there are also unique set of proteins called co-activators, which actually serve as a bridge between transcriptional activators and the basal transcription machinery. In fact, one of the first co-activators, which are actually identified in yeast cells, is a protein called mediator, which is a multi-subunit complex. So, for many yeast transcription activators, this mediator served as a co-activator and the mediator actually acted as kind of a bridge between the activators and the general transcription factors for enhancing transcriptional activation.

Mediator is an essential component of the RNA polymerase II general transcription machinery. It played a very crucial role in the activation and repression of many eukaryotic messenger RNAs in yeast cells. And, the saccharomyces cerevisiae mediator was the first co-activator to be identified and it has a very high molecular weight, and is again a multi-subunit complex composed of more than 20 distinct subunits. So, you see the complexity of eukaryotic gene activation. Already, the basal transcription itself consists about more than 50 different polypeptides. Then, I told you that you have other sequences, which play further important role in transcription activation. And, there are number of transcription factors with their own DNA binding domains and transactivation domains, which are essential for transcription activation. And now, I am telling you, there are another unique set of proteins called co-activators, which actually act as a bridge between eukaryotic activators and the basal transcription machinery. And, they again are diverse or many of them are actually complex proteins with a number of protein components.

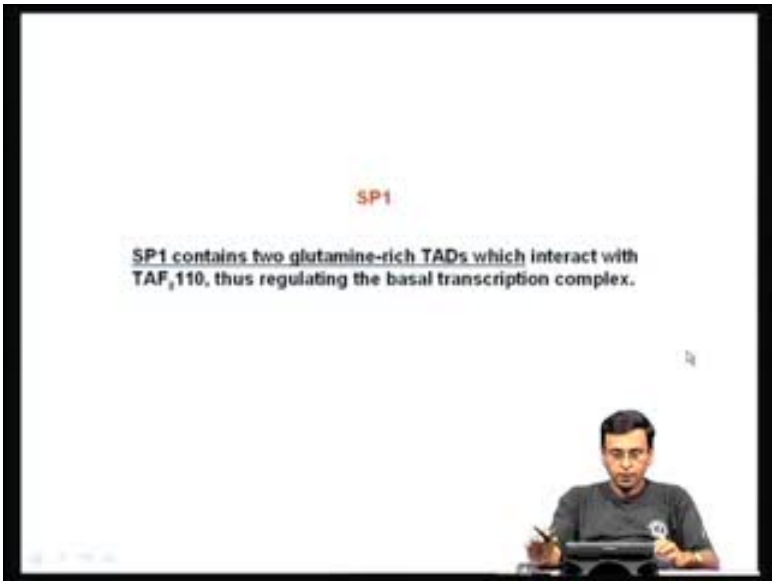
(Refer Slide Time: 32:09)



Over the years, especially in the 1980s or 1980s and even early part of 2000, people started looking at the DNA binding domains of various transcription factors; people also started looking at the transcription activation domain of various transcription factors. And, it became very clear that each transcription factor is unique. Although there are some common motifs have been identified like I told you that some of them contain acidic domains, some of them contain glutamine domains, and so on and so forth; many

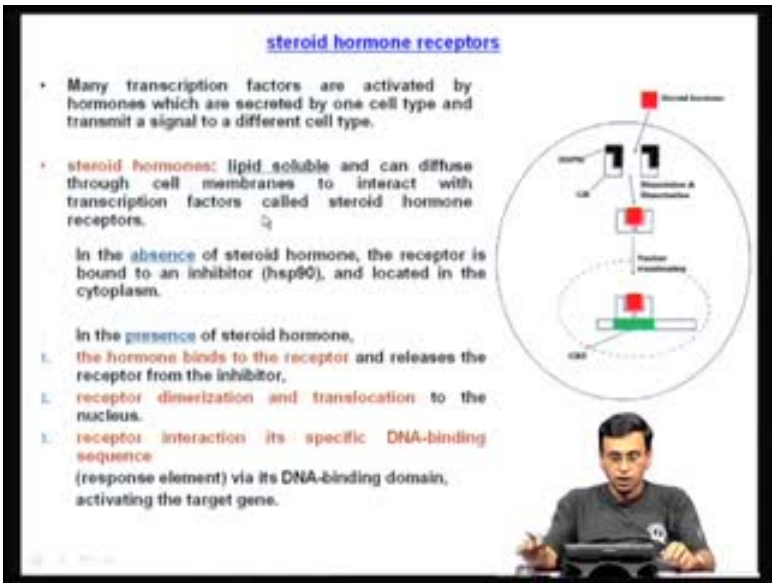
times it became very clear that as you start studying more and more transcription factors, it became more difficult to generalize. There are many of them, which are very unique. So, what I am doing is to actually just give you a very brief list of about six different transcription factors, and briefly tell you what kind of transcription activation domains these proteins actually possess and how people went on studying these things.

(Refer Slide Time: 33:02)



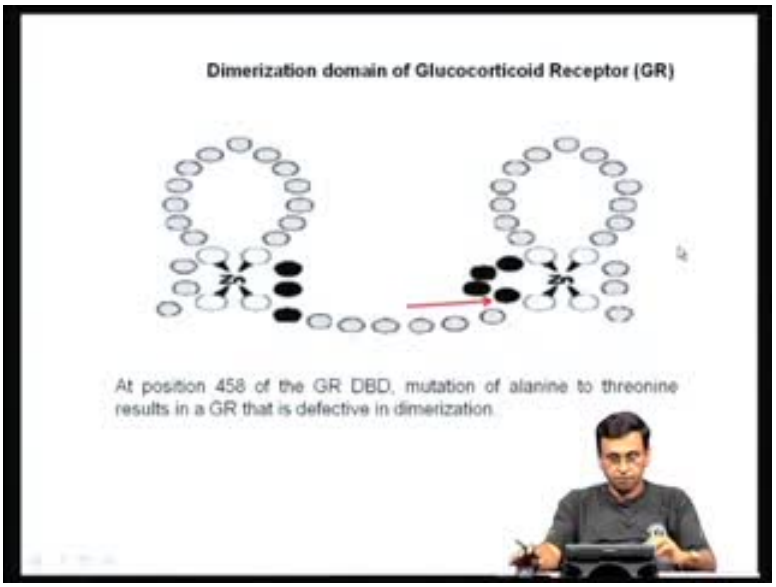
Let us take for example, SP1. Now, SP1, as I told you already, contains a glutamine-rich transcription activation domain. And, a series of experiments actually demonstrated that the transcription activation domain of SP1 actually interacts with one of the TBP versus TA factors of TF2D, called TAF to 110. A 110 kilo (( )) TAF protein of the TFD complex actually facilitates the interaction between the SP1 and a general transcription machinery. So, TAF,110 is a co-activator for SP1. This is what I identified.

(Refer Slide Time: 33:36)



If you come to steroid hormones, again the situation is very complex. Just to recapitulate, I told you already, steroids hormones are lipid soluble molecule; they can actually diffuse into the cells, but in the absence of steroid hormone, the glucocorticoid receptor is actually present in complex with heat shock protein 90, which I have shown here. The white component is the GR and the black component is the hsp90 here. And, once the steroid hormone diffuse into the cell and binds to the glucocorticoid receptor, it promotes dimerization of the two monomers of the GR and also dissociates the hsp90 from the glucocorticoid receptor. And, as a result of this, the nuclear localization signal present in the ligand binding domain is exposed, and therefore, the receptor now translocates in the nucleus, binds to what are called as glucocorticoid response elements in the promoters, and activates the transcription of the various genes. So, this is another example, where a ligand actually modulates the activity of the transcription factor.

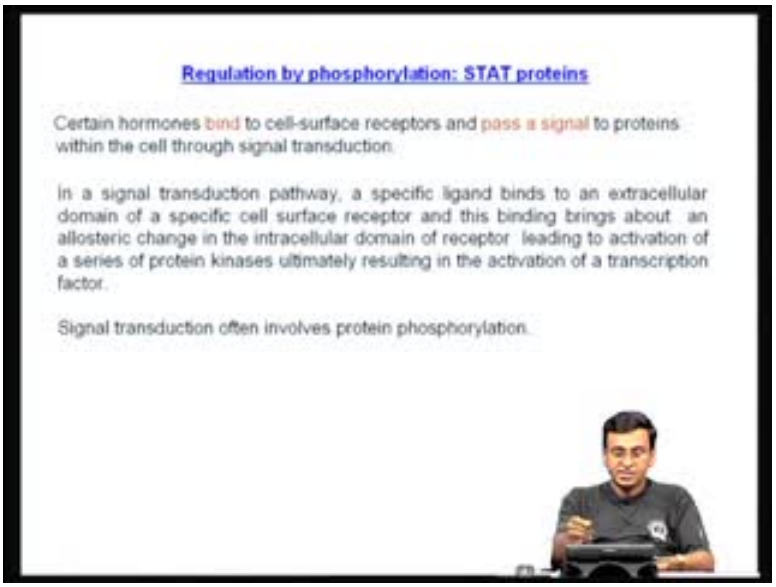
(Refer Slide Time: 34:31)



Very interestingly, in the case of glucocorticoid receptor, in addition to the DNA binding domain and transcription activation domain, a very important domain called a dimerization domain was actually identified. And in fact, this was found to be localized in the second zinc finger of the DNA binding domain. Remember, in the previous class, I told you when you were studying the DNA binding domains, the zinc finger in the glucocorticoid receptor is of C2-C2 type; it contains C2-C2 type of a zinc finger, and it can consists of two zinc fingers, which are actually involved in the DNA binding. And, I also told you, the amino acids in the base of the first zinc finger play a very important role in DNA recognition. And, by actually mutating these amino acids, you can actually change the DNA binding specificity of glucocorticoid receptor.

But, now, a amino acid, specific amino acid, especially the amino acid 458 present somewhere at the base of the second zinc finger of the glucocorticoid receptor, for example, if you mutate this amino acid, which is actually alanine, if you now mutate to threonine, the receptor can no longer dimerize. So, by simply making one point mutation in dimerizing the second zinc finger of glucocorticoid receptor, you can generate a dimerization defective glucocorticoid receptor. So, in addition to DNA binding domain, transcription activation domain, the DNA binding domain of GR also contains a very important dimerization domain. And, there are also cases where dimerization domains are actually present in the transcription activation domains of other transcription factors.

(Refer Slide Time: 35:55)



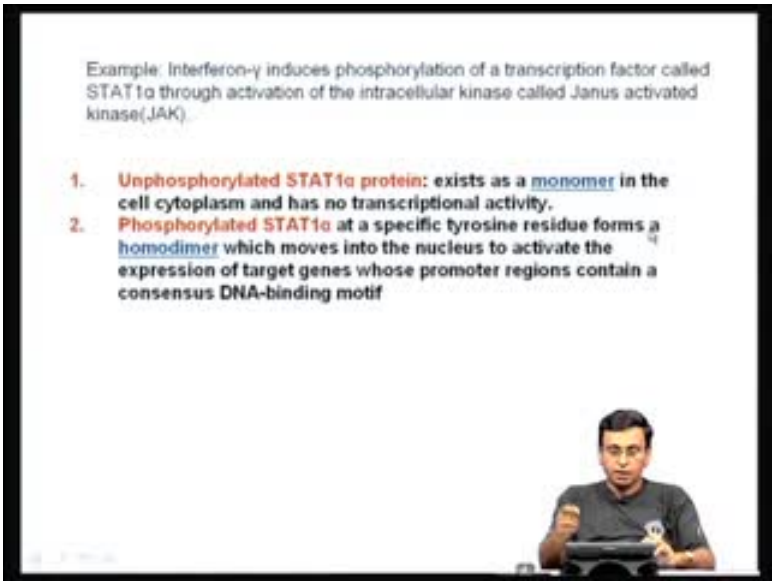
I will now give one example where the transcription activation domain function can be modulated by phosphorylation. Now, there are proteins called as STATs. These are called a signal transducers and activators of transcription. Again, I am just giving a very brief overview here; as the course progresses, at a later stage, we are going to take each one of these examples and study in detail how actually all these transcription factors function. Now, I am giving you only a very brief overview just to give you an idea of the diversity of regulation that is happening in case of eukaryotic transcription factors and the variety by which the function of transactivation domains can be regulated.

See in the previous case, the glucocorticoid hormone actually diffused into the cell and binds to the intracellular receptor. But, there are many other hormones, which actually do not enter. For example, growth factors; there are many growth factors, which are polypeptides. They cannot just like that enter into the cell; they actually interact with specific cell surface receptor. For example, if you take epidermal growth factor, it actually goes and binds to a surface receptor called as epidermal growth factor receptor. And, when these probe hormones, these polypeptide hormones bind to the cell surface receptor, it results in what is called as an initiation of a signal transduction cascade.

What do you mean by signal transduction cascade? A signal transduction pathway is nothing but when a specific ligand binds to an extracellular domain of a specific cell surface receptor, the binding brings about an allosteric change. Just like the

glucocorticoid receptor binding the glucocorticoid receptor in the (( )) brought out an allosteric change or a conformational change in the receptor, so that it can no longer bind hsp90, and therefore, it can dimerize and go into the nucleus, in the same manner, when certain polypeptide hormones bind to cell surface receptors, this ligand receptor interaction induces a conformational change in the cell surface receptor. And, this conformational change ultimate a result in the activation of a series of protein kinases and ultimately results in the phosphorylation of specific transcription factors, and they then go and bind to specific response elements, and activate transcription. So, phosphorylation is one of the key mechanisms in many cases; in many signal transduction pathways, by its specific transcription factors, can be activated.

(Refer Slide Time: 38:13)



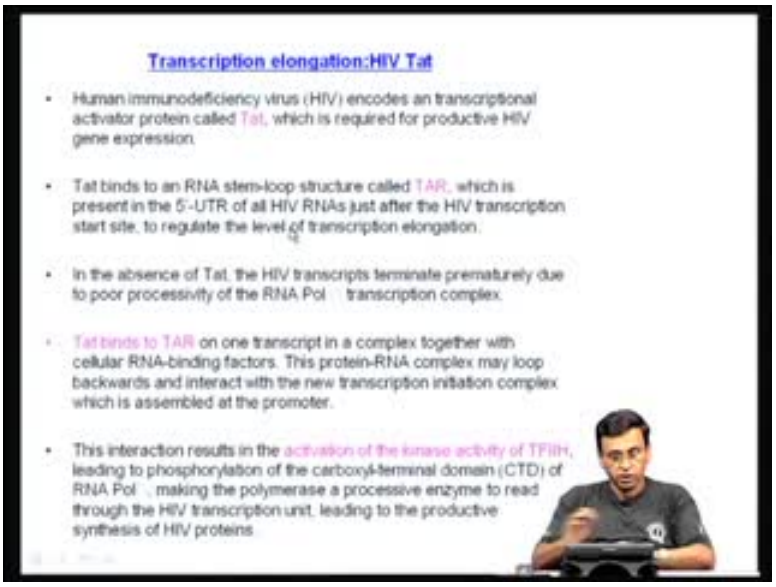
I will give you one example; a very important cytokine, which is involved in immured response, is interferon-gamma. Now, when interferon-gamma is added to the STATs, the interferon-gamma actually goes and binds to the specific cell surface receptor. And, this binding to cell surface receptor triggers the phosphorylation of a protein called STAT1. Now, the STAT1 is actually in an unphosphorylated form is present as a monomer and is actually present in the cytoplasm.

Now, once the interferon goes and binds to the cell surface receptor, it activates a protein kinase called Janus activated kinase or JAK kinase. And, this JAK kinase now goes and phosphorylates the STAT protein. And, once the STAT protein is phosphorylated

especially in the tyrosine residues – the JAK kinase is a tyrosine kinase, phosphorylates specific tyrosine residues of the STAT1 protein, and this results in the formation of homodimer. So, phosphorylation of the STAT1 protein results in the homodimerization of the STAT1. Now, this homodimer goes into the nucleus, bind to specific response elements called an interferon response element, and activates various genes, which actually manifests interferon-gamma response. So, here the function of the transcription factor is modulated by phosphorylation. And, this is actually because of a signal transduction cascade that initiates the cell surface, and ultimately, culminated in the phosphorylation of a transcription factor. So, there are two signal transduction factors.

In cases like glucocorticoid receptor, the signaling molecules actually enter the cell and then modulate the function of an intracellular transcription factor. There is another signal transduction pathway, where the molecules stay outside the cell, interact with specific cell surface receptor. And, this interaction results in initiation of a series of phosphorylation cascade; ultimately, culminating in the activation of specific transcription factor inside the nucleus.

(Refer Slide Time: 40:03)



There is another very interesting example. For example, in the case of HIV, it has a protein called as TAT protein. HIV as you know is a retrovirus and it is a very complicated retrovirus. Normally, retroviruses contains only three proteins called gag, pol and envelop. But, HIV is a complex retrovirus; there are number of other accessory

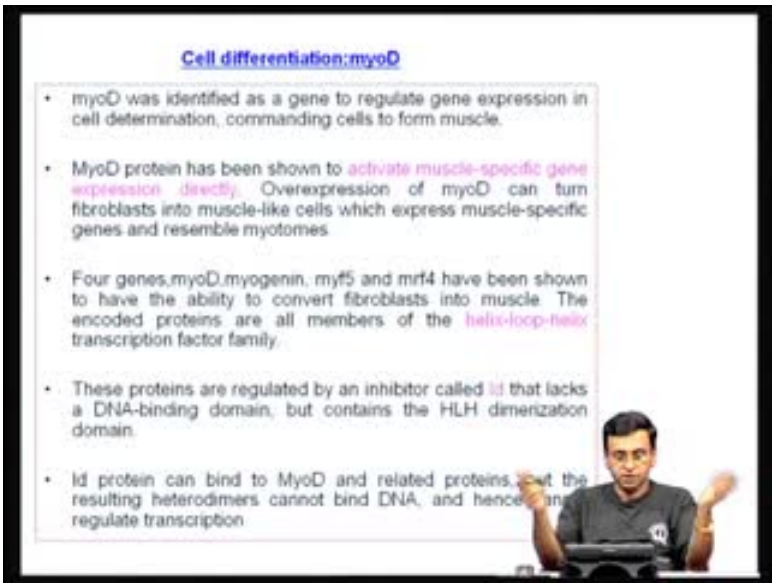
proteins in addition to these three proteins. And, one of the very important protein that HIV contains, which is relevant to gene regulation topic, which is being discussed, is a protein called Tat. Now, this Tat is actually required for productive HIV gene expression. If the HIV genome has to be expressed efficiently, you require the Tat protein.

Now, let us see what the function of Tat protein is. It tells out, the Tat protein actually binds to an RNA stem-loop structure called TAR, which is actually present in the 5 prime UTR of all HIV RNAs after the HIV transcription start site, to regulate the level of transcription elongation. So far, we have been discussing only about regulation of transcription initiation; here, there is a transcription factor, which actually binds to an RNA and affects the regulations (( )) of the transcription elongation.

It turns out when there is no Tat, the HIV transcripts terminate prematurely due to poor processivity of RNA polymerase to complex. So, the transcription is initiated, but it does not proceed further. Therefore, it gets stuck. So, the HIV RNA will not get transcribed efficiently, because elongation does not take place; initiation takes place and the RNA polymerase does not move further.

Now, in such a case when Tat is expressed, the Tat actually binds to this TAR region in the 5 prime region of this HIV RNA. And, because of this, it recruits certain cellular RNA-binding proteins to the nation-transcript DNA complex. And, as a result of this, the kinase activity of the TFIIH is activated, and this results in the phosphorylation of the c-terminal domain of the RNA polymerase II, which if you remember in the first class, phosphorylation of CTD is actually the signal for transcription elongation. It is the signal by which the RNA polymerase can now leave the initiation site and proceed further. So, here is a viral for transcription activation protein, which actually binds the specific sequence in the viral RNA, and then, recruits a number of host transcription factors; ultimately, resulting in the activation of the RNA polymerase to phosphorylation, so that the initiation complex can now leave the promoter and the viral RNA can now be transcribed. So, here you have a regulation, having (( )) at the stage of transcription elongation.

(Refer Slide Time: 42:37)

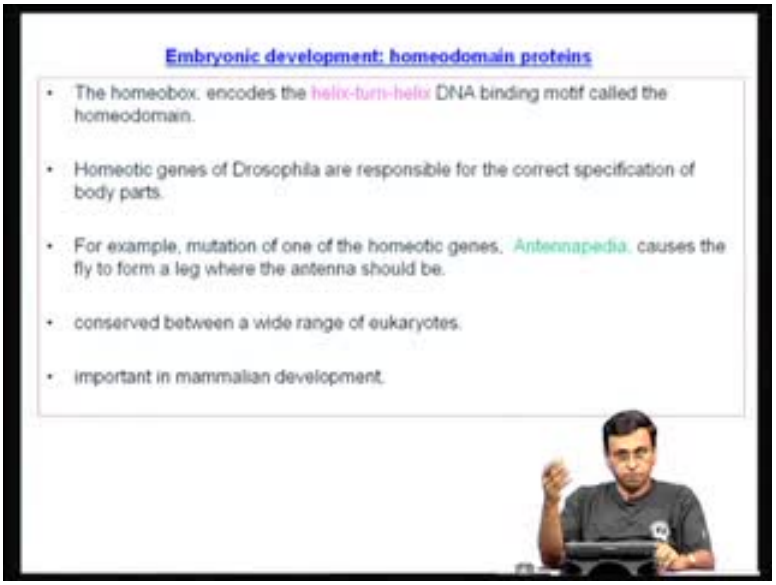


Another very interesting example I can tell you is a protein called myoD. And, if you remember my previous class, myoD actually contains a very interesting DNA binding domain called helix-loop-helix protein. The helix-loop-helix is again a very important dimerization motif, where it actually brings the two subunits together, so that it can actually go and bind to the DNA of the specific DNA sequence.

The myoD was first identified as gene that regulates the expression of cell determination, and required for muscle differentiation. If you want the fiberblasts-differentiated skeletal muscle, you actually require this particular transcription factor. In fact, myoD protein has been shown to activate muscle-specific gene expression program. So, if a particular cell type has differentiated skeletal muscle, the myoD transcription factor plays a very important role. So, if you now overexpress myoD, it can actually turn fibroblasts into muscle-like cells, which express now specific muscle gene markers. So, myoD can actually change the differentiation program of fibroblast to into that of a muscle cell. That is because myoD can go and bind to various genes, which are actually involved in the muscle differentiation. In addition to myoD, there are three other proteins called myogenin, myf5, mrf4; they have all been shown to have the ability to convert fibroblast into muscle. So, these are all transcription factors, which play a very important role in muscle differentiation program. And, they all belong to this family of helix-loop-helix transcription factor family.

Now, very importantly, there is a protein called Id, which actually lacks a DNA-binding domain, but it contains only the HLH dimerization domain. And, when this Id protein binds to myoD or related proteins, the heterodimers cannot bind DNA, and therefore, cannot regulate transcription. This is actually called as dominant negative kind of a regulation. That means here you have a protein, which has only dimerization domain, but does not have the DNA binding domain. So, the heterodimer that is formed actually is defective, and therefore, it cannot act as a transcription factor.

(Refer Slide Time: 44:37)



Other very important example is what is called as the embryonic development. There are a bunch of transcription factors, which are called as homeodomain protein. Again, I discussed in the previous class that these homeodomain proteins actually contain motifs related to the helix-turn-helix motifs, which are actually present in the prokaryotic transcription factors. So, this homeobox, which actually contains a helix-turn-helix DNA binding motif, which in this case is often referred to as the homeodomain, this transcription factor, which can be homeodomain, play a very important role during development and differentiation.

For example, in the case of Drosophila, a specific homeotic gene called antennapedia; if you now mutate or if you now delete this particular region, mutate this – usually, you have antenna in the head of the Drosophila fly; instead of antenna, you actually get legs. So, this (( )) homeo assays, that is, there are actually transcription factor, which can

change the developmental program or differentiation program such that instead of formation of an antenna, a leg is formed in the head. So, the whole organ transformation is actually controlled by a single transcription factor.

We will study little bit more about some of these things in the later classes when we start discussing about regulation of gene expression during embryonic development; very interesting aspects are come out of it. Just to mention here, that there are transcription factors like homeodomain containing proteins or homeobox genes, which actually can transform one organ to another organ if you have mutations in these transcription factors. So, I think, what I told you so far in this particular class, I just gave you a brief idea about what kind of transcription activation domains are actually present in various eukaryotic transcription factors. I told you, there are many transcription factors, which contain acidic activation domains; some of them contain proline-rich domains; some of them contain glutamine-rich domains. And, a very key observation that came out of all these strategies is that the transcription activation domain function can be separated from the DNA-binding domain function. And, in fact, there are number of experiments that have been carried out, wherein the transcription activation domains of one transcription factor were linked to the DNA-binding domain of another transcription factor, and such chimeric transcription factors have been used to understand the mechanism of both DNA-binding as well as transcription activation functions.

And, I also told you, one of the important outcomes of all these experiments have been to actually develop what is called as a yeast two hybrid assay, wherein a protein is used as a byte, and it is linked to the DNA-binding domain of one particular transcription factor. And proteins, which are likely to interact with these particular protein is linked to the transcription activation domain of other transcription factors. And, we now introduced, to co-express these two proteins together, the transcription activation domain is brought in close proximity to the DNA-binding domain; or, tethering of these two domains result in the activation of a reporter gene, and therefore, you can actually conclude that this is the protein that is actually interacting with your protein of your interest. So, the yeast two hybrid system was a very important outcome of understanding the modular nature of DNA-binding domain and transcription activation domain of various eukaryotic transcription factors.

And, also in these the examples, I gave you various examples like steroid hormone receptors, homeobox proteins, and the yeast GCN4 and GAL4, and so and so forth, just to give you an idea about the variety in the mechanism by which transcription activation domains function. And, the other important point I have told you in this class is that in order for these transcriptional activators or the transactivation domains to function or to activate transcription, in many cases, you require what are called as co-activators. So, these co-activators actually mediate and/or act as mediators between the transcriptional activators and the general transcription machinery.

Now, what we will do in the next class is to tell a little bit more about the discovery of co-activators. And, as we proceed further in the future classes… So far, I have been discussing only about DNA; now, I have been talking as if the gene expression is all taking place on a naked DNA. Remember, in eukaryotic cells, DNA is packet in the form of a chromatin. In fact, the most interesting aspect of this lecture series we will enter when we start looking at gene regulation that takes place in chromatin templates. We are now going to bring in at the stage histones into the picture. So far, we have not talked about histones at all; so far, we talked about core promoter elements; we talked about general transcription factors; then, we talked about transcription factors and upstream activation sequences; then, we discussed about the DNA binding domain and transcription activation domain of the eukaryotic transcription factors.

In this class, I also brought you the concept of co-activators, which act as mediators between the transcriptional activators and general transcription factors. But, so far, all these discussions centered around as if the DNA is a naked template, but which is not the case. In eukaryotes, DNA is compacted in the form of a chromatin and histones are actually studded all over the DNA. So, maybe about two or three classes later, we will now start discussing how actually all these promoter elements, namely the co-promoter elements, enhancer elements, the activator proteins, co-activators, all these functions have to be now take place on a chromatin template. And, that is when we are now going to bring histones and we are going to study what is called as the histone code hypothesis; how modification of histones actually play a very important role in the regulation of gene expression. So, I think we still have not entered the most interesting aspect of this lecture series.

And, as we now start studying chromatin templates, we will actually understand a very important concept; maybe how all these co-activators and transcription activators, the mechanism, which all these things actually activate transcription is actually by a modulating the function of histones. The transcriptional activators actually facilitate loose binding of histones; the transcription repressor facilitates tight binding of the histones. And, it is by modifying the histone DNA binding activity, that entire gene regulation takes place in eukaryotes. I think we will discuss some of this aspect as we proceed further in the couple of more classes. But, in the next class, we are going to discuss little bit about co-activators and co-repressors; and, how these were identified; and, what kind of examples are actually present till today in the case of eukaryotic gene regulation.

I think I will stop here.